



Quaderni del Dipartimento di Giurisprudenza
dell'Università di Torino

L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale

A cura di Roberto Cavallo Perin

Con il coordinamento editoriale di Isabella Alberti



UNIVERSITÀ DEGLI STUDI DI TORINO

QUADERNI DEL DIPARTIMENTO DI GIURISPRUDENZA
DELL'UNIVERSITÀ DI TORINO
20/2021

Comitato scientifico dei Quaderni del Dipartimento di Giurisprudenza dell'Università di Torino

Manuela Consito, Francesco Costamagna, Eugenio Dalmotto, Riccardo de Caria, Edoardo Ferrante, Domenico Francavilla (coordinatore), Valerio Gigliotti, Matteo Losana, Valeria Marcenò, Lorenza Mola, Luciano Olivero, Francesco Pallante, Margherita Salvadori, Giovanni Torrente

L'amministrazione pubblica con i *big data*:
da Torino un dibattito sull'intelligenza
artificiale

a cura di
Roberto Cavallo Perin

con il coordinamento editoriale di Isabella Alberti



UNIVERSITÀ DEGLI STUDI DI TORINO

Opera finanziata con il contributo del Dipartimento di Giurisprudenza dell'Università di Torino

La presente opera è stata sottoposta a revisione da parte di una Commissione di Lettura di docenti del Dipartimento nominata dal Comitato Scientifico della Collana in conformità al Regolamento delle pubblicazioni del Dipartimento di Giurisprudenza dell'Università di Torino.

Quaderni del Dipartimento di Giurisprudenza dell'Università di Torino

L'amministrazione pubblica con i big data: da Torino un dibattito sull'intelligenza artificiale,
a cura di Roberto Cavallo Perin

© 2021 - Università degli Studi di Torino
Via Verdi, 8 – 10124 Torino
www.collane.unito.it/oa/
openaccess@unito.it

ISBN: 9788875901806

Prima edizione: marzo 2021

Grafica, composizione e stampa: Rubbettino Editore



Quest'opera è distribuita con Licenza Creative Commons Attribuzione - Non commerciale
- Non opere derivate 4.0 Internazionale

Indice

Lunedì 20 maggio - I sessione
Algoritmi e diritto

<i>Roberto Cavallo Perin</i> Pubblica amministrazione e <i>data analysis</i>	11
<i>Carlo Tasso</i> Attori, processi, meriti e responsabilità nell'utilizzo di algoritmi di Intelligenza Artificiale: il caso del <i>Machine Learning</i>	19
<i>Ugo Pagallo</i> <i>Big data, open data e black box society</i>	49
<i>Andrea Simoncini</i> Intelligenza artificiale e futuro delle libertà costituzionali	55
<i>Agustí Cerrillo i Martínez</i> <i>Accountability</i> delle decisioni algoritmiche	61
<i>Fabiana Di Porto</i> Opacità algoritmica e trasparenza delle decisioni amministrative	69
<i>Renato Grimaldi</i> <i>Big data</i> e processi decisionali nella pubblica amministrazione: il traffico monitorato sulla piattaforma Yucca del CSI-Piemonte	73

Martedì 21 maggio - II sessione
Big data e attività conoscitiva della pubblica amministrazione

<i>Francesco Merloni</i> Data analysis e capacità conoscitive delle pubbliche amministrazioni	107
<i>Enrico Carloni</i> Qualità dei dati, big data e amministrazione pubblica	117
<i>Rosa Meo, Mirko Lai, Paolo Pasteris</i> Machine learning per la pubblica amministrazione	131
<i>Fulvio Costantino</i> Gli open data come strumento di legittimazione delle istituzioni pubbliche?	149
<i>Matteo Falcone</i> La funzione conoscitiva nella rivoluzione dei dati	183

Martedì 21 maggio - III sessione
Interoperabilità delle banche dati e funzione amministrativa

<i>Elena D'Orlando</i> Algoritmi e organizzazione dell'amministrazione locale: come declinare il principio di adeguatezza affrontando la complessità	193
<i>Alessandra Pioggia</i> Il Fascicolo sanitario elettronico: opportunità e rischi dell'interoperabilità dei dati sanitari	215
<i>Marco Aldinucci</i> L'infrastruttura necessaria per creare interoperabilità tra pubbliche amministrazioni	225

<i>Benedetto Ponti</i> L'amministrazione come fornitore e come fruitore di dati personali pubblici: sono praticabili soluzioni basate sulla <i>Big Data Analytics/Machine Learning</i> ?	233
<i>Gherardo Carullo</i> Interoperabilità dei dati e riflessi organizzativi: il caso della conservazione digitale	251
<i>Marina Caporale</i> Dalla <i>smart citizenship</i> alla cittadinanza digitale	261
<i>Isabella Alberti</i> La partecipazione procedimentale per legittimare gli algoritmi nel procedimento amministrativo	285
Postfazione	299
Bibliografia	301
Notizie sugli Autori	321

Lunedì 20 maggio - I sessione
Algoritmi e diritto

Pubblica amministrazione e *data analysis*

ABSTRACT: Il contributo intende mettere in evidenza la complessità del tema: sorvolando velocemente sulle lentezze dell'informatizzazione della pubblica amministrazione, l'articolo si sofferma sui diritti di cittadinanza digitale per dare loro un'interpretazione innovativa, di cui si afferma la capacità di fondare la pretesa del cittadino a che l'organizzazione pubblica sia anzitutto adeguata alla funzione che le è stata assegnata. La riflessione prosegue sollecitando una riflessione sulla relazione tra l'accessibilità dei dati della pubblica amministrazione (*open data*) e attività conoscitiva della stessa, considerando che in ciò trovi fondamento la pretesa dell'individuo a che si conduca un'istruttoria sufficiente e completa, che non ammette scelte compiute senza una previa consultazione dei dati in proprio possesso. Infine, il contributo si sofferma sullo studio dell'attività della pubblica amministrazione, vincolata o discrezionale, per mostrare come entrambe si prestino a essere supportate dall'intelligenza artificiale, sia nella fase istruttoria, sia in quella decisionale, fondando un nuovo modo di intendere la categoria dell'annullabilità del provvedimento amministrativo, sotto il profilo dell'eccesso di potere, della violazione di legge e dell'incompetenza.

1. *L'informatizzazione della pubblica amministrazione e i diritti di cittadinanza digitale*

Gli obiettivi dell'informatizzazione della pubblica amministrazione erano già stati individuati almeno 26 anni or sono: a) nel miglioramento dei servizi; b) nella trasparenza dell'azione amministrativa; c) nel potenziamento dei supporti conoscitivi per le decisioni pubbliche; d) nel contenimento dei costi dell'azione amministrativa (d.lgs. 12 febbraio 1993, n. 39, art. 1).

Lo sviluppo dei sistemi informativi automatizzati della pubblica amministrazione doveva rispondere ai criteri di integrazione ed interconnessione, rispettando standard definiti in armonia con le norme comunitarie e in collegamento con il sistema statistico nazionale.

La riforma non solo aveva posto i principi essenziali per una digitalizzazione dell'amministrazione pubblica, ma dato base legale all'atto amministrativo digitalizzato, offrendo con legge una copertura generale

e salvo eccezioni. “Gli atti amministrativi adottati da tutte le pubbliche amministrazioni sono di norma predisposti tramite i sistemi informativi automatizzati” (d.lgs. 39/1993, art. 3, co. 1° ora abrogato) liberando così dalla firma autografa – dunque anche dal supporto cartaceo – ogni assunzione di paternità indispensabile a definire l'imputabilità dell'atto medesimo ad una data persona fisica, fosse colui che impersonava l'organo monocratico o il presidente o segretario di un organo collegiale.

“Nell'ambito delle pubbliche amministrazioni l'immissione, la riproduzione su qualunque supporto e la trasmissione di dati, informazioni e documenti mediante sistemi informatici o telematici, nonché l'emanazione di atti amministrativi attraverso i medesimi sistemi, devono essere accompagnati dall'indicazione della fonte e del responsabile dell'immissione, riproduzione, trasmissione o emanazione. Se per la validità di tali operazioni e degli atti emessi sia prevista l'apposizione di firma autografa, la stessa è sostituita dall'indicazione a stampa, sul documento prodotto dal sistema automatizzato, del nominativo del soggetto responsabile” (d.lgs. 39/1993, art. 3, co. 2°).

Tutto il diritto sull'amministrazione digitale, sull'utilizzazione delle banche dati, pone un problema che noi amministrativisti riconduciamo alla capacità o adeguatezza (art. 118, co. 1°, Cost.) delle organizzazioni, quali elementi di professionalità, di beni, di investimenti che debbono esistere perché si possano congruamente svolgere le funzioni, dunque quale capacità l'organizzazione debba avere per potersi vedere legittimamente assegnare le funzioni (adeguatezza dell'ente), non molto diversamente dalla capacità tecnico-professionale o economico-finanziaria, che si richiede ai terzi che vogliono divenire destinatari dell'esternalizzazione di funzioni e servizi delle pubbliche amministrazioni.

Sono elementi (requisiti, forme di organizzazione, professionalità, ecc.) che definiscono la capacità di un ente pubblico – in taluni casi una pluralità di enti o anche solo un ufficio – che deve possedere l'organizzazione ancora prima di poter divenire titolare o di essere delegato all'esercizio di una funzione o servizio.

Quanto agli individui, molti indicano “i diritti di cittadinanza digitale” come diritti di quarta o quinta generazione, ma se vogliamo assumerli davvero come diritti soggettivi in senso proprio il pensiero corre ai diritti della personalità (artt. 2 e 13, s. Cost.), la cui efficacia è *erga omnes*, diritti assoluti la cui protezione è di carattere costituzionale, oggi inteso come diritto positivo, ma che taluno riafferma come diritti naturali. All'indicata qualificazione dei diritti di cittadinanza digitale come diritti assoluti della personalità non pare di ostacolo la circostanza che legislativamente

gli stessi – ed in particolare quando hanno ad oggetto i dati personali – richiedano una prestazione del titolare del trattamento, alla rettifica, alla portabilità, che solo in parte riecheggia le note classificazioni del fare, non fare, o patire, poiché obbligo di prestazione che è già noto per il diritto al nome e alla propria identità personale (art. 22, Cost.).

Ciò che colpisce è piuttosto il particolare carattere innovativo che questi diritti di cittadinanza digitale hanno, poiché pare riconosciuto all'individuo e ai gruppi organizzati un diritto potestativo alla conformazione dell'organizzazione pubblica o privata di chi è titolare del trattamento e perciò della banca dati.

Certo l'ordinamento conosce da tempo limiti di capacità speciale che superano quelli generali a tutti comuni: così l'imprenditore che debba dotare la sua organizzazione di particolari requisiti finanziari o professionali (così le società di capitali o per l'esercizio di attività bancaria, ecc.), il cui venire meno determinano la perdita di determinati benefici (personalità giuridica, di poter svolgere l'attività, per i professionisti avere conseguito la laurea e l'iscrizione all'albo professionale).

Con i diritti di cittadinanza digitale ad assumere rilievo non è solo una questione di legittimazione o decadenza dalla possibilità di svolgere lecitamente una determinata attività, ma un vero e proprio diritto soggettivo a poter agire per una determinata conformazione dell'organizzazione, non appena l'amministrazione abbia assunto a qualsiasi titolo un dato che riguardi l'individuo come parte della propria organizzazione inserendolo e gestendolo nella banca dati, con un effetto di sistema che pare di estremo interesse.

Assegnato a ciascuno il diritto alla cittadinanza digitale, con potere di ottenere la rettifica dei propri dati, o la portabilità, l'esercizio dello stesso non soddisfa solo un suo interesse individuale ma anche quello della intera collettività di riferimento, perché la conformazione dell'amministrazione vale atto di certificazione *erga omnes* – anche non privilegiata e tipizzata *ex lege* – sul dato medesimo.

Da qui un corollario sull'effettività dei diritti di libertà che, in casi come quello in esame, si risolve nell'effettività degli ordinamenti che li garantiscono e quindi - a rovescio - sul fatto che l'individuo si senta protetto dall'ordinamento, anche per ciò che concerne l'esattezza di dati anche particolari che lo riguardano, legando la sua soddisfazione individuale all'effettività dell'ordinamento che lo riconosce e protegge.

2. *L'attività conoscitiva della pubblica amministrazione*

Da tempo si è compreso che l'attività conoscitiva della pubblica amministrazione è giuridicamente rilevante, anche se un tempo era intesa più come attività prodromica all'emanazione di un atto amministrativo, all'organizzazione di un servizio pubblico, alla stipulazione di un contratto. Oggi è più evidente che essa è rilevante di per sé, anche quando non è finalizzata. L'attività di ricerca o di conoscenza e di certificazione non sono attività necessariamente finalizzate: ne è un esempio la ricerca scientifica delle Università pubbliche o l'attività svolta dall'ISTAT che è al servizio di tutti, non è invece servente ad una singola funzione o servizio pubblico.

Oggi diventa rilevante la questione dell'*open data*, del superamento del tradizionale modo di pensare l'accesso alle informazioni della pubblica amministrazione perché c'è una rilevanza in sé del dato.

La disciplina sull'accesso della l. n. 241 del 90 è stata intesa quasi come servente l'attività d'esercizio di funzioni e servizi, di recente sulla stessa esecuzione di contratti, nonostante l'accesso sia stato distinto tra possibili partecipi e terzi (l. 241 del 1990, art. 10 e art. 22 e ss.). In tal senso lo sforzo di configurare un diritto della personalità, di ciascuno, ad avere informazioni in ragione di una decisione altrui, significa rievocare alcune disposizioni della Costituzione (artt. 2 e 13) ma al tempo stesso l'art. 41 Carta dei diritti fondamentali dell'Unione europea sul diritto ad una buona amministrazione.

Un diritto di accesso generalizzato credo invece che abbia un fondamento completamente diverso ove la Costituzione prevede che la scienza è libera e libero ne è l'insegnamento (art. 33, co. 1°, Cost.). Se la scienza ha da essere libera, deve essere libero anzitutto qualsiasi accesso ai dati, ai fatti di conoscenza, che danno qualità ad ogni successiva valutazione e innovazione delle persone individue o delle organizzazioni.

La scienza libera implica non solo che sia libero colui che la produce, ma anche colui che la riceve, perché colui che la riceve può pretendere da colui che la esercita che sia libero nello svolgerla.

Se per definizione costituzionale la scienza ha da essere libera, non solo per chi la produce, ma anche per chi la riceve, vi è un diritto soggettivo pieno a poter vantare che la scienza sia libera e per ciò solo accessibile.

Un diritto di azione di chiunque (artt. 24 e 33, Cost.) nei confronti di qualsiasi banca dati, volto non solo a sapere ciò che è in possesso dell'amministrazione, o per sapere quali scelte la stessa intende assumere, ma un diritto di azione a sapere dall'amministrazione pubblica ciò che è nella sua disponibilità, al fine di poter assumere le scelte che io stesso

voglio compiere (dove e come investire, dove acquistare una casa e dove risiedere, dove istruirmi, dove curarmi, dove passare la mia vecchiaia).

L'azione d'impugnazione contro gli atti amministrativi nei confronti della pubblica amministrazione consente sin da fine ottocento di dedurre il difetto o l'insufficiente istruttoria che precede l'emanazione degli atti stessi che ora – per quanto sopra detto – può avvenire anche con riferimento all'ottenimento dei dati finalizzati alle scelte individuali di ciascuna persona. Più tradizionale è invece affermare che l'Università che decida su suoi percorsi di studi, senza avere consultato le banche dati, sui flussi di studenti degli ultimi anni mi pare determini di per sé un vizio di eccesso di potere per insufficiente istruttoria nella deliberazione d'attivazione dei corsi medesimi. Del tutto illogico è quell'amministrare che si permetta di assumere qualsivoglia decisione senza avere conosciuto ed analizzato (*data analysis*) le banche dati che sono in suo possesso.

Da ciò la necessità di un aggiornamento delle banche dati della pubblica amministrazione; banca dei dati che deve anche essere consultabile, proprio perché si ritiene che la stessa sia al servizio delle scelte pubbliche e dei terzi al fine di operare le loro scelte.

Un diritto individuale ai dati e alla valutazione che degli stessi le amministrazioni hanno manifestato a compimento delle proprie finalità istituzionali; ciò vale anzitutto per le università e gli enti di ricerca, ove il diritto alla conoscenza dei dati assume un particolare significato e che si ritrova come valore essenziale protetto in Costituzione. Quando s'afferma che la scienza è libera si enuncia un precetto forte e non equivocabile: il diritto dei professionisti a fare scienza ma anche di coloro che la ricevono e che su di essa fondano le proprie scelte.

Fallite le riforme fatte degli anni Novanta di cambiare le amministrazioni in astratto è venuto il tempo di riformare l'amministrazione in concreto, a partire dal concreto amministrare, di cui occorre capire le più intime logiche istituzionali, che le amministrazioni tramandano in sé medesime reinterprestando la tradizione.

Non interessa solo il dato in sé, ma il dato istituzionale dell'amministrazione, come rivelato nelle decisioni passate, che il *data analysis* può rivelare, per tradizioni, per etica, o tipo d'organizzazione che appartiene ad un certo tipo di amministrazione (onore: art. 54, co. 2°, Cost.), e che si rileva nelle sue decisioni, nei suoi comportamenti, come esito di quell'istituzione.

3. *Intelligenza artificiale tra attività discrezionale e attività vincolata*

Nell'attività vincolata le variabili degli algoritmi – che identificano classi di input – sono rappresentate dall'interpretazione delle disposizioni (costituzionali, di legge, di regolamento, ecc.) e dai criteri che in quelle disposizioni sono indicati: ad es. l'età, la residenza, il titolo di studio, le esperienze lavorative, che definiscono l'emanazione del provvedimento amministrativo in senso sia procedimentale sia sostanziale.

Nell'attività discrezionale invece le variabili sono individuate come classi di input ricavate dall'analisi delle casistiche pregresse affinché l'algoritmo possa assumere una decisione.

Gli algoritmi possono essere utilizzati sia nella fase istruttoria, sia nella fase integrativa dell'efficacia, in modo non molto diverso da come si pensa che possa avvenire per l'attività vincolata. Nella fase decisionale l'apporto algoritmico si può sostanziare in una proposta del contenuto discrezionale del provvedimento elaborata sulla base della combinazione dei precedenti applicativi e dei presupposti che ricorrono in concreto, combinando il contenuto discrezionale del provvedimento che l'amministrazione intende adottare e il fine di interesse pubblico concreto che l'amministrazione intende soddisfare con quella misura.

Definiti i precedenti che si intende assumere come rilevanti anche per le decisioni future (*data lake*) l'apporto dell'algoritmo è compendiato dalla verifica dell'osservanza del principio di proporzionalità: tra presupposti in concreto e contenuto discrezionale che intende assumere l'amministrazione, ivi compresi i limiti che il giudice ha dichiarato nel caso d'impugnazione del provvedimento.

In entrambe la precisione del modello è influenzata dalla quantità e dalla varietà della casistica messa a disposizione dello stesso, ossia in base alla rilevanza e alla rappresentatività della banca dati e del *data lake* di riferimento, più che a seguito di un giudizio di validità sul provvedimento in sé considerato. Con l'algoritmo assume rilevanza la validità sistemica dell'atto amministrativo che perciò è detto digitale. L'algoritmo a fini decisionali deve quindi poggiare sulla scelta accurata delle variabili che derivano dalla casistica settoriale di riferimento.

Ogni nuovo caso giurisprudenziale è dall'algoritmo classificato – secondo un'approssimazione probabilistica – in una delle qualificazioni normative ed è in ragione di tale calcolo probabilistico da cui scaturisce il provvedimento digitale.

L'interpretazione per la violazione di legge è l'operazione intellettuale indispensabile per passare da una norma generale ed astratta al caso

concreto, poiché la prima rappresenta pur sempre una sintesi della molteplicità dei casi, sicché la definizione delle variabili può non aver tenuto conto di casi meno frequenti o lo stesso enunciato normativo può, in tempi differenti, imporre di ridefinire il peso tra le varie classi.

Il controllo giurisdizionale si svolge dunque come sindacato sull'esito ottenuto dall'algoritmo, in ragione delle variabili che l'amministrazione ha selezionato, fermo restando che l'attività di selezione e di inserimento delle stesse può essere letterale, come nel caso di numeri indicati direttamente dalla norma, oppure un'interpretazione più complessa che richiede di valutare le variabili e i pesi ad esse assegnati.

Del pari è il ragionamento per l'incompetenza, che non configura una diversa formalizzazione algoritmica, trattandosi pur sempre di una speciale violazione di legge.

Quanto al vizio di eccesso di potere – al di là dell'ipotesi infrequente dello sviamento – lo stesso s'afferma normalmente pel tramite delle figure sintomatiche che limitano le scelte dell'amministrazione pubblica che perciò debbono essere in sé logiche e comprensibili.

Il controllo giurisdizionale si svolge dunque come sindacato sull'esito ottenuto dall'algoritmo, in ragione delle figure sintomatiche che sono tratte dalla casistica in materia (di concessione di servizi pubblici, d'incarichi dirigenziali, di procedure concorsuali, ecc.), che l'algoritmo del giudice ha selezionato dalla sua più che centenaria giurisprudenza.

Diversi sono necessariamente gli algoritmi dell'amministrazione, da quelli della giurisdizione: entrambi si fondano su una casistica o precedenti rilevanti, entrambi non negano la mediazione o il controllo umano a seguito di osservazioni o proposte di partecipazione al procedimento amministrativo o in contraddittorio nel processo depositando anche in tale sede documenti e memorie.

Certo l'algoritmo del giudice sull'attività vincolata avrà l'ultima parola, sicché il contraddittorio serve a promuovere soluzioni nuove volte a derogare motivando, o a modificare l'algoritmo in tutto o in parte perché riconosciuto ingiusto o divenuto tale. La nostra conoscenza oggi è ancora superficiale. La conoscenza per il tramite dell'algoritmo della casistica rilevante ci impone di vedere le soluzioni sistemiche che si sono imposte in dati periodi storici, che possiamo confermare o ritenere non più tollerabili.

Machine learning e reti neurali consentono di fare emergere il contesto, quindi la *ratio decidendi* con cui la decisione è stata presa in migliaia e migliaia di casi, che la macchina instancabilmente e imparzialmente riporta alla luce, mettendo in evidenza la ricchezza

dei casi e delle soluzioni, i caratteri identitari ma anche i pregiudizi culturali, gli errori o il malcostume, con una ricchezza di elementi che ci lascerà stupefatti.

Attori, processi, meriti e responsabilità nell'utilizzo di algoritmi di Intelligenza Artificiale: il caso del *Machine Learning*

ABSTRACT: Il presente lavoro illustra in generale lo scenario di sviluppo e di successivo utilizzo dei sistemi software di Intelligenza Artificiale, mettendo in evidenza le innovazioni e le differenze con lo scenario ormai più consolidato relativo allo sviluppo di soluzioni informatiche più tradizionali. In particolare, ci si sofferma sui sistemi più moderni, basati su tecniche di Machine Learning, individuando le varie competenze informatiche necessarie ed i ruoli e meriti relativi, mettendo altresì in risalto le varie responsabilità nella fase di sviluppo e di utilizzo. Il lavoro si conclude con una riflessione generale sulle prospettive dell'Intelligenza Artificiale.

1. *Introduzione*

Lo scenario 'tradizionale' di utilizzo dell'ICT (e del software in particolare) prevede che, relativamente alle specifiche procedure che si desidera informatizzare, venga progettato e sviluppato un algoritmo¹ che permette l'esecuzione automatica, precisa e deterministica di tutti i passaggi e di tutte le elaborazioni necessarie per portare a termine le procedure di interesse. In questo lavoro, con il termine *informatica tradizionale*, ci si riferisce ai sistemi software sviluppati in accordo con tale visione e che si presentano perfettamente adeguati ad automatizzare problematiche più o meno complesse di archiviazione, calcolo e trasferimento di dati, caratterizzate da procedure ben definite, secondo ben precisi modelli (spesso rappresentati con un linguaggio

1. Un *algoritmo* è un procedimento che risolve un dato problema *P* attraverso un numero finito di istruzioni elementari, chiare e non ambigue. L'algoritmo è alla base della *programmazione*, ossia il processo che gli esperti informatici (programmatore) eseguono per tradurre le operazioni specificate nell'algoritmo in una loro rappresentazione mediante un *linguaggio di programmazione*, ossia un linguaggio formale che può essere eseguito da un computer al fine di risolvere specifiche istanze del problema originale *P* che si desidera automatizzare.

matematico) sviluppati nell'ambito delle relative discipline tecnico/scientifiche e, in taluni casi, addirittura definite 'per legge', dovendo essere rispettate specifiche norme.

Con lo sviluppo dell'*Intelligenza Artificiale (IA)*² questo scenario cambia radicalmente: l'oggetto dell'IA, infatti, è un qualcosa di molto meno definito e conosciuto, ossia i processi cognitivi tipici della mente e dell'intelligenza umana³, tuttora scarsamente conosciuti dal punto di vista scientifico (in particolare dalle neuroscienze) e certamente non rappresentabili mediante modelli precisi⁴. Questo cambiamento è ancora più radicale nel caso di una delle tecniche che al giorno d'oggi è costantemente richiamata nella stampa e nei media, ossia il Machine Learning (ML), che ha come scopo apprendere automaticamente strategie risolutive e previsionali in base all'esperienza. Ne consegue quindi che nell'IA, e nel ML in particolare, si sfuma l'aspetto deterministico che caratterizza le applicazioni ICT tradizionali, ossia basate su algoritmi ben definiti che provengono dagli studi svolti nei settori applicativi considerati e che definiscono in modo deterministico i passi di elaborazione da compiere.

Ciò accade perché modellare processi cognitivi complessi non può essere fatto, allo stato attuale delle conoscenze scientifiche, in modo altrettanto deterministico: non sempre sono chiari i passi da compiere, le procedure risolutive spesso risultano incomplete (non coprono tutta la casistica), non sono noti tutti gli elementi cognitivi rilevanti da includere in un modello. Ne consegue che l'accuratezza dei modelli predittivi sviluppati mediante algoritmi di ML difficilmente può raggiungere (se non mai o solo in casi banali) il livello del 100% e si deve sempre considerare la possibile presenza di errori. In sostanza, viene meno quella tipica caratteristica di non discrezionalità e non indeterminatezza dell'informatica tradizionale in cui ogni elaborazione è precisamente e completamente definita e determinata.

Inoltre, per la realizzazione di applicazioni di IA e ML, lo scenario tradizionale di utilizzo e dello sviluppo di applicazioni ICT

2. S. RUSSEL - P. NORVIG, *Artificial Intelligence: A Modern Approach*, 3rd Ed., Pearson 2010.

3. Ci si riferisce qui agli obiettivi che vennero attribuiti all'IA quando tale disciplina venne inizialmente proposta. Nel seguito di questo lavoro (Sez. 2) è chiarito come successivamente alle sue origini, le sotto-aree e le applicazioni dell'IA si sono estese anche a aree 'meno cognitive'.

4. Precisione intesa nel senso che il modello rappresenta in modo adeguato tutti gli elementi rilevanti del processo cognitivo considerato.

richiamato precedentemente si presenta inadeguato poiché gli attori ed i processi coinvolti per l'informatizzazione sono molteplici, più strutturati ed articolati.

Lo scopo di questo lavoro è quindi delineare, almeno in sintesi e con un linguaggio accessibile, gli elementi principali dello scenario IA, con particolare riferimento agli algoritmi di ML, soffermandosi su:

- a) processi svolti nelle varie fasi, partendo dall'invenzione dell'algoritmo, allo sviluppo di modelli previsionali ed al loro successivo sfruttamento da parte degli utenti finali;
- b) attori coinvolti nelle varie fasi e loro ruolo in termini di chi fa cosa, cosa ottiene, di cosa è responsabile;
- c) rischi, ossia cosa può andare male e perché.

L'obiettivo finale è quindi fornire una base di lavoro corretta per analizzare se, quando e con quali precisi obiettivi utilizzare i sistemi di IA e ML.

2. *L'Intelligenza Artificiale*

Nonostante solo da qualche anno si senta parlare molto di Intelligenza Artificiale, le origini di questa disciplina risalgono alla metà degli anni '50. Fu in una scuola estiva presso il Dartmouth College nell'estate del 1956 che John McCarthy introdusse per la prima volta il termine *Artificial Intelligence (AI)*⁵. L'intento della scuola era quello di capire se i processi cognitivi ed altri aspetti dell'intelligenza umana potessero essere simulati dal computer, superando la visione di una macchina che era in grado solo di rappresentare ed elaborare dati numerici. Da allora l'IA si sviluppò ed attraversò fasi alterne di successi e delusioni (dovute soprattutto ad aspettative eccessive), fino ad arrivare al grande interesse e fervore odierni. Alla focalizzazione iniziale centrata sui processi di tipo cognitivo, si affiancarono nel corso degli anni successivi numerosi specifici filoni di interesse: la robotica, la visione artificiale, la percezione, la manipolazione, il riconoscimento di immagini, i metodi di ragionamento automatico, l'apprendimento automatico, per citarne solo alcuni. Le tecniche utilizzate furono mutate da moltissime discipline, tra cui, in primis,

5. J. MOOR, *The Dartmouth College Artificial Intelligence Conference: The Next Fifty years*, in *AI Magazine* 27(4), 2006, 87 ss.

l'informatica, ma anche la psicologia sperimentale, la matematica e la statistica, la linguistica computazionale ed altre ancora. E parallelamente si moltiplicarono a dismisura le aree applicative, che ormai oggi abbracciano un po' tutte le attività umane.

Per rendere più chiare le argomentazioni proposte in questo lavoro, restringiamo il significato del termine IA alla sua accezione originale: una disciplina dell'Informatica che studia i metodi e le tecniche per progettare e costruire sistemi software che forniscano prestazioni che solitamente vengono considerate tipiche ed esclusive dell'intelligenza umana. Secondo questa visione, le tematiche che furono affrontate all'inizio dell'evoluzione dell'IA furono l'elaborazione del linguaggio naturale (per comprendere il significato di un testo o per tradurre da una lingua all'altra), il problem solving (che modella strategie generali di ragionamento utilizzate dalla mente umana), la dimostrazione automatica di teoremi, la pianificazione, la diagnosi e la progettazione. Dopo un primo periodo di grande interesse, gli scarsi risultati ottenuti portarono (intorno al 1970) ad una fase di ripensamento, che in sostanza portò da un lato ad affrontare problematiche meno vaste e generali, focalizzando singoli problemi più specifici e dall'altro lato incentivò una ricerca più fondamentale, incentrata sulla conoscenza e sui processi cognitivi che l'uomo utilizza per ragionare sui problemi e risolverli e sulla loro rappresentazione mediante il computer.

I frutti non tardarono e alla fine degli anni '70 si iniziarono a costruire con successo i primi sistemi intelligenti, chiamati anche *Sistemi Esperti* o *Sistemi Basati sulla Conoscenza* (*Knowledge-Based Systems - KBS*)⁶, sistemi in grado di rappresentare la conoscenza umana in un determinato dominio (solitamente ristretto), al fine poi di utilizzarla per ragionare su uno specifico problema e trovarne la soluzione: si tratta di sistemi che automatizzano l'esecuzione dei cosiddetti processi *knowledge-intensive*, dove ciò che conta sono le capacità di ragionamento dell'esperto umano, che non si rifà ad alcun modello preciso (né tantomeno matematico), ma utilizza soprattutto la propria esperienza ed i criteri che in anni di lavoro ha sintetizzato (anche inconsciamente!) per risolvere quella specifica classe di problemi. I KBS mirano a simulare queste strategie di risoluzione degli esperti umani e da ciò consegue anche che la metodologia utilizzabile

6. G. GUIDA - C. TASSO, *Design and Development of Knowledge-Based Systems: from Life Cycle to Development Methodology*, John Wiley and Sons, Chichester, UK 1994.

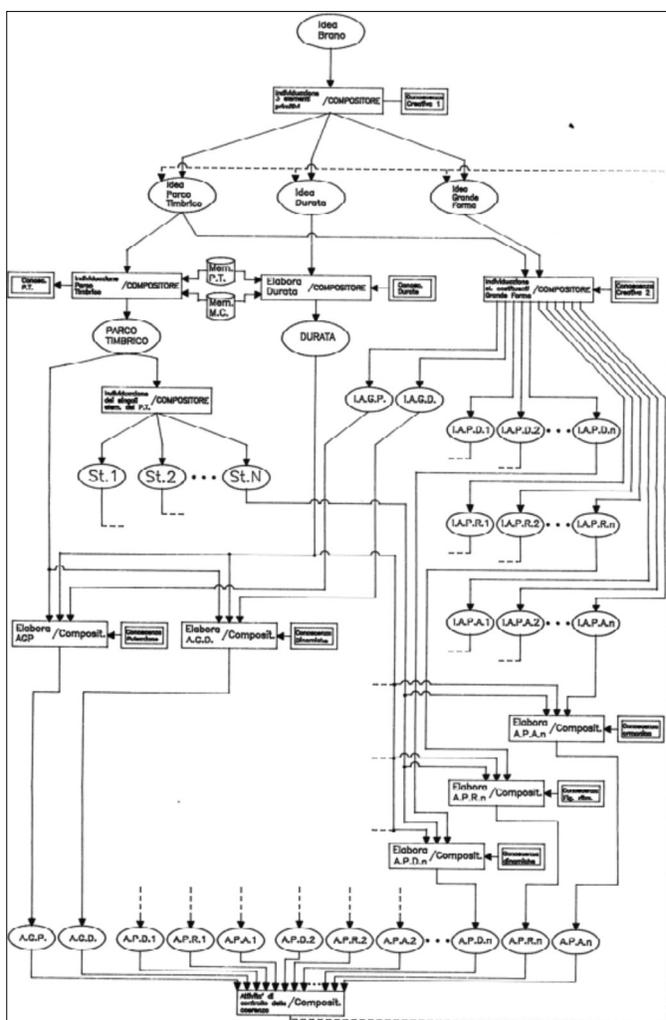
per lo sviluppo di tali sistemi KBS è diversa da quella utilizzata per i sistemi informatici tradizionali, come è illustrato più in dettaglio nella prossima sezione.

2.1 *L'approccio dei Sistemi Basati sulla Conoscenza*

Per progettare un KBS si deve procedere prima di tutto cercando di capire come ragiona l'esperto umano per la soluzione dei problemi in un determinato dominio. Più in particolare, si costruisce un *Modello Concettuale del Dominio*, una rappresentazione che descrive tutti i passi del processo (cognitivo) di ragionamento che l'esperto umano segue per risolvere il problema considerato. In Figura 1 è rappresentato il Modello Concettuale della composizione musicale di uno specifico compositore di musica contemporanea⁷.

7. Si tratta in particolare del compositore contemporaneo Nicola Cisternino. Il modello è stato realizzato con l'aiuto del M^o Marco Maria Tosolini ed è documentato in: L. PERESSON, *Sviluppo di un Modello Concettuale nella Composizione di Brani Musicali Contemporanei con Tecniche di Acquisizione e di Rappresentazione della Conoscenza*, Tesi di Laurea A.A. 1992-93, Università di Udine.

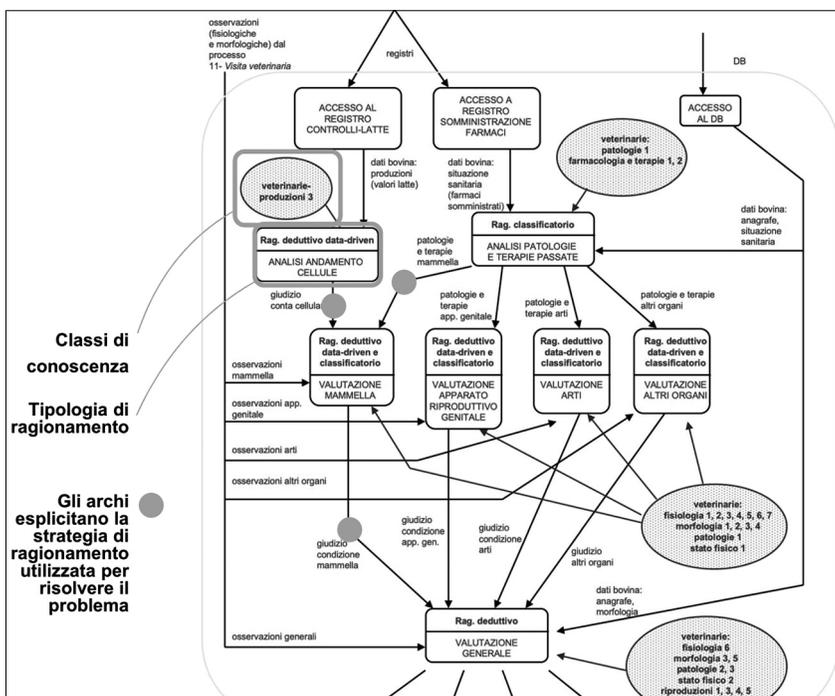
Figura 1 - Rappresentazione parziale del Modello Concettuale relativo al Processo di Composizione Musicale del compositore Nicola Cisternino



Ciascun passo del ragionamento viene analizzato e capito in dettaglio e viene descritto individuando le varie entità coinvolte nel ragionamento, i tipi di ragionamento, la strategia generale. In Figura 2 è rappresentato un altro modello concettuale, relativo in questo caso al processo di valutazione veterinaria realizzato periodicamente dal veterinario di un'azienda

zootecnica⁸. Il ragionamento, in questo caso, ha lo scopo di classificare lo stato di salute di una bovina, al fine di individuare la scelta migliore da fare tra un numero finito di possibilità: ad esempio, per una bovina sana, (i) continuare a produrre latte ovvero (ii) fecondarla. Anche in questo caso l'analisi è spinta al massimo livello, e per ciascun passo di ragionamento sono indicati i dati su cui esso opera, le classi di conoscenza utilizzate, il tipo di ragionamento svolto ed il risultato ottenuto ad ogni passo. Tutti i ragionamenti sono poi connessi ed inseriti nella struttura generale del processo (specificando mediante archi da dove si parte, che passi si seguono via via, ecc.), che rappresenta la strategia di ragionamento affinata in anni di operatività dall'esperto veterinario.

Figura 2 - Parte del Modello Concettuale del processo di ragionamento di un Veterinario



8. S. CHINELLATO, *Sperimentazione di una Metodologia per lo Sviluppo di Sistemi Basati su Conoscenza - Un'Applicazione in Campo Zootecnico*, Tesi di Laurea A.A. 1994-95, Università di Udine.

Una volta costruito il Modello Concettuale del Domino, si passa alle fasi più tecniche di progettazione esecutiva e successiva implementazione, ossia si traducono i vari passaggi specificati nel modello nei moduli software la cui esecuzione fornirà la capacità (automatica) di risolvere lo specifico problema affrontato. Nell'esempio veterinario di cui in Figura 2, si forniscono in input al sistema tutti i dati relativi ad una bovina, ed il sistema, ragionando su tali dati ed utilizzando le specifiche conoscenze e gli specifici passi di ragionamento identificati nel Modello Concettuale, produce in output la decisione ritenuta migliore per quello specifico caso.

È importante però notare che la parte più difficile ed importante dell'intero processo di sviluppo è senz'altro la prima, che esplora le capacità cognitive dell'esperto umano. Si tratta di un lavoro spesso molto lungo ed impegnativo, che utilizza dapprima le cosiddette *tecniche di acquisizione della conoscenza* (in parte mutate dalla psicologia sperimentale) per *elicitar*e conoscenza dall'esperto umano e poi le *tecniche di rappresentazione della conoscenza e ragionamento*, per caratterizzare come l'esperto svolge i singoli passi di ragionamento mediante le varie classi di conoscenza. I due Modelli Concettuali rappresentati nelle precedenti Figura 1 e 2 hanno richiesto molti mesi di incontri con gli esperti dei relativi domini, e ciò principalmente perché, solitamente, per il loro lavoro, essi non sono abituati a spiegare come ragionano (che è ciò che di fatto interessa agli informatici per realizzare il sistema software intelligente), ma sono invece molto abili a trovare le soluzioni, a valutare ed eliminare ipotesi, a focalizzarsi sugli aspetti rilevanti, ecc. Inoltre, va messo in evidenza che avere anni di esperienza, da parte degli esperti umani, ha due conseguenze per certi versi divergenti:

- a) certamente ripetere il processo di soluzione del problema affrontando una vastissima casistica nel corso degli anni porta l'esperto ad essere sempre più abile ad individuare la miglior soluzione del problema;
- b) d'altro canto, però, la ripetizione nel tempo porta anche a fare emergere nel ragionamento dell'esperto dei meccanismi di automatismo, diventando sempre più veloce nell'impostare e risolvere il problema e più capace nel trovare la soluzione. E tali automatismi sono spesso acquisiti ed utilizzati inconsciamente.

Proprio a causa di quest'ultimo aspetto, la conoscenza, col tempo, diventa *tacita*⁹ ed è utilizzata in modo automatico e quindi diventa assai

9. Una conoscenza è tacita quando fa parte del bagaglio conoscitivo dell'esperto, ma viene utilizzata in modo inconscio nei ragionamenti, quasi automatico. Tipicamente è il

lungo e complesso per gli informatici acquisire dagli esperti umani tutti gli elementi che servono per costruire il Modello Concettuale. È certo tuttavia che se si riesce nell'impresa di costruzione del Modello Concettuale, il sistema KBS che verrà poi realizzato sarà in grado di fornire delle prestazioni 'intelligenti'. Il fatto poi che il Modello Concettuale sia un'esplicita e fedele rappresentazione delle strategie utilizzate dall'esperto farà sì che il processo di risoluzione automatico simulerà il comportamento dell'esperto, seguendone strategie e linee di ragionamento. Quest'ultimo aspetto è molto importante perché conferisce a tale approccio una *trasparenza* ed una *tracciabilità* che in molti casi sono fondamentali per l'accettabilità da parte dei futuri utenti del sistema: in altri termini, il sistema KBS non solo risolve il problema, ma è anche in grado di spiegare ed illustrare come ha ottenuto la soluzione, quali elementi ha esaminato, a che risultati intermedi è giunto, quali inferenze ha fatto, che ipotesi ha scartato e perché, e così via. Volendo sintetizzare, potremmo dire che l'approccio dei sistemi basati sulla conoscenza è faticoso ed assai impegnativo, ma è molto gratificante e ricco di risultati!

2.2 *Il Machine Learning come approccio alternativo alla costruzione di sistemi intelligenti*

Negli anni '80, '90 ed inizi 2000 la tecnologia dei KBS si è affinata e metodologicamente stabilizzata¹⁰ e le applicazioni si sono moltiplicate in moltissimi settori. Va detto però che l'approccio KBS non era l'unico studiato dai ricercatori poiché molti di essi stavano sperimentando ed approfondendo una strada alternativa per la costruzione di sistemi intelligenti. In estrema sintesi, questi sostenevano che un'altra possibilità poteva essere quella di far esaminare al computer una vastissima casistica dello specifico problema affrontato, in modo tale da riuscire a mettere direttamente in relazione i dati del problema con le soluzioni; in tal modo si sarebbero evitate le oggettive e notevoli difficoltà dell'analisi delle conoscenze e delle strategie di ragionamento degli esperti umani. Tale approccio è quello denominato *Apprendimento Automatico*, o *Machine Learning (ML)*¹¹.

risultato dell'esperienza e di molteplici ripetizioni di un certo tipo di ragionamento. Di solito l'esperto umano non è pronto a verbalizzarla, ossia fa difficoltà a dare risposte alla domanda "come hai ragionato per produrre tale risultato?". Per acquisirla, sono spesso necessarie tecniche mutuata dalla psicologia clinica e/o sperimentale.

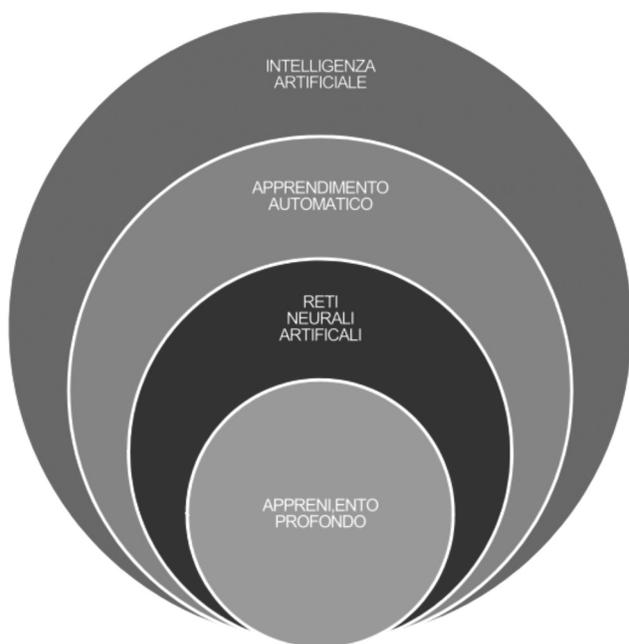
10. G. GUIDA - C. TASSO, *Design*, cit.

11. T.M. MITCHELL, *Machine Learning*, McGraw-Hill 1997.

In realtà, l'idea del ML era stata proposta già negli anni '50 e basti ricordare tra tutti il lavoro di Rosenblatt sul *perceptron*¹². Tuttavia, per vari motivi, il ML non ebbe il successo sperato e si dovrà attendere il primo decennio degli anni 2000 per vivere la stagione di molti successi e di grande fermento che stiamo vivendo ancora oggi. Nelle prossime sezioni saranno illustrati alcuni dei motivi di questo ritorno.

Il ML quindi va inquadrato come un settore dell'IA, che studia algoritmi che partendo da dati che descrivono esperienza e casistica passata, sintetizzano automaticamente nuova conoscenza utile per risolvere automaticamente nuova casistica futura. In Figura 3 sono illustrate schematicamente le relazioni tra IA, ML e le altre tipologie più specifiche di tecniche di ML che verranno trattate nelle prossime sezioni.

Figura 3 - Relazioni di inclusione tra la disciplina dell'IA ed i settori del ML, delle Reti Neurali e del Deep Learning.



12. F. ROSENBLATT, *The perceptron: A probabilistic model for information storage and organization in the brain*, in *Psychological Review* 65(6), 1958, 386 ss.

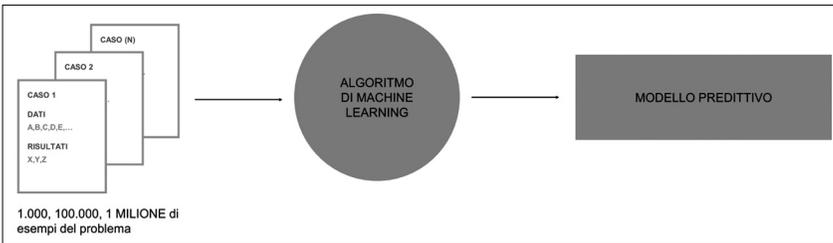
3. Metodologia di sviluppo dei sistemi di Machine Learning

Anche nel caso dell'utilizzo del ML, come già illustrato per i KBS, la costruzione di un sistema intelligente segue una metodologia di lavoro del tutto diversa dagli approcci dell'Informatica Tradizionale.

Sostanzialmente il ciclo di vita di un sistema di ML è strutturato nelle due seguenti fasi principali:

a) *Fase 1 - Fase di Addestramento o Training*. Come illustrato in Figura 4, si utilizza un algoritmo di ML che analizza un insieme (denominato *Training Set*) di moltissimi esempi di casi del problema considerato, descritti dai dati (denominati *attributi o feature*) che li caratterizzano e dalle relative soluzioni/decisioni/risultanze. L'algoritmo è in grado di costruire automaticamente un nuovo modulo software, capace di affrontare casi simili, non già visti in precedenza. Il modulo software viene indicato con molti possibili termini. Termini comuni sono *Modello Predittivo* (o più raramente *Previsionale*), *Modello Classificatorio* o semplicemente *Modello*.

Figura 4 - Training e produzione del Modello Predittivo



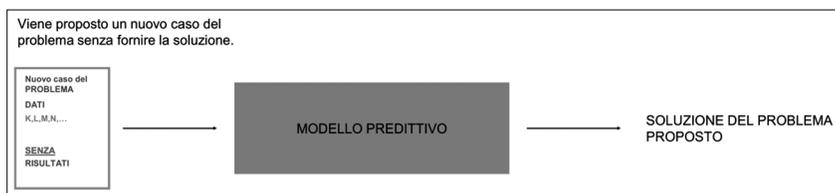
Tornando all'esempio veterinario, si tratterà di raccogliere una casistica numerosa di valutazioni già svolte dal veterinario su specifiche bovine, in cui ciascun caso è caratterizzato da tutti i dati sulla bovina e comprende anche la decisione che in precedenza è già stata presa su quel caso dall'esperto veterinario. L'algoritmo di ML esamina tutta la casistica, analizza le relazioni fra i dati di input e le soluzioni, identifica i parametri più significativi, ... e produce in output il Modello Predittivo. In un certo senso, nella fase di training si 'fa rivivere' alla macchina l'esperienza del veterinario che anno dopo anno ha esaminato una vasta casistica costruendo, caso dopo caso, una sua conoscenza esperta.

Appare abbastanza intuitivo, che maggiore è il numero di casi osservati e migliore sarà il modello prodotto. Si noti inoltre come in Figura 4 siano indicati numeri esemplificativi molto elevati relativamente a quanti casi possono essere sottoposti all'algoritmo: tali valori non devono spaventare, e in tal senso, si consideri un esempio fra tutti, ossia il caso di *Amazon.com*, famoso portale di commercio elettronico che analizzando miliardi di interazioni dei suoi utenti con il portale stesso, 'impara' modelli che riescono a prevedere con elevata accuratezza il comportamento ed i gusti degli utenti, al fine di raccomandare loro certi acquisti o per inviare loro specifica pubblicità ritagliata sulle loro preferenze.

La Fase 1 è preliminare all'utilizzo del modello prodotto, che avviene nella fase successiva.

b) *Fase 2 - Utilizzo del Modello Predittivo*. Il Modello Predittivo risultante dalla Fase 1, se costruito correttamente, sarà in grado di prevedere (da cui il nome *predittivo*) con buona precisione quale sarebbe la decisione del veterinario se fosse lui ad esaminare quel caso: come illustrato in Figura 5, se forniamo in input al modello un nuovo caso (questa volta senza la soluzione!), il modello individua la soluzione del caso.

Figura 5 - Utilizzo del Modello Predittivo per risolvere nuovi casi del problema



In altri termini, dall'analisi dei dati e della soluzione di ciascun esempio fornito all'algoritmo di ML nella Fase di Training, il Modello Predittivo ha 'capito' come trovare la soluzione 'direttamente' dai dati del problema. Questa tipologia di ML riguarda le così dette tecniche *supervisionate*, che cioè richiedono di fornire alla fase di training sia i casi sia le relative soluzioni. Non è però l'unico approccio possibile: gli algoritmi *non supervisionati* includono altre tecniche che sono in grado di esaminare grandi masse di dati andando a riconoscere raggruppamenti significativi, correlazioni fra gruppi di dati, similitudini e anomalie. In questo lavoro ci riferiamo principalmente solo agli algoritmi supervisionati.

Come ben si capisce, l'approccio ML si presenta del tutto diverso dall'approccio KBS ed è opportuno approfondire ulteriormente queste prime nozioni generali sull'approccio ML, al fine di inquadrare correttamente alcune caratteristiche importanti di questo nuovo scenario: nelle prossime sezioni sono trattate alcune tecniche specifiche di ML, che tra l'altro sono quelle che attualmente stanno riscuotendo il maggior successo e la maggior diffusione.

4. *Le Reti Neurali Artificiali e il Deep Learning*

Si è già menzionato il lavoro di Rosenblatt¹³, che nel 1958 propose ciò che viene considerato il primo modello di Neurone Artificiale (denominato *perceptron*). In realtà, Rosenblatt si ispirava a lavori ancora precedenti, fra cui la proposta di McCulloch e Pitts¹⁴ che quindici anni prima avevano proposto un modello (da loro definito logico) sul come funziona il sistema nervoso. Dagli anni '40 e 50' ci volle circa mezzo secolo affinché le loro idee (ancorché a livello ancora molto iniziale) raggiungessero un livello di perfezionamento tale da essere considerate oggi una delle attuali tecnologie più *disruptive*.

Nelle prossime tre sottosezioni vengono esaminate più in dettaglio.

4.1 *Reti Neurali Artificiali*

Un *Neurone Artificiale (NA)*¹⁵ è una unità di calcolo molto semplice, che riceve dei dati in ingresso, li elabora e produce un risultato in uscita: il suo funzionamento è ispirato a quello dei neuroni del nostro cervello. Un neurone biologico è collegato (in cascata) ad altri neuroni biologici che gli trasmettono dei segnali elettrici; se il neurone riceve una quantità di elettricità al di sopra di una certa soglia, trasmette a sua volta un segnale elettrico ai neuroni che lo seguono nella catena. I punti di contatto, denominati *sinapsi*, tra un neurone ed i neuroni successivi, possono essere più o meno 'forti', modulando quindi il passaggio della quantità di elettricità fra un neurone ed il successivo: una sinapsi che lascia passare meno elettricità può interrompere il flusso elettrico mentre al contrario una sinapsi che

13. F. ROSENBLATT, *The perceptron*, cit.

14. W.S. MCCULLOCH - W.A. PITTS, *A logical calculus of the ideas immanent in nervous activity*, in *Bull. math Biophysics* 5, 1943, 115 ss.

15. C.C. AGGARWAL, *Neural Networks and Deep Learning: A Textbook*, Springer 2018.

favorisce il passaggio di elettricità instaura una connessione fra i diversi neuroni in cascata. Attualmente si ritiene che nell'uomo l'apprendimento sia ottenuto dall'instaurarsi di legami sinaptici più o meno forti fra le decine di miliardi di neuroni presenti nel nostro cervello. In altri termini, semplificando, l'apprendimento umano è il processo per cui, a seconda degli stimoli e delle sollecitazioni ricevute, si aprono/rafforzano o si chiudono/indeboliscono le catene di connessione tra neuroni, che sono alla base delle nostre attività cognitive.

Un NA in un certo senso replica il funzionamento del neurone biologico: lo si illustra qui di seguito in modo essenziale facendo delle opportune (e numerose) semplificazioni. Una *Rete Neurale Artificiale* (RNA) (o *Artificial Neural Network*) è costituita da tanti NA, tra loro connessi e suddivisi in tre categorie: *neuroni di input*, *neuroni di output* e *neuroni interni*, quest'ultimi organizzati in strati, detti *strati nascosti*. Ciascun NA è connesso ai neuroni che lo precedono e ai neuroni che lo seguono. I dati in input alla RNA vengono ricevuti dai neuroni di input ed il risultato della computazione dell'intera rete viene prodotto dai neuroni di output che non hanno altri neuroni in cascata e che quindi forniscono il risultato finale della computazione. Ciascun neurone n opera al seguente modo:

1. n riceve in ingresso una serie di valori numerici che sono stati prodotti da ciascuno dei neuroni che lo precedono e sono ad esso connessi;
2. n pesa ciascuno di tali valori in ingresso con altrettanti *pesi numerici*; tali pesi, col loro valore, esaltano o indeboliscono l'impatto su n da parte dei corrispondenti neuroni di provenienza, connessi a n ;
3. infine n somma i valori così pesati, e se il risultato è maggiore di una certa soglia, manda in avanti tale risultato verso tutti i neuroni connessi che lo seguono, altrimenti manda il valore zero ai neuroni successivi.

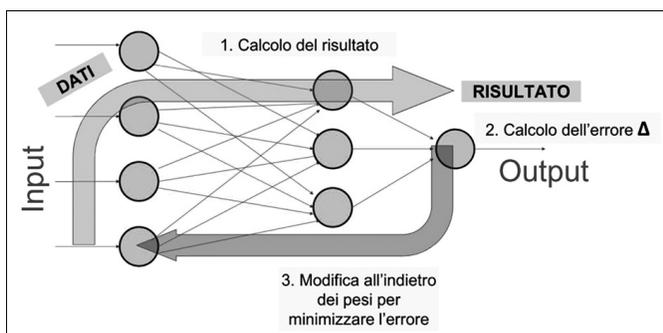
Il processo inizia dallo strato dei neuroni in ingresso e procede fino ad arrivare ai neuroni terminali di output (che forniscono la soluzione). Nell'esempio della valutazione veterinaria, potremmo immaginare uno strato di neuroni di input che riceve tutti i dati relativi alle bovine e un neurone di output che rappresenta con un valore binario (0 o 1) la scelta migliore (tra le due possibili).

Per capire più in dettaglio come le due fasi generali presentate nella Sezione 3 si applichino al caso delle RNA, sono illustrati in Figura 6 gli elementi principali relativi al training di una RNA esemplificativa, costituita da 4 neuroni di input, uno strato nascosto costituito da tre neuroni, ed un unico neurone di output. Il processo di Training è realizzato in modo

incrementale mediante molti cicli iterativi (chiamati *epoche*) che ripetono i seguenti tre passi sui casi inclusi nel Training Set:

1. Calcolo 'in avanti' (*feed-forward step*). Gli input, costituiti dai dati di un singolo caso del Training Set, sono forniti in ingresso ai neuroni di input che operano secondo la procedura precedentemente indicata. Di solito, nella prima iterazione si utilizzano dei pesi scelti a caso. In base a tali pesi, viene propagato il risultato dallo strato di input, ai vari strati nascosti successivi, fino allo strato di output, ottenendo un certo risultato che, in generale, sarà diverso dalla soluzione del problema indicata nel caso incluso nel Training Set (se non altro perché si parte da un insieme di valori dei pesi scelti a caso!).
2. Calcolo dell'*errore* Δ . Viene calcolata la differenza Δ tra il risultato ottenuto e la soluzione che si sarebbe dovuto ottenere, nota poiché inclusa nella descrizione di ciascun caso del Training Set.
3. Ritorno 'all'indietro' (*back-propagation step*). Si procede all'indietro, progressivamente, uno strato alla volta, modificando opportunamente i valori dei pesi appena utilizzati, al fine di ridurre (minimizzare, se possibile azzerare) l'errore Δ . Si inizia dall'ultimo strato, a seconda del valore dell'errore Δ si modificano tutti i pesi relativi ai neuroni di quello strato, facendo in modo che l'errore che risulterà dall'applicazione dei pesi modificati sia il più piccolo possibile. A questo punto si procede in modo analogo con il penultimo strato, e così di seguito fino allo strato dei neuroni di input.

Figura 6 - Processo di Training di una Rete Neurale Artificiale



I tre passi sono poi ripetuti utilizzando i *nuovi pesi* appena calcolati per tutti gli strati. La procedura viene ripetuta molte volte tenendo conto di tutti i casi del Training Set e modificando i vari pesi di ciascun neu-

rone in modo da minimizzare l'errore Δ . Più precisamente il processo di minimizzazione mira a ridurre l'errore medio relativamente a tutti i casi del training set. I pesi presenti nei vari neuroni della rete sono quindi il parametro su cui si agisce per addestrare una RNA al fine di risolvere con accuratezza le singole istanze del problema: in altri termini, durante il processo di training la RNA 'impara' modificando i valori dei pesi.

Questa caratteristica differenzia ulteriormente una RNA da un KBS: infatti, mentre un KBS è in grado di spiegare il proprio comportamento ripercorrendo i ragionamenti fatti, una RNA può, al massimo, mostrare i vari pesi numerici che sono stati prodotti progressivamente nel processo di training (nel passo di back-propagation), e quindi non riesce a fornire alcuna spiegazione intuitiva/qualitativa¹⁶ sul perché si è ottenuto un certo risultato. Questa situazione viene indicata con il termine di *opacità* delle reti neurali.

4.2 Deep Learning

Una tipica domanda che viene alla mente a chi si avvicina alle reti neurali per la prima volta riguarda il motivo per cui un modello di calcolo inizialmente proposto negli anni '40 e '50 abbia raggiunto un livello di maturità (e di gran successo) solo molti decenni più tardi. In effetti, sia le RNA che più in generale il ML sono sempre stati oggetto di studio, ricerca e insegnamento (in parecchie Università, anche italiane), fin dagli anni '70, ma nello scenario generale dell'IA sono rimaste a lungo relegate in un settore piuttosto secondario. Per capire le ragioni del fenomeno, si illustrano alcuni ulteriori sviluppi delle RNA.

Col diffondersi della sperimentazione delle reti neurali, in particolare nel settore della *visione artificiale*, le RNA hanno dimostrato fin dagli anni '80 di essere uno strumento molto potente, capace di fornire ottime prestazioni in termini di riconoscimento automatico delle immagini (riconoscere se in un'immagine è presente una parte che riguarda un certo oggetto, riconoscere un volto, ecc.). Per risolvere problematiche più complesse si iniziò ben presto ad aumentare il numero di NA delle reti, con molti strati nascosti che contenevano molti neuroni ed in effetti, rendendo la rete più complessa, le sue prestazioni miglioravano.

Inoltre, con l'evoluzione della digitalizzazione in tutti i settori applicativi e con l'aumento del numero di dati digitali disponibili, si iniziò a parlare di *big data*, ossia archivi che contengono enormi masse di dati digitali, tipicamente provenienti dall'automazione operativa di varie tipologie

16. Si intende qui: capibile dall'uomo, cognitivamente plausibile.

di processi e procedure, sia in ambito pubblico che privato. Si comprese subito che l'utilizzo di tecniche di ML sui *big data* poteva offrire grandi opportunità di *knowledge discovery*, ossia della possibilità di sintetizzare conoscenza su fenomeni che fino a quel momento erano molto complessi e complicati da analizzare.

Però ben presto emerse un problema critico: il training della rete, soprattutto con Training Set molto voluminosi (*big data*), molti strati nascosti e moltissimi neuroni, divenne ben presto troppo oneroso in termini di risorse di calcolo necessarie: una fase di training poteva necessitare potenzialmente di un tempo di mesi o anni di computazione. Ed ecco che quindi le RNA rimasero per anni sullo sfondo, pur attirando ancora molti ricercatori, che 'ostinatamente' credettero nelle potenzialità straordinarie dell'approccio¹⁷. Arriviamo a questo punto agli anni '90 e ai primi quindici anni del nuovo millennio: emergono due fattori determinanti:

- a) i ricercatori inventarono nuove tecniche e meccanismi per snellire e rendere più efficiente la fase di Training;
- b) lo sviluppo di piattaforme hardware sempre più potenti permise di eseguire in tempi molto brevi i calcoli che precedentemente richiedevano tempi lunghissimi (e quindi inaccettabili). Da una decina d'anni sono disponibili delle schede GPU¹⁸, particolarmente efficienti per le elaborazioni con reti complesse.

17. È d'obbligo qui ricordare i tre vincitori dell'Edizione 2018 del premio Turing Award: Yoshua Bengio (Professore all'Università of Montreal e fondatore della start-up Element AI), Geoffrey Hinton (di Google Brain e Professore Emerito dell'Università di Toronto) e Yann LeCun (di Facebook e professore alla New York University). Il Turing Award è considerato il Premio Nobel per l'Informatica e l'edizione 2018 è stata attribuita ai tre ricercatori con la seguente motivazione: "ACM named Yoshua Bengio, Geoffrey Hinton, and Yann LeCun recipients of the 2018 ACM A.M. Turing Award for conceptual and engineering breakthroughs that have made deep neural networks a critical component of computing. ...*omissis*... Working independently and together (*ndr*: dalla fine degli anni '80!), Hinton, LeCun and Bengio developed conceptual foundations for the field, identified surprising phenomena through experiments, and contributed engineering advances that demonstrated the practical advantages of deep neural networks." La motivazione è ripresa dal sito ufficiale dell'ACM, American Association of Computing Machinery (awards.acm.org/about/2018-turing).

18. Una GPU (*Graphical Processing Unit*) è una scheda elettronica che viene utilizzata al posto delle classiche CPU (*Central Processing Unit*) su cui si basano i computer. Le GPU, nate inizialmente per garantire adeguata efficienza nella gestione grafica dei Video Game, si è dimostrata utilissima ed adeguata anche per un'efficiente esecuzione delle moltissime operazioni elementari necessarie nella fase di Training del Deep Learning.

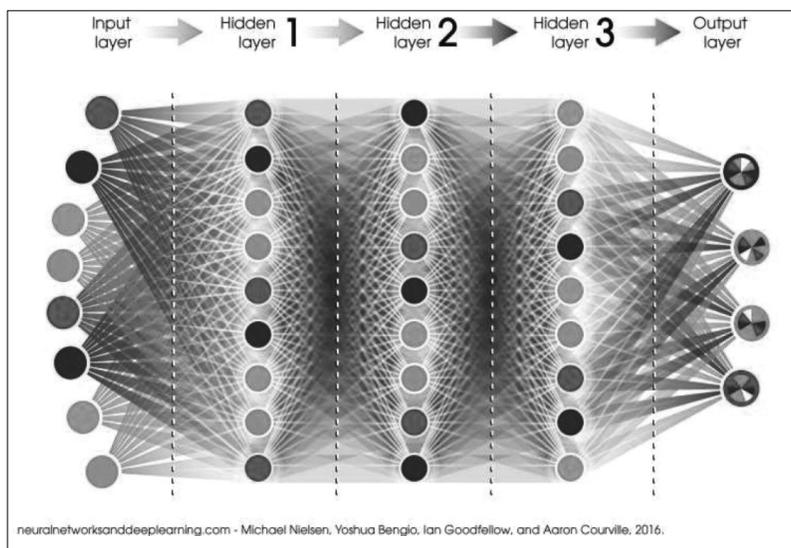
Questi fattori portarono alla sperimentazione e diffusione delle cosiddette *Reti Profonde* (o *Deep Learning*), ossia reti neurali con moltissimi strati nascosti (da cui il termine *profonde*), moltissimi neuroni di input ed interni (decine di migliaia o addirittura milioni!); è facile stimare che il numero di connessioni che connette tra loro i vari neuroni dei vari strati della rete può raggiungere l'ordine di grandezza dei miliardi! I Training Set inoltre possono arrivare a contenere migliaia o milioni di casi!

Nonostante il notevolissimo peso computazionale delle Reti Profonde, i fattori a) e b) di cui sopra permisero di risolvere con successo problemi mai affrontati prima. Tra questi, ricordiamo il *Progetto ImageNet*¹⁹, nel campo della computer vision, che ha contribuito alla diffusione e al successo del Deep Learning.

Quindi, il Deep Learning, di cui si sente tanto parlare, non è altro che un sotto settore specifico delle RNA, che riguarda le reti profonde, con moltissimi neuroni e strati nascosti (come ad esempio quella esemplificata in Figura 7). Senza entrare in ulteriori dettagli, vogliamo qui ricordare un ulteriore aspetto molto interessante: una Rete Profonda con i suoi numerosi strati nascosti, riesce a rappresentare il dominio di interesse mediante gerarchie di concetti, in cui ogni concetto è definito in relazione a concetti più semplici: le varie rappresentazioni (ad esempio di un'immagine o di un testo in linguaggio naturale) sono relative a diversi gradi di astrazione del dominio e sono distribuite sui vari strati nascosti della rete, rappresentando elementi più elementari in prossimità dei neuroni di input e progressivamente elementi sempre più astratti negli strati nascosti successivi.

19. *ImageNet* (image-net.org/) è un progetto che prevede lo svolgimento di una gara annuale fra ricercatori che competono al fine di provare la bontà dei propri algoritmi per il riconoscimento di immagini. Il progetto dispone di un data base di oltre 14 milioni di immagini di 20.000 categorie diverse. Nell'edizione 2012 vinse un algoritmo basato su una Rete Neurale profonda di tipo Convolutivo eseguita mediante GPU, ottenendo una performance molto superiore ai soggetti umani (spesso è stato usato il termine *superhuman*), e questo evento segnò anche l'entrata nei media e l'inizio del grande successo del Deep Learning.

Figura 7 - Illustrazione esemplificativa di una Rete Profonda



Il successo raggiunto con le applicazioni basate sulle Reti Profonde è tale che spesso, quando si parla dei sorprendenti successi dell'Intelligenza Artificiale, ci si riferisce in effetti a quanto si è stati capaci di fare con le tecniche di Deep Learning.

Nel Laboratorio di Intelligenza Artificiale dell'Università di Udine²⁰ da anni si svolgono numerose ricerche sul Machine Learning e sulle reti neurali (tradizionali e profonde). In particolare, le tecniche di Deep Learning sono utilizzate per analizzare testi scritti in linguaggio naturale (si lavora principalmente sulla lingua italiana, inglese e araba). In Figura 8 viene mostrato il risultato di un'analisi semantica di un testo tecnico di informatica, che estrae dal testo i concetti ritenuti più rilevanti e significativi per rappresentarne il contenuto²¹. Queste tecniche sono state utilizzate nel Laboratorio di Intelligenza Artificiale per costruire il profilo

20. L'AI Lab dell'Università di Udine è stato fondato nel 1984 dall'autore di questo lavoro, primo Laboratorio di Intelligenza Artificiale fondato in una Università Italiana.

21. M. BASALDELLA - G. SERRA - C. TASSO, *The Distiller framework: Current state and future challenges*, in *Digital Libraries and Multimedia Archives. 14th Italian Research Conference on Digital Libraries* (a cura di G. Serra e C. Tasso), CCIS, Springer Verlag, 2018, 93 ss.

personalizzato degli interessi di un utente, al fine di sfruttarli in motori di ricerca semantici che tengano conto degli specifici concetti che l'utente gradisce ricevere²².

Figura 8 - Risultato del sistema Distiller, sviluppato nell'AI Lab dell'Università di Udine, per l'estrazione di concetti da testi in linguaggio naturale. Sulla destra i concetti ritenuti più significativi

A UML Class Diagram Analyzer
 Tiago Massoni, Rohit Gheyi, and Paulo Borba
 Informatics Center, Federal University of Pernambuco, Brazil

Abstract. Automatic analysis of UML models constrained by OCL invariants is still an open research topic. Especially for critical systems, such tool support is important for early identification of errors in modeling, before functional requirements are implemented. In this paper, we present ideas on an approach for automatic analysis of UML class diagrams, according to a precise semantics based on Alloy, a formal object-oriented modeling language. This semantics permits the use of Alloy's tool support for class diagrams, by applying constraint solving for automatically finding valid snapshots of models. This kind of automation helps the identification of inconsistencies or under-specified models of critical software, besides allowing checking of properties about these models.

1 Introduction

As in other engineering fields, modeling can be a useful activity for tackling significant problems in software development. As a de-facto standard, the Unified Modeling Language (UML) [1] plays a significant role. In particular, the development of high-quality critical systems can benefit from features offered by a standard visual modeling notation like UML, since traditional code-driven approaches are highly risky and often error-prone for such complexity. Further, business rules can be precisely expressed by Object Constraint Language (OCL) [2] invariants attached to class diagrams, enabling the specification of constraints over complex states of critical systems. As the use of OCL for critical systems substantially grows, the lack of standardized formal semantics for the language does not stimulate the development of tools supporting analysis and verification of OCL expressions in UML models. When using class diagrams for modeling critical systems, the absence of tool support restrains the developer's task, since subtle structural modeling errors may be considerably hard to detect. For instance, inappropriate OCL invariants may turn a model over-constrained, or even inconsistent (i.e. the model allows no implementation). Likewise, models lacking important constraints may allow incorrect implementations.

Alloy,
 Alloy
 Analyzer,
 UML
 OCL
 OCL
 Invariants,
 Critical
 Systems,
 UML Class
 Diagram,
 Invariants,
 Snapshots

5. Campi di applicazione di ML e DL e potenziali vantaggi per la PA

Dovendo rimanere ad un livello molto generale, le tipologie di problemi affrontabili mediante il ML e il DL ricadono essenzialmente nelle seguenti categorie principali:

- 1) *Problemi di Classificazione*: analisi di dati e assegnazione ad una classe, valutazione di una qualità/tipologia/ecc., analisi di immagini o video, analisi di testi in linguaggio naturale, riconoscimento di situazioni, scelta di una decisione, ecc.

22. D. DE NART - C. TASSO, *A Personalized Concept-Driven Recommender System for Scientific Libraries*, in *Procedia Computer Science* 38, 2014, 84 ss.

- 2) *Problemi di Clustering*: identificazione di raggruppamenti di entità simili o in qualche relazione fra loro. Ovviamente, la situazione interessante si ha quando la quantità di dati a disposizione è talmente grande che un'analisi umana eseguita in modo manuale non riuscirebbe ad identificare alcuna relazione particolare: si pensi ad esempio ad un'analisi di grandi basi di dati di casi giudiziari, o di dichiarazioni fiscali, o di comportamenti, ecc. Per questa classe di problemi si utilizzano algoritmi di tipo non supervisionato.
- 3) *Problemi di Previsione (forecasting)*: di un esito, di un livello di rischio, dell'andamento di un parametro in un istante futuro, di una quantità, di un livello di soddisfazione, di un fattore di utilità, ecc.

Risolvendo tali categorie di problemi, si possono affrontare una variegata moltitudine di moltissime diversificate applicazioni, fra le quali (un elenco esaustivo sarebbe impossibile):

- a) Sistemi di Supporto alle decisioni (ad esempio amministrative nell'ambito della Pubblica Amministrazione).
- b) Sistemi di supporto alla attività Diagnostica (sia medica che non).
- c) Sistemi di Raccomandazione²³, che partendo dall'analisi del comportamento di navigazione degli utenti di un sito Web, personalizzano l'interazione e raccomandano contenuti e prodotti in modo adattativo con messaggi pubblicitari targhettizzati.
- d) Sistemi di videosorveglianza, che per motivi di sicurezza analizzano i comportamenti di persone in luoghi pubblici identificando situazioni di rischio.
- e) Analisi semantica di testi: per valutare il sentiment (positivo, negativo o neutro) di un'opinione online, per analizzare le argomentazioni che giustificano una certa conclusione, per cercare di scoprire fake news, per analizzare la reputazione di una persona, di una marca o di un prodotto, ecc.
- f) Sistemi intelligenti di Cybersecurity, per monitorare la rete aziendale ed identificare attacchi informatici.
- g) Sistemi di Traduzione automatica da una lingua all'altra.
- h) Analisi di immagini e video.
- i) Sistemi di Guida Autonoma.
- j) Manutenzione e Medicina predittiva.
- k) Composizione Automatica di Brani Musicali stilisticamente simili ad un certo genere o ad un certo autore.

23. C.C. AGGARWAL, *Recommender Systems - The Textbook*, Springer 2016.

Volendo restringere l'attenzione al settore della Pubblica Amministrazione²⁴, numerosi sono i potenziali vantaggi di una informatizzazione basata su tecniche di ML e DL, e più in generale di Intelligenza Artificiale²⁵. Eccone alcuni:

- semplificare la gestione delle relazioni con il cittadino,
- migliorare servizi, soddisfazione dei cittadini e qualità della vita,
- migliorare efficacia e precisione dei provvedimenti,
- aumentare velocità ed efficienza,
- migliorare e ottimizzare i processi decisionali,
- semplificare la Normativa,
- facilitare l'identificazione di frodi,
- ridurre il contenzioso,
- accedere in modo 'semantico' alla normativa,
- analizzare semanticamente gli atti,
- identificare anomalie e relazioni non evidenti all'interno di grosse basi di dati.

6. *Machine Learning: competenze coinvolte, meriti e responsabilità*

Per quanto illustrato finora possiamo certamente affermare che i sistemi intelligenti sviluppati tramite tecniche di Intelligenza Artificiale (KBS o Machine Learning o Deep Learning) permettono di realizzare sistemi software che forniscono prestazioni simili alle nostre capacità cognitive e portano dei significativi vantaggi aggiuntivi rispetto all'Informatica Tradizionale. È però anche necessario ed opportuno capire quali nuovi problemi e quali criticità possono emergere con questi approcci innovativi.

Tutti i metodi che abbiamo visto si basano, in modo più o meno profondo e diretto, su un'analisi della conoscenza del dominio considerato. Nel caso dei KBS ci si basa su esperti di dominio che forniscono dati e informazioni sul loro modo di risolvere gli specifici problemi di interesse. Nel caso del ML è necessario capire con l'aiuto di chi conosce il dominio quali sono i dati da collezionare per costruire i Training Set necessari per

24. Per un approfondimento relativo alla Pubblica Amministrazione: D.-U. GALETTA - J.G. CORVALAN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, Rischi e Sfide della Rivoluzione Tecnologica in Atto*, in *Federalismi.it* 3, 2019.

25. AGID - AGENZIA PER L'ITALIA DIGITALE, *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino - Versione 1.0* (a cura della Task force sull'Intelligenza Artificiale dell'Agenzia per l'Italia Digitale), 2018.

l'addestramento dei modelli predittivi. I due processi appena menzionati sono essenziali per la riuscita ed il successo del progetto di costruzione del sistema intelligente: ma chi ci assicura che il Modello Concettuale che abbiamo utilizzato per costruire un KBS e che abbiamo 'elicitato' dall'esperto del dominio sia sempre in grado di affrontare tutti i vari casi del problema considerato che si potranno verificare in futuro? Oppure, nel caso del ML, chi ci garantisce che i dati (e gli attributi) che ci sono stati forniti sono effettivamente quelli discriminanti per riuscire a costruire un modello predittivo accurato nella generalità dei casi? E se da un lato nel caso dei KBS c'è sempre la possibilità di valutare il risultato prodotto dal sistema grazie alle caratteristiche di trasparenza e tracciabilità dell'approccio (cfr. Sezione 2.1) che permettono di analizzare i passi di ragionamento svolti e gli elementi considerati, va riconosciuto come nel caso del ML (ed ancor di più nel caso delle reti complesse del Deep Learning), l'opacità dell'approccio (cfr. Sezione 4.1) non permette alcuna analoga valutazione, ed il risultato dell'utilizzo del modello predittivo va preso in quanto tale, senza alcuna spiegazione che permetta di riconoscere eventuali situazioni in cui il modello potrebbe essere inadeguato (e quindi soggetto ad errori). In altri termini, come accennato nella Sezione introduttiva, il sistema intelligente può non raggiungere un'accuratezza del 100%!

Nel seguito, approfondiamo il caso del ML analizzando in modo più dettagliato lo scenario completo che parte dall'invenzione di un algoritmo di ML per arrivare agli effetti dell'utilizzo del modello classificatorio. In Figura 9 è illustrato il *workflow* che porta a costruire ed utilizzare un sistema intelligente basato su un approccio di ML (le considerazioni sono del tutto analoghe nel caso del Deep Learning). In Figura sono anche evidenziati (cerchiati in nero) gli 'attori' coinvolti nelle varie attività, ossia coloro che hanno l'onere di costruire il sistema e, se le cose vanno bene, ne hanno anche il merito. Sono previsti 4 stadi successivi che per omogeneità con quanto illustrato in Sezione 3 sono indicati con il termine di Fasi:

- I. Fase 0: *invenzione di algoritmi di ML*. Solo per completezza si è considerata anche l'attività che i *ricercatori* di Intelligenza Artificiale svolgono quando propongono e mettono a punto algoritmi di Machine Learning. L'output di tale fase è costituito dalla disponibilità di algoritmi di Machine Learning (ne esistono centinaia), con le loro specifiche caratteristiche, la documentazione (articoli scientifici) relativa ai vari aspetti importanti, dalle prestazioni e dai vantaggi alle limitazioni e alla applicabilità degli specifici algoritmi.

II. *Fase 1: Training del Modello Classificatorio*. Questa fase si innesca quando si decide di realizzare un sistema basato su ML in un certo specifico dominio applicativo. Gli attori coinvolti sono:

- a. *Esperti del dominio*: da essi ci si aspetta che conoscano il problema che si vuole affrontare, gli obiettivi del progetto, i dati e gli attributi importanti relativi al fenomeno che si vuole modellare e su cui si vogliono risolvere specifiche istanze del problema di interesse. Inoltre, devono anche sapere quali dati non sono rilevanti e qual è il patrimonio di dati disponibile in formato digitale nell'ambito dell'organizzazione in cui essi operano e in cui si vuole sviluppare il Modello Classificatorio.
- b. *Data scientist* (o *data engineer*). Si tratta dei tecnici informatici che devono essere in grado di analizzare il problema in relazione agli obiettivi, che sanno scegliere gli algoritmi di ML da utilizzare (di solito più d'uno), che sanno sperimentare diversi algoritmi alternativi e sanno selezionare i più adatti, che sanno preliminarmente analizzare e valutare i Dati ricevuti dagli esperti di dominio per evidenziarne delle eventuali carenze, che sanno costruire i Modelli Classificatori mediante gli algoritmi selezionati, che sono in grado di selezionare gli attributi più discriminanti per il problema da risolvere, che sanno validare il Modello Classificatorio prodotto, ossia sanno valutare la bontà e l'accuratezza del modello ed i limiti delle scelte progettuali.

Questa fase ha come risultato un Modello Classificatorio (possibilmente costituito da diverse parti realizzate con diverse tecniche di ML) che è stato opportunamente validato.

III. *Fase 2: Utilizzo del Modello Classificatorio*. Si tratta della fase in cui il Modello viene inserito nell'operatività secondo gli obiettivi del progetto. Il modello opera su dati reali ed esegue il compito per cui è stato realizzato, producendo la soluzione di singole istanze del problema.

IV. *Fase 3: Processo decisionale basato sulla soluzione fornita dal Modello Classificatorio*. Questa è la fase finale, in cui l'output prodotto dal modello viene sfruttato nell'ambito applicativo.

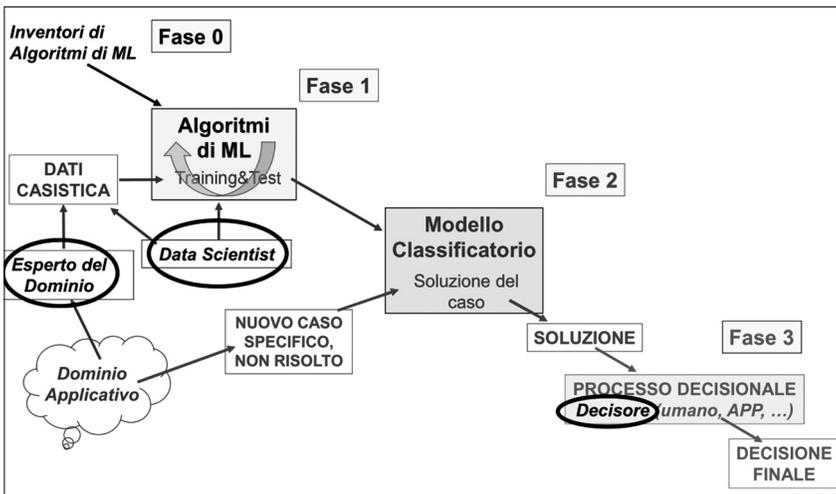
Si consideri l'esempio di un Modello Classificatorio che deve produrre la valutazione di un livello di rischio. Per semplicità, ci siano tre possibili valori di output: *stato normale*, *possibili problemi* e *rischio elevato*. A seconda del valore che sortisce, si tratta di attuare quanto è stato previsto che venga fatto in ciascuno dei tre casi.

Il passaggio dall'aver a disposizione l'output del modello all'effettiva attuazione di quanto previsto, prevede quindi un'ulteriore decisione finale relativa a (i) attuare ciò che consegue dal modello ovvero (ii) decidere di procedere diversamente. Tale decisione introduce due ulteriori possibili attori:

- a. *Decisore umano*. Un umano si prende la responsabilità di operare strettamente secondo quanto il modello predice, ovvero decide di procedere diversamente.
- b. *Decisore automatico*, ad esempio una APP, possibilità sempre più comune in tantissimi ambiti (dalla finanza con i trading system, alla sicurezza, alla guida autonoma dei veicoli, ecc.). Il decisore automatico procede in accordo all'output del modello, innescando le procedure previste.

In una o l'altra delle due modalità, qui si conclude l'intero processo per cui si è deciso di costruire il sistema di ML.

Figura 9 - Workflow per la costruzione e l'utilizzo di un sistema software intelligente basato su ML



Si analizza ora cosa 'può andare male', ossia, in ultima analisi, perché può essere presa una decisione finale errata. In Figura 10 sono illustrate le principali cause che qui di seguito esaminiamo, a ritroso partendo dalla fase finale, cercando anche di individuarne le responsabilità:

- a) *Decisione finale errata*. La decisione presa dal decisore non è adeguata alla situazione e può causare dei danni (anche molto gravi!). Ci sono due possibilità, a seconda del decisore finale:
1. *Caso del decisore umano*. Indipendentemente dal fatto che il modello classificatorio abbia prodotto una soluzione corretta o errata, il decisore umano ha la responsabilità della decisione. Quello che si può chiedere ad un sistema intelligente che analizza la situazione e produce in output una valutazione, è di fornire ulteriori informazioni, in particolare sulle modalità utilizzate per arrivare alla soluzione, che aiutino il decisore umano a valutare ulteriormente se quanto prodotto dal Modello Classificatorio è effettivamente accettabile o utilizzabile o da revisionare, ecc. Come messo in evidenza precedentemente, solo l'approccio KBS ha una buona trasparenza ed è in grado di fornire tali spiegazioni a supporto del risultato. Ciò non è invece possibile, allo stato attuale, con le Reti Neurali ed in particolare con il Deep Learning a causa delle loro opacità. Va detto tuttavia che in ambito di ricerca il problema è noto e si stanno studiando delle possibili modalità per rendere tali approcci più trasparenti.
 2. *Caso del decisore automatico* (ad esempio un APP). Ovviamente, nel caso il modello produca una soluzione corretta, non insorgono problemi. Se invece la soluzione prodotta è errata, al di là del perché sia errata, nasce subito una serie di questioni: “chi e perché ha deciso che la APP procedesse automaticamente ed autonomamente?”, “c'erano tutte le condizioni per fidarsi sempre dell'automaticità?”, e così via. In questo caso, oltre a permanere il problema dell'errore del Modello Classificatorio, si prefigura anche una responsabilità per chi ha deciso di lasciar agire in autonomia la APP. Per capire quanto gravi possano essere situazioni di questo tipo, si considerino le due seguenti modalità di sbagliare da parte del Modello Classificatorio:
 - *Falso Positivo*. Il modello, ad esempio nel caso di un sistema di diagnosi di patologie potenzialmente gravi, sbaglia e produce un risultato positivo relativamente alla presenza della patologia: l'utente si spaventerà, ma nel momento in cui procederà ad accertamenti potrà essere rassicurato scoprendo di essere sano e di non essere soggetto ad alcuna patologia.
 - *Falso Negativo*. Il Modello produce un'indicazione di assenza della patologia, ma in realtà la patologia è presente. L'utente pensa di non preoccuparsi, invece era proprio il caso di farlo: le conseguenze potrebbero essere molto gravi!

- b) *Il Modello produce una soluzione errata perché ha una bassa qualità.* Significa (i) che la accuratezza (una delle metriche che si utilizzano per capire quanto bene o male opera un Modello Classificatorio in termini di casi risolti con successo o di casi non risolti) è inferiore al 100% e (ii) che il caso affrontato ricade nei casi non risolti correttamente. Quale è il livello di accuratezza minimo da richiedere ad un modello classificatorio? A questa domanda si può rispondere solo considerando il contesto applicativo! A seconda dei casi si può accettare un livello d'errore più o meno elevato: un sistema di raccomandazione di pubblicità personalizzata può sbagliare senza procurare alcun danno diretto, ma ciò non è ammesso per un sistema di diagnosi medica o di guida autonoma!

Come precedentemente ricordato, il rilascio finale del Modello Classificatorio, alla fine della fase di training, deve subire un processo di validazione. Una validazione frettolosa o tecnicamente inadeguata (le tecniche ci sono!) è responsabilità dei tecnici informatici.

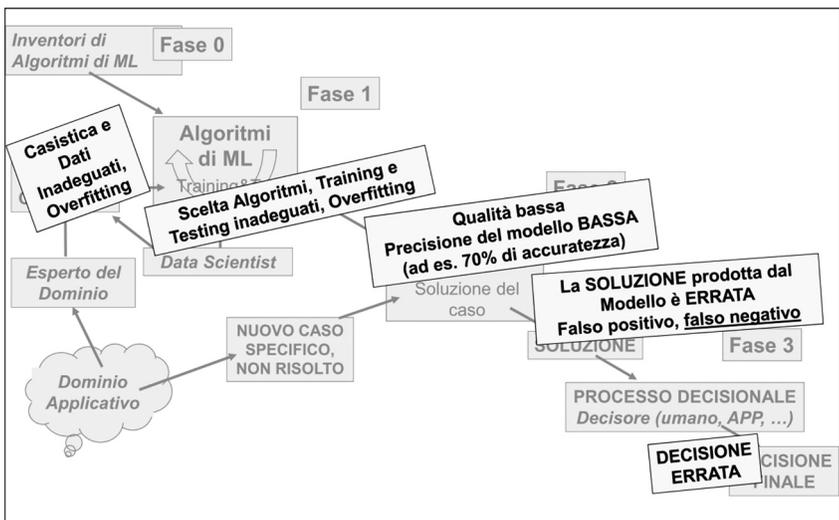
- c) *Varie cause di non funzionamento corretto del Modello Classificatorio.* Si entra qui nei dettagli tecnici, in quanto diversi sono i motivi tecnici durante la fase di Training che possono portare ad un modello 'non perfetto'. Senza approfondire tecnicamente, possiamo menzionarne alcune: attività preliminare di preparazione e valutazione dei Dati inadeguata, scelta degli algoritmi di ML non ottimali, tecniche e modalità di testing del modello classificatorio inadeguate o insufficienti, Training Set con pochi dati, overfitting²⁶, architettura inadeguata (nel caso del Deep Learning), e così via. In tutti questi casi la responsabilità è dei tecnici informatici.
- d) *Casistica inadeguata.* Il contributo dell'esperto di dominio è fondamentale (cfr. quanto illustrato al precedente pt. II lett. a). L'esperto di dominio deve fornire dei Dati adeguatamente numerosi ed equilibrati (i.e. non polarizzati) sulle varie casistiche e che includano tutti gli aspetti che direttamente o indirettamente caratterizzano il fenomeno in oggetto. È assai difficile che i Data Scientist possano sopperire a carenze dell'esperto di dominio, anche se comunque essi possono

26. *L'overfitting* è un fenomeno che può insorgere per diversi motivi, tra cui fra gli altri le modalità di training (si forza troppo l'algoritmo di ML ad adattarsi ai dati del Training Set, perdendo generalità) oppure un Training Set *polarizzato*, ossia contenente un numero non equilibrato di alcune tipologie di casistiche. Ad esempio, troppi casi di una casistica e troppo pochi di un'altra casistica causano la perdita della capacità di riconoscere le casistiche meno rappresentate e portano a prediligere le casistiche sovra-rappresentate.

condurre delle analisi preliminari dei dati che possono segnalare anomalie, da utilizzare poi con gli esperti per procedere a revisioni. Nel caso di carenze nell'attività appena richiamate a carico dell'esperto di dominio, appare corretto individuare specifiche responsabilità dell'esperto di dominio, e solo in casi più eccezionali, si ritiene che possa essere ragionevole imputare responsabilità al Data Scientist, sempre che comunque egli abbia svolto un'accurata attività di valutazione preliminare dei dati ricevuti.

- e) *Algoritmi di ML e DL.* Non appare realistico attribuire alcuna responsabilità, se le cose vanno male, ai ricercatori che hanno inventato gli algoritmi di ML e DL, lavorando questi in un ambito di ricerca, nel quale si lavora in laboratorio, c'è sempre un margine di ignoto e non c'è mai la pretesa di aver scoperto/proposto l'algoritmo perfetto e universale. Gli algoritmi utilizzati sono noti dalla letteratura ed è compito di coloro (Data Scientist) che affrontano gli specifici contesti applicativi selezionare i migliori tra quelli adeguati o capire che il problema non è risolubile o è risolubile solo in modo parziale.

Figura 10 - Possibili cause di una decisione finale errata



7. Conclusioni

In questo lavoro si è cercato di mostrare come le tecniche di IA e di ML estendono il campo d'azione tradizionale delle tecnologie digitali, permettendo di affrontare problemi nuovi sia dal punto di vista qualitativo (cioè relativo alle funzionalità e ai compiti che vengono affrontati) che quantitativo (ossia riguardante i grandi volumi – *Big Data* – di dati che si riescono ad elaborare). È quindi importante capire che siamo di fronte ad un nuovo e diverso tipo di informatizzazione, non affrontabile con le tecnologie ICT tradizionali: si possono affidare alla macchina nuovi compiti ma si deve essere pronti ad accettare delle prestazioni non ottimali. Va quindi sempre ben valutato come integrare questi nuovi sistemi nell'ambito di un processo decisionale, valutando problematiche di trasparenza, *accountability* e *liability*.

Nel caso del ML lo scenario di attori e processi che include lo sviluppo del Modello Predittivo, il suo utilizzo ed il conseguente sfruttamento dei risultati sono più articolati e coinvolgono molteplici e nuove figure professionali rispetto al caso dell'informatica tradizionale, con la necessità di competenze scarsamente reperibili oggi sul mercato.

Un commento finale va rivolto a coloro che immaginano futuri scenari apocalittici in un mondo in cui l'Intelligenza Artificiale sta entrando in tutti gli aspetti della nostra vita. Chi scrive ritiene che tali preoccupazioni siano eccessive e di fatto siano spinte dai numerosi ed eclatanti successi dell'IA di questi ultimi anni. Va però fatto notare che allo stato attuale delle conoscenze e dei risultati della ricerca non ci sono evidenze che rendano tali preoccupazioni effettivamente realistiche, giustificate o vicine nel tempo.

Infine, va anche ricordato che sarà sempre un uomo il responsabile di un possibile uso malvagio di queste tecnologie: nella misura in cui l'uomo utilizzerà questi strumenti nel modo corretto non dovremo preoccuparci.

Ugo Pagallo

Big data, open data e black box society

ABSTRACT: Il contributo intende affrontare tre dei principali temi su cui si snoda la rivoluzione attuale, a partire dalla Direttiva sull'informazione nel settore pubblico. Il presente lavoro descrive i concetti di *big data*, di *open data* e di *black box society*, prestando particolare attenzione a mettere in luce gli opportuni distinguo, si da chiarirne il significato e agevolare la comprensione del linguaggio della rivoluzione digitale. La preliminare definizione di questi concetti da un punto di vista filosofico è essenziale per la comprensione della sfida tecnologica e digitale, incentrata sul concetto di dato. Il presente lavoro prende infine in considerazione il tema della tendenziale imperscrutabilità di alcune scelte effettuate dai decisori pubblici o privati (*black box society*), conseguenza dell'impossibilità di garantire la perfetta veridicità e correttezza dei dati, nonché della difficoltà nel comprendere i meccanismi decisionali sottesi ai sistemi di intelligenza artificiale sollevano.

1. *Introduzione*

L'obiettivo di questo intervento è chiarire i rapporti tra '*big data*', '*open data*' e '*black box society*', in rapporto alla normativa sull'informazione nel settore pubblico (in particolare, le direttive PSI o 'Public Sector Information').

A questo fine, conviene procedere brevemente dalla definizione dei concetti in gioco. In primo luogo, con '*big data*', ci si riferisce al volume, varietà e velocità di trattamento dei dati. Avendo a mente che le nostre società ben possono definirsi società dipendenti dall'uso di tecnologie della informazione e comunicazione – a loro volta fondate sul trattamento algoritmico dei dati – è invalsa negli ultimi anni la metafora dei dati come nuovo petrolio del XXI secolo. Alla metafora occorre aggiungere come sia cambiato il motore della macchina, ossia il potere computazionale nel trattamento dei dati, e come siano inoltre stati affinati i programmi informatici che mettono a frutto tale potere con i relativi *big data*. Nel caso del potere computazionale, basti pensare alla cd. 'legge di Moore', per cui la potenza di elaborazione dei dati raddoppia ogni anno e mezzo; nel caso dei programmi informatici, il campo degli '*adversarial networks*' ne è un ottimo esempio.

Per ciò che riguarda invece la nozione di *open data*, essa va ricondotta all'idea d'accessibilità dei dati. La finalità dell'apertura è disciplinata giuridicamente sia attraverso licenze di diritto d'autore, come le licenze creative commons, sia attraverso forme di accessibilità ai dati compatibili con la tutela tradizionale dei diritti di proprietà intellettuale. Un esempio è dato dalla pubblicazione degli articoli scientifici in formato 'open access'. Nel settore pubblico, come diremo, tale apertura è disciplinata su scala europea con le direttive PSI.

Infine, la tesi che la nostra sia la “società della scatola nera”, secondo la fortunata formula di Frank Pasquale, rimanda alla nozione della segretezza degli algoritmi. A ben riflettere, decisioni sempre più cruciali per la nostra vita, dal credito bancario al lavoro, all'istruzione e via dicendo, sono prese da algoritmi che ora per ragioni tecniche, ora giuridiche, rimangono 'segreti'. Una ragione di tale segretezza risiede nella difficoltà di spiegare i risultati cui arriva un procedimento automatizzato, in modo più o meno intelligente, in ragione di un dato *input* di dati. Sul fronte giuridico, c'è poi un dibattito assai animato che riguarda, in Europa, l'art. 22 GDPR con riferimento al diritto alla spiegazione. Questo dibattito, negli Stati Uniti d'America, ha soprattutto riguardato la sfera dei diritti e garanzie in ambito penale. Emblematico è stato il caso *Loomis vs Wisconsin* (2016), in cui si è discusso fino a che punto la decisione di libertà vigilata per il signor Loomis potesse essere presa da un sistema esperto – COMPAS, nel caso di specie – piuttosto che dal giudice naturale. Salomonicamente, la Corte Suprema del Wisconsin ha ritenuto legittimo l'impiego del sistema esperto come mero supporto o aiuto alla decisione del giudice. Tuttavia, il punto critico è stato che l'avvocato del signor Loomis, quando ha richiesto di esaminare il codice del sistema esperto, sulla base del quale è stata respinta la domanda del proprio assistito, ha visto rigettata la richiesta per motivi di tutela della proprietà intellettuale. A quanto mi risulta, non abbiamo avuto fin qui casi analoghi in Italia (salvo qualche accenno della sesta sezione della Cass. Pen, 27 Novembre 2018, n. 14395). Sembra non di meno ragionevole sospettare che il trattamento automatizzato dei dati – sorta di oracolo consultato dai giudici – finirà per incidere sul modo in cui la Corte europea dei diritti umani ha inteso fin qui il rapporto tra il principio della parità delle armi processuali e quello della tutela della vita privata, ossia tra l'articolo 6 e l'art. 8 della CEDU. (Per chi fosse interessato, i dettagli nel mio saggio con Serena Quattrococo, *The Impact of AI on Criminal Law, and its Twofold Procedures*, in *Research Handbook on the Law of Artificial Intelligence*, a cura di W. Barfield e U. Pagallo, Elgar, Cheltenham, 2018, 385-409).

2. *La trasparenza e l'apertura dei dati nel settore pubblico*

Questa classe di problemi emersi nel diritto penale e civile, in linea di principio, non rileva nel diritto amministrativo, dove le parole d'ordine sono state (e tuttora sono) trasparenza e apertura dei dati. Questi principi si pongono, dunque, in antitesi ai dettami della *'black box society'*.

Come ricordato da Roberto Cavallo Perin nel suo intervento in questo volume, già il d.lgs. 12 febbraio 1993, n. 39, art. 1 elencava tra gli obiettivi dell'informatizzazione della pubblica amministrazione (P.A.), sia il miglioramento dei servizi sia la trasparenza e il contenimento dei costi. Naturalmente, di *'big data'* si sarebbe cominciato a parlare soltanto nei primi anni Duemila; ma, ora come allora, il "potenziamento dei supporti conoscitivi per le decisioni pubbliche", ai sensi del su citato art. 1, appare cruciale per chi, come la P.A., dispone di una massa sterminata di dati e informazioni: i *'big data'*, appunto.

La posta in gioco riguarda la possibilità d'innescare un circolo virtuoso che, partendo dalla raccolta e uso dei dati, giunga al trattamento e conoscenza dei flussi informativi in ragione di principi di trasparenza e accessibilità, fino al riuso di quei dati e alla scalabilità dell'intero processo. Per dirla in inglese, *'a win-win scenario'*.

Molti tuttavia hanno avvertito un problema, a proposito del nuovo regolamento europeo sulla protezione dei dati (GDPR). A questo riguardo, è importante evitare incomprensioni, distinguendo accuratamente tra privacy e protezione dati. Tanto è vero che la Carta dei diritti fondamentali dell'Unione europea si occupa della privacy all'art. 7 e della protezione dei dati personali all'art. 8. Mentre il diritto alla privacy mira a tutelare l'opacità delle persone – secondo l'immagine della filosofa tedesca Hannah Arendt – viceversa, il diritto alla tutela dei dati personali mira a garantire la trasparenza con la quale i dati sono raccolti, usati e riusati. La trasparenza della P.A. può perciò andare d'accordo con il quadro di garanzie predisposto per la tutela dei dati personali, perché in entrambi i casi ciò che è in gioco è il principio della trasparenza. Più che un gioco a somma zero, il rapporto tra trasparenza della P.A. – secondo un principio che possiamo far risalire al filosofo tedesco Immanuel Kant – e trasparenza nel campo della protezione dei dati suggerisce ancora una volta un *'win-win scenario'*.

2.1 *La direttiva sull'apertura dei dati e il riutilizzo dell'informazione pubblica*

La tesi è del resto confermata dalle direttive europee sull'informazione nel settore pubblico, o 'Public Sector Information' (PSI). L'idea è che aprendo (open access), i tanti dati del settore pubblico (*big data*), non soltanto si rafforzino i principi di trasparenza e accessibilità che vigono nel settore; ma, tale apertura dei *big data* pubblici si presenta anche come la chiave di volta per potenziare lo sviluppo dell'economia e della ricerca scientifica. L'esempio, in questo caso, è dato da un settore strategico, come quello dell'intelligenza artificiale (IA). Ambiti quali il data mining, il machine learning, o il deep learning si nutrono di dati.

La prima direttiva PSI risale a quasi vent'anni fa (2003/98/CE), seguita nel 2013 dalla seconda direttiva PSI (2013/37/CE). Quest'ultima è destinata a rimanere in vigore fino al 17 luglio 2021, quando sarà sostituita dalla nuova direttiva 2019/1024/CE. Sono tre i motivi principali che hanno spinto il legislatore europeo a emendare ancora una volta sé stesso. In primo luogo, c'è (stata) la volontà di spingere verso un'apertura by default dei dati pubblici, nel senso che il regime standard per i dati nel settore pubblico, specie per quanto riguarda musei, archivi e biblioteche, dovrebbe essere quello della piena accessibilità (salvo, ovviamente, le eccezioni al principio, disciplinate con regimi residuali). In secondo luogo, la disciplina dell'apertura dei dati deve tenere in conto il costo di questo processo, tarato sui costi marginali di riproduzione. In terzo luogo, risulta cruciale il formato in cui i dati devono essere rilasciati, vale a dire un formato automatizzabile. Insieme agli standard giuridici, gli standard informatici giocano un ruolo determinante nelle politiche digitali dei legislatori più accorti.

Preme qui segnalare un aspetto di queste politiche digitali (anche, ma non solo, per la P.A.): i dati sono fondamentali per la ricerca scientifica. Come detto, abbiamo bisogno di dati per sviluppare uno dei settori ritenuti strategici dalla Commissione europea: il campo dei sistemi di intelligenza artificiale. E siccome abbiamo bisogno di dati, uno dei modi più efficaci per raggiungere questo obiettivo passa attraverso l'apertura dei dati nel settore pubblico; opportunità rivolta al mondo della ricerca e le imprese di riusare questi dati come meglio credono. Un esempio rende bene quest'idea, semplice ma potente. Quando dieci anni fa i servizi di traduzione automatica di Google fecero un salto di qualità, per molti versi impressionante, l'efficacia di 'Google translate' la si è dovuta al fatto che la potenza computazionale della multinazionale americana aveva saputo

sfruttare in chiave statistica il tesoro dei testi normativi dell'Unione europea tradotti in 24 lingue diverse. Avremo mai qualche impresa italiana capace di sfruttare l'apertura dei dati nel settore pubblico in modo analogo?

3. Conclusioni

Tirando le fila del discorso, riassumiamo i risultati della nostra riflessione sul ruolo che gli open e i *big data* hanno nell'odierna normativa sull'informazione nel settore pubblico, ai tempi delle società algoritmiche, o *'black box society'*. Questa prospettiva sulle direttive PSI, succedutesi in quasi vent'anni, consente di chiarire cos'è cambiato, o meno, dai tempi del d.lgs. 12 febbraio 1993, n.39, ricordato in precedenza.

Innanzitutto, una novità: la mancata disponibilità di dati rappresenta spesso un grave ostacolo per la ricerca scientifica e in particolare per il settore dell'intelligenza artificiale. L'apertura dei dati, a partire dal settore pubblico, può rappresentare un volano per questo tipo di ricerca e sviluppo tecnologico, ritenuto cruciale, anche ma non solo, dalla Commissione europea.

In secondo luogo, l'apertura dei dati può essere volano per lo stesso sviluppo economico. Siccome tale sviluppo passa attraverso il potenziamento della ricerca scientifica e tecnologica che si avvale (e si basa) sul trattamento algoritmico dei dati, lo sviluppo economico non può che giovare degli obiettivi perseguiti dalle direttive PSI.

In terzo luogo, l'apertura dei dati puntella il tradizionale principio della trasparenza per la P.A., laddove abbiamo visto che trasparenza e tutela dei dati personali non sono principi incompatibili, ma ben possono rinforzarsi l'un l'altra.

Infine, in quarto luogo, l'apertura dei dati e il relativo principio di trasparenza si profilano come pilastri della P.A. contro i dettami della *black box society*. Informatizzando i *big data* della P.A. attraverso tecniche open, l'obiettivo è il miglioramento dei servizi, l'efficienza e il contenimento dei costi, attraverso un uso intelligente del digitale nel settore pubblico. Trasparenza e apertura dei dati ne sono la chiave.

Andrea Simoncini

Intelligenza artificiale e futuro delle libertà costituzionali

ABSTRACT: L'articolo sviluppa la prospettiva di uno studioso di diritto costituzionale, che delinea le perplessità sollevate dagli algoritmi sotto il profilo del rispetto delle libertà fondamentali. Il presente lavoro evidenzia come oggi l'assunzione di decisioni sia spesso relegato alle tecnologie, sicché le decisioni assunte appaiono spesso in contrasto con i principi del diritto costituzionale. Esse appaiono non comprensibili, perché non conoscibile è la logica, il nesso seguito dall'intelligenza artificiale; sono inevitabili ed esclusive, benché sia vigente una regola generale (art. 22 GDPR) che vieta l'assunzione di decisioni secondo modalità esclusivamente algoritmiche e infine non garantiscono il principio di non discriminazione. Il contributo, infatti, presta attenzione al ruolo svolto dalla base conoscitiva su cui l'algoritmo decide poiché essa è rappresentativa della realtà, non depurata dalle distorsioni che sono presenti nella società, con il rischio che l'algoritmo perpetui le discriminazioni sociali.

1. *Introduzione*

Affronterò il tema delle relazioni tra l'impatto che hanno le tecniche di intelligenza artificiale e le tematiche del diritto costituzionale. Vorrei partire dalla stringente attualità: Google ha deciso di togliere la licenza a Huawei perché il 15 maggio Trump ha emanato un *executive order* con il quale ha deciso di mettere sotto sicurezza l'information and technologies americana, vietando la possibilità di utilizzare o di fornire parte delle apparecchiature alla Huawei. Prima c'era stato un altro episodio di *trade war* (l'arresto della CFO di Huawei), ma come mai? Negli ultimi cinque anni c'è stata una scalata delle aziende cinesi nel settore tecnologico, che ha prodotto nell'amministrazione americana già dal giugno 2018 una forte preoccupazione per l'aggressione da parte della Cina, che sta investendo anch'essa (esattamente come gli Usa) nell'intelligenza artificiale, nei veicoli autonomi, nella realtà aumentata, nella blockchain: sono tutte tecnologie che interessano molto la difesa. Il campo di applicazione è profondamente complesso e crescente dal punto di vista dell'impatto.

2. *Le decisioni algoritmiche*

Oggi un numero crescente di decisioni che incidono sulle libertà individuali e collettive sono prese da intelligenze artificiali e non da uomini (o a volte gli uomini ne sono solo parte). Sempre più noi ci troviamo di fronte a decisioni che incidono anche molto profondamente sull'area della libertà costituzionale, decisioni che sono prese da meccanismi automatizzati. Un passaggio è fondamentale per il diritto costituzionale: fino ad oggi abbiamo inquadrato le tecnologie come strumenti per eseguire decisioni che gli umani hanno preso: qui il cambio di scenario è profondo. In questi casi non abbiamo tecnologie che si limitano ad assumere decisioni che sono state preventivamente impartite dagli umani, ma sono le tecnologie stesse a prendere queste decisioni: questo è il cambio di paradigma. Questo influsso della tecnologia sulle decisioni umane avviene in modo diretto cioè perché noi espressamente deleghiamo a dei sistemi automatizzati di prendere una decisione. È il caso delle decisioni algoritmiche (caso *Compas*), di decisioni amministrative prese per via algoritmica, cioè abbiamo un soggetto umano (Pubblica amministrazione, compagnia, società di assicurazione) che decide di affidare una decisione, o parte della stessa o una parte significativa di una decisione ad un sistema automatizzato. Ma c'è anche un altro modo che queste stesse tecnologie hanno di influenzare una decisione umana che è indiretta, ma non per questo meno importante. Il fatto che oggi sistemi tecnologici automatizzati sono il meccanismo ordinario attraverso il quale noi acquisiamo informazioni sulla realtà. Un accenno: su quale base conoscitiva oggi noi acquisiamo informazioni? Nel 99% della base conoscitiva c'è una piattaforma informatica, es. informazioni su politica, economia, questo percorso indiretto, perché la tecnologia è ciò che ci trasferisce l'informazione, c'è un monopolio dell'informazione che è il motore di ricerca Google. Questa forma indiretta è la base delle *fake news*, di Cambridge Analytica.

2.1 *Il caso Loomis e la mobilità degli insegnanti nella c.d. Buona Scuola: USA e Italia a confronto*

Soffermiamoci sul caso in cui un soggetto umano, un'amministrazione, decida attraverso un processo automatizzato (caso *Loomis* già citato). La vicenda italiana attiene invece alla Legge n. 107/2015 la cd. riforma della buona scuola: brevemente, molti docenti erano da sistemare in sedi di servizio, allora ci si affida ad un algoritmo prodotto da una società: i risultati sono stati irragionevoli. Il sindacato (Gilda Unams) dei do-

centi chiede l'accesso che, però, il MIUR nega sostenendo che il codice sorgente in questione non è un documento amministrativo perché è un software, inoltre esso appartiene ad una società ed è per questo protetto dall'ingegno.

Il TAR Lazio recide le due argomentazioni, da un lato perché la legge 241/90 e del CAD hanno portato ad una evoluzione del concetto di documento amministrativo. Ma l'*obiter dictum* interessante si riferisce al fatto che il TAR Lazio riconosce che è possibile utilizzare un algoritmo perché ci troviamo di fronte ad una amministrazione vincolata, qui non c'è nessuno spazio di discrezionalità, in questo caso bisognava applicare tutte regole che sono predeterminate, non c'è da imparare nulla della realtà, c'è una legge generale che è il concorso, poi ci sono i casi particolari da sussumere. Inoltre, siccome si riconosce un diritto di accesso pieno, di tipo amministrativo e non civico, non c'è proprietà intellettuale che si può opporre. Il TAR Lazio del 2018 torna su questo punto con una presa di posizione più forte: quand'anche ci trovassimo di fronte ad un'amministrazione vincolata, c'è una parte dell'attività amministrativa ("solo l'attività dianoetica dell'uomo può svolgere questa azione amministrativa") e "pro futuro in un'ottica conformativa, la pubblica amministrazione deve predisporre, unitamente a strumenti telematici di semplificazione" (un altro filone di contenzioso sono le piattaforme automatiche di ricezione delle domande, che quando sono sbagliate vengono automaticamente respinte dalla piattaforma automatica stessa), quindi quando si prevede solo una procedura informatica, bisogna comunque che ci sia un'altra procedura umana cui poter ricorrere, nel caso in cui quella informatica o non funzioni o abbia un mal funzionamento.

Il Consiglio di Stato del 2019 da un lato afferma che l'uso delle procedure algoritmiche è molto importante nei casi in cui l'attività è routinaria, c'è da semplificare (laddove sia definito l'atto generale), il Consiglio di Stato di fatto appoggia l'idea dell'uso di queste procedure solo in ipotesi di attività vincolata. Non si possono lasciare spazi applicativi discrezionali, di cui l'elaboratore è privo, ma deve prevedere una soluzione definita per tutti i casi possibili, anche i più improbabili (non stiamo parlando però degli algoritmi di cui ci occuperemo). Qui si pone il problema che voglio affrontare, soprattutto dal lato del diritto costituzionale che si occupa di proteggere le libertà dai poteri.

Il problema è la comprensibilità/conoscibilità di questi algoritmi, perché noi ci troviamo di fronte ad un sistema che prende decisioni, ma non possiamo ricostruire la logica, il nesso, l'iter, tutto quello strumentario che il giudice amministrativo ha creato per poter controllare un atto giuridico.

Su questo è molto importante considerare il GDPR, di cui l'articolo 22 pone dei principi molto forti ovvero il diritto di ciascuno a conoscere l'esistenza di un processo decisionale automatizzato che lo riguardi (informazioni significative sulla logica utilizzata): ognuno ha diritto di sapere se la decisione è stata presa in maniera automatica.

3. I principi del processo decisionale automatizzato

Il diritto a non essere sottoposto ad una decisione che produca effetti rilevanti basata unicamente su un processo automatizzato poggia su alcuni principi:

1. il principio di conoscibilità. Dal Prof. Tasso abbiamo colto come i nuovi sistemi di intelligenza artificiale non è detto che siano comprensibili in termini di logica deterministica (“i nuovi sistemi algoritmici, soprattutto di apprendimento neurale arrivano alla conclusione, predicono con una accuratezza elevatissima un evento, senza far la fatica di spiegare perché”): questo è un problema se sulla base di quella previsione si prende un atto capace di incidere sulle libertà di una persona.
2. Non esclusività: non potrà mai esistere una decisione presa esclusivamente sulla base di un algoritmo, ma questo è un principio difensivo debole, perché le eccezioni sono ampie (art.22, comma 2), tra cui c'è la presenza del consenso, ma questa è una *fiction* perché anche se c'è consenso, non vuol dire che non c'è lesione del diritto. Infine, la forza pratica degli algoritmi perché nella misura in cui inseriamo in un processo decisionale un procedimento previsionale, questo tenderà a catturare la decisione perché solleva dalla necessità di dover motivare, perché ho già una motivazione fatta, perché se motivo per discostarmi, mi assumo le responsabilità e sarà sempre più difficile.
3. Il principio di non discriminazione, scomparso dall'art. 22 GDPR, che più aveva fatto discutere il *Compas case*. Gli algoritmi si basano su *set* di dati esistenti, quella forma di apprendimento che fanno, lo fanno sull'essere, non su un essere depurato delle distorsioni che sono state create dalla società; per esempio, il tasso di recidiva in USA si calcola sul tasso di *imprisonment* che c'è oggi (che dice che l'85% sono afroamericani, cosicché a parità di condizione, un afroamericano aveva un indice di possibilità di reiterare il reato molto più alto di altri). Per noi è un errore inaccettabile perché è la vecchia fallacia di determinare il dover essere dall'essere. Altri esempi di *bias* preso da

Google *translate*: la traduzione dal turco “il dottore” era “*he is a doctor*” e “infermiera” era tradotto come “*she is a nurse*”: il traduttore era fondato sulle ricorrenze e la ricorrenza in tema di dottore è maschile e quella di infermiera è femminile, così da incidere sulla traduzione.

4. *Conclusioni: il costituzionalismo precauzionale*

Necessario è inserire nel processo di programmazione dell'algoritmo la decisione umana, quindi il vero problema è il recupero di certe forme di tutela dentro il processo di programmazione, per questo il futuro è il dialogo tra giuristi, computer science, tecnologi: nel momento in cui l'algoritmo è costruito, fare i comitati etici (come ha provato a fare Google), fare le leggi, tutto ciò che interviene dopo che questi enormi meccanismi economici sono messi in moto. Ecco perché ci deve essere un costituzionalismo precauzionale; ci deve anche essere un costituzionalismo ibrido che sappia dialogare con le diverse scienze.

Accountability delle decisioni algoritmiche*

ABSTRACT: Il contributo si concentra sui limiti che le pubbliche amministrazioni nazionali devono porre all'introduzione dell'intelligenza artificiale nell'esercizio della funzione amministrativa. La pubblica amministrazione deve essere trasparente nell'utilizzo delle stesse e responsabile per le decisioni così assunte. Il presente lavoro si concentra sulla necessità che il percorso logico che ha portato alla decisione debba essere reso noto, insieme ai dati che sono stati utilizzati, ai risultati che si attendevano e a quelli che si sono ottenuti. L'articolo pone dunque l'accento sulla necessità di un ruolo rafforzato della motivazione, affinché possa essere rafforzata la fiducia dei cittadini nei confronti dell'operato dell'amministrazione algoritmica.

1. *La construcción de la gobernanza inteligente en tiempos de la inteligencia artificial*

Las administraciones públicas utilizan cada día más la inteligencia artificial en el desarrollo de sus actividades y en la prestación de los servicios públicos. Aunque no lo están haciendo aún con la misma intensidad que el sector privado, cada vez son más las administraciones públicas que utilizan la inteligencia artificial para analizar datos (por ejemplo, para predecir el riesgo de incendios en edificios o identificar los locales que deben ser objeto de inspección), procesar el lenguaje natural (por ejemplo, para revisar las solicitudes formuladas por la ciudadanía o detectar irregularidades y fraudes en las licitaciones públicas), reconocer imágenes (por ejemplo, para contar personas en espacios públicos o identificar posibles

* Este texto fue preparado para su presentación en el Primo Convegno Nazionale Coordinamento Dottorandi in Diritto Amministrativo. L'amministrazione pubblica con i *Big data* Torino, 20-21 maggio 2019 a partir del artículo Cerrillo i Martínez, Agustí. (2019). How can we open the black box of public administration? Transparency and accountability in the use of algorithms. *Revista Catalana de Dret Públic*, (58), 13-28. <https://doi.org/10.2436/rcdp.i58.2019.3277>.

infractores), tomar o facilitar la toma de decisiones (por ejemplo, para adjudicar ayudas públicas, para determinar las calles donde enviar policías a patrullar, para identificar las escuelas que pueden tener un mayor índice de abandono escolar o sobre el tratamiento para una determinada enfermedad) o para personalizar los servicios públicos (por ejemplo, para prestar servicios de información, asesoramiento y atención ciudadana).

De hecho, la extensión del uso de la inteligencia artificial entre las administraciones públicas está empezando a dibujar un nuevo modelo de gestión pública, la gobernanza inteligente, basado en el uso intensivo de los datos a través de algoritmos y en una mayor colaboración con la ciudadanía a través del uso de las tecnologías de la información y la comunicación para obtener mejores resultados y una mayor apertura y rendición de cuentas a la ciudadanía.

La gobernanza inteligente puede generar numerosos beneficios tanto a las administraciones públicas como a la ciudadanía en entornos complejos y contextos multifactoriales, en los que se debe tener en cuenta grandes volúmenes de datos para poder tomar decisiones de calidad. Sin embargo, la gobernanza inteligente también puede generar numerosos riesgos. Confiar ciegamente en la bondad de los algoritmos puede llevar a decisiones públicas erróneas, sesgadas, discriminatorias o que vulneren los derechos fundamentales. Por ello, hay que garantizar que la toma de decisiones a través de algoritmos y, en general, la gobernanza inteligente se fundamente en los principios de la buena gobernanza como la transparencia, la participación, la rendición de cuentas y la eficacia para garantizar su legitimidad y efectividad.

2. *La opacidad de los algoritmos*

Los algoritmos son considerados como cajas negras (*black box*). Por ello, el uso de algoritmos por las administraciones públicas puede comportar que estas sean menos transparentes. Además, la opacidad de los algoritmos puede impedir que las administraciones puedan rendir cuentas por su actuación.

La opacidad de los algoritmos puede estar motivada por causas técnicas, jurídicas u organizativas.

En primera instancia, la opacidad de los algoritmos puede estar motivada por causas técnicas. En concreto, la opacidad puede ser fruto de la complejidad de los algoritmos o de su carácter dinámico. Hay algoritmos

en los que es difícil saber cómo funcionan o qué datos tienen en cuenta para producir un determinado resultado.

Desde un punto de vista jurídico, la opacidad de los algoritmos puede deberse a la existencia de normas o cláusulas contractuales que limiten el acceso a su código fuente para proteger otros bienes o derechos como el secreto empresarial, la propiedad intelectual e industrial, la protección de los datos personales o la seguridad pública. La opacidad también puede justificarse para proteger la confidencialidad o el secreto del proceso de toma de las decisiones públicas con el fin de evitar el conocimiento de decisiones en formación.

Finalmente, desde el punto de vista organizativo, la opacidad puede ser el resultado de la inexistencia de información sobre los algoritmos. En unos casos porque su código fuente no está en manos de las administraciones públicas que los utilizan. En otros casos porque las administraciones públicas no han formalizado la decisión de utilizar un algoritmo ni los fines ni las fuentes o los resultados obtenidos a través de los algoritmos. En última instancia, la opacidad también puede perseguir evitar la elusión de su aplicación.

3. *La transparencia de los algoritmos*

La transparencia es un principio de las administraciones públicas que también debe guiar el uso de los algoritmos en su actividad.

Para evitar la opacidad de los algoritmos a la que hemos hecho referencia en el apartado anterior es necesario que las administraciones públicas adopten determinadas medidas que faciliten su transparencia. Estas medidas deben permitir no sólo el acceso a los algoritmos sino también la comprensión de su funcionamiento y del fundamento de las decisiones que se toman a través suyo.

En primer lugar, para garantizar la transparencia de los algoritmos es necesario que cualquier persona pueda acceder a su contenido. El contenido del algoritmo puede ser conocido a través del acceso al código fuente pero también a través del acceso a la documentación que contenga una descripción del algoritmo, de los datos que utiliza, de cómo funciona, de los modelos que utiliza y de los resultados que produce y cómo los produce.

Las administraciones públicas pueden dar a conocer los algoritmos que utilizan a través de su difusión proactiva. Al respecto, por ejemplo, el artículo 5 Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la

información pública y buen gobierno dispone que las administraciones públicas “publicarán de forma periódica y actualizada la información cuyo conocimiento sea relevante para garantizar la transparencia de su actividad relacionada con el funcionamiento y control de la actuación pública”. Sin embargo, no podemos desconocer que la legislación de transparencia no prevé una obligación específica de difundir información relativa a los algoritmos como sí lo hacía la Ley 30/1992, de 26 de noviembre, de régimen jurídico de las administraciones públicas y del procedimiento administrativo común que disponía que las administraciones públicas debían difundir públicamente las características de los programas y aplicaciones informáticas que fueran a usar y que hubieran sido aprobadas con carácter previo (artículo 45.4). Actualmente, las administraciones públicas no están difundiendo información relativa a los algoritmos que están utilizando. La simple búsqueda en los portales de transparencia de las administraciones públicas lo confirma.

Las administraciones públicas también pueden dar a conocer los algoritmos que utilizan a través del acceso a la información. Al respecto, podemos observar cómo la Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno reconoce el derecho de todas las personas a acceder a la información pública. De acuerdo con esta norma, la información pública es cualquier contenido o documento, cualquiera que sea su formato o soporte, que esté en poder de una administración pública que haya sido elaborado o adquirido en el ejercicio de sus funciones forma parte del objeto del derecho de acceso. Por lo tanto, los algoritmos, así como toda la información vinculada a ellos, son susceptibles de ser objeto de una solicitud de acceso a la información pública. Así lo ha entendido también la Comisión de Garantía del Derecho de Acceso a la Información Pública (GAIP) en relación a la reclamación interpuesta por un profesor de secundaria a quien se le denegó el acceso al algoritmo matemático que se empleaba para designar los miembros de los tribunales correctores de las pruebas de acceso a la universidad. En su resolución de 21 de septiembre de 2016, de estimación de las reclamaciones 123/2016 y 124/2016, la GAIP pone de manifiesto que el algoritmo, aunque pueda encontrarse en lenguaje matemático o informático, no deja de ser un tipo de información que estando en poder de la Administración pública constituye información pública a los efectos del ejercicio del derecho de acceso que reconoce la legislación vigente.

En segundo lugar, es importante recordar que el acceso a la información puede ser limitado cuando pueda suponer un perjuicio a diversos bienes y derechos identificados por el legislador como la seguridad nacional

y la seguridad pública, los intereses comerciales, la propiedad intelectual e industrial o la confidencialidad o secreto en los procesos de toma de decisión. También debemos tener en cuenta que en aquellos casos en los que los algoritmos hayan sido adquiridos de terceras personas puede ser de aplicación lo dispuesto en la legislación de contratos del sector público y, por tanto, que el contratista declare su confidencialidad. En cualquier caso, la aplicación de estos límites no es automática y se debe valorar el perjuicio que pueda suponer el acceso al código fuente del algoritmo o a la información que se haya podido generar sobre su diseño y funcionamiento.

Sin embargo, a pesar del reconocimiento del derecho de acceso a la información pública, los pocos estudios empíricos que existen muestran como el ejercicio de este derecho no está siendo un mecanismo efectivo para acceder a los algoritmos.¹

Además, el acceso a los algoritmos es necesario pero no suficiente para garantizar su transparencia. El acceso al código fuente del algoritmo no lo hace necesariamente más transparente. Hay que garantizar también la comprensión de su funcionamiento. Para garantizar la comprensión del algoritmo, será útil que las administraciones públicas faciliten una descripción del algoritmo en lenguaje natural.

4. *La rendición de cuentas (accountability) sobre el uso de los algoritmos*

La gobernanza inteligente exige que las administraciones públicas no sólo sean transparentes sino también que rindan cuentas por las decisiones que toman a través algoritmos. La rendición de cuentas se refiere a la obligación de explicar y justificar las decisiones adoptadas por los algoritmos. A través de la rendición de cuentas, las administraciones públicas justifican el uso de los algoritmos en la toma de decisiones y explican cómo funcionan, qué datos utilizan, qué resultados se esperan y qué resultados se han obtenido.

La rendición de cuentas facilita el control de la actividad desarrollada por las administraciones públicas a través de algoritmos. También permite la detección de errores y sesgos. En última instancia, la rendición de

1. Por ejemplo, BRAUNEIS ROBERT - GOODMAN ELLEN P. (2017), *Algorithmic transparency for the smart city*. *Yale Journal of Law & Technology*, 20, 104-176, o FINK KATHERINE (2017), *Opening the government's black boxes: freedom of information and algorithmic accountability*. *Information, Communication & Society*, 1-19.

cuentas fortalece la confianza ciudadana en las administraciones públicas cuando utilizan algoritmos en la toma de decisiones.

Más allá de poder acceder y conocer el algoritmo es necesario que tanto el diseño de los algoritmos por las administraciones públicas como su aplicación se haga en el marco de un procedimiento que garantice los principios de transparencia y participación y los derechos de los interesados de acuerdo con lo dispuesto en la legislación vigente respecto a la elaboración de disposiciones de carácter general y de actos administrativos.

La rendición de cuentas también se puede realizar dando una explicación sobre el funcionamiento del algoritmo y sobre las razones que justifican las decisiones adoptadas por el algoritmo. Sin embargo, a nuestro entender esta explicación debe ir más allá de la motivación de los actos administrativos prevista en la Ley 39/2015, de 1 de octubre. Tanto en cuanto a los actos administrativos que se deben motivar como al contenido de la motivación. En última instancia, debemos tener presente que el uso de los algoritmos en la toma de decisiones públicas cambia la forma en que éstas se generan y, por tanto, cómo se motivan los actos administrativos. Por ello, también se debe poder explicar cómo la administración pública ha llegado a aquella decisión, cuál ha sido el proceso, qué datos se han tenido en cuenta o qué objetivos se perseguían.

Sin embargo, no existe en la actualidad una obligación explícita en esta dirección. Por este motivo, sería necesario avanzar hacia la definición de una obligación de las administraciones públicas de explicar el algoritmo, sus finalidades y funcionamiento así como los datos utilizados que permitan su conocimiento por parte de los empleados públicos y, en última instancia, su control por parte de la ciudadanía. Paralelamente, se debería reconocer un derecho a obtener una explicación del uso de los algoritmos para las administraciones públicas.

En esta dirección, con numerosas limitaciones como han puesto de relieve distintas autoras, el Reglamento (UE) 2016/679 del Parlamento Europeo y del Consejo, de 27 de abril de 2016, relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales y a la libre circulación de estos datos y por el que se deroga la Directiva 95/46/CE (Reglamento general de protección de datos) reconoce un derecho a la explicación de los tratamientos automatizados en el ámbito de la protección de los datos personales.

Asimismo, las propuestas realizadas para reconocer la trazabilidad y la explicabilidad como principios éticos que deben regir el funcionamiento de los algoritmos (por ejemplo, entre otros, por el Grupo de expertos de alto nivel sobre inteligencia artificial creado por la Comisión Europea).

5. Conclusiones

La gobernanza inteligente genera nuevos riesgos para las administraciones públicas. En estas páginas hemos centrado nuestra atención en la opacidad de los algoritmos como uno de los principales riesgos con que se enfrentan las administraciones públicas cuando utilizan algoritmos en la toma de decisiones. También hemos podido reconocer la importancia de la transparencia y de la rendición de cuentas como principios para darles respuesta.

Sin embargo, a lo largo de las páginas anteriores hemos podido constatar algunas de las limitaciones que presentan los mecanismos de transparencia y rendición de cuentas cuando las administraciones públicas utilizan algoritmos en la toma de decisiones. Si bien, es evidente que estos no han sido regulados para dar respuesta a la gobernanza inteligente, las administraciones públicas tampoco están impulsando los cambios suficientes para ser más transparentes ni tampoco para rendir cuentas mejor cuando utilizan algoritmos. Además, la complejidad que caracteriza toda la información relativa a los algoritmos y su funcionamiento limita significativamente el papel que puede tener la ciudadanía para controlar la actividad de las administraciones públicas a través de los algoritmos.

A la vista de las limitaciones de los mecanismos de transparencia y rendición de cuentas que se han observado a lo largo del trabajo, consideramos que para lograr una buena gobernanza inteligente podría ser útil, sino necesario, que las administraciones públicas se dotasen de un marco institucional que permitiese garantizar la seguridad -tecnológica y jurídica-, la transparencia y la responsabilidad en el uso de los algoritmos por las administraciones públicas. Este marco institucional también debería impulsar la innovación tecnológica en las administraciones públicas con el fin de transformar su funcionamiento y sus relaciones con la ciudadanía (por ejemplo, prestando servicios públicos de manera personalizada y proactiva). En última instancia, debería ser un nodo de interacción entre todos los actores implicados en el desarrollo de la gobernanza inteligente.

Opacità algoritmica e trasparenza delle decisioni amministrative

ABSTRACT: L'articolo affronta il tema della *governance* e della regolazione delle piattaforme, quali nuovi attori sulla scena dell'innovazione tecnologica, ponendo particolare attenzione al "potere autoregolatorio" di cui godono che si esplica nella definizione di regole sia di interazione tra la piattaforma stessa e i soggetti che con essa si rapportano, sia di controllo ed esclusione dall'offerta di servizi che esse erogano.

1. *Introduzione*

Mi concentrerò sul tema della *governance* e regolazione delle piattaforme e di come queste -avendo poteri di regolazione- debbano a loro volta essere oggetto di regolazione. Questo è attualmente motivo di dibattito internazionale tra studiosi della concorrenza e di teoria della regolazione.

Le piattaforme che sono a noi più note (Google, Alibaba, Booking, Airbnb) dispongono di potere autoregolatorio; cioè autodefiniscono regole perché fissano l'architettura del sistema nel quale noi operiamo, le regole di interazione tra la piattaforma stessa e i soggetti commerciali che vi operano e tra gli utilizzatori attraverso la definizione di *format* contrattuali che limitano la nostra autonomia contrattuale; inoltre stabiliscono forme di controllo e sanzioni che possono anche portare all'esclusione dalla piattaforma. Questi poteri di regolazione sono forme di autoregolazione: due orientamenti del Consiglio d'europa e del Parlamento europeo che mettono proprio in mostra il fatto che i sistemi di *machine learning* hanno una crescente capacità di predire le scelte ma anche di influenzare le emozioni, i pensieri e di alterare e anticipare il corso delle azioni talvolta in maniera subliminale, per questo sollecitano gli Stati membri ad adottare strumenti di *empowerment* che promuovano la consapevolezza del pubblico e la trasparenza delle decisioni algoritmiche intervenendo sulle piattaforme e attraverso strumenti basati anch'essi sull'intelligenza artificiale.

2. *Le ragioni dell'intervento regolatorio sulle piattaforme*

Quali sono le domande e l'interesse generale che giustificerebbe un intervento regolatorio sulle piattaforme; che tipo di intervento regolatorio sarebbe auspicabile e che produca maggiore trasparenza e secondo quale modello di *governance*?

L'autoregolazione delle aziende prima citate trova la sua base in “*Code Law*” di Lessig ove si sostiene che le regole tecniche sono diritto, questo è stato poi portato avanti e sviluppato nella formula di “intermediari regolatori” (come all'inizio i certificatori e ora esteso anche alle piattaforme). Sundararajan e Cohen sostengono che le piattaforme sono un anello imprescindibile della catena regolatoria e non possiamo percepirli come un soggetto da regolare, proprio perché dispongono di una capacità di conoscere i soggetti da regolare. Ad esempio: Airbnb conosce meglio di qualsiasi amministrazione locale i locali che vengono dati in affitto e per questo diventa il soggetto che meglio di altri può raccogliere l'imposta relativa agli affitti su persone per brevi periodi (questo succede per esempio a Lisbona), usata anche come *enforcement* della regolazione (Danimarca e Londra), cioè di quella normativa che prevede alcuni limiti che per gli affitti brevi non possano esserci più di un certo numero di persone nell'abitazione. Fare controlli di questo tipo è possibile per mezzo di queste piattaforme che regolano milioni di dati sul territorio. C'è una delega di funzione regolatoria alla piattaforma. Le decisioni di tipo regolatorio sono granulari, personalizzate perché sono definite mediante algoritmi e alimentate da dati e si basano sui sistemi reputazionali che definiscono il successo di queste piattaforme; sono anche strumenti informativi verso i terzi. Il problema sta nella possibilità che le piattaforme hanno di regolare queste informazioni: non c'è solo un problema di trasparenza ma anche di manipolazione, di discriminazione.

La conoscenza della logica dell'algoritmo non è sempre la soluzione, perché il consenso – base legittimante delle decisioni algoritmiche – è raccolto in maniera accidentale. Inoltre, l'incertezza sulla titolarità del dato che viene prodotto nelle piattaforme o attraverso gli strumenti che utilizziamo è un altro problema e, ancora, una volta prodotto, chi lo può riutilizzare? Chi può accedere? A volte le piattaforme negano l'accesso.

3. *Proposte di regolazione per la trasparenza delle decisioni algoritmiche*

Quali sono le proposte in campo per avere la trasparenza delle decisioni algoritmiche?

1. Lasciar fare: queste decisioni sono efficienti e hanno portato innovazione; questa è la strada di lasciar fare all'autoregolazione e di promuoverla (c'è un po' di intervento eteronomo ed è un po' la strada dei codici di condotta, presa dal GDPR e dal regolamento 1807/2018 sulla *free flow* dei dati non personali- termine molto dibattuto perché nell'epoca delle banche dati e dell'*analytics* è possibile risalire dai dati non personali a quelli personali)
2. Co-regolazione: ha un grado di coinvolgimento del regolatore maggiore perché si assume che debba essere raggiunto un certo interesse pubblico; per garantire maggiore trasparenza i modelli potrebbero essere due. Uno tradizionale, del legislatore europeo nella proposta di regolamento sulla trasparenza e correttezza delle relazioni tra piattaforme ed utenti business dell'aprile 2018. Una parte di *hard regulation* indirizzata alle piattaforme con riguardo a quali dati devono fornire e una di *self-regulation* di invito a creare degli organismi di mediazione per la risoluzione delle controversie e un gruppo di monitoraggio.

Il secondo mira a dare alle piattaforme un obiettivo di interesse generale, es. eliminare le disuguaglianze nelle applicazioni per gli affitti di alloggi, attraverso una regolazione sapiente o basata sui dati; inoltre la piattaforma deve fornire al regolatore i propri dati cosicché chiunque possa fare delle verifiche sull'attuazione dell'obbligo che le è stato dato. In questo modo, soggetti terzi e università possono fare ricerche valutando sia l'interesse pubblico perseguito sia il controllo sui *dataset*: questa *audit* collettiva probabilmente sarà contrastata.

La posizione contraria invece (Sundararajan) riconosce l'importanza di dare alle piattaforme un obiettivo di interesse generale, ma non appoggia l'idea di dare libero accesso ai dati. *Data driven delegation*, a cui si potrebbero aggiungere i controlli per mezzo della API cioè *Application programming interfaces*: il regolatore potrebbe fare dei controlli a campione e valuta come vengono usati i dati. Così non ci sarebbe accesso ai dati grezzi della piattaforma ed essa rimarrebbe proprietaria di tutti i dati in suo possesso (proposta di tipo più proprietario). Così non si svilupperebbe nessuna competenza sulle autorità pubbliche.

Uno strumento di *empowerment* potrebbe essere quella di fare delle *disclosure* di carattere informativo basate sugli algoritmi, quindi delle *di-*

sclosure algoritmiche basate su un procedimento che sia partecipato tanto dalle piattaforme quanto dagli utilizzatori business, quanto dai consumatori sotto il coordinamento del regolatore. Questo servirebbe a correggere dei *bias* che ci sono perché mancano frequenti controlli (*testing*). Quindi all'esito di questi *testing* partecipati, abbiamo delle informative che sono differenziate a seconda dei bisogni informativi dei destinatari e in più sono partecipate e proporzionate, anche mirate agli interessi informativi dei destinatari che possono cambiare.

Renato Grimaldi*

Big data e processi decisionali nella pubblica amministrazione: il traffico monitorato sulla piattaforma Yucca del CSI-Piemonte

ABSTRACT: Il contributo evidenzia come lo sfruttamento consapevole di grandi quantità di dati pubblici e accessibili (ad esempio, attraverso la piattaforma Yucca) permetta una conoscenza approfondita della realtà. In ragione di ciò, la sociologia può beneficiare della molteplicità di dati a disposizione e può altresì essere funzionale per il governo del territorio: l'azione amministrativa risulterebbe efficace perché mirata e consapevole. Il lavoro, infine, sottolinea che questa metodologia nasconde dei rischi, che sono connessi al fatto che la ricerca è influenzata dalla base di dati scelti dal sensore e non dalle domande del ricercatore.

1. *Premessa*

L'origine e la diffusione dei *Big data* è collegata all'espansione di tecnologie; nel 2011 il numero di dispositivi connessi alla Rete ha superato quello della popolazione mondiale mentre nel 2025 si stima che cinquanta miliardi di sensori, ognuno in grado di produrre un'enormità di dati, saranno connessi alla Rete. Dunque le informazioni digitali prodotte ogni giorno da fonti diverse, mette a disposizione un numero di dati praticamente illimitato per la comprensione di fenomeni sociali, economici e di consumo. In Piemonte, la piattaforma Yucca di Smart Net del CSI-Piemonte¹ gestisce più di 800 milioni di dati che provengono da oggetti (*Internet of Things*) o da persone (*Internet of People*). Come esempio di uso dei *Big data* nei processi decisionali nella pubblica

* Renato Grimaldi è direttore della Scuola di Scienze Umanistiche dell'Università degli Studi di Torino ed è professore ordinario di Metodologia della ricerca presso il Dipartimento di Filosofia e Scienze dell'Educazione (DFE). Questo lavoro è stato svolto presso il Laboratorio di simulazione del comportamento e robotica educativa "Luciano Gallino" (di cui è coordinatore scientifico) che è stato costituito sul Progetto di Eccellenza vinto dal medesimo Dipartimento.

1. Yucca di Smart Data Net è un'iniziativa della Regione Piemonte in collaborazione con CSI-Piemonte (www.smartdatanet.it/yucca.html).

amministrazione, si è preso in esame (come vedremo) uno dei sensori monitorati da Yucca.

Tale indagine è stata utilizzata anche a fini didattici nel corso on-line di *Elementi di sociologia e ricerca sociale* (nella parte trattata da chi scrive) e quindi nel MOOC² realizzato nell'ambito del progetto Start@UniTO³. È questo il motivo per cui il lettore, in questo contributo, noterà un impianto rivolto in gran parte a studenti e ricercatori di metodologia e tecniche della ricerca sociale.

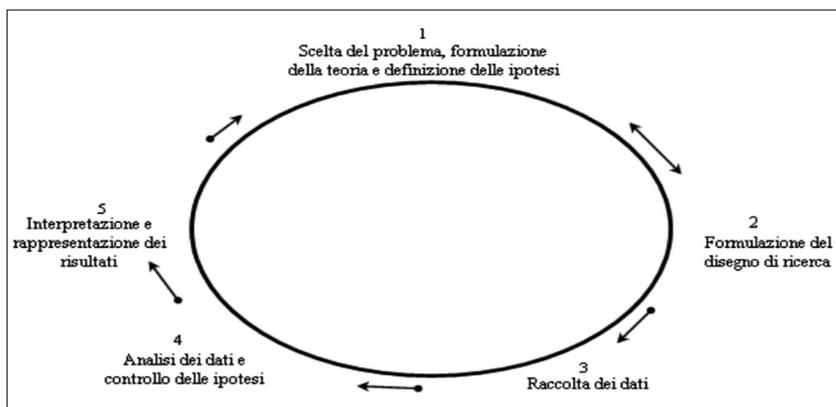
Occorre sottolineare che i dati utilizzati in tale ricerca sono reali, così da poter procedere in concreto all'esplicitazione del ragionamento scientifico che guida e sottende tutto il percorso di sviluppo di un'indagine empirica. L'approccio sociologico inoltre costituisce la base per la progettazione e le decisioni demandate al sistema politico di una città metropolitana come quella torinese. Prima di entrare nel merito dell'indagine, è fondamentale considerare le *fasi* che compongono la ricerca empirica, esposte nella Fig. 1; infatti si può notare che la prima operazione in una ricerca empirica è di scegliere attentamente il *problema*.⁴ In questo caso, si esamina il traffico automobilistico in città.

2. Per MOOC (Massive Open Online Courses) s'intende un corso di formazione a distanza, rivolto a un numero elevato di studenti.

3. Desidero ringraziare Valentina Rosso che nel suo lavoro di tesi di laurea magistrale *Dall'informazione alla conoscenza. Big data e nuove generazioni* (Dipartimento di Filosofia e Scienze dell'Educazione, 2016-17, relatore prof. R. Grimaldi) ha contribuito a completare il lavoro che avevo svolto per Start@UniTO portandolo da un approccio descrittivo a un approccio inferenziale, anche con il ricco apparato di note qui presente, ripreso in particolare dal manuale *Metodi formali e risorse della Rete* (Grimaldi, 2005).

4. La prima fase – *scelta del problema e formulazione di teorie e ipotesi* – consente al ricercatore di definire il quadro concettuale e l'oggetto d'indagine. In questa fase viene formulata una teoria di partenza, ossia un asserto o sistemi di asserti sottoponibili a controllo empirico, e vengono derivate le ipotesi da controllare (che sono rigidamente delimitate nel tempo e nello spazio) mediante un processo di deduzione.

Fig. 1 - Le fasi della ricerca



Fonte: Grimaldi, 2005, p. 65.

Anche il *quadro concettuale* è essenziale nel progettare la ricerca in questione. Ovviamente in questo esempio si trattano solo gli aspetti di metodo e non si ha l'intento di esporre quanto la letteratura scientifica propone in merito al problema scelto. Ci si limita a dire che esso è rilevante perché il Comune di Torino potrebbe decidere (con una ipotetica delibera) di garantire finanziamenti aggiuntivi per la manutenzione del manto stradale ad alcune circoscrizioni. Il requisito richiesto sarebbe quello di dimostrare che – in alcune zone del loro territorio – si registra un traffico giornaliero superiore alle 10 autovetture al minuto (e che tale livello venga superato come minimo in *metà* della giornata); inoltre che nel 20% della giornata, ossia perlomeno *nel 20% dei minuti* in cui è costituita una giornata, sia transitato come minimo 1 veicolo pesante.

Si presuppone, in particolare, che la Circoscrizione 5 (dove si trova il tratto stradale di cui monitoreremo il traffico), interessata a ricevere tale finanziamento, si faccia committente a sua volta di un bando per un lavoro sociologico atto a misurare il traffico sul proprio territorio in modo da poter partecipare ai finanziamenti comunali.

L'ipotetico gruppo di sociologi che decide di partecipare al presunto bando della Circoscrizione 5, nel progetto che presenta, per quanto attiene la rilevazione dei dati, stabilisce di utilizzare i *Big data* che si trovano nella piattaforma Yucca del CSI-Piemonte. In particolare, sceglie di avvalersi delle informazioni fornite dal sensore posizionato all'angolo tra corso

Mortara e via Livorno, che registra, per ogni minuto della giornata, il numero di veicoli che vi transitano. Il sensore informa pure se tra questi è passato almeno un veicolo pesante (quest'ultima informazione – come si vedrà – non è reale ed è stata immaginata e introdotta a solo scopo didattico). Diciamo subito che la Circoscrizione 5 affida al nostro gruppo di sociologi la ricerca sui flussi sul suo territorio, dato che le altre proposte ricevute, nel nostro esempio, hanno costi di rilevazione molto maggiori.

2. *Dalla teoria al metodo*

Per quanto riguarda l'applicazione nel concreto dello schema delle fasi della ricerca, si può notare che la prima azione compiuta dal gruppo di lavoro è l'attenta scelta del problema: in questo caso, il traffico automobilistico nella città torinese. Il quadro concettuale, che orienta il progetto di ricerca, in questo esempio viene tralasciato per dare risalto agli aspetti di metodo. Occorre, ad esempio, osservare che il fatto di anticipare alcuni elementi della fase 2, cioè quella della formulazione del disegno della ricerca, nel momento in cui si progetta la fase 1, evidenzia che tra di esse vi è una continua interazione (come mostra la doppia freccia collocata appunto sulla Fig. 1).

La ricerca presentata per il bando comunale segue le fasi esplicitate nella Fig. 1, di conseguenza si può sostenere che i ricercatori si avvalgono nel loro lavoro di una *procedura formale* con l'utilizzo del *metodo statistico*, basato su una *matrice dei dati* (*casi x variabili*).

Dopo aver individuato il problema, si può formulare la *teoria*, ricordando che essa consiste in un asserto⁵ o complesso di asserti in cui vengono messi in relazione concetti non (ancora) definiti operativamente e quindi non empiricamente controllabili.

In concreto, la nostra teoria recita:

T₁. Il traffico giornaliero è mediamente superiore a 10 veicoli al minuto; tale valore medio è superato almeno nella metà della giornata; il passaggio di automezzi pesanti interessa più del 20% della giornata.

Prima però di derivare le ipotesi, occorre definire alcuni elementi che fanno parte del *disegno della ricerca* e che consentiranno di formulare "ipotesi ben formate". In effetti è fondamentale definire *l'unità di analisi*

5. Per asserto si intende una costruzione mentale passibile di essere pensata come vera o falsa.

che si vuole studiare, che nel nostro caso è il *minuto* in cui è divisa la giornata. In questa ricerca, dunque, l'unità di analisi non è l'individuo, non è un'unità ecologica⁶ (ad esempio l'Asl o il quartiere), ma è una frazione della giornata che ha come *proprietà* il *numero di veicoli* che vi transitano (si suppone inoltre – come già detto – che il sensore sappia riconoscere se tra i veicoli che passano in un minuto sia transitato almeno *un veicolo pesante*).

L'*ambito* riguarda invece i limiti spaziali e temporali della ricerca che, nello specifico, si configurano in: Corso Mortara angolo via Livorno, Circoscrizione 5, Torino, lunedì 16 gennaio, dalle ore 00:00 alle 23:59. Inoltre, i *referenti* dell'indagine, ovvero «ciò a cui si pensa o di cui si parla nel lavoro di ricerca», sono i 1.440 minuti della giornata, che vanno dal minuto 00:00 al minuto 23:59 e che costituiscono anche la nostra *popolazione*⁷ di riferimento.

Se si prende in considerazione la *portata*, il disegno della ricerca prevede la raccolta di tutti i minuti in questione; quindi non si estrae alcun campione⁸ (mentre si potrebbe ad esempio rilevare solo il primo minuto di ogni ora, quindi avere 24 casi, oppure estrarne mediante i numeri casuali un numero suggerito dalle formule di campionamento, ad esempio 300). Avendo quindi a disposizione tutta la *popolazione* dei 1.440 minuti della giornata scelta, potremmo limitarci per l'analisi alla *statistica descrittiva*; in questo lavoro invece utilizzeremo anche la *statistica inferenziale* dato che supporremo di ottenere risultati che attribuiremo anche ad altri giorni della settimana e ad altro periodo. Assumiamo quindi che *lunedì 16 gennaio 2017* sia una giornata *rappresentativa* di tutti i giorni dell'anno, anche se in questo contesto non abbiamo gli elementi per poterlo affermare (ricordiamo ancora una volta che pur lavorando su dati reali, il progetto e la ricerca in oggetto hanno una valenza didattica). Quindi la *portata* della ricerca *non* è limitata al 16 gennaio 2017, corso Mortara angolo via Livorno ma si assume che essa sia estesa a tutto il 2017 e a tutta la Circoscrizione 5.

6. Si parla di unità ecologica quando l'unità corrisponde a un aggregato a base territoriale.

7. In statistica per popolazione s'intende un insieme di individui/elementi referenti di un'indagine, di solito indicato con N; quando si fa ricerca le inferenze si fanno dal campione alla popolazione dalla quale esso è estratto.

8. Si distingue il campione probabilistico da quello non probabilistico; il primo è un campione in cui i soggetti sono scelti sulla base delle probabilità note, il secondo è un campione in cui gli oggetti o gli individui sono inclusi senza tenere conto della loro probabilità di selezione.

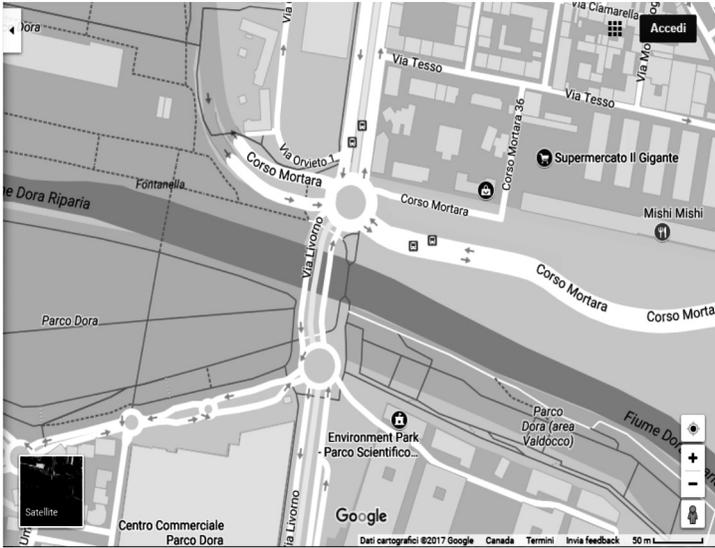
Per quanto riguarda la *raccolta dei dati*, si è già accennato di aver scelto l'impiego dei *Big data* per reperire informazioni utili per la ricerca; in particolare, si è scelto il sensore: f07f098e-2be4-50df-9753-cdc3905f4c41, "Road 3. Vehicle counter on roundabout of corso Mortara and via Livorno"⁹ che Google Maps individua nel luogo riportato dalle Fig. 2 e Fig. 3.

Fig. 2 - Rotonda in cui è collocato il sensore di corso Mortara, angolo via Livorno visibile in Street View di Google



9. www.google.ca/maps/place/Via+Livorno,+10149+Torino+TO,+Italia/@45.0894968,7.6671797,17z/data=!4m13!1m7!3m6!1s0x47886db07a26f20b:0xc6b340dedffb3a3!2sCorso+Mortara,+Torino+TO,+Italia!3b1!8m2!3d45.0901567!4d7.6727476!3m4!1s0x47886db070c392f3:0x4e798d17e44051a4!8m2!3d45.0905691!4d7.6700031.

Fig. 3 - Cartina stradale visibile in Street View di Google che indica il luogo in cui è situato il sensore di corso Mortara, angolo via Livorno



3. Formulazione e controllo delle ipotesi

In generale, si può definire l'ipotesi come un asserto in cui è stata data una definizione di tipo operativo di tutti i concetti richiamati, quindi – a differenza della teoria – è empiricamente controllabile. Inoltre, il suo ambito di pretesa validità è prettamente circoscritto nel tempo e nello spazio.

Di seguito, sono espone le varie ipotesi formulate dal gruppo di ricerca con le relative analisi statistiche al fine di controllarle, positivamente o negativamente.

3.1 Ipotesi monovariata (o descrittiva) con variabile cardinale

H_1 . La media del numero di veicoli transitati è maggiore di 10 (veicoli/minuto).

H_{1bis} . La media del numero di veicoli transitati è maggiore di 10 (veicoli/minuto) in almeno il 50% dei minuti della giornata.

È importante sottolineare che per definire una “ipotesi ben formata”, non è più necessario menzionare i referenti, l'unità di analisi e l'ambito

spazio-temporale (cfr. par. 2). Vediamo ora la *definizione operativa* dei concetti richiamati nell'ipotesi H_1 , per cui *numero di veicoli* è la variabile cardinale¹⁰ enumerata/conteggio che indica il numero di veicoli transitati al minuto, come fornito dal sensore in oggetto.

I dati della ricerca sono stati elaborati con Spss (*Statistical Package for the Social Sciences*) e, in questo caso, si è ottenuta la distribuzione di frequenza di Tab. 1, ossia il numero di casi che rientrano in una data modalità.

Tab. 1 - Distribuzione dei minuti della giornata per numero di veicoli che vi transitano

N veicoli/minuto		Freq. (minuti)	Percent	Cumulative Percent
Valid	0	77	5,3	5,3
	1	53	3,7	9,0
	2	47	3,3	12,3
	3	44	3,1	15,3
	4	34	2,4	17,7
	5	42	2,9	20,6
	6	43	3,0	23,6
	7	38	2,6	26,3
	8	58	4,0	30,3
	9	57	4,0	34,2
	10	61	4,2	38,5
	11	67	4,7	43,1
	12	61	4,2	47,4
	13	63	4,4	51,7
	14	83	5,8	57,5
	15	66	4,6	62,1
	16	64	4,4	66,5
	17	69	4,8	71,3

10. Le variabili cardinali comprendono: quelle metriche (come il peso e l'età), quelle ad intervalli (come la temperatura o l'anno di nascita dove il punto 0 è convenzionale), quelle enumerate (come il numero di libri posseduti) e quelle quasi cardinali (come le variabili che provengono da scale auto-ancoranti, ad esempio il differenziale semantico, e da scale auto-grafiche).

	18	58	4,0	75,3
	19	35	2,4	77,8
	20	48	3,3	81,1
	21	34	2,4	83,5
	22	35	2,4	85,9
	23	36	2,5	88,4
	24	17	1,2	89,6
	25	22	1,5	91,1
	26	23	1,6	92,7
	27	13	,9	93,6
	28	12	,8	94,4
	29	7	,5	94,9
	30	14	1,0	95,9
	31	12	,8	96,7
	32	10	,7	97,4
	33	9	,6	98,1
	34	2	,1	98,2
	35	4	,3	98,5
	37	2	,1	98,6
	39	3	,2	98,8
	42	3	,2	99,0
	44	1	,1	99,1
	45	1	,1	99,2
	46	2	,1	99,3
	47	1	,1	99,4
	48	1	,1	99,4
	50	1	,1	99,5
	51	1	,1	99,6
	52	1	,1	99,7
	53	2	,1	99,8
	56	1	,1	99,9
	65	1	,1	99,9
	67	1	,1	100,0
	<i>Total</i>	<i>1440</i>	<i>100,0</i>	

La Tab. 1 mostra che 0 veicoli (prima colonna, cella in alto) sono “transitati” in 77 minuti della giornata (seconda colonna), e tali 77 minuti sono pari al 5,3% (terza colonna) dei 1.440 totali (vedi cella del totale in fondo); dunque nel 5,3% della giornata sotto quel sensore non sono transitati veicoli). L’ultima colonna fornisce la percentuale cumulata¹¹; se si vuole ad esempio sapere qual è il valore mediano, si cerca il valore 50% e si trova in corrispondenza il valore 13 (prima colonna). La mediana vale dunque 13 e tale misura indica che metà della giornata (metà dei 1.440 minuti totali) vede il passaggio di meno di 13 veicoli al minuto e l’altra metà il passaggio superiore a 13 veicoli al minuto; tali frazioni della giornata non sono contigue ma sparse nelle 24 ore. Dunque ciascuna delle righe di questa distribuzione risponde alla domanda: in quanti *cas*i (*minuti*) della giornata (vedi seconda colonna delle frequenze), è transitato un *numero* di veicoli pari a x (prima colonna)?

Dato che le ipotetiche linee guida del Comune richiedono (per finanziare un intervento) un traffico giornaliero superiore alle 10 autovetture al minuto in almeno metà della giornata, con il semplice calcolo percentuale (che desumiamo dalla colonna delle percentuali cumulate) pari a $100 - 38,5 = 61,5\%$ dove 38,5% è la quota di minuti della giornata in cui passano meno o pari 10 veicoli, possiamo concludere che nel 61,5% della giornata transitano più di 10 veicoli e quindi quella zona rientra nei criteri fissati dal Comune per il finanziamento. L’ipotesi H_{1bis} è dunque controllata *positivamente*.

Tab. 2 - Media, mediana, moda, deviazione standard, massimo, minimo, somma della variabile veicoli transitati al minuto

N	Valid	1440
	Missing	0
Mean		13,51
Median		13,00
Mode		14
Std. Deviation		9,12

11. La frequenza cumulata rappresenta il numero dei casi che su quella variabile hanno dato un valore pari o inferiore di quel valore. Dividendo frequenze semplici e frequenze cumulate per il numero totale dei casi, si hanno le percentuali semplici e le percentuali cumulate.

Minimum	0
Maximum	67
Sum	19457

La Tab. 2 sintetizza mediante indici statistici la distribuzione di frequenza della Tab. 1 e informa che, nei 1.440 minuti della giornata, c'è stato il transito di una *media*¹² di 13,51 veicoli (al minuto), la *mediana*¹³ (come appena visto) è 13 veicoli, la *moda*¹⁴ è 14 (ovvero sono 83 pari al 5,8% i minuti della giornata in cui sono passati 14 veicoli e questa è la frequenza di minuti più elevata). Per seguire questi ragionamenti occorre sempre tornare al fatto che l'unità di analisi è il minuto di cui è composta la giornata e che il numero di auto transitate sono una proprietà.

La *deviazione standard*¹⁵ (o *scarto tipo*) è pari a 9,12 (veicoli) e indica che c'è una certa dispersione di valori attorno alla media, ossia che il traffico tende ad essere differente nelle ore della giornata. Il *minimo*, il valore più basso della distribuzione, è 0 (in alcuni minuti – 77 per la precisione, come si è già visto – passano 0 veicoli) mentre il *massimo*, il valore più alto, è 67 (si può vedere che vi è 1 solo minuto della giornata in cui sono transitati 67 veicoli). La *somma* di veicoli transitati è pari a 19.457; questa statistica è stata calcolata dato che la variabile è di *quantità* (secondo la definizione di alcuni studiosi, ipoteticamente la proprietà è trasferibile da un caso a un altro e quindi si può calcolare a somma).¹⁶

La distribuzione di frequenza del numero di veicoli transitati nei 1.440 minuti del 16 gennaio 2017 (Tab. 1) è invece rappresentata dall'istogramma nella Fig. 4.

12. La media aritmetica rappresenta la somma dei valori di tutti i minuti della giornata su quella data variabile divisa per il numero dei casi.

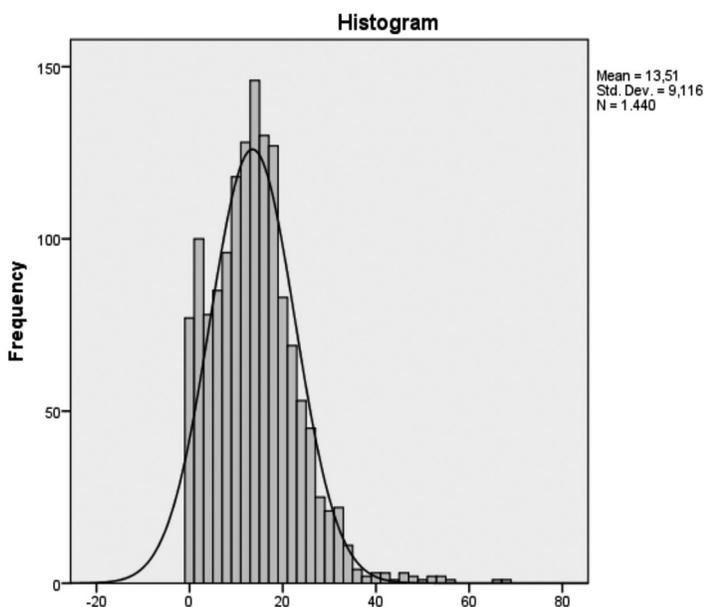
13. La mediana è il punto che divide in due parti uguali la distribuzione ordinata dei casi.

14. Per moda s'intende la categoria più frequente, ossia quella con la frequenza più alta (*ibidem*).

15. La deviazione standard (o scarto tipo) è un indice di dispersione per la misura della variabilità dei dati in una distribuzione di frequenza.

16. Ad esempio, l'età non è di quantità in quanto non posso trasferire – neppure ipoteticamente – anni della vita di un individuo a un altro nella stessa matrice dei dati; il reddito è di quantità, così pure posso pensare che veicoli che passano in un certo minuto della giornata possano transitare in un altro minuto della stessa giornata.

Fig. 4 - Distribuzione di frequenza del numero di minuti della giornata per numero di veicoli che vi transitano (istogramma)



Nella Fig. 4 si può distinguere anche solo con approssimazione il valore medio, la moda, la mediana (l'area del grafico risulta in tal caso divisa in due parti uguali dal valore 13), la deviazione standard, il minimo e il massimo. È stata sovrapposta sul grafico la distribuzione normale, calcolata con la stessa media e deviazione standard ottenute dalla ricerca empirica (ossia 13,51 e 9,116); questa comparazione consente di vedere anche solo graficamente se la distribuzione segue questo schema statistico casuale¹⁷.

17. Per quanto riguarda la distribuzione di frequenza riportata in Fig. 4, essendo monomodale abbiamo calcolato i coefficienti di Skewness=0,993 e di Kurtosis=2,549; quest'ultimo valore (maggiore di zero) indica che la forma della distribuzione è più "appuntita" rispetto alla normale calcolata con medesima media e deviazione standard (v. ancora Fig. 4) mentre il valore di Skewness pari circa a +1 mostra una leggera coda a destra, verso i valori positivi dell'asse delle ascisse. Quando la distribuzione ha i due coefficienti pari a zero, la distribuzione ha l'andamento della curva normale. Nel nostro caso la distribuzione empirica si discosta da quella teorica soprattutto per quanto riguarda i valori negativi rispetto alla media.

Tab. 3 - Intervallo di fiducia calcolato sulla media del numero di veicoli transitati al minuto

N. di veicoli transitati al minuto	Mean		13,51
	95% Confidence Interval for Mean	Lower Bound	13,04
		Upper Bound	13,98

Per controllare l'ipotesi H_1 bisogna quindi ricorrere a indici statistici calcolati sempre da Spss sulla variabile appena vista. L'ipotesi H_1 è stata controllata *positivamente* (v. Tab. 3), dato che la media del numero di veicoli/minuto è pari a 13,51 (intervallo di confidenza/di fiducia¹⁸ pari a 13,04 – 13,98; quindi con una probabilità del 95% il valore reale della media è compreso nell'intervallo di fiducia) e quindi superiore a 10. Con questo calcolo (intervallo di fiducia) siamo entrati nell'approccio inferenziale dato che consideriamo il nostro database non più come una popolazione ma come un campione e quindi che il 16 gennaio 2017 sia rappresentativo degli altri giorni dell'anno.

L'analisi appena condotta ha fatto comprendere che – per quanto i criteri espressi dalle ipotesi H_1 e H_{1bis} – la Circostrizione 5 può accedere ai finanziamenti del Comune.

3.2 *Ipotesi monovariata (o descrittiva) con variabile categoriale non ordinata*

L'obiettivo della ricerca è anche quello di controllare il passaggio di veicoli pesanti (bus, camion, etc.). Sulla base delle linee guida del Comune di Torino (che prende in considerazione le arterie interessate almeno per il 20% dei minuti della giornata da traffico pesante) e basandosi sempre sulla teoria di partenza, si è formulata la seguente ipotesi:

18. In generale, il calcolo degli intervalli di fiducia consente di accettare o respingere le ipotesi sulla media di una variabile cardinale oppure sulla proporzione di una modalità di una variabile categoriale e quindi di sottoporre a controllo empirico ipotesi descrittive (monovariate). La stima intervallata ha come finalità l'individuazione di un intervallo entro il quale si trova, con una certa probabilità, il parametro da stimare. I limiti dell'intervallo sono chiamati limiti di fiducia e variano in base alla numerosità del campione estratto. Il livello di fiducia della stima è dato da $1-\alpha$ (solitamente è uguale al 95% o al 99%) e rappresenta la probabilità che l'intervallo contenga il valore reale del parametro che si vuole stimare.

H_2 , Il transito di almeno un veicolo pesante è presente in più del 20% dei casi (ossia è presente in più del 20% dei 1.440 minuti della giornata).

La variabile *transito di almeno un veicolo pesante* (variabile categoriale non ordinale e dicotomica) si definisce operativamente indicando con 1 il fatto che il sensore ha registrato nel minuto il passaggio di almeno un bus, camion, etc., segnalando con 2 il fatto che il sensore non ha registrato nel minuto il passaggio di almeno un bus, camion, etc., ma solo di veicoli non pesanti.

Bisogna precisare ancora una volta che, mentre gli altri dati sono stati tutti registrati effettivamente dal sensore in oggetto, questa variabile e i dati caricati nella matrice sono stati inventati allo scopo di disporre di una variabile in più per consentire il controllo di questa ipotesi.

La distribuzione di frequenza (Tab. 4) e il grafico a barre (Fig. 5) che la rappresenta indicano che 645 minuti dei 1.440 totali della giornata – pari dunque al 44,8% del totale – sono interessati dal passaggio di almeno un veicolo pesante.

Tab. 4 - Distribuzione di frequenza dei minuti della giornata in cui è transitato almeno un veicolo pesante.

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	almeno un veicolo pesante	645	44,8	44,8	44,8
	solo altri tipi di veicoli	795	55,2	55,2	100,0
	Totale	1440	100,0	100,0	

Fig. 5 - Distribuzione di frequenza dei minuti della giornata in cui è transitato almeno un veicolo pesante (grafico a barre)



Si può concludere che l'ipotesi H_2 è stata controllata *positivamente* in quanto sono 44,8% i minuti della giornata interessati dal passaggio di almeno un veicolo pesante, quota superiore al 20% prefissato.¹⁹

3.3 Ipotesi bivariata tra una variabile categoriale ordinata e una cardinale enumerata

Il gruppo di ricerca dei sociologi intende ora comunicare al Comune che il traffico non è indipendente dalle ore della giornata e desidera fornire quali sono i momenti di criticità. Formula quindi la seguente ipotesi:

H_3 : *Esiste una relazione tra le ore della giornata e il numero di veicoli che transitano al minuto.*

In questo caso a partire dalla variabile *minuti della giornata* (i 1.440 che il sensore etichetta a partire da 00:00 fino a 23:59) e mediante un procedimento di ricodifica, si è ottenuta una nuova variabile che mette assieme 24 classi di 60 minuti (ossia si sono ricostruite le 24 ore della giornata):

19. In questo caso non abbiamo fatto ricorso alla statistica inferenziale dato che i risultati forniti dall'approccio descrittivo sono stati ritenuti sufficientemente robusti.

00 da 00:00 a 00:59

01 da 01:00 a 01:59

...

23 da 23:00 a 23:59

Si sceglie ora di considerare tale variabile come *categoriale ordinata* (anche se può essere definita *cardinale enumerata*, in quanto risponde ai requisiti di ordine e distanza tra le classi/modalità). Ovviamente la variabile *numero di veicoli* è già presente nella matrice dati, poiché già declinata operativamente in precedenza. Come tecnica di analisi si è optato per l'analisi della varianza²⁰ tra la variabile indipendente *ore della giornata* e la variabile dipendente *numero di veicoli*. La Tab. 5 mostra chiaramente che i veicoli transitati sotto il sensore superano il valore di 10 veicoli/minuto nelle seguenti ore (prima e seconda colonna): 0 (cioè dalle 00:00 alle 00:59), 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22 e 23. Quindi in 18 ore della giornata (in tre quarti delle 24 ore) il traffico è superiore ai limiti di soglia. La terza colonna contiene il numero di minuti per ora (ovviamente sono sempre 60), quindi la deviazione standard. La quinta colonna suggerisce il numero di veicoli transitati nella rispettiva ora; il maggior numero di veicoli (colonna Sum), 2.001, si rileva tra le ore 08:00 e le 08:59. L'ultima colonna riporta il *coefficiente di variazione*,²¹ calcolato con il rapporto tra deviazione standard e la media; tale normalizzazione consente di comparare la dispersione attorno alla media nelle varie ore; esso informa che la minore dispersione, pari a 0,22, si ha tra le ore 09:00 e 09:59 (ossia che, in ciascuno dei 60 minuti, il numero di veicoli che transita si avvicina al valore medio, ossia 22,47), mentre la maggiore dispersione, pari a 1,45, si ha tra le 04:00 e le 04:59 (nei minuti che costituiscono quest'ora si contano numeri di auto che possono essere molto diversi dal valore della media, che vale 1,3).

20. L'analisi della varianza consente di studiare le relazioni tra una variabile cardinale e una variabile categoriale.

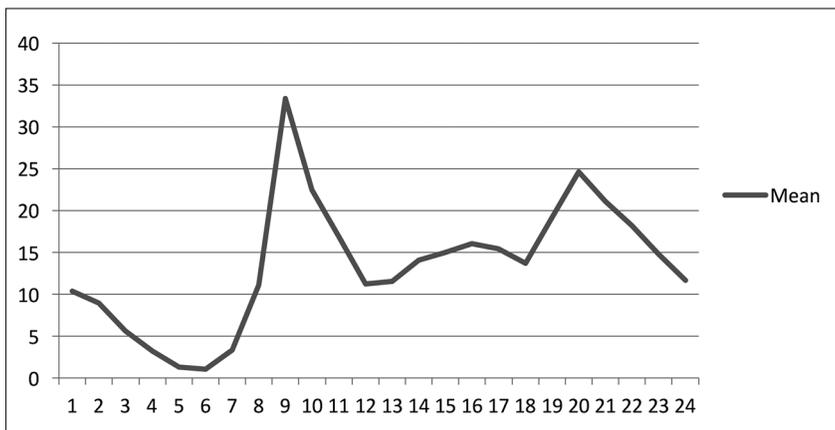
21. Il coefficiente di variazione è un indicatore statistico di dispersione relativa, che serve per indicare la variabilità di un fenomeno in termini comparativi. È usato per confrontare la variabilità dei fenomeni senza prendere in considerazione l'unità di misura.

Tab. 5 - Media del numero di veicoli transitati al minuto nelle 24 ore del 16 gennaio 2017

Ore della giornata	Mean	N	Std. Deviation	Sum	Coeff. Variazione
0	10,40	60	4,43	624	0,43
1	8,97	60	3,96	538	0,44
2	5,6	60	3,51	336	0,63
3	3,23	60	2,35	194	0,73
4	1,30	60	1,89	78	1,45
5	1,07	60	1,22	64	1,14
6	3,33	60	2,78	200	0,84
7	11,12	60	5,21	667	0,47
8	33,35	60	13,20	2001	0,40
9	22,47	60	4,87	1348	0,22
10	16,85	60	5,80	1011	0,34
11	11,22	60	3,57	673	0,32
12	11,53	60	3,90	692	0,34
13	14,05	60	5,09	843	0,36
14	15,00	60	4,58	900	0,31
15	16,05	60	4,99	963	0,31
16	15,43	60	4,96	926	0,32
17	13,72	60	4,98	823	0,36
18	19,18	60	5,74	1151	0,30
19	24,65	60	5,90	1479	0,24
20	21,12	60	5,65	1267	0,27
21	18,20	60	4,83	1092	0,27
22	14,8	60	5,12	888	0,35
23	11,65	60	4,56	699	0,39
<i>Total</i>	<i>13,51</i>	<i>1440</i>	<i>9,12</i>	<i>19457</i>	<i>0,67</i>

La Fig. 6 illustra l'andamento della media del numero di veicoli al minuto che transitano nelle 24 ore.

Fig. 6 - Andamento della media del numero veicoli/minuto per ore della giornata (16 gennaio 2017)



Tab. 6 - Intensità della relazione tra numero di veicoli transitati al minuto e ore della giornata: Eta quadrato

	Eta	Eta Squared
N. di veicoli transitati al minuto * Ore della giornata	,826	,682

Dalla Tab. 6 si evince che *Eta quadrato*²² corrisponde a 0,682: tale coefficiente varia tra 0 e 1, nello specifico 0 significa che non esiste alcuna relazione e 1 che sussiste una relazione perfetta. Nel nostro caso *Eta quadrato* informa che il 68,2% della variabilità di *numero di veicoli transitati al minuto* (variabile dipendente) è spiegata/riprodotta dalle *ore della giornata* (variabile indipendente). A conclusione di questa analisi, si può sostenere che l'ipotesi H_3 è stata controllata *positivamente*, in quanto

22. Il coefficiente *Eta quadrato* si utilizza per quantificare la forza della relazione tra le due variabili.

esiste una relazione tra le *ore della giornata* e il *numero di veicoli*; abbiamo constatato inoltre che la media dei veicoli/minuto supera il valore 10 in 18 ore della giornata.

3.4 *Ipotesi bivariata tra variabile categoriale ordinata e variabile cardinale enumerata*

L'andamento del grafico della Fig. 6 offre uno spunto per formulare la seguente ipotesi:

H_4 : *Esiste una relazione tra i flussi di traffico e il numero di veicoli al minuto; in particolare la media dei veicoli/minuto è maggiore nelle fasce orarie prossime all'ingresso e all'uscita dal lavoro (dalle 06:00 alle 08:59 e dalle 18:00 alle 19:50).*

Per ottenere la variabile *flussi di traffico*, si è lavorato sulla variabile *ore della giornata* e si sono creati, mediante una procedura di ricodifica, dei periodi della giornata, distinguendo quando l'andamento della funzione grafica illustrata sopra (vedi Fig. 6), passa da monotona decrescente e monotona crescente e viceversa:

- 1 00 – 05
- 2 06 – 08
- 3 09 – 11
- 4 12 – 15
- 5 16 – 17
- 6 18 – 19
- 7 20 – 23.

Si inizia quindi con il periodo/flusso di traffico dalle ore 00:00 sino alle 05:59 dove il traffico è decrescente, poi col periodo 2 dalle 06:00 fino alle 08:59 quando cresce e così via per tutti i 7 periodi. Come informano la Tab. 7, la Tab. 8 e la Fig. 7, effettivamente i periodi della giornata dalle 06:00 alle 08:59 e dalle 09:00 alle 11:59 della mattina e dalle 18:00 alle 19:50 del tardo pomeriggio vedono il traffico più intenso. I risultati informano pure che anche le ore della mattina fino a mezzogiorno, dopo l'ingresso normale al lavoro, sono molto trafficate. Pure in questo caso si procede con l'analisi della varianza.

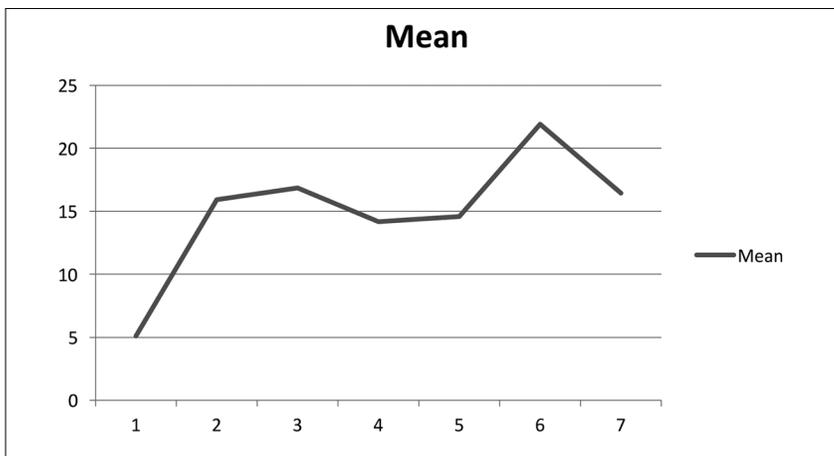
Tab. 7 - Distribuzione di frequenza dei minuti per flusso di traffico (orari)

Flusso traffico (orari)		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	0 - 5	360	25,0	25,0	25,0
	6 - 8	180	12,5	12,5	37,5
	9 - 11	180	12,5	12,5	50,0
	12 - 15	240	16,7	16,7	66,7
	16 - 17	120	8,3	8,3	75,0
	18 - 19	120	8,3	8,3	83,3
	20 - 23	240	16,7	16,7	100,0
	<i>Total</i>	<i>1440</i>	<i>100,0</i>	<i>100,0</i>	

Tab. 8 - Numero dei veicoli/minuto transitati nelle 24 ore del 16 gennaio 2017 per flusso di traffico (orari)

Flusso di traffico (orari)	Mean	N	Std. Deviation	Sum	Coeff. Variazione
0 - 5	5,09	360	4,75	1834	0,93
6 - 8	15,93	180	15,21	2868	0,96
9 - 11	16,84	180	6,66	3032	0,40
12 - 15	14,16	240	4,93	3398	0,35
16 - 17	14,58	120	5,03	1749	0,34
18 - 19	21,92	120	6,41	2630	0,29
20 - 23	16,44	240	6,16	3946	0,37
<i>Total</i>	<i>13,51</i>	<i>1440</i>	<i>9,12</i>	<i>19457</i>	<i>0,67</i>

Fig. 7 - Andamento della media del numero di veicoli/minuto per flusso di traffico (orari)



Tab. 9 - Intensità della relazione tra numero di veicoli transitati al minuto e flussi di traffico: Eta quadrato

	Eta	Eta Squared
N. di veicoli transitati al minuto * Flussi di traffico	,574	,329

Dalle informazioni riportate in Tab. 9 possiamo affermare che esiste una relazione tra il *flusso di traffico* e *transito di veicoli al minuto*, poiché *Eta quadrato* pari a 0,329 indica che il *flusso traffico* spiega il 33% della variabile *transito dei veicoli al minuto*. In particolare, la media dei veicoli/minuto è maggiore nelle fasce orarie prossime all'ingresso (dalle 06:00 alle 08:59) e all'uscita dal lavoro (dalle 18:00 alle 19:50), ma anche nella seconda metà della mattinata (dalle 09:00 alle 11:59). Le fluttuazioni della media del transito al minuto dei veicoli, non è dunque dovuta a fattori casuali, ma è sistematica. L'ipotesi H_4 è stata controllata *positivamente*.

3.5 *Ipotesi bivariata tra variabile categoriale ordinata e variabile categoriale non ordinata dicotomica*

All'amministrazione della Città di Torino è utile conoscere se il passaggio dei veicoli pesanti è in relazione con le ore della giornata e se è concentrato nelle ore del mattino e della sera, quando i bus e i camion escono dai depositi per il lavoro oppure si allontanano dalla città: ciò consentirebbe di avere informazioni per procedere alla manutenzione del manto stradale. Si formula quindi la seguente ipotesi:

H₅. Esiste una relazione tra flusso di traffico (periodi della giornata) e transito di almeno un veicolo pesante/minuto; in particolare questi si concentrano nel periodo 06:00 – 08:59 e 18:00 – 19:59.

In questo caso la tecnica utilizzata è la tabella di contingenza²³ (tabella a doppia entrata) in cui la variabile indipendente (*flusso di traffico*) è posta sulle righe, per cui è stata riportata la percentuale di riga. La variabile dipendente (*transito di almeno un mezzo pesante nei rispettivi minuti*) è messa in colonna (v. Tab. 10). Come si può facilmente osservare, la comparazione tra le percentuali di riga (tra di loro e soprattutto con quella totale) depone a favore di un'indipendenza tra le due variabili; in altre parole, cambiando riga (ovvero passando da una modalità all'altra della variabile indipendente) non varia la distribuzione (percentuale) calcolata sulle modalità della medesima variabile indipendente. Ciò fa supporre una mancanza di relazione tra le due variabili; dunque il periodo della giornata non influenzerebbe un maggiore o minore transito di veicoli pesanti. Si ricorda ancora che la variabile *transito veicoli pesanti* è stata creata al solo scopo didattico e che i dati nelle celle sono stati inseriti a caso; il risultato mostrato nella tabella ha proprio confermato la casualità dei numeri collocati nelle rispettive celle e quindi della mancanza di sistematicità nella relazione con la variabile flusso di traffico.

23. La tabella di contingenza viene utilizzata per analizzare la relazione tra due variabili categoriali.

Tab. 10 - Minuti della giornata ripartiti tra il passaggio di almeno un veicolo pesante in funzione del flusso di traffico (classi di ore)

			Passaggio di almeno un veicolo pesante		Total	
			camion	altro		
Flusso di traffico (classi di ore)	0 - 5	Count	165	195	360	
		%	45,8	54,2	100,0	
	6 - 8	Count	81	99	180	
		%	45,0	55,0	100,0	
	9 -11	Count	81	99	180	
		%	45,0	55,0	100,0	
	12 - 15	Count	106	134	240	
		%	44,2	55,8	100,0	
	16 - 17	Count	53	67	120	
		%	44,2	55,8	100,0	
	18 - 19	Count	53	67	120	
		%	44,2	55,8	100,0	
	20 - 23	Count	106	134	240	
		%	44,2	55,8	100,0	
	Total		Count	645	795	1440
			%	44,8	55,2	100,0

Tab. 11 - Test Chi-quadrato tra il passaggio di almeno un veicolo pesante e il flusso di traffico (classi di ore)

X quadro Gdl Sign.

Pearson Chi-Square	,278 ^a	6	1,000
N of Valid Cases	1440		
a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 53,75			

Ma passiamo ora al controllo della relazione mediante il metodo statistico inferenziale. La significatività del *Chi quadrato*²⁴ pari a 1,00 informa che le differenze tra le percentuali di riga (in particolare la differenza tra frequenze osservate e quelle attese secondo l'ipotesi di indipendenza), dipendono per il 100% dal caso (vedi Tab. 11, dove Sign=1,000); quindi siccome tale probabilità è alta (in questo caso il massimo) significa che le frequenze osservate non si discostano significativamente da quelle attese, calcolate con l'ipotesi di indipendenza (o volendo, che le percentuali di riga non si discostano significativamente tra loro e con la percentuale marginale sempre di riga). Quindi più è bassa tale probabilità, più tali differenze non sono imputabili al caso, ma alla presenza di una relazione sistematica. Nello specifico, solo se tale significatività fosse inferiore al 5% (cioè con significatività del Chi quadrato inferiore a 0,05) si potrebbe pensare che il caso incide meno del 5% e dunque che la relazione esista.

Tab. 12 - Coefficiente V di Cramer di intensità della relazione

		Value	Approximate Significance
	Cramer's V	,014	1,000
N of Valid Cases		1440	

Il calcolo di V di Cramer²⁵ pari a 0,014 conferma quanto appena detto (vedi Tab. 12); l'intensità della relazione è praticamente 0. Si specifica che il valore di V di Cramer può variare tra 0 e 1. Infine, si può asserire che l'ipotesi H_5 è stata controllata *negativamente*, poiché non esiste relazione tra *flusso di traffico* (periodi della giornata) e il *transito di almeno un veicolo pesante* al minuto. E questo risultato era immaginabile dato che – come abbiamo detto

24. La distanza tra la condizione osservata e quella ipotetica viene calcolata mediante l'indice X quadro (X^2), che è pari alla sommatoria delle differenze tra le frequenze osservate (O_i) e le frequenze attese (A_i), elevate al quadrato e rapportate alle frequenze attese. A ciascun valore di X quadro corrisponde una data probabilità, definita dalla distribuzione di probabilità *chi quadro*, caratterizzata da un dato numero di *gradi di libertà* (gdl).

25. Uno degli indici più utilizzati è V di Cramer, il quale normalizza X quadro dividendolo per il suo massimo (dato dal numero dei casi moltiplicato per il minimo tra il numero di righe e di colonne meno uno) e ne estrae la radice quadrata, per annullare l'effetto dell'elevazione al quadrato delle differenze tra frequenze osservate e attese.

– questa variabile è stata costruita al solo scopo didattico e quindi i valori sono stati inseriti in modo del tutto casuale; e la statistica ha evidenziato proprio ciò.

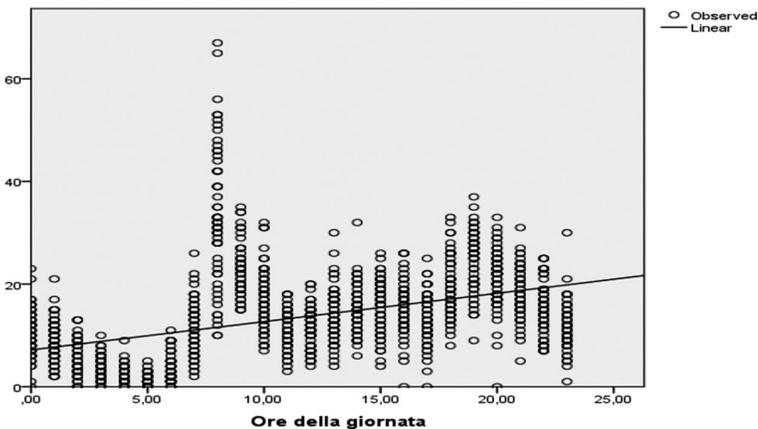
3.6 Ipotesi bivariata tra due variabili cardinali

A questo punto dell'analisi i ricercatori si chiedono se sia possibile fornire al Comune di Torino un modello matematico che spieghi e descriva l'andamento del numero di veicoli che transitano al minuto, in funzione delle 24 ore della giornata. Sarebbe un importante risultato della ricerca e un ottimo strumento di progettazione da parte dei tecnici del Comune. A tal fine, si è formulata la seguente ipotesi:

H_0 : Esiste una relazione lineare tra ore della giornata e numero di veicoli che transitano al minuto; in particolare al crescere del numero delle ore della giornata (da 0 a 23, variabile indipendente) tende a crescere il numero di veicoli che transitano al minuto (variabile dipendente).

Si vuole sottolineare che in questo caso si è considerata la variabile in oggetto come cardinale²⁶ enumerata, dato che risponde ai criteri di ordine e distanza tra le classi/modalità (come già accennato precedentemente).

Fig. 8 - Minuti della giornata del 16 gennaio 2017 rappresentati per numero medio di veicoli transitati in funzione delle ore della giornata (diagramma a dispersione)



26. Nel caso di una relazione tra due variabili cardinali si individua il coefficiente di correlazione che indica l'intensità della medesima.

Il diagramma a dispersione della Fig. 8 (*scattergram*) posiziona i punti – che corrispondono ai 1.440 minuti rilevati nella giornata del 16 gennaio 2017 – nel quadrante formato in ascissa dall’asse delle ore della giornata e in ordinata dal numero di veicoli transitati al minuto. Ad esempio il punto più alto del grafico evidenzia il minuto dell’ottava ora della giornata in cui sono transitati 67 veicoli, valore che abbiamo rilevato come massimo in Tab. 2. Come si può osservare, i dati si distribuiscono non a caso, ma con una certa regolarità, confermando la relazione già scoperta precedentemente. Ora però è opportuno vedere se questo schema può essere incorporato in un modello lineare.

Si è cercato quindi di calcolare una statistica in grado di indicare al gruppo di ricerca che i dati possono essere rappresentati da un modello matematico lineare. Come si può notare nella Tab. 13, il coefficiente di correlazione di Pearson²⁷ è 0,418 e informa che esiste una relazione tra le due variabili. Nello specifico, il coefficiente di correlazione varia tra -1 e +1: -1 significa che esiste una relazione perfetta indiretta; +1 indica la relazione perfetta diretta, mentre attorno allo 0 non esiste alcuna relazione lineare.

Tab. 13 - Coefficiente di correlazione di Pearson tra N. di veicoli transitati al minuto e Ore della giornata

		N. di veicoli transitati al minuto	Ore della giornata
N. di veicoli transitati al minuto	Pearson Correlation	1	,418**
	Sig. (2-tailed)		,000
	N	1440	1440
Ore della giornata	Pearson Correlation	,418**	1
	Sig. (2-tailed)	,000	
	N	1440	1440
** Correlation is significant at the 0.01 level (2-tailed)			

27. Il coefficiente di correlazione indica un’eventuale relazione di linearità tra due variabili cardinali.

Tab. 14 - Riepilogo del modello e stime dei parametri

Dependent Variable: N. di veicoli transitati al minuto							
Equation	Model Summary					Parameter Estimates	
	R Square	F	df1	df2	Sig.	Constant	b1
Linear	,175	305,247	1	1438	,000	7,177	,551
The independent variable is Ore della giornata							

R quadro²⁸ è 0,175 e indica che il 17,5% della variabilità della variabile dipendente *numero di veicoli* transitati al minuto è spiegata (linearmente) dalla variabile indipendente *ore della giornata* (vedi Tab. 14). Se si riprende la Fig. 8, si può notare che è stata calcolata la retta che meglio interpola i 1.440 punti del diagramma a dispersione: tale retta è quella sovrapposta sul diagramma stesso. Spss ne ha calcolato i parametri che sono risultati i seguenti:

$$Y = a + bX$$

$$\text{Numero veicoli al minuto} = 7,177 + 0,551 * (\text{ore giornata})$$

Se si volesse calcolare tramite questo modello qual è il *numero* (medio) *di veicoli* al minuto che transitano tra le ore 12:00 e le 12:59, si ottiene:

$$\text{Numero veicoli al minuto} = 7,177 + 0,551 * 12 = 7,177 + 6,612 = 13,789$$

Se si ritorna a vedere la Tab. 5, si nota che il numero esatto è 11,53 veicoli/minuto: la differenza tra 13,789 e 11,53 pari a 2,259 è il “prezzo” pagato da chi utilizza questo modello, semplice, utile, economico (in quanto incorpora due colonne di 1.440 numeri), ma che solo in parte rappresenta la realtà dei fatti (illustrata compiutamente dal diagramma a dispersione, che mette in luce l’andamento *non* lineare del fenomeno).

Esiste quindi una relazione lineare tra le variabili *ore della giornata* e *numero di veicoli* al minuto; tale relazione però è poco intensa, infatti R quadro

28. R -quadro – o coefficiente di determinazione – rappresenta una misura che, partendo dalla retta di regressione, sintetizza in un unico valore di quanto la grandezza analizzata si discosta mediamente da tale retta.

indica che solo il 17,5% della variabilità della variabile dipendente *numero di veicoli* transitati al minuto è spiegata (linearmente) dalla variabile indipendente *ore della giornata*. Quindi il modello lineare matematico riportato sopra:

$$\text{Numero veicoli al minuto} = 7,177 + 0,551 * (\text{ore giornata})$$

non spiega sufficientemente bene il fenomeno in questione, in quanto già dal diagramma a dispersione si rivela essere non lineare. In questo caso si preferisce dire che l'ipotesi H_0 è stata controllata *negativamente*.

In conclusione, se si riprendono le fasi della ricerca (Fig. 1) da cui si è partiti, si nota che si è chiuso il cerchio e, sulla base del controllo delle ipotesi, si può riformulare la teoria di partenza, corroborata con le analisi:

T₁. Esiste una relazione tra il traffico giornaliero e le ore della giornata. Tale transito è superiore a 10 veicoli al minuto in almeno metà della giornata; esso è maggiore nelle fasce orarie prossime all'ingresso (dalle 06:00 alle 08:59) e all'uscita dal lavoro (dalle 18:00 alle 19:50), ma anche nella seconda metà della mattinata (dalle 09:00 alle 11:59). Il traffico in funzione delle ore della giornata (pur essendo in relazione) non è riproducibile in modo accurato con un modello lineare. Il traffico pesante interessa più del 20% della giornata ed è indipendente dai flussi (orari) della giornata.

4. Punti di forza e di debolezza

Per quanto riguarda gli aspetti positivi della ricerca, la rilevazione mediante *Big data* ha consentito di avere informazioni utilizzando le 3 V proprie dei *Big data*: Velocità (di raccolta dei dati), Volume (ossia la quantità) e Varietà (impiego di un sensore).²⁹ Un altro elemento favorevole è dato dal fatto che si è lavorato su di una popolazione (tutti i 1.440 minuti della giornata oggetto d'indagine) e non su di un campione (anche se abbiamo poi assunto che la giornata in esame potesse essere rappresentativa di tutto l'anno), con tempi rapidissimi e con un costo decisamente basso (rispetto alla necessità di posizionare dei ricercatori all'incrocio di corso Mortara, angolo via Livorno). Si sottolinea inoltre che, avendo utilizzato i *Big data*, molto probabilmente il gruppo di ricerca avrebbe sbaragliato i concorrenti partecipanti al bando della Circostruzione 5. Tale raccolta di dati ha infatti consentito di presentare un preventivo di molto inferiore a

29. Anche se le V attribuite in letteratura ai *Big data* ora sono aumentate, noi ci limitiamo in questo caso a riportarne le prime tre.

quello dei (supposti) competitori, che invece avrebbero dovuto ipotizzare di mandare sul posto i rilevatori con cronometro e blocco note.

I punti di debolezza riguardano invece altri aspetti: si sono utilizzati solamente i dati forniti dal sensore e non si è potuto decidere in anticipo cosa rilevare, tanto che si è “inventata” – a scopo didattico – la variabile che informa se nel minuto oggetto di osservazione è transitato almeno un mezzo pesante. In effetti, in una ricerca reale non si sarebbe potuto rispondere a questa domanda di lavoro basandosi solo sui *Big data*.

Un altro fattore negativo riguarda la scelta del giorno di rilevazione. Di solito al lunedì (dopo le ore 00:00) si accumulano i rientri da una serata collegata alla domenica, giorno festivo. Tali valutazioni fanno parte del disegno della ricerca, così come il fatto che lunedì 16 gennaio 2017 sia sufficiente per monitorare il traffico di un periodo temporale e di uno spazio più ampio. L'indagine prevede infatti che il 16 gennaio 2017 possa essere considerata una giornata tipo - in cui si è misurato il traffico in quella zona della circoscrizione - e magari che i dati misurati dal sensore in oggetto possano valere per tutta la Circoscrizione 5.

In effetti, i 1.440 minuti di quella giornata sono stati dunque considerati come un campione del traffico che si presenta nei minuti di tutto l'anno nell'area in questione. Si è impiegata la statistica inferenziale per estendere la *portata* della ricerca a tutto il 2017 e a tutta la Circoscrizione 5.

Tab. 15 - Porzione di matrice dei dati (ricavata dai 1.440 casi-minuti totali riferiti al giorno 16 gennaio 2017) costruita col sensore f07f098e-2be4-50df-9753-cdc3905f4c41 e mediante ricodifiche ottenute con Spss

<i>data</i>	<i>ora-minuto</i>	<i>numero veicoli</i>	<i>enumerazione dei 1.440 minuti</i>	<i>ora-minuto (ricodificata in ore)</i>	<i>giornata flussi</i>	<i>transito veicoli pesanti</i>
16-01-2017	00:00	8	1	0,00	1,00	1,00
16-01-2017	00:01	13	2	0,00	1,00	1,00
16-01-2017	00:02	12	3	0,00	1,00	1,00
16-01-2017	00:03	9	4	0,00	1,00	2,00
...
16-01-2017	23:57	17	1438	23,00	7,00	2,00
16-01-2017	23:58	8	1439	23,00	7,00	2,00
16-01-2017	23:59	5	1440	23,00	7,00	2,00

In conclusione, si ritiene importante fornire una porzione della matrice dei dati (vedi Tab. 15) così come si presenta dopo il prelevamento dei *Big data* forniti dal sensore in oggetto e dopo il lavoro di ricodifica svolto con Spss.

Nella prima colonna è memorizzata la data, quindi l'ora-minuto di rilevazione, il numero di veicoli transitati nel minuto, l'ora-minuto trasformata in sequenza numerica (da 1 a 1.440, utile per le operazioni di ricodifica che seguono), l'ora-minuto (ricodificata nelle 24 ore della giornata), l'ora-minuto (ricodificata in 7 classi di flusso di traffico). L'ultima colonna contiene infine la variabile dicotomica creata al solo scopo didattico e con dati tutti inventati (valore 1 = nel minuto in questione è transitato almeno un mezzo pesante; valore 2 = nel minuto in questione sono transitati veicoli altri, ossia nessun veicolo pesante). Nello specifico, i dati contenuti nelle prime tre colonne sono forniti dal sensore in questione, mentre gli altri sono oggetto delle operazioni di ricerca.

5. *Alcune riflessioni*

In questo esempio di ricerca, l'analisi dei *Big data* è stata focalizzata sulle richieste dell'ipotetico bando del Comune di Torino per poter ricevere fondi da impiegare nella ricostruzione del manto stradale se fossero transitati più di 10 veicoli al minuto per almeno metà della giornata e per almeno un quinto anche da mezzi pesanti.

In realtà, la grande potenzialità dei *Big data* - come si è già più volte menzionato - è quella di poter riutilizzare i medesimi dati in altri contesti o per altre finalità. Ad esempio, i dati estratti dalla piattaforma Yucca potrebbero essere utili per monitorare il traffico vicino alle scuole. Se si ipotizza infatti che in corso Mortara ci sia una scuola, l'insieme di tutti i dati sul traffico potrebbero servire a comprendere quanti veicoli transitano in media nell'orario di entrata e di uscita degli studenti.

A tal proposito, si potrebbe utilizzare la rielaborazione statistica dell'ipotesi H_4 : *esiste una relazione tra il flusso di traffico e il numero di veicoli al minuto; in particolare la media dei veicoli/minuto è maggiore nelle fasce orarie prossime all'ingresso e all'uscita dal lavoro (dalle 06:00 alle 08:59 e dalle 18:00 alle 19:50)*.

In effetti, costruendo la variabile *flusso di traffico*, è possibile individuare quanti veicoli passano nella fascia oraria di entrata e di uscita degli studenti dalla scuola. Sono tutte informazioni che il Comune potrebbe utilizzare per migliorare i servizi: ad esempio, nelle ore di maggior traffico

sarebbe utile posizionare in prossimità della scuola alcuni vigili e volontari del traffico, in modo da far attraversare il corso in una condizione di maggior sicurezza per gli studenti.

Un ulteriore impiego dei *Big data* raccolti dalla medesima piattaforma potrebbe essere quello di monitorare indirettamente l'inquinamento atmosferico nella città torinese, qualora si potesse estendere l'esperimento ad altre aree. Dalle analisi statistiche si potrebbero individuare precisi archi temporali in cui il transito di veicoli aumenta e quindi, di conseguenza, cresce anche il livello delle polveri sottili. In base a tali indicazioni, il Comune potrebbe perciò ideare e proporre campagne di sensibilizzazione indirizzate a promuovere l'utilizzo di trasporti pubblici e/o mezzi meno inquinanti, soprattutto nelle ore in cui si registra un maggior afflusso di traffico, come avviene in corso Mortara, angolo via Livorno, dalle 06:00 alle 08:59, dalle 09:00 alle 11:59 e dalle 18:00 alle 19:50.

Tutti queste informazioni potrebbero andare a implementare una *machine learning* capace di apprendere dai dati e impiegare tecniche di intelligenza artificiale per una migliore gestione e organizzazione del territorio.

Martedì 21 maggio – II sessione
Big data e attività conoscitiva
della pubblica amministrazione

Data analysis e capacità conoscitive delle pubbliche amministrazioni

ABSTRACT: L'articolo si pone dal punto di vista delle pubbliche amministrazioni anzitutto per indagare il ruolo che le banche dati pubbliche assumono nell'attività conoscitiva. L'articolo prosegue concentrandosi sulla capacità tecnica informatica della pubblica amministrazione, rilevando che la frequente tendenza a esternalizzare il servizio di creazione dei sistemi informatici ha aumentato la scarsa conoscenza del mercato del settore informatico, con ciò determinando una inadeguatezza organizzativa interna.

1. *Premessa*

Il tema che mi è stato assegnato dice “*Data analysis e capacità conoscitiva della pubblica amministrazione*”. In realtà non mi occuperò di *data analysis* nella prospettiva dell'impiego degli algoritmi, ma seguirò un approccio che cerca di valutare quale è la condizione attuale delle pubbliche amministrazioni sotto il profilo della gestione dei dati a fini conoscitivi. Conoscenze che l'amministrazione deve acquisire per la migliore cura dell'interesse pubblico che la legge le affida. Conoscenze da mettere a disposizione dei cittadini, come beni comuni, funzionali alla garanzia del loro diritto di sapere.

Problematiche, queste, che si possono considerare rientranti nella “informatica tradizionale”, nella quale pure si utilizzano degli algoritmi. La vera differenza, in realtà, è il passaggio da una gestione dei dati nella quale si costruiscono banche dati per esigenze conoscitive delle pubbliche amministrazioni, partendo da sistemi che già predeterminano i dati utili ai fini conoscitivi e quindi raccolgono i dati già avendo in mente il risultato conoscitivo, ad una gestione dei dati con l'intelligenza artificiale (IA) nella quale da un lato aumentano a dismisura i dati disponibili, accessibili, interrogabili e dall'altro l'uso degli algoritmi e dei modelli di intelligenza artificiale dovrebbe consentire una diversa capacità di connessione tra i dati e di approfondimenti conoscitivi; che tali approfondimenti poi siano più o meno utili, siano strumentali alla decisione, siano rilevanti di per sé,

o siano essi stessi una decisione, questi sono solo i possibili esiti di questo lavoro di raccolta e di elaborazione.

2. *Le esigenze conoscitive delle pubbliche amministrazioni. Le banche dati pubbliche*

Il mio approccio sarà quello di mettermi dal punto di vista delle pubbliche amministrazioni partendo da una domanda: le pubbliche amministrazioni sono oggi in grado di svolgere i compiti che già dovevano assolvere ventisei anni fa, con il decreto legislativo del 12 febbraio 1993 n. 39?

Se rispondiamo positivamente ci troviamo nella condizione di valorizzare i dati conoscitivi pubblici nella prospettiva della *data analysis*. In caso contrario si pone il problema di colmare il gap conoscitivo accumulato.

Partiamo da cosa succede intorno alle banche dati pubbliche. L'esperienza di cui mi sono occupato in questi ultimi anni è quella relativa alla Banca Dati Nazionale dei Contratti Pubblici. Questa è una banca dati che presenta le caratteristiche tipiche dell'informatica tradizionale: la banca dati è alimentata da dati elaborati dalle singole amministrazioni, in particolare dai Responsabili Unici del Procedimento (RUP). I RUP, cioè coloro che iniziano la procedura di gara, estrapolano i dati dagli atti adottati dalle amministrazioni e li trasmettono alla banca dati attraverso un complicato meccanismo di passaggio che vede protagonisti gli Osservatori regionali sui contratti pubblici. Una procedura di trasmissione che, di fatto, ostacola la sollecita costruzione della banca dati nazionale, ma di questo non tratteremo.

Il primo problema di questa banca dati, così come accade per molte delle banche dati di questa tipologia, è relativo alla sua garanzia di affidabilità, in quanto è l'intervento umano che traduce il dato contenuto nei documenti ufficiali adottati dalle pubbliche amministrazioni in dati trasmessi per la banca dati. Chi ha studiato la Banca Dati Nazionale dei Contratti Pubblici ha rilevato come sia presente un tasso non proprio irrilevante di imprecisione nella banca dati, dovuto alla trasmissione "umana", soggettiva, dei dati.

Il secondo problema che presenta la banca dati è quello di chiedere alla stessa amministrazione gli stessi dati che sono chiesti da altre banche dati, per motivi diversi o vicini. Le banche dati del MEF, della Ragioneria dello Stato, varie banche dati chiedono dati relativi ai contratti pubblici e ciascuna di queste amministrazioni chiede al titolare del dato una elaborazione diversa, non sempre coincidente con le altre. La difficoltà risiede nel

mettere insieme le amministrazioni interessate ai dati relativi ai contratti pubblici per avere dalle amministrazioni titolari una sola trasmissione di dati, nel rispetto del principio dell'unico invio che il nostro ordinamento oramai ha adottato: non puoi far gravare inutilmente sulle pubbliche amministrazioni l'obbligo di fare tante elaborazioni dello stesso dato quante sono le banche dati da costruire. Occorrerebbe costruire una banca dati il cui contenuto sia sufficiente a soddisfare le esigenze conoscitive di tutte le amministrazioni interessate. Stiamo comunque parlando di un lavoro che si pone come obiettivo quello di migliorare un'impostazione che continua ad essere di tipo tradizionale, di predeterminazione dei dati rilevanti ed utili. Il lavoro che si sta facendo, con lentezza e coordinato da un funzionario dell'ANAC, è quello di definire quali sono i dati che devono essere raccolti, utili allo svolgimento delle funzioni relative ai contratti pubblici. La banca dati, oltre a presentare dei difetti di sovrabbondanza di dati, ha anche un difetto di mancanza di dati veramente rilevanti. C'è un problema nell'impostazione preventiva, dunque, per cui tu devi sapere prima quali sono le tue esigenze conoscitive e costruisci una determinata banca dati difficile da gestire e che poi richiede la collaborazione di un numero elevato di soggetti che devono alimentarla quotidianamente.

C'è stata anche un'ipotesi che l'ANAC ha valutato con la Fondazione Ugo Bordoni sulla possibilità di passare al modello dell'intelligenza artificiale. Questa soluzione non è del tutto banale perché si tratterebbe di consentire all'algoritmo di estrapolare direttamente dagli atti relativi ai contratti pubblici, così come adottati dalle amministrazioni, i dati rilevanti, i più utili a fini conoscitivi, superando l'obbligo della trasmissione dei dati da parte delle amministrazioni titolari, le quali dovrebbero limitarsi ad adottare gli atti contenenti i dati che servono ad alimentare la banca dati. Visto che stiamo costruendo una banca dati pubblica e non stiamo costruendo un algoritmo che *random* va alla ricerca delle informazioni con la loro significatività, l'obiettivo è costruire una banca dati che abbia una significatività precisa e predeterminata, i cui dati siano affidabili. Per fare questo il problema diventa quello di fissare delle regole sulle modalità di costruzione della banca dati che, a questo punto, sarebbero regole tecniche di confezionamento dell'atto amministrativo, grazie alle quali devo essere sicuro che quell'atto amministrativo contenga i dati che mi interessano. È un obiettivo sfidante. Non ci illudiamo, però, che l'algoritmo risolva di per sé il problema, perché l'algoritmo, in questo caso, pone problemi significativi di attuazione concreta.

3. *La gestione dei dati: i servizi informatici*

Passiamo ad altri due argomenti.

Quando parlo di gestione dei dati faccio riferimento ai servizi informatici. Che cosa intendiamo per servizi informatici che l'amministrazione deve rendere? Intendiamo molte cose: l'informatizzazione dei processi e dei procedimenti amministrativi; l'erogazione dei servizi online; la costruzione del sito web dell'amministrazione; la costruzione di piattaforme conoscitive; la raccolta e la conservazione dei dati al fine di garantirne la qualità e l'uso pubblico; il supporto e la conoscenza dell'esistenza delle banche dati pubbliche, in quanto molti ignorano l'esistenza stessa delle banche dati nelle pubbliche amministrazioni; la fornitura di strumenti che consentono all'amministrazione di accedere alle banche dati; tutti i servizi legati alla trasparenza.

Quindi non ci sono solo esigenze conoscitive, ma anche esigenze legate alla fornitura di un servizio al cittadino. In materia di trasparenza, per esempio, senza tecnologie informatiche, gli obblighi di pubblicazione sarebbero praticamente inesigibili. Affinché si realizzi il modello degli obblighi di pubblicità, le tecnologie informatiche sono indispensabili e questo è un ulteriore servizio che grava sull'amministrazione. Per non parlare poi dell'utilità delle tecnologie informatiche per l'effettivo esercizio del diritto all'accesso FOIA.

Come vengono gestiti i servizi informatici nelle pubbliche amministrazioni? Tipicamente attraverso processi di esternalizzazione, di affidamento a soggetti esterni.

Di per sé l'avvalimento di servizi qualificati resi da soggetti privati esterni non deve essere demonizzato. Molte operazioni informatiche, anche di sviluppo di sistemi applicativi, comportano attività che è inutile organizzare all'interno delle amministrazioni.

Nell'esternalizzazione di queste attività, però, deve sempre essere garantito l'interesse pubblico curato dall'amministrazione, l'interesse ad avere e utilizzare un servizio che soddisfi le esigenze conoscitive dell'amministrazione.

L'amministrazione deve, quindi, sapere che cosa ha bisogno di conoscere, quali dati e informazioni soddisfano i suoi bisogni, quali sono le alternative tecniche a disposizione per costruire un servizio informatico efficace ed economico.

Ora, queste condizioni raramente si verificano nel nostro sistema amministrativo.

Considerando la condizione media delle pubbliche amministrazioni, anche quelle di grandi dimensioni (per non parlare di quelle piccole), le

pubbliche amministrazioni, quando usano questo meccanismo di esternalizzazione, anche se è presente nella loro organizzazione un ufficio deputato ai servizi informatici, non sono assolutamente in grado di stabilire quali sono i loro bisogni informatici.

I funzionari con competenze informatiche raramente possiedono una autonoma conoscenza del mercato del settore e finiscono per risultare, in realtà, quasi totalmente dipendenti dalla proposta che gli fa il fornitore. C'è una fase, dunque, addirittura preliminare alla definizione dell'oggetto dell'affidamento del servizio, nel quale l'amministrazione deve brancolare e trovare un fornitore che gli dia il servizio. In questa fase è elevato il rischio che sia il fornitore a dire alla pubblica amministrazione ciò di cui essa ha bisogno. Questo è l'inversione preoccupante di rapporti prodotta dall'asimmetria informativa che caratterizza questo mercato.

Vi è, a mia conoscenza, un rapporto di esternalizzazione che possiamo considerare virtuoso: quello stabilito da molte amministrazioni centrali, ma in particolare dalle nostre Agenzie fiscali, con la SOGEI, in quanto questa sembra essere un soggetto che non ha imposto alle amministrazioni prodotti già confezionati, ma ha costruito i prodotti insieme alle amministrazioni, sulla base delle loro esigenze informative e informatiche. Il modello *in house*, in questo caso, può funzionare perché può garantire alle amministrazioni la presenza di un soggetto qualificato con cui confrontarsi. Ma, se la SOGEI costituisce un esempio di *in house* potenzialmente virtuoso, cosa possiamo dire delle *in house* di tanti piccoli comuni che costituiscono in proprio tante piccole società per soddisfare le proprie esigenze informatiche, con il rischio di basso livello di qualità del servizio erogato? Quindi il problema che ho brevemente sottolineato si inserisce più in generale nel problema delle capacità tecniche delle pubbliche amministrazioni e di come noi rafforziamo all'interno delle amministrazioni la capacità tecnica informatica.

4. *I modelli di gestione dei servizi informatici*

4.1 *I servizi informatici come servizi (privati) resi all'amministrazione*

Il servizio è reso dall'impresa privata all'amministrazione, che utilizza le sue prestazioni al fine di organizzare le proprie esigenze conoscitive, erogando direttamente i servizi resi al cittadino. Nel caso, ad esempio, del

sito WEB, questo è di norma gestito dall'amministrazione con il supporto esterno del fornitore. Non ci sono contatti tra fornitore e cittadini.

Affidamento come appalto di servizi, soggetto al codice dei contratti pubblici, da affidare con gara. Ma la procedura competitiva, di per sé, non assicura la buona riuscita degli affidamenti, a causa della posizione subordinata delle amministrazioni, all'asimmetria informativa di cui sopra si è detto, che produce gravi criticità, quanto alla difficoltà di individuazione delle esigenze informatiche già in sede di programmazione e di progettazione del servizio, con conseguente difficoltà di confezione del bando di gara e di vigilare adeguatamente sull'esecuzione del contratto.

Molto spesso i contratti informatici aggiudicati sono contratti "aperti" nei quali la controprestazione dell'impresa fornitrice viene individuata nel corso del contratto e non preventivamente, con ciò aumentando la dipendenza informativa degli uffici dell'amministrazione rispetto all'impresa.

4.2 I servizi informatici come servizio pubblico

Si ha servizio pubblico quando un servizio, assunto dall'amministrazione, viene affidato ad un soggetto esterno che è poi responsabile dell'erogazione del servizio ai cittadini. Si consideri il caso dell'affidamento a concessionario di servizi on line, ovvero l'affidamento del servizio di accesso ai dati pubblicati (o di fornitura dei dati richiesti con l'accesso FOIA)

Le criticità che abbiamo visto nell'affidamento di servizi privati resi all'amministrazione sono più significative nel caso di erogazione del servizio pubblico sia nella fase della assunzione/configurazione dei contenuti del servizio da erogare, sia nella fase della scelta del contraente, sia nella fase della vigilanza sulla corretta erogazione del servizio.

5. La capacità tecnica informatica nelle pubbliche amministrazioni

Alla domanda iniziale dobbiamo, quindi, dare una risposta prevalentemente negativa. Le nostre pubbliche amministrazioni, salvo eccezioni virtuose, non sono in grado di gestire adeguatamente i servizi informatici, né direttamente, né indirettamente.

È allora necessaria un'opera di ripensamento complessivo dell'organizzazione pubblica del settore, che parta da alcune nozioni fondamentali, per poi giungere a innovative soluzioni organizzative.

Sotto il primo profilo partiamo dalle nozioni di *capacità conoscitiva* e *capacità tecnica informatica*: larga coincidenza. Conoscere significa sa-

pere cercare le informazioni utili. La ricerca è ormai sempre condizionata dall'uso di tecnologie informatiche.

Qualunque sia la scelta sulla modalità di gestione del servizio informatico è imprescindibile il rafforzamento della capacità tecnica informatica all'interno delle pubbliche amministrazioni. In modo estremamente schematico si può affermare che per capacità tecnica informatica si devono intendere almeno quattro competenze/capacità, riferite tanto a strutture organizzative quanto alle persone che in quelle strutture operano: a) saper identificare i propri bisogni conoscitivi; b) saper progettare soluzioni informatiche utili (sulla base delle tecnologie disponibili); c) saper individuare i fornitori di servizi informatici in grado di soddisfare le proprie esigenze; d) saper vigilare sulla corretta fornitura dei servizi resi.

La gestione diretta del servizio da parte dell'amministrazione non esclude la necessità che alcune attività possano essere affidate, con appalto, a fornitori esterni, ma la gran parte delle attività conoscitive è svolta con risorse proprie dell'amministrazione.

Nello schema individuato questo significa che la parte pubblica deve saper organizzare al proprio interno la gran parte delle competenze che configurano la capacità tecnica informatica.

La capacità tecnica informatica in caso di affidamento del servizio. Siamo nelle ipotesi in cui la gran parte dei servizi informatici, siano essi servizi resi all'amministrazione o servizi resi ai cittadini, è affidata a soggetti privati esterni. Decisiva la fase della individuazione del fabbisogno conoscitivo, della progettazione dei servizi. Altrettanto importante la fase della gara informatica (bando e aggiudicazione) e della vigilanza sulla corretta fornitura del servizio.

L'amministrazione non deve organizzare al proprio interno dell'intero servizio, ma deve poter contare su un nucleo di competenze professionali proprie tale da guidare l'intero processo, dalla individuazione del fabbisogno alla vigilanza sul contratto.

6. *Come si rafforza la capacità tecnica informatica*

Bisognerebbe ragionare sia su quale capacità tecnica serve per le parti di attività che comunque restano interne alle amministrazioni, come organizzo una capacità tecnica per fare direttamente alcuni servizi, sia come organizzo la capacità tecnica della pubblica amministrazione nell'interlocuzione con il mercato, perché il problema che mi sembra il più rilevante è che la gran parte delle pubbliche amministrazioni non

sanno dove sono i dati che interessano loro. In una società nella quale le pubbliche amministrazioni operano quasi esclusivamente sulla base delle conoscenze, non disporre di conoscenze significa che le pubbliche amministrazioni sono in difficoltà.

Entriamo in un campo di proposte, di ipotesi organizzative di come si può organizzare una capacità tecnica informatica. Considerando quello che abbiamo appena, detto sicuramente un presidio interno a ciascuna pubblica amministrazione sarebbe indispensabile, ma è pensabile un presidio di livello minimo, in grado comunque di svolgere quei compiti minimi che abbiamo elencato distribuito in 8000 comuni, in migliaia di istituzioni scolastiche, che sono tutte stazioni appaltanti, che sono tutti soggetti che acquistano prodotti informatici? Così com'è il sistema amministrativo non è assolutamente una strada percorribile. Almeno a livello comunale si dovrebbero trovare dei meccanismi propri, per cui queste funzioni minime siano costruite intorno a soggetti aggregati. Su 8000 comuni italiani, 7500 sono sotto i 15 mila abitanti e a questa soglia organizzare il livello minimo è già molto difficile. Tra i 15.000 e i 20.000 mila abitanti qualcosa si potrebbe fare.

Stiamo parlando di un livello minimo, poi ci dovrebbe essere un livello tecnico di qualità intermedio (e qui subentra il tema più complesso delle funzioni di area vasta) che svolga quelle funzioni, anche di livello superiore, che un presidio minimo non saprebbe svolgere. Si pensi ad esempio a come guidare le pubbliche amministrazioni alla ricerca delle banche dati che sono di loro interesse; a come utilizzare i motori di ricerca e le banche dati, che sono funzioni apparentemente elementari che però sono ignorate nella gran parte dei casi. Un livello intermedio che faccia per conto delle pubbliche amministrazioni un progetto di servizio informatico, il bando e la gestione della gara e la vigilanza sui servizi resi.

A completare il quadro, quale soggetto che faccia da punto di riferimento dell'intera rete costituita dai soggetti con qualificate competenze informatiche, il livello che dovremmo come di eccellenza: un grande servizio tecnico informatico nazionale. È vero che i grandi servizi tecnici nazionali sono scomparsi dall'orizzonte del nostro ordinamento, però non c'è dubbio che servirebbe un soggetto di questo genere, con una grande ampiezza di competenze e con efficaci poteri di intervento, perché finora i vari soggetti che si sono susseguiti, AIPA, CNIPA, AgID, non sono stati in grado di svolgere questo compito.

Di cosa ha bisogno una pubblica amministrazione nello scenario che è stato configurato ieri? Una pubblica amministrazione ha bisogno di un soggetto di ricerca, di qualità elevata, che non solo sia aggiornato sulle

innovazioni tecnologiche, ma che produca innovazione di ricerca tecnologica a vantaggio del pubblico, che si ponga, ad esempio in materia di algoritmi, dal punto di vista dell'esigenza pubblica e non soltanto a rimorchio, cercando di seguire e compensare le asimmetrie con il sistema privato.

Occorre un luogo di pensiero, di ricerca qualificata che è anche ricerca fondamentale (*ex art. 33, I comma Cost.*), di ricerca applicata nel quale si ragioni nei termini di quali sono gli strumenti utili usando le nuove tecnologie, l'intelligenza artificiale, ponendosi dal punto di vista delle necessità di una pubblica amministrazione.

Ho cercato di adempiere al mio compito e concludo dicendo che per ragionare di questi temi dobbiamo innanzitutto analizzare quali sono le condizioni della pubblica amministrazione e comprendere che oggi porsi il problema dell'intelligenza artificiale rispetto ad un'amministrazione che è ben lontana dall'aver completato il percorso dell'informatica tradizionale, non è certamente una fuga in avanti, ma non ci si può occupare di questo profilo, cioè di come costruiamo le conoscenze tecniche informatiche delle amministrazioni del futuro, senza rimontare lo svantaggio che abbiamo accumulato.

Enrico Carloni

Qualità dei dati, *big data* e amministrazione pubblica

ABSTRACT: Il contributo concentra l'attenzione sul tema della qualità dei dati pubblici, ritenendo questa caratteristica una condizione necessaria per una corretta valorizzazione del patrimonio conoscitivo pubblico. L'articolo evidenzia che al tempo dei *big data* il concetto di qualità amplia il proprio ambito applicativo fino a comprendere caratteristiche prima estranee come la velocità, la varietà dei dati fino a porre al centro del dibattito la caratteristica della veridicità. La progressiva tendenza all'acquisizione di dati diretta e automatica di informazioni da parte di pubbliche amministrazioni nei confronti di altre banche dati pubbliche, senza l'intermediazione di un funzionario, infatti, rende urgente sia l'elaborazione di standard qualitativi dei dati pubblici, che garantiscano la minor circolazione di errori tra i dati, sia l'individuazione di una struttura organizzativa che definisca ruoli e responsabilità per il rispetto di questi standard.

1. *Premessa*

Il tema della qualità dei dati presenta continui elementi di attualità e si mostra, d'altra parte, in continua evoluzione sulla base di coordinate anzitutto tecnologiche, oltre che normative. Se la qualità rimanda all'insieme delle caratteristiche attese, e quella che viviamo è la società dei dati, si comprende bene come sulla questione della qualità dei dati si concentrino una serie di aspettative, di esigenze, di trasformazioni. Nondimeno, il tema della *data quality* è raramente oggetto di attenzione nella letteratura e negli studi sull'amministrazione pubblica italiana¹ per quanto non possa essere ignorato nelle riflessioni sull'innovazione amministrativa collegata alla rivoluzione digitale.

In termini complessivi, credo che sia utile muovere dalla consapevolezza che ci troviamo di fronte ad uno scenario che si presenta su due

1. Si v. in merito E. CARLONI, *La qualità dei dati pubblici*, in B. PONTI (a cura di), *Il regime dei dati pubblici. Esperienze europee e ordinamento nazionale*, Rimini, 2008, 253 ss.; ID., *La qualità delle informazioni pubbliche. L'esperienza italiana nella prospettiva comparata*, in *Rivista trimestrale di diritto pubblico*, 2009, 67 ss.

piani: mentre ci confrontiamo con una serie di sfide della contemporaneità molto forti e il sistema giuridico già si sta interrogando sulla loro portata e sulle loro implicazioni (*big data*, decisioni algoritmiche: basti pensare alla giurisprudenza recente del giudice amministrativo sulle decisioni automatizzate ed al dibattito scientifico sulle implicazioni dell'applicazione di algoritmi e intelligenze artificiali in ambito pubblico), ci troviamo al tempo stesso dinnanzi ad una pubblica amministrazione che è spesso (e prevalentemente) distante dalla capacità di utilizzare appieno queste potenzialità.

È ancora presente nella memoria la dichiarazione dell'allora ministra Bongiorno, che in un intervento pubblico formulava icasticamente un giudizio ingeneroso sulla pubblica amministrazione italiana affermando che essa era (nel 2018) "all'anno zero della digitalizzazione pubblica"²: ne discende che ci muoviamo su piani distanti ed il rischio è quello di non tenere conto dello scarto tra le sfide e la pratica³ dell'amministrazione contemporanea. Da un lato siamo ancora su frontiere che sono di molti anni addietro e dall'altra parte ci confrontiamo con uno scenario tecnologico del quale, quotidianamente, avvertiamo l'accelerazione e cogliamo, spesso anche solo intuitivamente, le implicazioni.⁴

Dunque, è in un contesto in cui la riflessione deve muoversi tra un accumulo di ritardi e improvvise accelerazioni che si colloca la questione dell'utilizzo dei *big data* da parte delle amministrazioni e, per quello che più ci interessa, della qualità dei dati pubblici come esigenza avvertita, in termini rinnovati, nel sistema conoscitivo pubblico.

Il tema della qualità dei dati rimanda ad una nozione generale di qualità che è "l'insieme delle caratteristiche di un'entità, idonee a soddisfare le esigenze esplicite ed implicite".⁵ Va subito detto che nello scenario dei

2. Questa la dichiarazione della allora Ministra per la pubblica amministrazione, Giulia Bongiorno, formulate dal palco dell'EY Digital Summit, il 4 ottobre 2018, ma poi ripresa e riformulata anche in altre circostanze.

3. Diffusamente, si v. in questo senso Commissione parlamentare di inchiesta sul livello di digitalizzazione e innovazione delle pubbliche amministrazioni, *Relazione sull'attività svolta - La digitalizzazione nella pubblica amministrazione italiana: analisi degli errori e valutazione delle priorità, dall'efficacia degli strumenti all'importanza del capitale umano*, Approvata dalla Commissione nella seduta del 26 ottobre 2017.

4. Diffusamente sul punto ci sia consentito rinviare a E. CARLONI, *Algoritmi su carta. Politiche di digitalizzazione e trasformazione digitale delle amministrazioni*, in *Diritto Pubblico*, 2019, 363 ss.; da una diversa prospettiva cfr. J.-B. AUBY, *Il diritto amministrativo di fronte alle sfide digitali*, in *Istituzioni del Federalismo*, 2019, 619 ss.

5. Così per la International Organization for Standardization: ISO (si v. ISO 8402, *Quality Management and Quality Assurance Vocabulary*). Una definizione, questa, frequentemente

big data le esigenze implicite, non previamente conoscibili nel momento in cui i dati vengono prodotti, raccolti ed oggetto di una prima elaborazione, aumentano in modo significativo e questo stressa la nozione di qualità dei dati. Ma non è questa l'unica sfida proposta dai nuovi paradigmi: nelle nuove dinamiche conoscitive la dimensione della "quantità" si intreccia con i requisiti richiesti alle dinamiche conoscitive basate su grandi moli di dati secondo i caratteri dei *big data*.⁶ Ne discende che da un lato si modifica l'idea di qualità (che incorpora, ad esempio, questioni come la varietà e la velocità di generazione, ed assume una caratterizzazione che si lega anche alla dimensione della mole di dati in questione), e dall'altro che un'esigenza di qualità intrinseca dei singoli dati, intesa come "veridicità", è sempre più spesso avvertita anche nelle riflessioni sul fenomeno della *big data analysis*.

2. La qualità dei dati e la digitalizzazione pubblica

Perché se la qualità, come insieme delle caratteristiche attese va interpretata alla luce delle potenzialità che si collegano ad un determinato utilizzo dei dati, è questa la prospettiva entro cui collocare le nostre linee di riflessione, quella di una nozione che muta al mutare delle condizioni anzitutto tecnologiche, ma anche normative, dell'azione pubblica.

Le diverse stagioni della digitalizzazione pubblica ci propongono, e richiedono, nozioni diverse di "qualità" dei dati.

Nel modello digitalizzazione rivolta all'erogazione di servizi, all'interno delle politiche di *e-Government* maturate in Italia a partire dall'Action Plan del Governo Amato di vent'anni addietro, la prospettiva dell'erogazione di "prestazioni" digitali comporta l'esigenza di una integrazione

richiamata a livello scientifico (si v. per tutti M.L. ABATE - K.V. DIEGERT *et al.*, *A Hierarchical Approach to Improving Data Quality*, in *Data Quality*, 1, 1998, 2-3) che seppure non pacificamente ammessa quale standard condiviso ("a standard definition of data quality does not exist in the current literature": *ivi*, 2) è utile a fornire un primo inquadramento del tema.

6. Cfr., in ambito pubblico, M. FALCONE, "Big data" e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica, in *Rivista trimestrale di diritto pubblico*, 2017, 601 ss.; tra i numerosi scritti più recenti, v. es. F. Costantino, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei "big data"*, in *Diritto Pubblico*, 2019, 43 ss.; per un inquadramento generale del tema, cfr. tra gli altri V. MAYER SCHOMBERGER - K. CUKIER, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, trad. it. R. Merlini, Milano, 2013; D. BOLLIER, *The promise and peril of Big data*, Washington, 2010.

tra le banche dati pubbliche, specie nell'ottica di servizi interoperabili e resi da amministrazioni di *front line* potenzialmente diverse da quelle che detengono le informazioni⁷: in quest'ottica di integrazione ed utilizzo del patrimonio informativo di altre amministrazioni si avverte l'esigenza, per l'amministrazione procedente, di poter fare affidamento su dati di qualità nel momento in cui li acquisiscono da banche dati spesso di altre amministrazioni.

La qualità dei dati si pone dunque in termini nuovi rispetto ai tradizionali principi di adeguatezza conoscitiva procedimentale e come sfida centrale⁸ nel momento in cui le amministrazioni sviluppano i propri processi conoscitivi e decisionali tramite l'acquisizione diretta, ed automatica, di informazioni raccolte altrove o precedentemente: questo determina la perdita di consapevolezza e controllo sulle informazioni utilizzate, e produce la generalizzazione di un'esigenza di qualità "standard" che prima si concentrava solo su alcune specifiche informazioni contenute in determinati archivi pubblici.

In termini generali, la sfida si pone quando perdiamo il filtro umano nell'accesso, formazione e trattamento dei dati rilevanti e quindi non possiamo più garantire "soggettivamente" la qualità del processo conoscitivo, che viene resa attraverso la mediazione del funzionario, ma dobbiamo assicurare una qualità oggettiva della pubblica amministrazione che verrà usata per finalità diverse. Questa esigenza aumenta man mano che aumenta l'integrazione tra le amministrazioni ed è in primo luogo un'esigenza interna: se le pubbliche amministrazioni utilizzano i dati, delle proprie e di altri soggetti pubblici in modo automatico, senza la mediazione di funzionari, devono poter fare particolare affidamento sulla qualità delle informazioni che utilizzano. Si tratta di una questione che si sta ponendo da tempo, ed è sottesa ad istituti come cooperazione applicativa, la de-

7. Cfr. es. già MINISTRO PER LE RIFORME E LE INNOVAZIONI NELLA PUBBLICA AMMINISTRAZIONE, *Verso il sistema nazionale di e-Government, Linee strategiche*, Roma, 2007 (già in innovazione.gov.it), dove si afferma che "in particolare è necessario [...] monitorare e accrescere la qualità delle banche dati", visto che "la costante evoluzione qualitativa dei dati (eliminazione duplicazioni, errori, incompletezze, ambiguità, inconsistenze, ...) [...] è la premessa indispensabile all'ampliamento della loro fruizione" (*ivi*, § 2.2).

8. In questi termini cfr. già U. GASSER, *Framing information Quality Governance Research*, in ID. (a cura di), *Information Quality Regulation. Foundations, Perspectives, Applications*, Baden Baden, 2004, 3 ss. In materia, con particolare attenzione alle problematiche di tipo informatico, C. BATINI - M. SCANNAPIECO, *Qualità dei dati. Concetti, metodi, tecniche*, Milano, 2008.

certificazione⁹, l'accesso interamministrativo¹⁰ alle banche dati, il regime delle basi di dati di interesse nazionale: tutte questioni che, in particolare, presuppongono il fatto che le pubbliche amministrazioni che vanno ad utilizzare dati presenti in sistemi informativi di altre pubbliche amministrazioni debbano poter contare sulla qualità di quelle informazioni.

Il dato di cattiva qualità genera errori e questo ha un costo (in termini di inefficienze, in termini di pregiudizio per i destinatari, in termini di cattiva qualità dei servizi, di disparità di trattamento, ecc.), e quanto più raccogliamo ed usiamo ed integriamo dati diversi, tanto più si pone un problema della qualità obiettiva dei dati. Proprio l'integrazione tra banche dati diverse pone infatti una questione di fondo: in un sistema "aperto"¹¹ il dato di cattiva qualità (la cui erroneità non è più riscontrabile "soggettivamente" dall'operatore, grazie alla conoscenza del fenomeno ed al suo rapporto con i dati generati), si può diffondere e radicare.

Nell'integrazione tra banche dati, l'errore frutto del dato di cattiva qualità può moltiplicarsi e compromettere la qualità di altre banche dati: se noi usiamo il dato della residenza da attribuire ad una persona fisica, lo possiamo ricavare (lo si trova nell'anagrafe dell'agenzia dell'entrate, nelle anagrafi comunali, altri contesti dove il cittadino lo dichiara) integrando banche dati diverse, noi dobbiamo selezionare quale è il dato vero che si può utilizzare quando vogliamo individuare con certezza la residenza e questo lo possiamo fare attraverso altri dati sul dato e cioè i metadati e cioè il dato più recente è quello più vero, oppure più vero è il dato che anche se meno recente è stato convalidato da un pubblico ufficiale, oppure in base a quale banca dati si trova. La selezione dei criteri di qualità determina, peraltro, un innalzamento della qualità complessiva od il suo abbassamento: questo dipende dai criteri sulla base dei quali sceglieremo di allineare i dati in caso di loro discordanza.

Questo progressivamente si pone per tutti i dati in un sistema aperto, l'errore non viene necessariamente scoperto ma anzi si può moltiplicare

9. Su cui diffusamente v. M. GNES, *La decertificazione. Dalle certificazioni amministrative alle dichiarazioni sostitutive*, Rimini, 2014.

10. In questo senso v. già M.P. GUERRA, *Circolazione dell'informazione e sistema informativo pubblico: profili giuridici dell'accesso interamministrativo telematico. Tra Testo Unico sulla documentazione amministrativa e codice dell'amministrazione digitale*, in *Diritto pubblico*, 2005, 525 ss.

11. "L'utilizzo di reti aperte accentua l'importanza delle problematiche di qualità dei dati": così già AUTORITÀ PER L'INFORMATICA NELLA PUBBLICA AMMINISTRAZIONE - AIPA, *Piano triennale per l'informatica della pubblica amministrazione, 2002-2004*, già in www.cnipa.gov.it, 8.

e quindi possiamo avere di fronte a noi un sistema di archiviazione e raccolta dei dati formidabile ma in cui sono presenti in misura diversa errori.

Allora posto che la qualità si lega alle aspettative che abbiamo sul dato e sull'uso che ne faremo, quanto più ammettiamo usi secondari dei dati¹², tanto più diventa difficile assicurare la qualità del dato stesso. In altri termini, il dato, seppure di qualità adeguata per il suo uso primario, può rivelarsi incompleto, non aggiornato, inadeguato e persino erroneo se colto nella prospettiva di un uso secondario non previsto. Si pone sempre più un problema di qualità che è legato ad usi secondari dei dati, non previsti e non prevedibili nel momento in cui raccoglievamo i dati stessi: quindi dobbiamo porci l'obiettivo di garantire caratteristiche oggettive dei dati, standardizzando i requisiti di qualità di modo che i nuovi potenziali utilizzi possano ragionevolmente fondarsi su dati di qualità (di una qualità media adeguata).

3. *I principi di qualità e la loro definizione*

Definire standard di qualità, in termini non puntuali e concreti (in ragione di ciò che mi serve in uno specifico procedimento), ma oggettivi ed astratti (come requisiti propri del dato pubblico in quanto suscettibile di diversi e successivi utilizzi) è tutt'altro che semplice.

Basti pensare che gli attributi che definiscono un dato di qualità sono potenzialmente numerosissimi.

Esistono specifici standard ISO di qualità: l'ISO 25012/2008 elencava 15 requisiti, ma poi lo standard è stato aggiornato perché i requisiti erano insufficienti e nello standard 25024 sono stati aumentati a 63.¹³ In letteratura possiamo rinvenire non meno di 180 diversi possibili requisiti di qualità dei dati¹⁴, tanto che la discussione si è spesso orientata nella ricerca

12. F. MERLONI, *Sull'emergere della funzione di informazione nelle pubbliche amministrazioni*, in ID. (a cura di), *L'informazione delle pubbliche amministrazioni*, Rimini, 2002, spec. 74; cfr. già A. MELONCELLI, *L'informazione amministrativa*, Rimini, 1983, 238 ss.

13. Nel novembre del 2015 è stato pubblicato dall'International Organization for Standardization lo standard ISO/IEC 25024 "Measurement of data quality", estensione dell'ISO/IEC 25012 "Data quality model" del 2008 già referenziato dall'AgID con riferimento ai dati della PA.

14. Così per R.Y. WANG - D.M. STRONG - L. GUARASCIO, *An Empirical Investigation of Data Quality Dimensions. A Data Consumer's Perspective*, working paper TDQM-94-01, MIT, Cambridge (Ma), 1984.

di una selezione di caratteristiche salienti di qualità¹⁵. In questo contesto è ricorrente, in particolare, il riferimento alle categorie proposte da Wang e Strong¹⁶, che sintetizzano in quattro caratteri una serie di dimensioni rilevanti: la accuratezza (credibilità, obiettività, affidabilità), la rilevanza (valore aggiunto, rilevanza, tempestività, completezza, appropriatezza), la rappresentatività (interpretabilità, comprensibilità, consistenza rappresentativa, sinteticità rappresentativa) ed infine l'accessibilità (facilità di accesso, sicurezza di accesso) dei dati stessi¹⁷.

Queste sono alcune delle caratteristiche delle qualità rilevanti ma non è soltanto in astratto che si può arrivare a definire i requisiti di qualità, posto che assume rilevanza anche un criterio di legalità: basti pensare ai requisiti di legittimità, minimizzazione, proporzionalità, correttezza ed aggiornamento che discendono dalla normativa in materia di protezione dei dati personali, od a quelli di completezza e chiarezza che si legano alla diffusione dei dati secondo quanto previsto dall'art. 6 del d.lgs. n. 33 del 2013.

Il soggetto che è chiamato a definire i requisiti di qualità delle banche dati di interesse nazionale è l'Agid (che in questo riprende una competenza a suo tempo valorizzata soprattutto dall'Aipa), che con una determinazione del 2013 è intervenuta rispetto alla qualità dei dati cercando di valorizzare alcuni requisiti di qualità (accuratezza, coerenza, completezza, attualità). Nelle Linee guida per la valorizzazione del patrimonio informativo pubblico, in particolare, la qualità dei dati è poi declinata in termini organizzativi, definendo ruoli e responsabilità per assicurarla¹⁸.

Alcuni dati sono raccolti in banche dati di interesse nazionale, previste dall'art. 60 del Cad per il quale valgono in modo particolare questi requisiti di qualità ed altri di credibilità, sicurezza ed accessibilità, comprensibilità,

15. Possono rintracciarsi, in letteratura, vari modelli di strutturazione delle caratteristiche di qualità (i c.d. *information quality frameworks*): cfr., ad esempio, il modello semplificato proposto da M. EPPER, *Managing Information Quality. Increasing the Value of Information in knowledge-intensive Products and Processes*, Berlin, 2003.

16. R.Y. WANG - D.M. STRONG, *Beyond accuracy. What data quality means to data consumers*, cit., 5 ss.; più recentemente, cfr. Y.W. LEE - D.M. STRONG - B.K. KAHN - R.Y. WANG, *AIMQ: A Methodology for Information Quality Assessment*, in *Information & Management*, 2002, 133-146.

17. Per più articolate definizioni di ciascuna di queste "dimensioni" richiamate, si v. il *framework* proposto da M.S. GENDRON - M.J. D'ONOFRIO, *Data Quality in the Healthcare Industry*, in *Data Quality*, 1, 2001, 4-5, sviluppando le categorie di Wang e Strong.

18. Cfr. ora AGID, *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico*, Revisione 1-2017, in agid.gov.it.

disponibilità, portabilità e reperibilità. Il modello che ha in mente l'Agid è cooperativo cioè sono dati di qualità che vengono integrati attraverso l'azione di più amministrazioni, però questo modello cooperativo è un modello che si sviluppa a fatica e in modo problematico.

Si avverte in modo particolarmente forte, nel sistema pubblico italiano, la difficoltà ad integrare le basi di dati delle diverse pubbliche amministrazioni.

Prevale, fatta salva l'eccezione di alcune importanti basi di dati di interesse nazionale, nel nostro ordinamento una logica *single agency approach* cioè ogni amministrazione tende a disciplinare, trattare, proteggere e gestire i propri dati in modo proprietario, diffidando dal trattamento integrato con altre amministrazioni. Una lettura particolarmente rigorosa e restrittiva della normativa sulla protezione dei dati personali si lega a questo approccio proprietario e nel complesso contribuisce in modo significativo a frenare l'integrazione e la comunicazione tra spezzoni diversi del patrimonio informativo pubblico.

L'evoluzione verso basi di dati integrate procede faticosamente ed in modo episodico, con passi avanti significativi essenzialmente con riferimento ad alcune (limitate) basi di dati avvertite come strategiche. Nel Piano per l'Informatica delle Pubbliche amministrazioni 2019-2020 si continua però ad investire in questa prospettiva¹⁹, proprio considerando l'importanza che oramai i dati assumono per lo svolgimento delle attività istituzionali delle altre amministrazioni.

4. *Le regole sulla qualità dei dati*

Non mancano previsioni legislative rivolte ad assicurare la qualità dei dati pubblici, ma queste normative sono frammentarie ed affrontano il tema da specifiche prospettive, a volte in termini non concordanti: è però evidente una tendenza ad un irrobustimento dei requisiti di qualità, con regolazioni solo apparentemente di settore ma a ben vedere dalla grande pervasività (come, per tutte, quella sulla protezione dei dati personali),

19. Agid, Piano triennale per l'informatica nelle pubbliche amministrazioni, Roma, 2019, § 5 (in Agid.gov.it): "nel contesto del modello strategico dell'ICT nella Pubblica Amministrazione, considerata la crescente rilevanza delle basi di dati di interesse nazionale anche per lo svolgimento delle funzioni istituzionali delle altre pubbliche amministrazioni (oltre a quelle proprie delle amministrazioni titolari) è opportuno attivare processi finalizzati a garantire un adeguato livello di qualità dei dati".

seppure non sempre pensate nella prospettiva del miglioramento dei processi conoscitivi dell'amministrazione e dell'integrazione del patrimonio conoscitivo pubblico.

In altri ordinamenti ci sono legislazioni dedicate alla qualità dei dati (in America, per esempio, c'è il *Data Quality Act* del 2001, sulla cui base sono state adottate delle importanti linee guida nel 2002²⁰) che puntano su alcuni requisiti di qualità che sono quelli di integrità, utilità ed obiettività dei dati rilevanti per l'assunzione di una decisione pubblica.

Nel nostro ordinamento troviamo regole sulla qualità dei dati in vari corpi normativi.

Nel d.lgs. n. 33 del 2013 questa qualità è declinata in una prospettiva "esterna": deve caratterizzare i dati pubblici in quanto destinati alla divulgazione verso l'esterno. I dati che provengano dalla pubblica amministrazione devono essere di qualità e questo comincia a spostare l'asse del ragionamento da una *data quality* interna al sistema pubblico, ad una *data quality* orientata all'*open data* cioè al fatto che la dimensione conoscitiva pubblica non è soltanto quella interna ma è anche, in modo significativo, quella esterna (la produzione di conoscenza tra i cittadini, i riutilizzi in ambito pubblico e privato di queste informazioni). Nel d.lgs. 33/2013, a questi fini, si prevede che le pubbliche amministrazioni garantiscono la qualità dei dati nei portali istituzionali, assicurando l'integrità e il costante aggiornamento, la completezza, la tempestività, la comprensibilità, l'omogeneità, la facile accessibilità nonché la conformità ai documenti originali²¹. Va segnalato che la rilevanza della qualità delle informazioni dell'amministrazione trasparente si lega in modo alla prospettiva del loro riutilizzo e quindi alla generazione di conoscenza possibile entro i paradigmi dell'*open data*²².

20. Su cui v. E. CARLONI, *La qualità delle informazioni diffuse dalle amministrazioni federali negli Stati Uniti*, in *Giornale di diritto amministrativo*, 2002, 1232 ss.

21. Sull'art. 6 del d.lgs. n. 33 del 2013, cfr. E. CARLONI, *I principi del codice della trasparenza* (artt. 1, commi 1 e 2, 2, 6), in B. PONTI (a cura di), *La trasparenza amministrativa dopo il d.lgs. 14 marzo 2013*, n. 33, Rimini, 2013, 29 ss.; D. MARONGIU, *Qualità, durata e modalità di erogazione della pubblicazione*, in B. PONTI (a cura di), *Nuova trasparenza amministrativa e libertà di accesso alle informazioni*, Rimini, 2016, 215 ss.

22. Diffusamente E. CARLONI, *L'amministrazione aperta*, Rimini, 2014; cfr. già B. PONTI, *Open Data and Transparency: A Paradigm Shift*, in *Informatica e diritto*, 2011, 305 ss.; più recentemente, sul ruolo complessivo degli *open data* nei processi decisionali pubblici, cfr. F. COSTANTINO, *Lampi. Nuove frontiere delle decisioni amministrative tra "open" e "big data"*, in *Diritto amministrativo*, 2017, 799 ss.; F. DI MASCIÒ, *Miti e realtà degli "open data" all'italiana*, in *Giornale di diritto amministrativo*, 2017, 399 ss.

Nel codice dell'amministrazione digitale (Cad)²³ il tema della qualità è molto presente anche se non troviamo un articolo espressamente dedicato alla qualità dei dati: il tema ricorre in più parti, in particolare allorché (art. 51, su "Sicurezza e disponibilità dei dati, dei sistemi e delle infrastrutture delle pubbliche amministrazioni") si prevede che "con le Linee guida sono individuate le soluzioni tecniche idonee a garantire la protezione, la disponibilità, l'accessibilità, l'integrità e la riservatezza dei dati e la continuità operativa, dei sistemi e delle infrastrutture".

Sempre alle Linee guida di Agid è affidato il compito di definire regole più stringenti in merito alle basi di dati di interesse nazionale, le quali "costituiscono, per ciascuna tipologia di dati, un sistema informativo unitario che tiene conto dei diversi livelli istituzionali e territoriali e che garantisce l'allineamento delle informazioni e l'accesso alle medesime da parte delle pubbliche amministrazioni interessate": in base a queste Linee guida, queste basi di dati sono realizzate e aggiornate assicurando "caratteristiche minime di sicurezza, accessibilità e interoperabilità" (art. 60, c. 2, Cad)²⁴.

D'altra parte, nel regolare l'utilizzo dei siti da parte delle pubbliche amministrazioni, lo stesso Cad prevede che siano rispettati i "principi di accessibilità, nonché di elevata usabilità e reperibilità, anche da parte delle persone disabili, completezza di informazione, chiarezza di linguaggio, affidabilità, semplicità di consultazione, qualità, omogeneità ed interoperabilità" ed, ancora, che siano "in particolare resi facilmente reperibili e consultabili i dati di cui all'articolo 54" (che poi sarebbero i dati disciplinati ora dal d.lgs. n. 33 del 2013). L'art. 53 del Cad contiene, in sostanza, una serie di principi (di qualità) che si rivolgono ora ai siti, ora ai dati, ora ad entrambi.

Non meno importante l'impatto della disciplina di matrice europea in materia di trattamento dei dati: in materia di protezione dei dati per-

23. Per un inquadramento sintetico dell'evoluzione della disciplina del d.lgs. n. 82 del 2005, cfr. F. CARDARELLI, *Il codice dell'amministrazione digitale*, in *Libro dell'Anno del diritto 2017*, Roma, 2017; F. MARTINES, *La digitalizzazione della pubblica amministrazione*, in *La rivista di diritto dei media*, 2018, fasc. 2; cfr. già E. CARLONI, *La riforma del Codice dell'amministrazione digitale*, in *Giorn. dir. amm.*, 2011, 5, 469 ss.

24. L'articolo 60 del Cad fornisce un primo elenco di Basi di dati di interesse nazionale e demanda ad Agid il compito di individuare (e pubblicare) l'elenco di tali basi di dati. Il comma 3-bis, in particolare, indica una prima serie di basi di dati di interesse nazionale: il Repertorio nazionale dei dati territoriali (RNDT); l'Anagrafe nazionale della popolazione residente (ANPR); la Banca dati nazionale dei contratti pubblici (BDNCP); il Casellario giudiziale; il Registro delle imprese; gli Archivi automatizzati in materia di immigrazione e di asilo; l'Anagrafe nazionale degli assistiti (ANA); l'Anagrafe delle aziende agricole.

sonali, in particolare, la qualità va intesa come rispetto di una serie di requisiti di legittimità dei trattamenti, così come previsto ora anzitutto dall'art. 5 del GDPR. Si tratta di garantire in primo luogo una qualità come "minimizzazione" (i dati devono essere "adeguati, pertinenti e limitati a quanto necessario rispetto alle finalità per le quali sono trattati") e questo fin dalla progettazione del trattamento ed in modo automatico (privacy "by design" e "by default")²⁵. Sempre ai sensi della normativa sulla privacy, va poi assicurata l'esattezza dei dati: i dati devono essere esatti, eventualmente corretti anche su richiesta dell'interessato ove errati, aggiornati. I dati, inoltre, devono essere "legittimamente conservati": conservati per un tempo non superiore a quello necessario al conseguimento della finalità del trattamento.

A questi regimi di qualità se ne affiancano altri, che discendono da regole quali quelle sulla sicurezza informatica (basti pensare al fatto che "la sicurezza delle reti e dei sistemi informativi comprende la sicurezza di dati conservati, trasmessi e trattati": considerando 46 della direttiva (UE) 2016/1148, "NIS", recante misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione), o da discipline di settore come quelle in materia ambientale, o statistica²⁶.

È, nel complesso, un regime frantumato, che va a coinvolgere dati che sono a loro volta raccolti in un sistema frantumato di gestione di dati da parte delle singole amministrazioni. Si avverte però chiaramente la pressione sul patrimonio informativo pubblico di questi diversi requisiti di qualità, tradotti in termini operativi ed organizzativi dalle politiche di digitalizzazione e quindi dall'azione degli uffici dedicati (in particolare, il responsabile per la transizione digitale), dall'impianto normativo ed organizzativo dell'anticorruzione e trasparenza (con il responsabile della prevenzione della corruzione e della trasparenza), da quello della protezione dei dati (con il *data protection officer*).

25. Sul nuovo approccio organizzativo cfr. es. S. CALZOLAIO, "Digital (and privacy) by default". *L'identità costituzionale della amministrazione digitale*, in *Giornale di storia Costituzionale*, 2016, 185 ss.

26. Da ultimo, in merito, cfr. D. NATALE - M.C. PAOLETTI - A. SIMONETTA, *La qualità dei dati e l'informazione statistica*, in *Rivista degli infortuni e delle malattie professionali*, 2012, 281 ss.; cfr. già F. GIGLIONI, *Le soluzioni istituzionali alla qualità dell'informazione statistica*, in *Diritto pubbl.*, 2005, spec. 1030 *et passim*; cfr. L. TORCHIA, *Autonomia dei soggetti e funzionalità del sistema: condizioni di qualità dell'informazione statistica*, in *Riv. ital. dir. pubbl. com.*, 1999, 643 ss.

5. *Big data, algoritmi e mutamenti nelle esigenze di qualità*

Fatto questo inquadramento, possiamo provare a riflettere sulle implicazioni che discendono dalle trasformazioni tecnologiche, all'interno di nuovi paradigmi dei *big data* e delle decisioni algoritmiche.

Dal punto di vista *big data* va in primo luogo rimarcato che mutano le esigenze di qualità: perché la qualità che serve ai *big data* è più una prospettiva quantitativa che non strettamente qualitativa. Dal momento che deve poter operare in una prospettiva anche di tipo predittivo, la logica dei *big data* tende ad utilizzare in finalità anzitutto classificatoria e conoscitiva l'insieme dei dati a disposizione. In questo senso, la "velocità" ed il "volume" fanno pregio sulla qualità dei singoli dati: Google sfrutta molto anche gli errori, non soltanto le informazioni corrette; esatta o meno, una certa *query* di ricerca è un dato che può essere significativo. Per quanto la "v" della "veridicità" si sia affiancata nelle riflessioni più recenti alle altre "v" che già da tempo qualificavano il fenomeno dei *big data*, è anzitutto l'insieme della mole di informazioni a costruire un presupposto delle "caratteristiche attese" (e quindi della "qualità") dei dati. Il che non sottintende però una totale disattenzione al dato "di qualità", quanto piuttosto l'accettazione di singoli dati anche errati la cui fallacità è però compensata dalla mole delle informazioni e dalla loro "qualità media".

La "*data veracity*" si riferisce infatti al set di dati (alla loro precisione e veridicità) piuttosto che alla singola informazione. Nel contesto dei *big data* assume un significato centrale: non si tratta solo della qualità dei dati stessi, ma di quanto siano affidabili l'origine, il tipo e l'elaborazione dei dati. Rimuovere distorsioni, anomalie o incoerenze, duplicazioni e volatilità sono, ad esempio, aspetti che contribuiscono a migliorare l'accuratezza dei *big data*.

La questione cambia invece nella prospettiva della decisione algoritmica e quindi nel momento in cui invece una decisione viene fondata su dati e non abbiamo piena contezza del percorso che viene seguito dagli algoritmi (se sono algoritmi intelligenti) nella definizione della scelta. In questo caso aumenta a maggior ragione il fatto di fondare questa decisione algoritmica su elementi conoscitivi robusti, ovvero su dati di qualità.

È in questo senso che pare orientarsi la giurisprudenza del giudice amministrativo, che condiziona l'ammissibilità delle decisioni amministrative automatizzate al rispetto di specifiche condizioni. Condizioni che vengono in parte mutate dalla normativa europea

sulla privacy²⁷, che richiede in particolare per trattamenti automatizzati il rispetto di specifiche regole di qualità. Il considerando 71 del GDPR prevede che a fronte di trattamenti automatizzati siano poste in essere “misure tecniche e organizzative adeguate al fine di garantire, in particolare, che siano rettificati i fattori che comportano inesattezze dei dati e sia minimizzato il rischio di errori e al fine di garantire la sicurezza dei dati personali secondo una modalità che tenga conto dei potenziali rischi esistenti per gli interessati”. Una formulazione, presa a riferimento dal giudice amministrativo, dal quale, a ben vedere, si può ricavare (oltre al principio di non discriminazione che è ben esplicitato nel prosieguo del considerando), un principio, di “correttezza/qualità” dei dati. Su queste basi, il Consiglio di Stato richiede di “rettificare i dati in “ingresso” per evitare effetti discriminatori nell’output decisionale” (sentenza n. 8472 del 2019²⁸).

A questo punto la sfida che abbiamo davanti, di garantire i diritti delle persone nei confronti della decisione algoritmica, diventa tutt’uno con la questione della qualità dei dati che vengono assunti a presupposto della decisione; cioè si sposta nella dimensione organizzativa della costruzione di un apparato conoscitivo adeguato, fatto di elementi conoscitivi sufficientemente robusti che poi possano essere sviluppati ed implementati da algoritmi su cui non avremmo più il controllo completo, potendo però fare affidamento sulla materia prima su cui hanno fondato il loro percorso decisionale.

La qualità dei dati si pone, in questa prospettiva, quale fondamentale momento di garanzia della decisione algoritmica: la trasparenza della logica decisionale (se non la rigida ripercorribilità dei singoli passaggi), unita ad una qualità degli elementi conoscitivi utilizzati, costituiscono alcuni dei nuovi paradigmi di riferimento nella fase di transizione che stiamo vivendo. Il tutto, peraltro, con l’esigenza di sciogliere una serie di nodi ancora confusi: quale l’idea di “qualità” cui fare riferimento (è sufficiente, ad esempio, la loro “correttezza”, di cui sembra parlare il giudice

27. Spunti in questo senso già in S. CIVITARESE MATTEUCCI, “*Umano troppo umano*”. *Decisioni amministrative automatizzate e principio di legalità*, in *Diritto Pubblico*, 2019, 5 ss.; sul trattamento automatizzato dei dati ai sensi del GDPR cfr. es. E. PELLECCIA, *Profilazione e decisioni automatizzate al tempo della “Black Box Society”: qualità dei dati e leggibilità dell’algoritmo nella cornice della “responsible research and innovation”*, in *Le Nuove leggi civili commentate*, 2018, 1209 ss.

28. Si tratta di una sentenza particolarmente importante in materia di utilizzo degli algoritmi nel procedimento amministrativo: Cons. St., sez. VI, 13 dicembre 2019, n. 8472 - Pres. Montedoro, Est. Ponte.

amministrativo, se largamente incompleti?) e, soprattutto, se la qualità vada intesa in senso “atomistico” (dati, singolarmente presi, corretti) o con riferimento al data-set nel suo insieme (come qualità media, media affidabilità, presenza di adeguate condizioni organizzative che ne assicurino la qualità). Il tutto, in ogni caso, a prefigurare una tendenza ad un sistema di *data quality by default*, che si colloca, al momento, ancora all’incrocio tra i diversi sistemi organizzativi e le diverse discipline di settore volte ad assicurare specifici requisiti (di qualità) dei dati.

Machine learning per la pubblica amministrazione

ABSTRACT: L'articolo si occupa dell'impatto dei *big data* sul patrimonio conoscitivo delle pubbliche amministrazioni a partire da una sperimentazione sulla banca dati dei contratti pubblici nazionali che ha coinvolto in prima persona l'Università di Torino e l'Autorità Nazionale Anticorruzione (ANAC). L'articolo illustra le varie fasi che un informatico o *data scientist* seguono per giungere all'utilizzo dei dati ai fini conoscitivi: l'iniziale approccio statistico volto a identificare le caratteristiche descrittive dei casi oggetto di studio, è seguito dall'approccio descrittivo volto a individuare le regolarità e correlazioni della base di dati a disposizione; dopo queste due fasi, il percorso di ricerca prosegue tramite l'approccio predittivo attraverso le tecniche di *machine learning*. L'articolo conclude promuovendo l'approccio di tipo prescrittivo come funzionale all'individuazione di decisioni che dovrebbero essere assunte sulla base dei dati a disposizione e che potrebbe suggerire delle buone pratiche future.

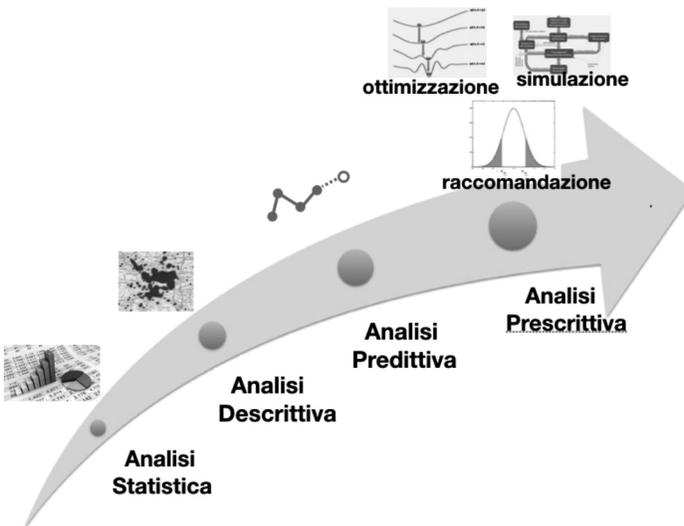
L'Università di Torino ha recentemente attivato una convenzione con l'Autorità Nazionale Anticorruzione (nel seguito denominata brevemente ANAC) per svolgere attività di ricerca nel campo dell'analisi dati e fornire un supporto conoscitivo sui grossi volumi di dati raccolti a riguardo delle gare e degli appalti banditi dalla pubblica amministrazione. L'importanza di utilizzare l'analisi dati per indirizzare l'azione di governo e di controllo del settore pubblico è ormai ampiamente riconosciuta e perseguita¹. Discutiamo qui le prime nostre esperienze e svolgiamo alcune considerazioni sul processo di elaborazione dati sviluppato e i risultati ottenuti.

1. SHAILY KUMAR, *The importance of analytics in the public sector*, 17th of April, 2018, downloaded at: <https://www.digitalistmag.com/future-of-work/2018/04/17/importance-of-analytics-in-public-secotr-06061094>, editor: 2019 SAP SE.

1. Caratteristiche di un percorso di analisi dati

Normalmente un percorso di analisi di dati è fatto da passi successivi^{2,3}. Un esempio del percorso è mostrato in Figura 1. I passi di tale percorso servono per orientarsi nell'ambito della base dati e per conoscere le caratteristiche dei dati. Il primo passo serve per capire come le caratteristiche descrittive degli esempi (o casi) sono distribuite nella popolazione dei casi. Il primo passo del processo è quindi di tipo statistico, nel quale si calcolano le distribuzioni dei valori delle caratteristiche. Al passo successivo si costruiscono e si mostrano dei grafici descrittivi che rendono immediatamente percepibili all'analista la presenza di alcune regolarità e correlazioni tra le caratteristiche. I passi successivi sono costituiti dai modelli di apprendimento automatico (*machine learning*), che servono a predire alcune delle variabili di interesse in funzione delle altre variabili. Questo passo (in Figura 1, è il terzo step) è di tipo predittivo.

Figura 1 - Il percorso di analisi dati



2. VIJAY KHATRI - BINNY M. SAMUEL, 2019. *Analytics for managerial work*. *Communications of the ACM* 62, 4 (March 2019), 100-100. DOI: <https://doi.org/10.1145/3274277>.

3. WATSON IoT, *Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics*, downloaded at: <https://www.ibm.com/downloads/cas/3V9AA9Y5>, editor: IBM Corp. 2017.

Infine il quarto step è ancora più avanzato ed è di tipo prescrittivo: spiega quali decisioni dovrebbero essere prese a fronte dei dati che sono stati analizzati e a fronte dei modelli che predicono le variabili interessanti. In quest'ultimo step si sfruttano modelli di simulazione, di ottimizzazione (per trovare il valore ottimale di alcune variabili di interesse, come i costi o le risorse impiegate) o tramite una raccomandazione che suggerisce quali dovrebbero essere questi valori in alcune condizioni di contesto. Infine concluderemo con alcune considerazioni.

2. *Rischi ed errori nel percorso di analisi dati*

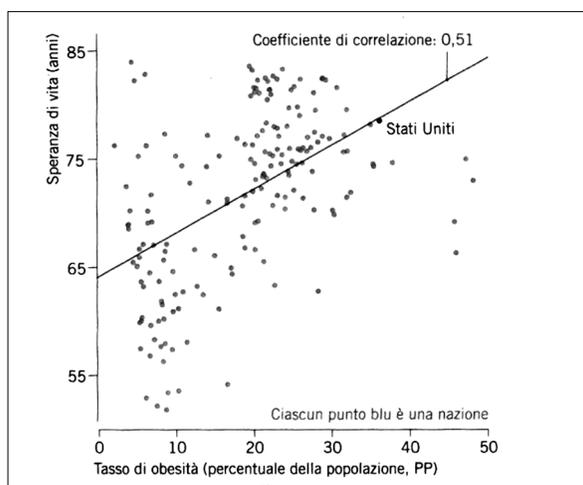
Nelle analisi di tipo statistico e di machine learning può accadere che si facciano degli errori a causa della scarsa attenzione riguardo ai rischi che sono insiti nell'analisi statistica⁴. Un errore comune, ad esempio, è assumere implicitamente che i dati siano stati raccolti al verificarsi di alcune condizioni, però irrealistiche. Oppure, utilizzare una variabile di input (determinata da fattori esterni) come se fosse una variabile di output, adatta a misurare il fenomeno di interesse. Un esempio calzante di questo errore è illustrato dall'economista di Stanford, Thomas Sowell, discutendo del modo in cui si conducono le analisi delle *performances* delle Università. Nell'analisi dei sistemi complessi si usano spesso indici *proxy*, che misurano indirettamente il fenomeno di interesse. Se erroneamente si scegliesse come indice di qualità delle Università la loro capacità di spesa (costi del personale, numero delle biblioteche, ecc.) si sarebbe scelto invece un indicatore della quantità dei finanziamenti di cui beneficiano le Università. Una misura della loro qualità andrebbe invece ricercata in relazione al cambiamento che inducono sugli studenti e che costituisce anche un fattore di controllo di molte altre variabili.

Un altro errore ancora è fare affidamento su una correlazione “spuria”, osservata tra due variabili e concludere che essa sia un indicatore di una relazione diretta tra le variabili, di tipo “causa-effetto”, quando invece le variabili potrebbero essere indipendenti ma avere una causa comune (sebbene nascosta o sconosciuta all'analista). Per fare un esempio di questa situazione mostriamo in Figura 2 il diagramma delle nazioni descritte attraverso la percentuale della popolazione con obesità e la speranza di vita (in anni) della popolazione in quelle nazioni: i dati sono reali, discussi

4. Alcuni esempi di errori comuni sono citati in questo blog: <https://www.datascience.com/blog/common-data-science-pitfalls>

nel n. 613 di “Le Scienze”⁵. Si osserva che esiste una correlazione positiva tra le due variabili a livello macroscopico, della popolazione. Tuttavia, il grafico non deve indurre a pensare che l’obesità sia una condizione che predispone un individuo ad una migliore speranza di vita, al contrario! Si celano due errori dietro questa errata interpretazione dei dati osservati. Il primo errore è che i dati individuali sono stati aggregati ed esaminati a livello nazionale. Se scegliessimo un minor livello di aggregazione (ad esempio analizzassimo gli individui in base al reddito, o per nazione) la correlazione tra le due variabili diverrebbe più debole o si invertirebbe all’aumentare del reddito. Nei paesi a basso reddito, il tasso di obesità è correlato positivamente con la speranza di vita perché l’obesità è un indicatore di reddito elevato, associato a sua volta all’accesso alle cure sanitarie e alla speranza di vita. Invece nei paesi ad elevato reddito, il tasso di obesità è negativamente correlato con la speranza di vita. Il secondo errore, quindi, è che le due variabili, speranza di vita e tasso di obesità, sono influenzate (ma in maniera inversa) da un fattore comune: il reddito. Mischiando in un unico dataset le due tipologie di paesi (a basso e alto reddito) i fattori predisponenti una lunga speranza di vita si confondono e danno luogo a correlazioni spurie.

Figura 2 - Correlazione tra tasso di obesità e speranza di vita nella popolazione delle nazioni



5. A. CAIRO, *L'obesità allunga la vita?*, in *Le Scienze*, n. 613, 88, settembre 2019.

3. *Alcuni risultati preliminari dell'analisi dati*

In questo intervento illustriamo alcuni esempi di risultati che abbiamo ottenuto partendo dalle analisi condotte grazie alla collaborazione tra ANAC e UNITO avviata nel mese di Gennaio 2019. Purtroppo, una quota rilevante di dati è stata acquisita solo successivamente ed abbiamo iniziato a studiare solo recentemente come sono costituiti i dati della banca dati nazionale dei contratti pubblici. Abbiamo a disposizione:

- il sistema informativo gare (SIMOG) sulle procedure di gara bandite dalle amministrazioni pubbliche.
- la banca dati SMARTCIG che riguarda solo i contratti sotto la soglia di 40000 euro.
- l'anagrafe delle stazioni appaltanti.

In questo contesto abbiamo avuto problemi a trovare i dati degli enti aggiudicatari delle gare, la cui presenza in banca dati ha forti discontinuità. Inoltre, un'informazione che risulterebbe utilissima da avere per "segmentare"⁶ la popolazione delle imprese per dimensione e tipologia sarebbe l'informazione del codice ATECO⁷ di classificazione delle attività economiche, che informa sull'area di attività dell'impresa. Il campo c'è ma è sempre vuoto. Per ottenerlo, ci si potrebbe rivolgere alle camere di commercio, tuttavia con costi e tempi prolungati.

4. *Analisi delle tipologie di procedura di selezione dell'impresa aggiudicataria*

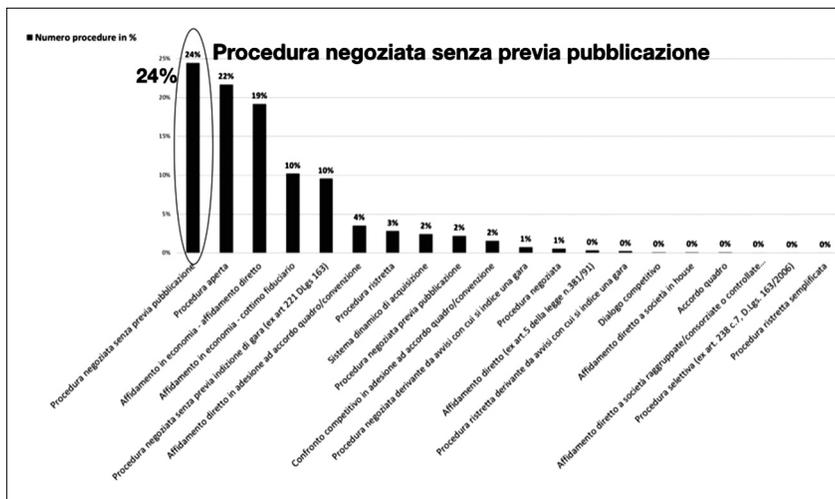
Stiamo costruendo dei supporti conoscitivi per supportare le pubbliche amministrazioni avendo iniziato ad elaborare i dati a disposizione. Ad esempio, relativamente alla tipologia di procedura di selezione dell'ente aggiudicatario nei bandi di gara degli appalti sotto soglia in un grande comune italiano, la Figura 3 mostra un diagramma a barre riassuntivo con alcuni indicatori utili per ANAC per rilevare le consuetudini operative. Si osserva come una quota rilevante delle procedure di selezione persegua

6. Segmentare una popolazione di esempi è analogo a partizionare gli esempi in insiemi omogenei per caratteristiche descrittive.

7. Il codice ATECO, approvato dall'ISTAT, è un identificatore che descrive un'attività economica tramite il suo macro-settore economico e l'articolazione in sottocategorie (a diverso livello di dettaglio) dei settori economici dell'attività stessa.

una operatività che prevede una negoziazione senza previa pubblicazione. Una percentuale così elevata potrebbe essere, a nostro avviso, un indicatore da considerare per indagare ulteriormente su fenomeni di collusione o bandi di gara sospetti o anomali.

Figura 3 - Tipologia delle procedure per la scelta del contraente in un grande comune



5. Opportunità di benchmarking

Un'ulteriore attività di analisi dati, di tipo prescrittivo, potrebbe confrontare i valori degli indicatori osservati con modelli di riferimento, per attuare un'attività di *benchmarking*. Un modello di riferimento potrebbe essere ottenuto in seguito all'aggregazione dei valori degli indicatori osservati in una tipologia di casi omogenei. Un modello siffatto descrive le attività tipiche di quel segmento di casi. Successivamente, caso per caso, si potrebbero verificare quali amministrazioni o enti si discostano da questi modelli di riferimento. L'obiettivo è quello di giungere ad una maggior efficienza della pubblica amministrazione e arrivare a supportare le amministrazioni aggiudicatrici.

Ad esempio il RUP, che deve definire un importo ragionevole del lotto a base d'asta in un bando di gara, potrebbe beneficiare di tali modelli di riferimento per avere un'idea di come sono distribuiti gli importi nei bandi

di gara simili sul territorio nazionale; oppure l'analista potrebbe capire se un ribasso, su un importo di gara, è eccessivo e sospetto oppure è legato ad una buona efficienza degli operatori economici che hanno risposto al bando di gara.

6. *Usa del Common Procurement Vocabulary*

In questo studio vediamo il caso del consumo di carta per ufficio. Prendiamo il caso di un grande ente e osserviamo cosa accade quando bandisce i bandi di gara. Nella descrizione del bando di gara si usa un vocabolario e i codici di *Common Procurement Vocabulary* (CPV)⁸ che costituiscono un'ontologia stabilita a livello europeo, comune a tutti i paesi membri per descrivere i beni, i lavori, le forniture che sono oggetto di gare pubbliche. È un'ontologia a quattro livelli gerarchici successivi, estesissima, composta da centinaia di migliaia di codici diversi. È utile e precisa per coprire tutto lo scibile, ma costituisce un ostacolo all'operatività del RUP che tendenzialmente ha poco tempo per analizzare tutti i dettagli e le raffinatezze di questa classificazione e può scegliere fino a 20 codici diversi per descrivere l'oggetto del bando di gara. Questa operazione di scelta del codice CPV potrebbe condurre a una scelta non esaustiva dei codici e potrebbe produrre un elemento di imprecisione nei dati. Inoltre il criterio con cui il RUP sceglie un codice per descrivere l'oggetto principale del bando di gara potrebbe non essere lo stesso usato dalle imprese che rispondono ai bandi per quel tipo di prodotti. Consideriamo l'esempio dell'acquisto di risme di carta: limitandosi ai codici CPV che contengono la parola "carta" nella descrizione del prodotto, si trovano moltissimi codici che cambiano classificazione e categoria⁹. Per generare problemi sarebbe sufficiente per una impresa scegliere codici diversi per non trovare tutti i bandi di gara pertinenti.

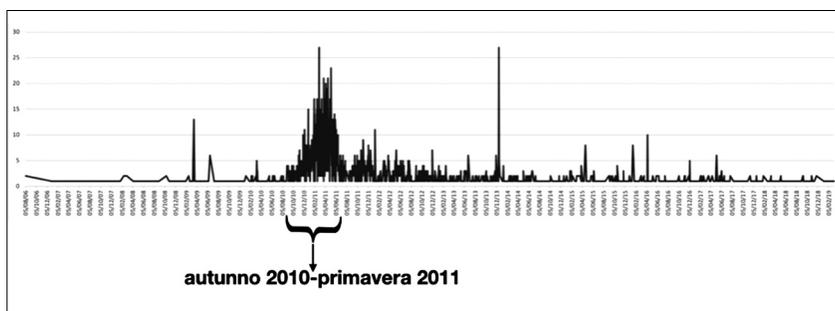
8. Il CPV è un sistema di classificazione dell'oggetto degli appalti che permette di unificare i riferimenti utilizzati negli appalti pubblici a livello della comunità europea. Si veda: <https://simap.ted.europa.eu/it/cpv>

9. Ad esempio si trovano codici di diverso gruppo, classe e categoria per "carta o cartone manufatto", "carta e cartone fotosensibili, termosensibili o termografici" e "carta e cartone trattati"; si trovano codici con diversa categoria associati a descrizioni che sembrano riferite allo stesso prodotto come per "carta per fotocopie" e "carta da stampa". La scelta corretta del codice da indicare dunque non sembra facile.

7. Analisi dell'acquisto di carta nel tempo

Un'altra analisi riguarda lo studio della distribuzione nel tempo dei bandi per l'acquisto di carta (all'interno del periodo temporale di riferimento dei dati a nostra disposizione, cioè dal 2012 al 2017). Mostriamo tale analisi in Figura 4.

Figura 4 - Serie storica del numero di procedure di acquisto carta bandite dagli enti



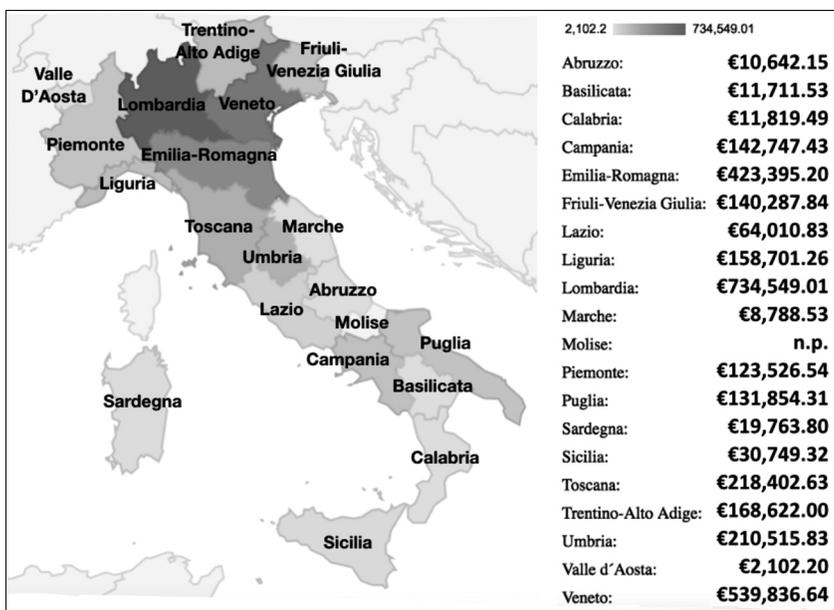
Si rileva la presenza di un periodo temporale, ossia l'autunno 2010 - e la primavera 2011, dove i bandi di gara per l'acquisto di carta sono molto concentrati rispetto al resto della serie storica: bisognerebbe approfondire nello specifico l'analisi per capire cosa è successo in quell'esatto periodo e verificare se c'erano scadenze di fondi, cambi di prezzi o di condizioni di acquisto ecc.

7.1 Analisi della distribuzione dell'acquisto di carta sul territorio

Successivamente si è analizzata una mappa descrittiva del comportamento, sul territorio italiano, di alcune pubbliche amministrazioni (nello specifico le Università) per l'acquisto del bene: "carta". Abbiamo prodotto una mappa del territorio nazionale sugli importi totali dei bandi di gara suddivisi per regione, nel periodo 2012-2017, utilizzando per selezionare i bandi di gara un insieme di 12 CPV la cui descrizione contiene la parola "carta". Il risultato è indicato in Figura 5. Si osserva che tipicamente le regioni più attive a livello economico nel nostro Paese assumono colori più scuri, indicanti maggiori importi¹⁰.

10. Un'eccezione è costituita dalla regione Molise alla quale, nel periodo indicato, non risultano presenti dati relativi a procedure di acquisto per il bene in questione.

Figura 5 - Importi totali per l'acquisto di carta da parte delle Università italiane



Tuttavia quelle regioni sono sicuramente ad alta densità di popolazione. Ci chiediamo cosa capiterebbe se *normalizzassimo* gli importi totali, aggregati per regione, per numero di abitanti. Il risultato fornisce la spesa pro-capite per abitante residente in regione. La mappa cambia completamente ed è mostrata in Figura 6 nella pagina seguente.

Qualora volessimo fare uno scoop giornalistico andremo a far vedere che ci sono il Trentino Alto Adige, l'Umbria e la Liguria che spendono in carta cifre pro capite di quasi 2 ordini di grandezza superiori a Lazio, Piemonte, Sardegna, Sicilia, Calabria e Abruzzo. Come mai? Abbiamo diviso gli importi totali per la quantità di popolazione, quindi le regioni più piccole, con minor popolazione, finiscono in alto nella classifica. Questo è un risultato dovuto alla normalizzazione fatta per popolazione regionale. Se invece normalizzassimo per il numero di studenti, ai quali si rivolge molta dell'attività universitaria (ad esempio per la stampa dei materiali di progettazione didattica, dispense, ecc.) notiamo che la classifica cambia leggermente (in Figura 7) ma rimane simile a quella precedente.

Figura 6 - Importi stanziati dalle Università italiane per l'acquisto di carta pro-capite

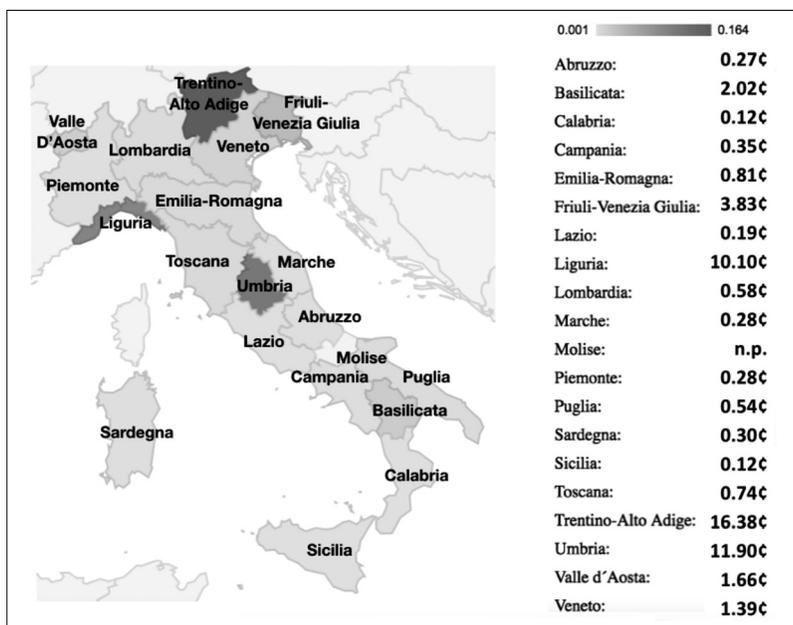


Figura 7 - Importi per acquisti di carta normalizzati per la popolazione studentesca



È stata determinante in questo caso la quantità di studenti per regione: nel risultato della classifica delle regioni (di Figura 7) vengono per prime le regioni che non hanno grandi Atenei, ma medio-piccoli. Infine, se normalizzassimo gli importi totali per numero dei docenti (in fin dei conti, coloro che stampano il materiale di progettazione didattica sono i docenti!), la mappa torna simile a quella iniziale (si veda la Figura 8): si osserva che le regioni con una spesa più alta in acquisto di carta sono le regioni del Nord-Est. Tuttavia, a un'analisi più accurata, risulterebbe che Trentino Alto Adige ed Umbria siano regioni i cui Atenei spesso forniscono servizi supplementari agli studenti, quali la carta per consentire di stampare il materiale didattico. Ciò giustificerebbe la spesa osservata sopra la norma.

La conclusione di questo studio è che anche partendo dagli stessi dati, gli importi totali stanziati dalle Università per l'acquisto di carta, a seconda della modalità con la quale si normalizzano gli importi delle varie regioni, la classifica delle regioni cambia. Se ne deduce che non vi è un metodo unico per costruire le statistiche che sia completamente attendibile: a seconda di come si preparano e si presentano i dati, il messaggio finale cambia!

Figura 8 - Importi per l'acquisto di carta normalizzati per il numero dei docenti

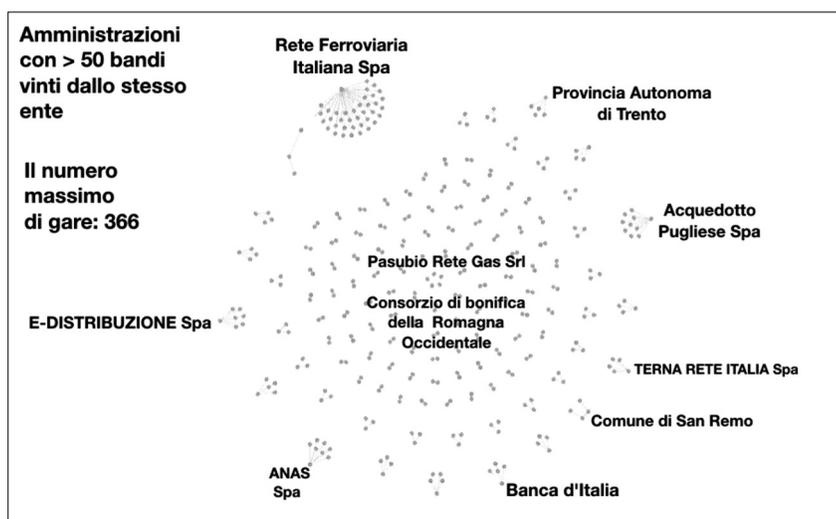


8. Il grafo delle gare di appalto dei lavori pubblici

Un altro obiettivo che ci siamo posti è costruire il grafo degli appalti pubblici che abbia come nodi le entità coinvolte nell'appalto (rispettivamente, l'amministrazione che bandisce e l'operatore economico o impresa che risponde alla gara di appalto). Se due entità del grafo sono collegate da un arco significa che una ha emesso un bando di gara e l'altra ha risposto (e si è aggiudicata) il bando. Mostrare il grafo completo non sarebbe chiaro a causa dell'alta densità di nodi con collegamenti reciproci. Aggiungendo dei filtri in base al numero delle volte che un'impresa aggiudicatrice si è aggiudicata un bando di gara, bandito dalla stessa amministrazione pubblica, il grafo si semplifica e l'immagine diventa più leggibile. Questa difficoltà a rendere comprensibili i risultati evidenzia le difficoltà di comunicare i *big data*.

Ad esempio imporre la soglia del numero di appalti vinti da una impresa a 50, significa che l'impresa che si è aggiudicata il bando di gara lo ha fatto almeno 50 volte nell'arco di tempo analizzato e con la stessa amministrazione pubblica o ente appaltante. Un numero esageratamente elevato potrebbe essere un indicatore di anomalia da approfondire. Il grafo corrispondente alla soglia impostata al valore 50 si osserva in Figura 9.

Figura 9 - Il grafo degli appalti pubblici (per numero di appalti banditi dallo stesso Ente e aggiudicati alla stessa Impresa, con soglia = 50)

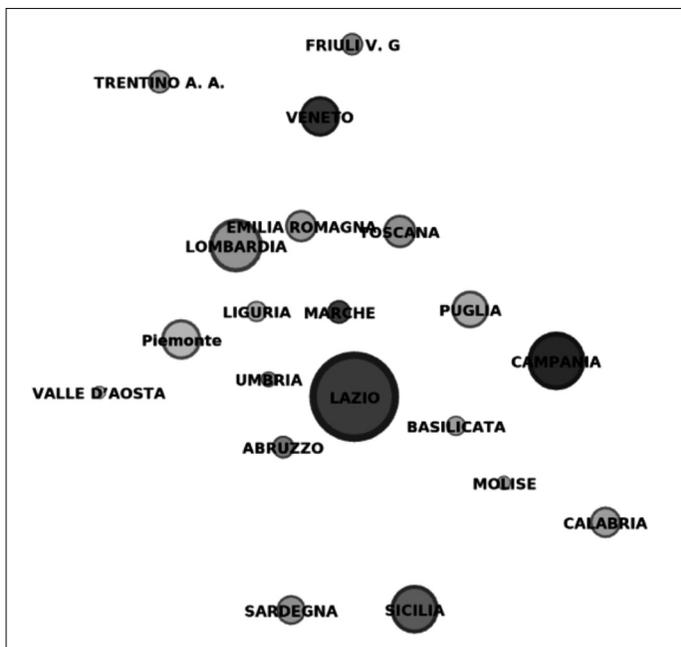


Il grafo riguarda le grandi opere ed emergono principalmente le grandi imprese come le Ferrovie dello Stato, ANAS e altri enti di distribuzione di energia. È curioso notare come nell'arco di 5 anni vi sia un'impresa che si è aggiudicata addirittura 366 gare bandite dallo stesso ente (6 al mese per 5 anni consecutivi!).

8.1 *Analisi del grafo degli appalti dal punto di vista territoriale*

Ora analizziamo il grafo dal punto di vista territoriale. Se noi aggregassimo in un macro-nodo gli enti e gli aggiudicatari residenti nella stessa regione, quei nodi si ingrandirebbero in proporzione. La dimensione dei nodi dipende anche dalla quantità di bandi pubblicati sul territorio regionale. Il risultato del grafo è mostrato in Figura 10.

Figura 10 - Il grafo degli appalti con aggregazione a livello regionale



Il nodo più grande è riferito alla regione Lazio, che ha il massimo numero di bandi; seguono Lombardia, Campania e Sicilia. I risultati fanno emergere un atteggiamento di tipo localistico nell'assegnazione dei bandi

ad imprese del territorio. Anche la vicinanza relativa dei nodi del grafo in figura ci dice quanto le amministrazioni di quei territori tendono ad affidare bandi ad enti di regioni limitrofe: più sono “vicini” i nodi e più sono esclusive le amministrazioni nell’assegnazione dei bandi ad imprese di quei territori. Ad esempio il Trentino tende a non assegnare appalti ad enti di altre regioni, come accade anche in Veneto e in Friuli Venezia-Giulia.

9. *Modelli predittivi*

I modelli di *machine learning* di tipo supervisionato richiedono che nel data set che viene analizzato e che chiamiamo *training set* (insieme di addestramento degli algoritmi ad apprendimento automatico), deve essere presente un *quid* di informazione che facilita l’algoritmo a capire la tipologia del bando di gara: regolare piuttosto che anomalo. Queste informazioni attualmente non sono a nostra disposizione. Per ottenerle dovremmo incrociare informazioni da banche dati diverse: per esempio, con riferimento al portale Internet della giustizia amministrativa¹¹, se riuscissimo ad incrociare le sentenze con i bandi di gara, troveremmo i bandi sui quali vi è stato contenzioso e potremmo andare ad etichettare quei bandi come sospetti. Di conseguenza si potrebbe costruire un modello predittivo lanciando un algoritmo di apprendimento automatico che cerchi di capire quali sono le caratteristiche tipiche di quei bandi anomali. L’insieme di tali caratteristiche costituirebbe un modello tipico e prototipale dei bandi di gara sospetti. Successivamente, sui bandi di gara in corso, si potrebbe applicare lo stesso modello per ottenere un indicatore che dica quali sono i bandi di gara sospetti, prima ancora che venga coinvolto il giudice amministrativo. Questo metodo è utile per avere dei modelli *proattivi*, ossia per sapere in anticipo e poter indagare sull’effettiva regolarità dei bandi oppure no. Al momento non abbiamo ancora avuto modo di incrociare la banca dati delle gare pubbliche con le sentenze amministrative, ma contiamo di farlo presto.

Come funziona l’apprendimento di un algoritmo automatico? Il *training set* contiene la descrizione di due tipologie di gare: quelle su cui l’esperto del dominio ha già detto che i bandi sono davvero regolari, e quelle davvero irregolari. L’algoritmo “macina” questi dati e produce il modello discriminativo che viene applicato ad una gara attuale sulla quale l’algoritmo predice la regolarità o meno.

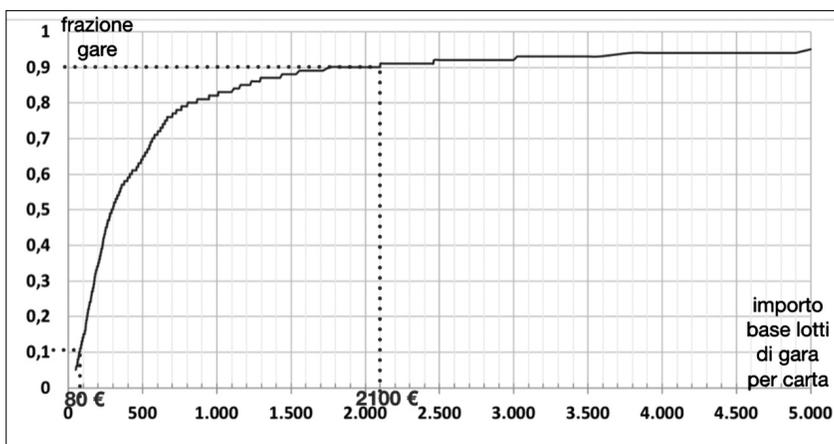
11. <https://www.giustizia-amministrativa.it/web/guest/dcsnpr>.

9.1 Esempio di analisi prescrittiva

L'analisi prescrittiva riguarda algoritmi di simulazione ed ottimizzazione su queste tipologie di dati al fine di suggerire delle buone pratiche nel futuro. A questo proposito facciamo un'ultima considerazione su un possibile modello prescrittivo che funge da suggerimento di un intervallo di prezzi raccomandati per un bando di gara: i prezzi non sono in realtà molto significativi perché non siamo riusciti a normalizzare gli importi per l'acquisto di risme di carta in modo unitario. Abbiamo quindi gli importi dei lotti dei bandi di gara, ma non sappiamo la quantità delle risme nei lotti e quindi abbiamo un *range* di valori molto ampio e non comparabile. Tuttavia sono stati usati questi dati per mostrare questo esempio paradigmatico di come avremmo potuto operare se avessimo potuto normalizzare questi lotti per quantità (numero delle risme di carta ordinate).

Abbiamo costruito la distribuzione dei prezzi di gara sugli ordinativi della carta. In Figura 11 mostriamo sulle ascisse l'importo base dei lotti dei bandi di gara e sull'ordinata corrispondente, mostriamo la frazione dei bandi di gara che hanno un importo fino a quel valore estremo. La cifra di 300 euro è il valore mediano¹² della distribuzione.

Figura 11 - Distribuzione degli importi per gli acquisti di carta



12. Il valore mediano di una distribuzione si trova a metà dell'ordinamento dei valori.

Se volessimo raccomandare un intervallo di prezzi, potremmo escludere le code della distribuzione (ad esempio, il 10% dei bandi di gara i cui importi dei lotti siano o troppo bassi o troppo alti). Potremmo cioè tenere la parte centrale della distribuzione che utilizza il rimanente 80% nelle gare ed avremmo un intervallo di prezzi raccomandati. In Figura 11 l'intervallo dei prezzi risulta variare da 80 euro a 2.100 euro.

Dovremmo fare attenzione che stiamo parlando di Università che sono molto diverse tra loro: si va dai piccoli Atenei che hanno 10000 iscritti al massimo, ai Mega Atenei, con oltre i 40000 iscritti. Se *segmentiamo* i bandi di gara in base alla taglia degli Atenei, cambiano le distribuzioni degli importi dei bandi di gara e gli intervalli di prezzi raccomandati. Per i piccoli Atenei si va dai 72 ai 1280 euro; se guardiamo ai Mega Atenei, l'intervallo dei prezzi raccomandato diventerebbe da 300 euro circa a 5000 euro circa. Questa segmentazione dei bandi di gara per tipologia di Ateneo è raccomandata per confrontare i bandi di gara pubblicati da Atenei comparabili tra di loro.

10. Conclusioni

I dati sono importantissimi nella nostra era. Sono importanti per migliorare la qualità dei servizi offerti dalle pubbliche amministrazioni e la loro efficienza.

Tuttavia, dobbiamo prestare attenzione all'esperienza fatta con le mappe descrittive del territorio nazionale. Nel caso di esempio della spesa delle Università per carta, abbiamo visto che, a seconda di come si realizza la normalizzazione dei dati, il messaggio comunicativo cambia completamente. Il ranking delle regioni virtuose per spesa cambia drasticamente. Quindi bisogna fare attenzione alla normalizzazione dei dati e alla aggregazione, anche su dati puliti. Occorre individuare un metodo unificato di normalizzazione, per tenere conto di tutti i fattori coinvolti nel fenomeno analizzato. È questo uno dei nostri obiettivi di lavoro in futuro.

Infine, bisogna prestare attenzione ai legami tra le variabili dei diversi modelli descrittivi ed esplicativi così da non confondere i legami tra le variabili correlate tra di loro con le relazioni di tipo causa-effetto, che sono dei legami dal significato profondamente diverso.

Ringraziamenti

Vorrei ringraziare i miei collaboratori, i dott. Mirko Lai e Paolo Pasteris, senza il cui supporto questo studio non sarebbe stato possibile e il gruppo di ricerca dei proff. Racca e Cavallo Perin.

Gli *open data* come strumento di legittimazione delle istituzioni pubbliche?*

ABSTRACT: L'articolo muove dalla considerazione che l'amministrazione, in quanto istituzione democratica, necessita del consenso dei consociati sul proprio operato e individua negli *open data* un importante strumento per raggiungere questo obiettivo. Infatti, gli strumenti tecnologici, unitamente all'accessibilità totale dei dati, permette di sviluppare un controllo sulla gestione della cosa pubblica e, conseguentemente, una maggiore condivisibilità delle scelte adottate. L'articolo prende in considerazione ulteriori effetti positivi realizzati dagli *open data*, come il coinvolgimento dei cittadini nei percorsi di cambiamento organizzativo, la responsabilizzazione dell'amministrazione e, non meno importante, la circolazione di conoscenza per favorire il progresso sociale e nuove forme di collaborazione.

1. Introduzione

Si è osservato, nell'affrontare il tema dei cd. *dati aperti* (*open data*)¹, che essi potrebbero fungere da strumenti di legittimazione delle istituzioni

* Articolo pubblicato in *Diritto e Società*, 3, 2019, 443-475.

1. La dottrina sull'*open government* è vasta: si v. M.J. AHN, *Whither E-Government? Web 2.0 and the Future of E-Government, Web 2.0 Technologies and Democratic Governance*, in *Pub. Adm. Inf. Tech.*, 2012, 169; L. AL-HAKIMCHE, *Global E-Government: Theory, Applications and Benchmarking*, Hershey, 2007; F. BANNISTER - R. CONNOLLY, *The Trouble with Transparency: A Critical Review of Openness in e-Government*, in *Policy and Internet*, vol. 3, 2011, 1; P.J. BIRKINSHAW - A. HICKS, *The law and public information on UK. quality, access and re-use*, *Dir. Pub.*, 2007, 959; L. BLOMGREN BINGHAM, *The next generation of administrative law: building the legal infrastructure for collaborative governance*, in *Wis. L. Rev.*, 2010, 297; A. BONOMO, *Informazione e Pubblica Amministrazione. Dall'accesso ai documenti alla disponibilità delle informazioni*, Bari, 2012; B. BUGARIC, *Openness and transparency in public administration: challenges for public law*, *Wiscon. Intern. law journ.*, 2004, 483; E. CARLONI, *La qualità delle informazioni diffuse dalle amministrazioni negli Stati Uniti*, *Giorn. dir. Amm.*, 2002, 1232; ID., *Nuove prospettive della trasparenza ammi-*

nistrativa: dall'accesso ai documenti alla disponibilita' delle informazioni, *Dir. pub.*, 2005, 573; ID., *La qualità delle informazioni pubbliche. l'esperienza italiana nella prospettiva comparata*, *Riv. trim. dir. pub.*, 2009, 155; ID., *L'amministrazione aperta. Regole strumenti limiti dell'open government*, Rimini, 2014; A. CHADWICK, *Web 2.0: New Challenges for the Study of E-Democracy in an Era of Informational Exuberance*, *I/S: A Journal of Law and Policy*, 2008-2009, 9; A. CLARKE - H. MARGETTS, *Governments and citizens getting to know each: other? Open, closed, and Big Data in public management reform*, in *Policy and Internet*, 2014, 393; B. COCCAGNA, *Libero accesso nelle politiche di open data: trasparenza, apertura e autoorganizzazione nel riutilizzo delle informazioni del settore pubblico*, in *Cib. dir.*, 2011, 129; C. COGLIANESE, *Citizen Participation in Rulemaking: Past, Present, and Future*, *Duke L. J.*, 55, 2005-2006, 943; ID., *The Transparency President? The Obama Administration and Open Government*, in *Governance*, 2009, 529; S. COLEMAN, *Making the E-Citizen: A Socio-Technical Approach to Democracy*, *Connecting democracy: online consultation and the flow of political communication*, a cura di S. Coleman e P.M. Shane, Cambridge (US), 2011, 380; E. DE BLASIO, *Democrazia digitale. Una piccola introduzione*, Roma, 2014; J.M. DE FIGUEIREDOT, *E-Rulemaking: Bringing Data to Theory at the Federal Communications Commission*, *Duke L. J.*, 55, 2005-2006, 969; M.C. DE VIVO - A. POLZONETTI - P. TAPANELLI, *Open data, Business intelligence e Governance nella Pubblica amministrazione*, *Inf. Dir.*, 2011, 239; F. DI DONATO, *Lo stato trasparente. Linked open data e cittadinanza attiva*, Pisa, 2010; F. DI MASCIÒ, *Miti e realtà degli "Open data" all'italiana*, *Giorn. dir. amm.*, 2017, 399; P. DUNLEAVY - H. MARGETTS et al., *New Public Management is Dead-Long Live Digital-Era Governance*, *J. Public Adm. Res. Theory*, 2006, 467; ID., *Digital Era Governance. IT Corporations, the State, and e-Government*, Oxford, 2006; ID., *The second wave of digital era governance*, *ssrn.com*, 2010; F. FAINI, *La strada maestra dell'open government: presupposti, obiettivi, strumenti*, *Cib. dir.*, 2013, 213; ID., *Trasparenza, apertura e controllo democratico dell'amministrazione pubblica*, *Cib. dir.*, 2014, 39; C.R. FARINA, *Achieving the potential: the future of federal e-rulemaking. Report of the committee on the status and future of federal e-rulemaking*, *Admin. L. Rev.*, 2010, 279; R. FARINA - P. MILLER et al., *Rulemaking in 140 characters or less: social networking and public participation in rulemaking*, *Pace Law Rev.*, 2011, 382; C.R. FARINA - M.J. NEWHART et al., *Rulemaking 2.0*, *U. Miami L. Rev.*, 2010-2011, 395; J. FISHENDEN - M. THOMPSON, *Digital government, open architecture, and innovation: why public sector it will never be the same again*, *Journ. Pub. Adm. Res. Theory*, 2012, 977; A. FUNG - M. GRAHAM - D. WEIL, *Full disclosure, The Perils and Promises of Transparency*, Cambridge, 2007; M. GASCO-HERNANDEZ (a cura di), *Open Government. Opportunities and Challenges for Public Governance*, New York, 2014; K. GAVELIN - S. BURALL - E. WILSON, *Open Government: Beyond Static Measures*, 2009; B. GOLDSTEIN - L. DYSON (a cura di), *Beyond Transparency*, S. Francisco, 2013; T. IM - W. CHO - G. PORUMBESCU - J. PARK, *Internet, trust in government, and citizen compliance*, *J. Public Adm. Res. Theory*, 2012, 741; B. KOHLER-KOCH - B. FINKE, *The institutional shaping of EU-society relations: a contribution to democracy via participation?*, *Journ. Civ. Soc.*, 2007, 205; H. LANDEMORE, *Inclusive Constitution-Making: The Icelandic Experiment*, *Journ. Pol. Phil.*, 2015, 166; D. LATHROP - L. RUMA, *Open Government: Collaboration, Transparency, and Participation in Practice*, Sebastopol, 2010; D. LINDERS, *From e-government to we-government: Defining a typology for citizen coproduction in the age of social media*, *Gov. inform. quart.*, 2012, 446; G. MANCOSU, *Trasparenza amministrativa e*

open data: *un binomio in fase di rodaggio*, *federalismi.it*, 2012; F. MARCELLI - P. MARSOC-CI - M. PIETRANGELO (A CURA DI), *La rete Internet come strumento di partecipazione politica. Una prospettiva giuridica*, Napoli, 2015; D. MARONGIU, *I dati delle pubbliche amministrazioni come patrimonio economico nella società dell'informazione*, *Inf. dir.*, 2008, 355; F. MARZANO, *La trasparenza nella Pubblica Amministrazione passa dall'Open Data o l'Open Data passa dalla trasparenza?*, *Inf. dir.*, 2011, 287; A. MCAFEE - E. BRYNJOLFSSON, *Big data: The management revolution*, *Harv. Bus. Rev.*, 2012, 60; K. MCNUTT, *Public engagement in the Web 2.0 era: Social collaborative technologies in a public sector context*, *Can. Publ. Adm.*, 2014, 49; F. MERLONI (a cura di), *Introduzione all'@government. Pubbliche amministrazioni e società dell'informazione*, Torino, 2005; F. MERLONI - R. CAVALLO PERIN (a cura di), *Al servizio della Nazione. Etica e statuto dei funzionari pubblici*, Milano, 2009; F. MINAZZI, *Il principio dell'open data by default nel codice dell'amministrazione digitale: profili interpretativi e questioni metodologiche*, *Federalismi.it*, 2013; G. MISURACA - F. MUREDDU - D. OSIMO, *Policy-making 2.0: unleashing the power of Big Data for public governance*, 171; K. MOSSBERGER - C. TOLBERT, *The effects of E-Government on trust and confidence in government*, *Pub. Adm. Rev.*, 2003, 66; T. NAM, *Suggesting frameworks of citizen-sourcing via Government 2.0*, *Gov. inf. Quart.*, 2012, 12; P.G. NIXON - V.N. KOUTRAKOU - R. RAWAL (a cura di), *Understanding E-Government in Europe: Issues and challenges*, Londra-New York, 2010; B.S. NOVECK, *Wiki Government: how technology can make government better, democracy stronger, and citizens more powerful*, Washington D.C., 2009; W. PARKS, *The Open Government Principle: Applying the Right to Know under the Constitution*, *George Wash. Law Rev.*, 1957, 1; P. PATRITO - F. PAVONI, *La disciplina del riutilizzo dei dati pubblici dal punto di vista del diritto amministrativo*, *Dir. inf.*, 2012, 87; A. PELED - N. KARINE, *Towards Open Data for Public Accountability: Examining the US and the UK Models*, *papers. ssrn.com*, 2015; O. PEREZ, *Open Government, technological innovation, and the politics of democratic disillusionment: (E-)Democracy from Socrates to Obama*, *I/S, Journ. law pol. inf. Soc.*, 2012, 66; P. PIRAS, *Servizi pubblici e nuove tecnologie*, *Dir. inf.*, 2006, 93; B. PONTI, *Il patrimonio informativo pubblico come risorsa. i limiti del regime italiano di riutilizzo dei dati delle pubbliche amministrazioni*, *Dir. pub.*, 2007, 996; ID. (a cura di), *Il regime dei dati pubblici. Esperienze europee e ordinamento nazionale*, 2007; C. RAIOLA, *Le consultazioni telematiche delle Autorità indipendenti. Gli effetti dell'AIR su tecniche e caratteristiche*, *osservatorioair.it*, 2012; C.G. REDDICK, *Citizen interaction and e-government: evidence for the managerial, consultative, and participatory models*, in *Transforming Government: People, Process and Policy*, 2011, 167; R. REINIKKA - J. SVENSSON, *The Power of Information in Public Services: Evidence from Education in Uganda*, *Journal of Public Economics*, 2011, 956; S. RICCI, *Note in tema di "riutilizzo dell'informazione pubblica" e diritto alla privacy*, in *federalismi.it*, 2005; D. ROBINSON - H. YU et al., *Government Data and the Invisible Hand*, *Yale J.L. & Tech.*, 2009, 160; N. RODEAN, *Participatory Democracy: Mechanism of Better Regulation in Europe*, *Acta U. Danubius Jur.*, 2011, 31; S. RODOTÀ, *Tecnopolitica*, 2, Roma-Bari, 2004, 112; D. RUSHKOFF, *Open source democracy. How online communication is changing offline politics*, Londra, 2003; L. SARTORI, *Open Government: what else?*, *Istit. Federal.*, 2013, 753; M. SAVINO, *The right to open public administrations in Europe: emerging legal standards*, *Sigma Paper n. 46*, Oecd, 2010; J. SHKABATUR, *Transparency with(out) accountability: Open Government in the United States*, *Yale Law Pol. Rep.*, 2012, 79; C.R. SUNSTEIN, *Infotopia. How many minds produce knowledge*,

pubbliche². Il tema appare meritevole di approfondimento, anche alla luce dell'esperienza dell'ultimo decennio e del varo di alcune riforme (tra cui in particolare quella europea sul riutilizzo dei dati).

2. *L'istanza di legittimazione delle istituzioni*

Anzitutto, ci si deve interrogare sul significato della legittimazione delle istituzioni pubbliche, rispetto alla quale gli *open data* si porrebbero come uno strumento.

Il tema è richiamato con riferimento a molteplici situazioni diverse: ad esempio, quando si affronta il nodo delle competenze e dei poteri delle autorità amministrative indipendenti³, così come circa il fondamento

Oxford, 2006; J. TAUBERER, *Open Government Data: The Book*, 2, 2014, opengovdata.io; B. UBALDI, *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives*, *Oecd Working Papers on Public Governance*, 22, 2013; B. VALTYSSON, *Democracy in disguise: the use of social media in reviewing the Icelandic Constitution*, *Media, Culture & Society*, 2014, 52; M.R. VICENTE - A. NOVO, *An empirical analysis of e-participation. The role of social networks and e-government over citizens' online engagement*, *Gov. Inf. Quart.*, 2014, 379; A.M. WARREN - A. SULAIMAN - N.I. JAAFAR, *Social media effects on fostering online civic engagement and building citizen trust and trust in institutions*, *Gov. Inf. Quart.*, 2014, 291; H. YU - G.D. ROBINSON, *The New Ambiguity of Open Government*, *Ucla Law Rev. Disc.*, 2012, 178; X. ZHAO - H.D. XU, *E-Government and Corruption: A Longitudinal Analysis of Countries*, *Int. J. Pub. Adm.*, 38, 2015, 410. Tra i lavori più recenti, si segnala R. SANNA, *Dalla trasparenza amministrativa ai dati aperti. Opportunità e rischi delle autostrade informatiche*, Torino, 2018, il quale si occupa del tema soprattutto a 205 ss.

2. E. MINARDI (a cura di), *Codici e paradigmi per rileggere lo sviluppo locale*, Teramo, 2011, 71.

3. Sul punto della legittimazione delle autorità indipendenti, si v. E. CHELI, *Le autorità amministrative indipendenti nella forma di governo*, *Associazione per gli studi e le ricerche parlamentari. Quaderno n. 11*, Torino, 2000, 130; G. AMATO, *Autorità semi-indipendenti e autorità di garanzia*, *Riv. trim. dir. pub.*, 1997, 645; D. CORLETTI, *Autorità indipendenti e giudice amministrativo*, in P. CAVALIERI - G. DELLE VEDOVE - P. DURET, *Autorità indipendenti e agenzie*, Padova, 2003, 114; G. GRASSO, *Le autorità amministrative indipendenti della Repubblica. Tra legittimità costituzionale e legittimazione democratica*, Milano, 2006; A. RIVIEZZO, *Autorità amministrative indipendenti e ordinamento costituzionale*, *Quad. cost.*, 2005, 338. Tra i vari contributi in S. CASSESE - C. FRANCHINI (a cura di), *I garanti delle regole. Le autorità amministrative indipendenti*, Bologna, 1996, in particolare M. CLARICH, *Autorità indipendenti. Bilancio e prospettive di un modello*, Bologna, 2005; N. LONGOBARDI, *Autorità amministrative indipendenti e sistema giuridico-istituzionale*, Torino, 2004; A. PAJNO, *L'esercizio di attività in forme contenziose*, 107. Si afferma che esse non hanno legittimazione democratica, in quanto i vertici, scelti sulla base delle competenze, non vengono eletti e ope-

dell'esistenza e dell'operato delle istituzioni europee⁴. Come si è osservato in dottrina, secondo una lettura puramente formale del termine *legittimazione*, le istituzioni sono sempre legittimate, in quanto conformi ad un parametro giuridico (che si tratti di fonte fondamentale, primaria, sovranazionale è tema che qui non interessa affrontare)⁵.

In altro senso, si intende che la legittimazione proviene, più in radice, dalla sovranità popolare, così che è la comunità, nella quale le istituzioni nascono e vivono, che le legittima. In senso statico la sovranità popolare fonda il potere delle istituzioni e giustifica la cogenza delle decisioni all'interno di un dato ordinamento. In un'ottica contrattualistica la legittimazione appare come un processo e si fonda su un consenso che si deve rinnovare incessantemente. Secondo questa lettura, diventano molte le variabili che incidono su di essa: tra i tanti, i rapporti di forza politici, le conflittualità sociali, le contingenze economiche, il ruolo dei mezzi di comunicazione di massa, gli sviluppi della cultura. Ogni elemento che possa minare la credibilità del potere finisce coll'indebolire la legittimazione, per cui la legittimazione costituirebbe un fondamento *effimero e fragile*, non governabile facilmente⁶.

Il tema è affrontato quando si discute della cd. crisi delle democrazie rappresentative, e, in particolare della partecipazione dei cittadini, intesa in senso lato⁷: le manifestazioni più evidenti sono ritenute da un lato l'astensione elettorale e il disinteresse verso i partiti politici (con una contestuale avversione nei confronti della politica); dall'altro una generale

rano sulla base di una legittimazione tecnocratica. Il riconoscimento di poteri normativi in loro favore, che sarebbe in contrasto con i principi della democrazia rappresentativa, avviene in virtù della succedanea esistenza di un procedimento partecipativo, in cui i soggetti interessati possano fare valere le proprie ragioni; altra giustificazione viene data dall'indicazione legislativa degli scopi di interesse pubblico da perseguire. La garanzia della partecipazione non può però essere sostituita dalla motivazione del provvedimento, e la motivazione stessa del provvedimento deve darne atto, a pena di illegittimità.

4. B. RITBERGER, *Building Europe's Parliament. Democratic Representation Beyond the Nation State*, Oxford, 2005; A. MANZELLA, *Parlamento Europeo e Parlamenti Nazionali come Sistema*, *Rivista AIC*, 1/2015; C. PINELLI, *Il deficit democratico europeo e le risposte del Trattato di Lisbona*, in *Rassegna Parlamentare*, 2008, 925; J.H. WEILER, *La Costituzione dell'Europa*, Bologna, 2003.

5. Q. CAMERLENGO, *Ritratto costituzionale della legittimazione politica*, *Rivista AIC*, 2016, 3.

6. Q. CAMERLENGO, *Ritratto costituzionale della legittimazione politica*, cit., 23.

7. M. LUCIANI, *Il paradigma della rappresentanza di fronte alla crisi del rappresentato*, Milano, 2001; ID., *Democrazia rappresentativa e democrazia partecipativa*, *Rivista AIC*, 2003;

G. PASQUINO, *I problemi della rappresentanza politica*, in *XXI Secolo*, 2009, Treccani.it.

sfiducia nei confronti dell'operato delle istituzioni, in quanto l'organizzazione e l'apparato politico sono accusati di tendenza all'autoreferenzialità e all'autoconservazione; ancora, contribuisce a rafforzare questo fenomeno l'emersione di centri transnazionali di produzione di regole, non sottoposti ai meccanismi di responsabilità (c.d. *accountability*) propri degli ordinamenti statali⁸.

Se il fondamento della legittimazione è nel consenso sociale, le istituzioni delegittimate dal declino dello stesso, a prescindere dall'effettiva investitura popolare *ab origine*, finirebbero per porsi in contrasto con il principio di democraticità, e risulterebbero incapaci di assolvere al loro ruolo di interpreti degli interessi comunitari⁹.

La prima indicazione che viene è che la legittimazione delle istituzioni democratiche e, fra queste, dell'amministrazione, non possa esaurirsi nel solo veicolo elettorale: la legittimazione si otterrebbe, in questo senso, con il consenso dei destinatari alle decisioni assunte dall'autorità. Il che però, si anticipa, pone il tema della variabilità della propensione dei consociati al confronto dialettico e alla adesione alle scelte del potere¹⁰.

L'agire amministrativo, a tal riguardo, partecipa dello stesso fenomeno di delegittimazione: essa ha le proprie cause, in aggiunta al fenomeno già indicato relativo alla pluralità di centri di potere, nel diffuso senso d'insoddisfazione per l'operato dei dirigenti, nella scarsità di risorse economiche e strumentali, nella difficoltà delle amministrazioni di offrire un servizio competitivo ed efficiente, effettuare valutazioni tecniche complesse e risolvere questioni complesse, nella permanente elevata conflittualità, nella difficoltà di offrire tutela ai beni comuni e agli interessi sensibili¹¹.

Il problema del *deficit* democratico, in particolare riguardo la legittimazione dell'operato dell'amministrazione, consiste nel fatto che il potere esecutivo, a causa dell'interventismo proprio del *welfare state*, ha esteso funzioni e influenza, senza un'adeguata compensazione attraverso il

8. Si vv. sul punto F. PATRONI GRIFFI, *Élite politiche e amministrative tra crisi della democrazia e riforme dell'amministrazione*, *Federalismi*, 2017 e, più in generale, L. VIOLANTE (a cura di), *La fiducia nelle società democratiche*, Soveria Mannelli, 2014.

9. S.J. GOULD, *Legitimacy*, in V. BOGDANOR (ed.), *The Blackwell Encyclopedia of Political Institutions*, Oxford-New York, 1987, 333.

10. I. MARINO, *Autonomie e democrazia. Profilo dell'evoluzione dell'autonomia e della sua ricaduta sul sistema giuridico*, in *Nuove autonomie*, 2007, 197 e ss.

11. Sulle cause di crisi, tra i molteplici contributi, si segnalano S. CASSESE, *Che cosa resta dell'amministrazione pubblica?*, *Riv. trim. dir. pub.*, 2019, 1 e M. D'ALBERTI, *Lezioni di diritto amministrativo*, 4°, Torino, 2019, 389.

controllo del potere legislativo o giudiziario¹²; inoltre, l'attuale scenario richiede ai poteri pubblici interventi rapidi, di ampia portata, e l'esercizio di un'ampia discrezionalità, il che rende impensabile un ritorno al primato assoluto della legge¹³. Risulta così in crisi il paradigma weberiano, radicato su una pubblica amministrazione gerarchica, professionale e politicamente neutrale, che rinviene il fondamento della propria azione nella legge (quest'ultima prodotta da un corpo rappresentativo eletto democraticamente) che deve applicare, in un quadro di legittimazione democratica indiretta¹⁴.

3. *La partecipazione (in senso lato) come fonte di legittimazione*

Appare plausibile che cittadini meglio informati siano maggiormente in grado di comprendere e accettare il fondamento delle decisioni che li riguardano e di meglio organizzare le situazioni in cui vivono; così come che siano meglio in grado di contribuire ai processi democratici, qualora siano chiamati a partecipare attivamente alla formazione di decisioni, politiche o alla soluzione di problemi pubblici.

Specifico valore legittimante può perciò essere assunto dalla previsione di efficaci forme di partecipazione¹⁵.

Si è rilevato come sia in corso un processo di graduale avvicinamento delle istituzioni ai cittadini. Se nel XIX secolo, in presenza di una ridotta partecipazione politica, il rapporto con le istituzioni pubbliche era caratterizzato dall'adempimento dei obblighi di leva militare e dal versamento dei tributi; con il riconoscimento del suffragio universale, e anche in virtù dello stesso, il processo di maggiore coinvolgimento dei cittadini si è intensificato, fino a che si discute di una cittadina amministrativa a fianco a quella politica, caratterizzata dalla garanzia della partecipazione immediata e continuativa dei cittadini ai processi decisionali pubblici ad ogni livello dell'ordinamento¹⁶. In questo senso, la partecipazione, collocata fin

12. B. BUGARIC, *Openness and transparency in public administration: challenges for public law*, cit., 483.

13. J. HIRSCH, *The state's new clothes: NGOs and the internationalization of states, Rethinking Marxism*, 2003, 237.

14. M. WEBER, *Economia e società*, trad. it., Milano, 1961, 271, 697.

15. Si v. sul punto M.R. SPASIANO, *La partecipazione al procedimento quale fonte di legittimazione dell'esercizio del potere: un'ipotesi ricostruttiva*, *Dir. amm.*, 2002, 283.

16. M. TIMO, *Cittadinanza amministrativa e democrazia partecipativa*, Osservatorio AIC, 2016. Tra i primi a cogliere le modificazioni del rapporto tra cittadini e amministrazione

dalla fine del Settecento nell'alveo della democrazia rappresentativa, ora, sulla scorta delle esperienze anglosassoni e latino-americane, è orientata alla partecipazione alle decisioni pubbliche¹⁷.

Segno di questa nuova concezione è, come noto, il d.lgs. 33/2013 (cd. decreto trasparenza), il quale, al novellato comma 2 dell'art. 1 (d. lgs. 97 del 2016), pone lo scopo di favorire forme diffuse di controllo sul perseguimento delle funzioni istituzionali e sull'utilizzo delle risorse pubbliche e di promuovere la *partecipazione al dibattito pubblico*¹⁸.

In senso critico, si osserva che nel nostro ordinamento la partecipazione è rimasta ristretta all'intervento dei privati nel procedimento amministrativo, regolato in maniera precisa dalla normativa, ed è accusata di non essere sufficientemente inclusiva e dialogica, e di far prevalere nei confronti dell'amministrazione un approccio difensivo rispetto a quello

A. ROMANO, *Il Cittadino e la pubblica amministrazione*, in *Il diritto amministrativo degli anni 80, Atti del XXX Convegno di studi della scienza dell'amministrazione*, Milano, 1987, 160.

17. M. TIMO, *Cittadinanza amministrativa e democrazia partecipativa*, cit.

18. Tra i contributi recenti sulla trasparenza, pubblicati successivamente alla pubblicazione del d. lgs. 97 del 2016, si v. C.P. SANTACROCE, *Conquiste e arretramenti nella trasparenza amministrativa che cambia*, *Dir. soc.*, 2017, 577; A. CAUDURO, *Il diritto di accesso a dati e documenti amministrativi come promozione della partecipazione: un'innovazione limitata*, *Dir. Amm.*, 2017, 601; S. FOÀ, *La nuova trasparenza amministrativa*, *Dir. amm.*, 2017, 65; G. GARDINI, *Il paradosso della trasparenza in Italia: dell'arte di rendere oscure le cose semplici*, *Federalismi.it*, 2017; E. CARLONI, *Il nuovo diritto di accesso generalizzato e la presunta centralità degli obblighi di pubblicazione*, *Dir. amm.*, 2016, 579; S. VILLAMENA, *Il c.d. FOIA [Freedom Of Information Act] (o accesso civico 2016) ed il suo coordinamento con istituti consimili*, *Federalismi.it*, 2016; M. SAVINO, *Il FOIA italiano. la fine della trasparenza di Bertoldo (Commento a d.lg. 25 maggio 2016, n. 97)*, *Giorn. dir. amm.*, 2016, 593; D.U. GALETTA, *La trasparenza, per un nuovo rapporto tra cittadino e pubblica amministrazione: un'analisi storico-evolutiva, in una prospettiva di diritto comparato ed europeo (Relazione al 62 Convegno di Studi Amministrativi, Varenna, 22-24 settembre 2016)*, *Riv. It. Dir. Pub. Com.*, 2016, 1019; M. GIORGIO, *Trasparenza e pubblicità dei dati delle pubbliche amministrazioni (Commento a delib. ANAC 28 dicembre 2016, n. 1310)*, *Giorn. dir. amm.*, 2018, 109. Sulla trasparenza in generale, per limitarci ai contributi che sono stati maggiormente tenuti presente, G. ARENA, *Trasparenza amministrativa*, *Diz. dir. pub.*, a cura di S. Cassese, VI, Milano, 2006, 5945; E. CARLONI, *La "casa di vetro" e le riforme. Modelli e paradossi della trasparenza amministrativa*, *Dir. pub.*, 2009, 790; R. CHIEPPA, *La trasparenza come regola della pubblica amministrazione*, *Dir. econ.*, 1994, 613; C. CUDIA, *Trasparenza amministrativa e pretesa del cittadino all'informazione*, *Dir. pub.*, 2007, 99; F. MANGANARO, *L'evoluzione del principio di trasparenza amministrativa*, *astrid-online.it*, 2009; F. MERLONI (a cura di), *La trasparenza amministrativa*, Milano, 2008; A. MOLITERNI, *Il principio di trasparenza: recenti tendenze e prospettive future*, *Riv. it. sc. giur.*, 2014, 475; A. NATALINI - G. VESPERINI (a cura di), *Il big bang della trasparenza*, Napoli, 2015; B. PONTI (a cura di), *La trasparenza amministrativa dopo il d.lg. 14 marzo 2013, n. 33*, Rimini, 2013.

collaborativo, sebbene un effetto collaborativo, persino se non voluto, si produca sempre¹⁹.

In questo senso, si è sostenuto che superare la partecipazione ristretta ai soggetti che possono essere incisi negativamente dall'azione amministrativa, potrebbe aiutare, in una veste inclusiva e nell'ottica della collaborazione, della riduzione del contenzioso, del miglioramento dell'istruttoria²⁰.

4. *Gli open data come strumenti per la legittimazione*

L'attenzione verso una maggiore trasparenza e partecipazione è cresciuta anche perché l'evoluzione tecnologica rende possibili nuove forme di coinvolgimento dei cittadini. In questa ottica va valutato se si possano collocare, come strumento principale per favorire questa partecipazione, gli *open data*.

Il fenomeno degli *open data* si colloca come la voce più importante della più ampia iniziativa internazionale *Open government Partnership*²¹, alla quale ha aderito anche il governo italiano, come altri 78 paesi.

Gli *open data* sono dati pubblici, che vengono raccolti nell'ambito dell'azione amministrativa²² e messi a disposizione per incentivare la par-

19. Sull'aspetto difensivo, si v. in questo senso P. LAZZARA, *I procedimenti ad istanza di parte. Dalla disciplina generale sul procedimento (l. 241/90) alla direttiva servizi (2006/123)*, Napoli, 2008. Sulla partecipazione la dottrina è copiosa: tra i molteplici contributi, in particolare i recenti contributi di M. D'ALBERTI, *Processi decisionali delle amministrazioni pubbliche. un'analisi comparata (Relazione al Convegno annuale dell'Associazione Italiana dei Professori di Diritto Amministrativo (AIPDA), Bergamo, 5-7 ottobre 2017)*, *Dir. pub.*, 2018, 61; R. FERRARA, *La partecipazione al procedimento amministrativo: un profilo critico (Relazione al convegno "La legge generale sul procedimento amministrativo"*, Torino, 29 maggio 2017), *Dir. amm.*, 2017, 209.

20. U. ALLEGRETTI, *Democrazia partecipativa*, *Enc. dir., Annali*, V, Milano, 2011, 271; R. BIFULCO, «*Democrazia deliberativa*», in *Enc. dir., Annali*, IV, Milano, 2011, 271. Su questi temi si vv. anche G. ARENA, *Introduzione all'amministrazione condivisa*, in *Studi parlamentari e di politica costituzionale*, 1997, 29; L. BOBBIO, *Dilemmi della democrazia partecipativa*, *Dem. dir.*, 2007, 11.

21. Gli obiettivi principali sono: una maggiore trasparenza dell'azione di governo, il sostegno alla partecipazione dei cittadini alle scelte pubbliche, la lotta alla corruzione e la promozione dell'innovazione nel settore pubblico.

22. Art. 50 del CAD: *Disponibilità dei dati delle pubbliche amministrazioni. I dati delle pubbliche amministrazioni sono formati, raccolti, conservati, resi disponibili e accessibili con l'uso delle tecnologie dell'informazione e della comunicazione che ne consentano la fruizio-*

tecipazione alla gestione della cosa pubblica. Secondo la definizione più autorevole, contenuta sul sito di riferimento, *Opendefinition.org*, *aperti* sono i *dati* che possono essere liberamente utilizzati, riutilizzati e ridistribuiti a chiunque, eventualmente soggetti alla necessità di citarne la fonte e di condividerli con lo stesso tipo di licenza con cui sono stati originariamente rilasciati. Generalmente le caratteristiche degli stessi sono identificate in tre formule: *disponibilità e accesso*; *riutilizzo e redistribuzione*; *partecipazione universale*²³. I dati devono essere rilasciati a un costo non superiore a quello sostenuto per la riproduzione e devono essere disponibili in un formato modificabile²⁴, in quanto tutti gli utenti devono potere riutilizzare facilmente i dati combinarli con altri *dataset*.

I dati dal punto di vista economico sono qualificabili come *commons*, in quanto si parla di beni non rivali, inesauribili e accessibili a chiunque; però anche essi sono soggetti a recinzione, in quanto possono essere, in virtù di accordi e prassi, oggetto di appropriazione esclusiva da parte di singoli²⁵.

ne e riutilizzazione, alle condizioni fissate dall'ordinamento, da parte delle altre pubbliche amministrazioni e dai privati.

23. Uno dei manifesti di riferimento, *The Fair Data Principles*, afferma che i dati di qualità debbano essere rintracciabili, accessibili, interoperabili e riutilizzabili: <https://www.force11.org/group/fairgroup/fairprinciples>. Sul punto, si veda come ha recepito questi tre principi l'ordinamento italiano, in quanto, *ex art. 1 lett. l-ter del CAD*, sono *dati di tipo aperti: i dati che presentano le seguenti caratteristiche: 1) sono disponibili secondo i termini di una licenza o di una previsione normativa che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato; 2) sono accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti ai sensi della lettera l-bis), sono adatti all'utilizzo automatico da parte di programmi per elaboratori e sono provvisti dei relativi metadati; 3) sono resi disponibili gratuitamente attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, oppure sono resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione salvo quanto previsto dall'articolo 7 del decreto legislativo 24 gennaio 2006, n. 36.*

24. *opendatahandbook.org*: in merito ai costi, la creazione dei *dataset*, il mantenimento dei dati e la rielaborazione necessaria per renderli accessibili a chiunque possono essere costosi, laddove invece il costo di produzione degli stessi è basso, dal momento che si tratta di dati di per sé aggiornati e raccolti in maniera continuata ed automatica (F. SCIACCHITANO, *Disciplina e utilizzo degli Open Data in Italia*, *Medialaws.eu*, 289). Sul punto, si v. F. WELLE DONKER, *Funding Open Data*, in *Open Data Exposed, Information Technology and Law*, a cura di B. van Loenen et al., 2018, 55.

25. Emblematicamente, viene richiamata la *tragedia dei beni comuni* anche con riferimento ai dati aperti: J. YAKOWITZ, *Tragedy of the Data Commons*, *Harv. Journ. Law Tech.*, 2011, 1.

Dal lato dell'offerta di *open data* non troviamo solo i soggetti pubblici: anche i cittadini stessi possono produrre, e comunque producono, dati (che possono decidere o meno di rendere disponibili, o che i soggetti pubblici che li detengono possono rendere pubblici, eventualmente dopo un processo di anonimizzazione, conformemente alla normativa e nel rispetto della tutela dei dati personali); in questo senso, è la detenzione dei dati da parte dei soggetti pubblici che rende possibile l'apertura.

Una data particolarmente significativa per la regolazione dei dati aperti è costituita dall'8 dicembre del 2009, quando viene emanata dal presidente degli Stati Uniti l'"*Open Government Directive*" per promuovere la trasparenza amministrativa, la partecipazione dei cittadini e la collaborazione tra pubbliche amministrazioni, imprese e associazioni²⁶. L'azione, si è osservato, mirava a responsabilizzare i cittadini e influenzare in modo trasparente le politiche pubbliche, realizzando così quella parificazione che promuove la legittimazione²⁷. Altra tappa fondamentale è costituita dal Maggio 2013, quando è stato firmato un *executive order* che dichiarava tutte le informazioni governative aperte e leggibili dalle macchine di *default*²⁸.

L'impatto di questa direttiva, frutto delle istanze dei movimenti per l'apertura dei dati che in quegli anni si andavano affermando²⁹, è stato globale, e lo strumento per promuovere *trasparenza, partecipazione e collaborazione* sono stati appunto i dati pubblici.

Le nuove tecnologie potrebbero così porsi come lo strumento in grado di far meglio fronte alle sfide della legittimazione e di consentire uno sviluppo più inclusivo e non gerarchico. I cittadini coinvolti accrescerebbero la loro fiducia nelle istituzioni³⁰, approfittando della loro aspirazione a

26. Nel successivo maggio del 2013 il Presidente degli Stati Uniti ha emanato il *Memoandum Open data (Executive Order Making Open and Machine Readable the New Default for Government Information)*, per definire le caratteristiche degli *Open Data*.

27. V. PATRUNO, *Dieci anni di open data: ora servono "dati vivi", forumpa.it*.

28. Altra rilevante indicazione contenuta nell'*Order* è quella che stabilisce che i dipartimenti esecutivi e le agenzie debbano gestire le informazioni come un bene, attraverso il suo ciclo, per promuoverne apertura e l'interoperabilità, a salvaguardare in maniera adeguata i sistemi e le informazioni: viene, in altri termini, precisato che l'apertura è un processo e non può essere ridotta al compimento di un unico atto.

29. S. CHIGNARD, *A brief history of Open data, parisinnovationreview.com*, 2014.

30. Thomas Jefferson, in una lettera a Samuel Kercheval Monticello del 12 luglio 1816: «*making every citizen an acting member of the government, and in the offices nearest and most interesting to him, will attach him by his strongest feelings to the independence of his country, and its republican constitution*», *let.rug.nl*.

volere partecipare alla gestione del bene pubblico³¹, peraltro mediante l'uso degli stessi strumenti che alimentano la sfiducia, se è vero che, quanto più i cittadini usano le ICT, tanto più diminuisce il loro rispetto nei confronti del governo³². Si è pensato quindi che le ICT possano offrire, mediante l'apertura dei dati, quella legittimazione che altrimenti non sarebbe possibile recuperare³³.

Ci sono poi i fini indiretti che si ritiene si possano perseguire: a fianco alle già citate legittimazione delle istituzioni, democratizzazione del rapporto tra cittadino e amministrazione e accrescimento della fiducia dei cittadini nei poteri pubblici si legherebbero il coinvolgimento dei cittadini nel cambiamento organizzativo e la responsabilizzazione dell'amministrazione³⁴ e, comunque, la diffusione di informazioni comporterebbe benefici per la popolazione in termini di conoscenza, formazione dell'opinione pubblica, progresso sociale.

Con gli *open data* si osserva che si possono infatti potenziare la trasparenza, la deliberazione, la partecipazione.

Nel primo senso, la trasparenza derivante dall'apertura dei dati può essere intesa secondo due modelli, che costituiscono due facce della stessa medaglia, ed avere quindi due funzioni.

Secondo un primo approccio, grazie ai dati aumentano le possibilità di controllo sull'operato del potere pubblico, anche senza che i cittadini debbano necessariamente essere proattivi, in quanto possono continuare ad attendere alle proprie occupazioni, e al tempo stesso evitare dei danni ai propri interessi o a quello generale; questa visione sottende un atteggiamento di sfiducia nei confronti dell'amministrazione.

Secondo un altro approccio, la trasparenza che fornisce maggiori informazioni e dati, e che realizza una democratizzazione dell'interpretazione degli stessi (che non è più riservata al Governo), è uno stimolo alla partecipazione, in un'ottica di alleanza tra amministrazione e cittadino³⁵.

31. L. SARTORI, *Open Government: what else?*, Istit. Federal., 2013; ID., *Open Government: what else?*, Istit. Federal., 2013, 753.

32. T. IM - W. CHO - G. PORUMBESCU - J. PARK, *Internet, Trust in Government, and Citizen Compliance*, *Journ. Publ. Admin. Res. Th.*, 2014, 741, i quali osservano che maggiore è il ricorso a *internet* dei cittadini, minore è la fiducia nelle istituzioni, e ipotizzano che il ricorso all'*e-government* possa invertire la rotta.

33. B. BUGARIC, *Openness and transparency in public administration: challenges for public law*, *Wiscon. Intern. law journ.*, 2004, 484.

34. Riguardo all'idea che il potere si fondi sulla gestione dei dati e che quindi il processo di apertura dei dati accrescerebbe la democraticità del sistema, si v. F. MERLONI, *Trasparenza delle istituzioni e principio democratico, La trasparenza amministrativa*, a cura di F. Merloni, Milano, 2008, 4.

35. Nella letteratura statunitense si fa riferimento alla trasparenza *fishbowl* (vaschetta del pesce) e *reasoned* (ragionata) (C. COGLIANESE, *The Transparency President? The Obama*

In questa ottica, la dottrina statunitense ha evidenziato come la disciplina del FOIA e quella sull'*open government* siano distinte, anche se strettamente correlate: se infatti la prima serve ad acquisire conoscenza sul potere pubblico, affinché i cittadini possano combattere la corruzione e mantenere la responsabilità sui controllori, l'approccio della seconda aggiunge al riferimento alla trasparenza, quanto alla partecipazione e alla collaborazione; in questo senso, il FOIA viene ritenuto uno strumento, una voce dell'*Open Government*, insieme al bilancio partecipato e alle petizioni elettroniche³⁶ ed è spesso correlato agli *open data*³⁷.

Simili valutazioni possono essere effettuate anche rispetto al nostro ordinamento: il d. lgs. 33 del 2013, cd. *trasparenza*, si occupa di *pubblicazione* dei dati, laddove la normativa contenuta nel codice dell'amministrazione digitale (cd. CAD), d. lgs. 82 del 2005, si occupa strettamente di *dati aperti*.

Nel secondo senso, il rilascio di dati e informazioni è fondamentale per un dibattito aperto finalizzato alla presa di una decisione³⁸. Rispetto a questo profilo, possono essere richiamate le osservazioni relative alla partecipazione³⁹, in particolare relative al dibattito pubblico.

Nel terzo senso, infine, dati e informazioni sono la materia prima per attivare forme di collaborazione, anche finalizzate a definire strategie e implementare politiche. Le possibili implicazioni di questa terza voce

Administration and Open Government, Governance, 2009, 529): la capacità di sorvegliare, rendere pubblici i comportamenti e gli incontri è oggi resa possibile dagli strumenti tecnologici; così come la capacità di rendere evidenti le ragioni che sottostanno ad una decisione è potenziata dall'informatizzazione, ma dipende soprattutto dalla maturità dell'ordinamento in generale.

36. Sul rapporto tra FOIA e *Open Government*, si v. B.S. NOVECK, *Rights-Based and Tech-Driven: Open Data, Freedom of Information, and the Future of Government Transparency*, *Yale Hum. Rts. & Dev. L.J.*, 2017, 1. Sulla esperienza statunitense sull'accesso e l'informazione, fondamentale è il lavoro di G. ARENA, *La "Legge sul diritto alla informazione" e la pubblicità degli atti dell'Amministrazione negli Stati Uniti*, *Pol. dir.*, 1978, 279.

37. Si v. il principio 5 della *Open Data Charter*, su cui *infra*.

38. Sul punto, si ampliano gli strumenti, ma valgono le considerazioni effettuate per istituti quali l'inchiesta pubblica e il dibattito pubblico: S. CASSESE, *La partecipazione dei privati alle decisioni pubbliche. saggio di diritto comparato*, *Riv. trim. dir. pub.*, 2007, 13; L. CASINI, *L'inchiesta pubblica. Analisi comparata*, *Riv. trim. dir. pub.*, 2007, 43; P. CHIRULLI, *Istituti di partecipazione e semplificazione amministrativa, Italiadecide. Rapporto 2015. Semplificare è possibile: come le pubbliche amministrazioni potrebbero fare pace con le imprese*, a cura di C. Pinelli, 2015, 311; G. PIZZANELLI, *La partecipazione dei privati alle decisioni pubbliche. Politiche ambientali e realizzazione delle grandi opere infrastrutturali*, Milano, 2010.

39. S. COGNETTI, *Quantità e qualità della partecipazione*, Milano, 2000; A. CROSETTI - F. FRACCHIA (a cura di), *Procedimento amministrativo e partecipazione. Problemi, prospettive ed esperienze*, Milano, 2002.

restano ancora oggi abbastanza indefinite: se generalmente, in riferimento alla partecipazione, si intende il contributo che il cittadino può dare all'attività autoritativa dell'amministrazione, nel caso della collaborazione si intende il contributo alla prestazione di servizi⁴⁰. L'espressione statunitense *citizen sourcing*, utilizzata per questo fenomeno, ha punti in comune con il nostro principio di sussidiarietà orizzontale: l'amministrazione si può avvalere di un ampio e indefinito numero di cittadini come risorsa⁴¹. I cittadini forniscono informazioni, segnalano inefficienze, interagiscono con i gestori di servizi (in ogni settore, dalla sanità alla mobilità, all'energia, all'istruzione, alla sicurezza, ai servizi sociali e culturali).

Si tratta di un modello in contrapposizione con quello dell'esternalizzazione invalso negli anni novanta del secolo scorso, in quanto non vi è un ricorso ai privati in sostituzione del pubblico⁴², ma questi, sono partecipanti, e non come utenti o consumatori⁴³, e il contributo è significativo sia per l'apporto concreto che per l'accrescimento della fiducia nelle istituzioni⁴⁴. Appare eccessivo ritenere che si realizzi così una condivisione della gestione dei servizi, o una "co-produzione"⁴⁵, ma forme di collaborazione possono riguardare l'elaborazione, l'esercizio, il monitoraggio (si

40. J.VO LUCKE - K. GROSSE, *Open Government Collaboration. Opportunities and Challenges of Open Collaborating With and Within Government*, *Open Government. Opportunities and Challenges for Public Governance*, cit., 189.

41. D. LINDERS, *From e-government to we-government: Defining a typology for citizen coproduction in the age of social media*, *Gov. inform. quart.*; T. NAM, *Suggesting frameworks of citizen-sourcing via Government 2.0*, *Gov. inf. quart.*, 12. Sul principio di sussidiarietà, si vv. V. CERULLI IRELLI, «Sussidiarietà (dir. amm.)», *Enc. giur.*, XII, Roma, 2003; F. GIGLIONI, *Il principio di sussidiarietà orizzontale nel diritto amministrativo e la sua applicazione*, *Foro amm Cds*, 2009, 2909; L. GRIMALDI, *Il principio di sussidiarietà orizzontale tra ordinamento comunitario e ordinamento nazionale*, Bari, 2006; G.C. SALERNO, *Servizi di interesse generale e sussidiarietà orizzontale fra ordinamento costituzionale e ordinamento dell'Unione europea*, Torino, 2010.

42. D. LINDERS, *op. cit.*, 6; T. NAM, *op. cit.*, 12.

43. S. RODOTÀ, *op. cit.*, XIV; M. MARGOLIS - G. MORENO-RIACHO, *E-Government, customers and citizens. Understanding E-Government in Europe: Issues and challenges*, a cura di G. NIXON, V.N. KOUTRAKOU e R. RAWAL, Londra-New York, 2010, 86; l'immagine del consumatore è anche in M.R. SPASIANO, *La partecipazione al procedimento quale fonte di legittimazione dell'esercizio del potere: un'ipotesi ricostruttiva*, cit., 287.

44. Su quest'ultimo aspetto, si vv. A.M. WARREN - A. SULAIMAN - N.I. JAAFAR, *Social media effects on fostering online civic engagement and building citizen trust and trust in institutions*, *Gov. Inf. Quart.*, 2014, 291; K. MOSSBERGER - C. TOLBERT, *The effects of E-Government on trust and confidence in government*, *Pub. Adm. Rev.*, 2003, 66.

45. S. MARTIN, *Engaging with citizens and other stakeholders*, *Public management and governance*, a cura di T. Bovaird e E. Löffler, Londra, 2009, 279.

pensi alla rilevazione della qualità) di un servizio e i contributi soprattutto possono essere utili ad una migliore gestione degli stessi, maggiormente plasmata sulle esigenze degli amministrati. In questo senso, la sussidiarietà non riguarda tanto la sostituzione del soggetto privato al pubblico, ma il coinvolgimento in nuove forme di collaborazione.

E, del resto, costituisce una forma di collaborazione anche la messa a disposizione dei dati al cittadino, affinché questi li possa usare per i fini più disparati⁴⁶, anche per sviluppare dei servizi⁴⁷, i quali, in quanto connessi a servizi pubblici, indirettamente contribuiscono al miglioramento di questi ultimi.

Queste idee sono diventate patrimonio comune e globale con la pubblicazione, un mese dopo l' *order* del 2013, della *Open Data Charter*, promossa dal G8, contenente, tra i propri principi, uno riferito al miglioramento della *governance* e al coinvolgimento dei cittadini⁴⁸.

46. D. LINDERS, *From e-government to we-government: Defining a typology for citizen coproduction in the age of social media*, cit., 3.

47. M. J. AHN, *Whither E-Government? Web 2.0 and the Future of E-Government*, in *Web 2.0 Technologies and Democratic Governance*, Pub. Adm. Inf. Tech., 2012, Vol. 1, Part 2, 169.

48. Principle 5 - *For Improved Governance and Citizen Engagement*.

We recognize that the release of open data strengthens the governance of and trust in our public institutions, reinforces governments' obligation to respect the rule of law, and provides a transparent and accountable foundation to improve decision-making and enhance the provision of public services. We recognize that open data encourages better development, implementation, and assessment of programs and policies to meet the needs of our citizens, and enables civic participation and better informed engagement between governments and citizens. We recognize that engagement and consultation with citizens and civil society and private sector organizations can help governments understand which types of data are in high demand, and, in turn, can lead to improved data prioritization, release, and standardization practices. We recognize that city or local governments are often the first point of interaction between citizens and government, and that these governments therefore have a crucial role in supporting citizen engagement on open data.

We will: Implement oversight and review processes to report regularly to the public on the progress and impact of our open data initiatives; Ensure that information published as a result of transparency or anticorruption laws is released as open data; Provide training programs, tools, and guidelines designed to ensure government employees are capable of using open data effectively in policy development processes; Engage with the Freedom of Information / Access to Information / Right to Information community to align the proactive release of open data with governments' obligation to release information on request; Engage proactively with citizens and civil society and private sector representatives to determine what data they need to effectively hold governments accountable; Respect citizens' right to freedom of expression by protecting those who use open data to identify corruption or criticize governments; Encourage

5. Gli open data nelle riforme europee tra efficienza e democraticità

A livello europeo⁴⁹ è stata di recente ribadita l'importanza dei dati pubblici⁵⁰, e anzi si è deciso di potenziarne il ricorso, attraverso la nuova direttiva 2019/1024/UE, relativa all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico, che ha sostituito la direttiva sul riutilizzo dell'informazione del settore pubblico (2013/37/UE, Direttiva PSI)⁵¹. Sebbene la direttiva si ponga come un aggiornamento e un poten-

the use of open data to develop innovative, evidence-based policy solutions that benefit all members of society, as well as empower marginalized communities

49. Ci si occuperà in questo lavoro dell'ordinamento europeo. Appare tuttavia importante segnalare che a metà aprile negli Stati Uniti è entrato in vigore l'*OPEN Government Data Act* (dove OPEN sta per *Open, Public, Electronic, and Necessary*), legge *bipartisan* che stabilisce l'obbligo di pubblicazione dei dati di tipo non sensibile prodotti dalle agenzie federali in formato leggibile dalle macchine e secondo licenze che impediscano qualsiasi forma di restrizione al loro utilizzo. Si istituisce la figura dello *Chief Data Officer*, con il compito di gestire la produzione e il rilascio dei dati per garantire l'adempimento delle amministrazioni agli obblighi normativi. Si sancisce il criterio dell'*open by default* per i dati governativi. La norma prevede anche l'istituzione dello *Chief Data Officer Council*, con il compito di consulenza del Governo per lo sviluppo di nuove politiche e strategie in materia di *open data*, presso il quale i vari *Officer* scambieranno esperienze e strategie. La normativa prevede il supporto dell'*Office of Management and Budget* statunitense per la formazione di inventari, che devono essere di consultazione più agevole per gli utenti. I dati dovranno essere pubblicati su un portale pubblico, liberamente accessibile a chiunque.

50. L'interesse dell'Unione europea è anche testimoniato dai diversi *report* sviluppati negli ultimi anni: tra questi i più rilevanti sono *The economic benefits of Open Data* che esamina i benefici dell'utilizzo degli *Open Data* dal punto di vista economico; *The future of Open Data Portals*, che si occupa della sostenibilità temporale delle iniziative; *Barriers in working with Open Data* che affronta gli ostacoli al pieno sfruttamento dei dati; *Recommendations for Open Data Portals: from setup to sustainability* che si occupa di in particolare di *governance*, finanziamento e valutazione e che, tra l'altro, ha raccomandato ai portali la predisposizione e la raccolta di indicatori; *Re-using Open Data* che indaga l'utilizzo commerciale dei dati da parte del settore privato; fino al più recente *Ensuring the economic sustainability of open data portals: understanding impact and financing*.

51. La Commissione europea aveva pubblicato nel 1998 il *Libro verde sull'informazione del settore pubblico nella società dell'informazione. L'informazione del settore pubblico: una risorsa fondamentale per l'Europa*. L'Italia peraltro è stata oggetto di incompleto e scorretto recepimento della Direttiva PSI 2003/98 in riferimento ai dati ipotecari e catastali (IP/09/425). Si vedano in particolare la Comunicazione *Dati aperti Un motore per l'innovazione, la crescita e una governance trasparente* COM (2011) 882 e la Comunicazione *Verso uno spazio comune europeo dei dati* SWD (2018) 125. In mezzo, la *Strategia europea per il mercato unico digitale in Europa* (2015), che ricomprende anche questa politica. In dottrina, si v. A. ALEMANNI, *Unpacking the principle of openness in EU Law. Transparency, participation and democracy*, *Eur. Law Rev.*, 2014, 72.

ziamento della precedente, non stravolgendone i contenuti, diversi punti sono meritevoli di sottolineatura.

In linea generale, il tema dell'apertura dei dati viene presa sul serio: in primo luogo, appare evidente sin dalla riformulazione del titolo che si vuole dare maggiore rilievo all'apertura dei dati⁵², che non era indicata nella precedente versione della direttiva.

In secondo luogo, la direttiva punta alla più ampia applicazione soggettiva della disciplina possibile: essa infatti punta a promuovere il riutilizzo delle amministrazioni nazionali e di ogni livello territoriale, estendendosi anche ai soggetti privati finanziati o controllati dai soggetti pubblici, compresi musei, biblioteche ed archivi (fatti salvi i settori dell'istruzione e della ricerca, nonché radiotelevisivo)⁵³. La direttiva si estende alle imprese pubbliche operatrici di servizi pubblici (acqua, energia, trasporti, servizi postali), sebbene si preveda la possibilità e non l'obbligo di rendere questi dati disponibili: tuttavia, una volta resi disponibili per il riutilizzo, essi rientreranno nell'ambito di applicazione della direttiva.

In terzo luogo, la direttiva, sul solco della precedente, afferma la gratuità dei dati, anche se, innovando rispetto al passato, presta grande attenzione ai dati di valore e al loro possibile potenziale commerciale. Infatti, da un lato segnala come tutti i contenuti del settore pubblico accessibili ai sensi delle norme nazionali debbano essere resi disponibili gratuitamente per il riutilizzo e stabilisce che gli enti pubblici non potranno imporre tariffe superiori ai costi marginali per il riutilizzo dei loro dati, tranne che in casi eccezionali (ad es. per i dati prodotti nell'ambito della ricerca finanziata con fondi pubblici o di dati di imprese pubbliche)⁵⁴. Nel caso di dati di ricerche oggetto di finanziamento pubblico, bisognerà sviluppare politiche di accesso pubblico e facilitare il riutilizzo di dati già contenuti in pubblici archivi⁵⁵. Da un altro lato, appunto, mostra grande attenzione per i *dataset* cd. ad alto valore (contenenti statistiche, dati geospaziali, meteorologici, ambientali, di mobilità): si tratta di dati a notevole poten-

52. L'*European Data Portal* stimava in 325 miliardi di euro il valore di mercato degli *Open Data* nel periodo 2016-2020, 100.000 i nuovi posti di lavoro e 1,7 miliardi di euro di risparmio per la Pubblica Amministrazione: *euopeandataportal.eu*.

53. Anche se per musei, archivi, biblioteche, in virtù del possibile elevato valore economico dei dati, è prevista la deroga ai principi di gratuità o di pretesa dei costi marginali (cons. 38, 49, 55, 65; artt. 1 comma 2 lett. j, 3 comma 2, 4 comma 3, 6 comma 2, 14 comma 4).

54. Gli Stati membri saranno tenuti a elaborare politiche per l'accesso aperto ai dati della ricerca finanziata con fondi pubblici, mentre a tutti i dati di tale natura, resi accessibili tramite archivi, saranno applicate norme armonizzate in materia di riutilizzo.

55. Cons. 27-28; art. 10.

ziale commerciale che, negli auspici dell'Unione, possono accelerare lo sviluppo di un'ampia gamma di prodotti e servizi di informazione a valore aggiunto, compreso lo sviluppo di Intelligenza Artificiale: la Commissione dovrà identificare questi *dataset*, che saranno soggetti a regole peculiari, e che dovranno essere forniti gratuitamente⁵⁶.

In quarto luogo, si tenta di impedire che il carattere dell'apertura dei dati venga meno: in merito ai rapporti tra pubblico e privato, si punta a rafforzare la trasparenza dei relativi accordi e impedire la conclusione di accordi che potrebbero portare a un riutilizzo esclusivo dei dati pubblici da parte dei privati⁵⁷.

E purtuttavia, con rilievo particolare per il tema in oggetto, se l'attenzione per i dati pubblici è sicuramente maggiore che in passato, non compare nel testo della direttiva alcun riferimento al tema della legittimazione delle istituzioni, inteso nel senso più ampio possibile: si può leggere solo in un passaggio che *rendere pubblici tutti i documenti generalmente disponibili in possesso del settore pubblico – concernenti non solo il processo politico ma anche quello giudiziario e amministrativo – rappresenta uno strumento fondamentale per ampliare il diritto alla conoscenza, che è principio basilare della democrazia*⁵⁸. Si ritornerà sul punto *infra*, in sede di considerazioni conclusive.

6. La scarsa diffusione e qualità degli open data in Italia

Il grande nodo, una volta introdotta la normazione sugli *open data*, riguarda, in ogni ordinamento, quantità e qualità dei *data set* disponibili.

A livello nazionale, il ruolo di incentivazione che hanno rivestito gli organi centrali in questi anni ha incrementato la quantità di dati resi disponibili⁵⁹, sebbene le iniziative italiane e il numero di *dataset* rilasciati

56. Art. 13.

57. Il tema è affrontato in diversi punti. Si v. in particolare cons. 48-51 e art. 12.

58. Cons. 43.

59. Nell'ultima rilevazione del Portale Europeo dei dati, l'*Open Data Maturity*, relativa al periodo luglio 2016-giugno 2017, l'Italia si è posizionata tra i *Trendsetters* e non più tra i *followers*: europeandataportal.eu. Nel *Global Open Data index*, curato dalla *Open Knowledge Foundation*, l'Italia è in 32° posizione (preceduta da Francia, Germania, Regno Unito, Stati Uniti), index.okfn.org. Nell'*Open Data Barometer*, curato dalla *World Wide Web Foundation*, l'Italia è al 15° posto, opendatabarometer.org. La prima adozione di un catalogo di dati risale al maggio 2010, nella Regione Piemonte; le iniziative di apertura dei dati sono state stimulate da iniziative quali *Spaghetti Open Data*, *Openpolis*, *Stati Ge-*

non siano paragonabili a quelli statunitensi o britannici, il numero dei dati è in crescita⁶⁰; in questo senso, si è notato che il ritardo nell'adozione dell'accesso civico generalizzato non ha favorito un significativo incremento dei dati⁶¹.

Il tema comunque è oggetto di attenzione, tanto che è stato inserito nel *Piano triennale per l'informatica nella pubblica amministrazione 2019-2021*⁶². I dati che sarebbe bene potere aprire sono quelli relativi alle opere culturali, quelli prodotti come parte di una ricerca, i dati finanziari, come bilanci e informazioni sui mercati finanziari, dati statistici e indicatori economici, dati relativi a clima ed ambiente, dati relativi all'organizzazione e all'attività degli enti pubblici.

A livello territoriale, mentre le Regioni hanno reso disponibili diversi *data set*⁶³, più preoccupante è invece la situazione nei Comuni, i quali possiedono una parte consistente dei dati di interesse pubblico, come quelli su trasporto pubblico, turismo, cultura e attività produttive. Non solo preoccupa lo scarso numero di dati da loro messi a disposizione⁶⁴, ma più in generale l'atteggiamento di disinteresse od ostilità nei confronti delle iniziative di apertura⁶⁵. Proprio perché non sono disponibili categorie di dati pubblicati da tutti i Comuni in modo uniforme è difficile potere utilizzare quelli che sono stati resi disponibili su più larga scala a livello

nerali dell'Innovazione; il portale italiano dei dati aperti è *dati.gov.it*, che è stato realizzato nell'ottobre del 2011.

60. *Portale europeo dei dati: europeandataportal.eu.*

61. Il portale *dati.gov.it* è stato realizzato da Agid (Agenzia per l'Italia digitale) per conto del Governo ed è possibile trovare i *dataset* rilasciati dalla Pubblica Amministrazione: ne sono stati pubblicati più di 27.000 (che era il target per il 2020), circa 450 amministrazioni pubblicano dati (il *target* per il 2020 era di 300), che possono essere scaricati gratuitamente, ci sono più di 50 portali, più che in Germania e Olanda (le informazioni su *dataset* e amministrazioni sono reperibili su *avanzamentodigitale.italia.it*).

62. *agid.gov.it*: si tratta del secondo Piano triennale (il primo ha coperto il triennio 2017-2019). I Piani triennali sono previsti dall'art. 1 comma 513 della legge 208 del 28 dicembre 2015.

63. L'Osservatorio eGovernment del Politecnico di Milano (*osservatori.net*): Molise e Calabria non dispongono di un portale di dati proprietari ma fanno affidamento sul portale Sciamlab; Valle d'Aosta e Liguria hanno una sezione dedicata agli *Open Data* all'interno del proprio portale istituzionale, le altre Regioni hanno un portale *Open Data* dedicato. Sono stati pubblicati 18.000 *dataset* sui portali regionali, anche se il 70% di Regioni del Nord Italia.

64. Solo il 40% dei Comuni ha pubblicato dati, e solo il 16% *Open Data*.

65. L'80% dei Comuni non riscontra alcun impatto positivo dalla pubblicazione di dati in formato aperto e il 55% li ritiene addirittura inutili o poco utili per la crescita del tessuto imprenditoriale.

nazionale. A ciò si deve aggiungere che i dati sono poco accessibili⁶⁶ e di bassa qualità⁶⁷: le amministrazioni scelgono *dataset* già rinvenibili sui siti, e dopo solo il primo anno negli Stati Uniti già si notava che la qualità dei *dataset* proposti era scemata⁶⁸.

Dietro la bassa offerta da parte delle amministrazioni, i noti problemi relativi alle risorse, sia in termini di finanze che umane, a livello locale⁶⁹ ma anche nazionale⁷⁰.

Se l'offerta è di scarsa qualità, la domanda è scarsa. Quanto alle imprese, l'utilizzo di *open data* proveniente da pubbliche amministrazioni viene effettuato da pochi operatori, anche se l'interesse appare considerevolmente in crescita⁷¹, e serve che anche le imprese investano sul punto⁷². Le aziende che operano con gli *open data* sono principalmente quelle che lavorano con altre aziende per l'estrazione dei dati, piuttosto che quelle che svolgono attività nei confronti dei cittadini⁷³.

In tema di riutilizzo dei dati, nell'ordinamento statunitense, i riutilizzatori principali sono i singoli cittadini con competenze tecniche; in seconda battuta pubbliche amministrazioni e aziende tecnologiche; i settori in cui vengono più utilizzati sono, secondo la regola delle tre B,

66. La maggior parte dei Comuni pubblica i dati nella propria sezione trasparenza (il 83%) o sul sito istituzionale in una sezione *ad hoc* (33%), solo l'8% sul sito *open data* della Regione e solo il 2% sul sito *open data* nazionale.

67. A. PELED - N. KARINE, *Towards Open Data for Public Accountability: Examining the US and the UK Models*, cit.

68. H. YU - G.D. ROBINSON, *The New Ambiguity of Open Government*, cit., 199.

69. Gli enti pubblici attribuiscono la colpa della scarsa offerta e qualità alla mancata disponibilità di competenze interne (50%), alla insufficienza di personale interno (42%), alle ridotte risorse economiche da dedicare (24%) e al poco interesse dei politici (23%).

70. C. LAI, *Maturità degli open data in Europa: Italia tra i trend setter, ma si può fare meglio*, 2019, *agendadigitale.eu*, fa riferimento in particolare alla mancanza della figura del *data manager*.

71. Il 77% considera strategico l'uso dei dati per la propria attività; il 45% vorrebbe conoscerli meglio; il 4 per cento delle imprese manifatturiere utilizza dati in formato aperto offerti dalle amministrazioni pubbliche; il 15% non conosce ma è interessato ad approfondire l'argomento e il 30% sta cercando di capirne le reali potenzialità.

72. Le imprese sostengono (il 70%) di essersi dotate di strumenti e competenze per il *data management*, ma in realtà il 68% non conosce l'esistenza di figure professionali come il *Big Data Analytics specialist*, il *Chief Data officer*, il *Data Scientist* o il *Big Data Architect*, figure indispensabili per la gestione degli *Open Data*.

73. Il 47% delle non molte aziende che hanno risposto all'indagine opera nel mercato B2B (verosimilmente per l'integrazione dati), il 36% opera con l'amministrazione (B2G), per lo più per il rilascio e la generazione di *open data*, e quindi non per il riuso, e solo il 17% opera nel mercato B2C, anche perché già diversi operatori sono sul mercato.

buses (trasporti)⁷⁴, *budgets* (spese della P.A.), *bullets* (reati)⁷⁵. In Italia, in particolare, i dati più utilizzati sono quelli di fonte *OpenStreetMap* e *Istat*, e ciò anche in quanto si tratta di dati di alta qualità, oggetto di manutenzione nel tempo e disponibili per e al territorio nazionale⁷⁶.

Il processo di rilascio dei dati è stato sinora generalmente unilaterale, così che gli operatori non hanno potuto interagire con le amministrazioni né sulla scelta dei *dataset* da mettere a disposizione, né sulla qualità dei dati offerti: Il mancato incontro tra la domanda e l'offerta ha riguardato, come si è detto, in particolare gli enti locali, ma in ogni caso sono eccezionali i casi in cui delle amministrazioni hanno attivato processi stabili di pubblicazione di dati interagendo anche con gli *stakeholder*⁷⁷. La disponibilità di informazioni non costruita partendo dalle richieste e dai bisogni di cittadini ed operatori costituisce in questo senso un ostacolo insormontabile alla effettiva partecipazione e collaborazione cui si aspira. Del resto, se la pubblicazione non risponde all'istanza dei cittadini, la supposta legittimazione non si può realizzare.

In questo senso la prima linea d'azione del piano è la definizione di una “*Strategia complessiva per la valorizzazione dei dati*”⁷⁸, in cui abbiano un ruolo rilevante i cittadini⁷⁹. Proprio però perché il dialogo presuppone

74. *Moovit*, ad esempio, agevola gli spostamenti degli utenti perché attinge alle informazioni fornite dalle aziende di trasporto cittadine e territoriali.

75. C. CARPINETO, *Open data PA, bilanciare rischi privacy e benefici: un nuovo approccio*, 2018, *agendadigitale.eu*.

76. I dati provengono dallo studio *Open Data 200 Italia*, il primo sulle imprese italiane che utilizzano *open data* per le proprie attività.

77. V. PATRUNO, *Open data nel piano triennale ICT PA: la prima strategia unica su riuso e valorizzazione*, *agendadigitale.eu*, 2018 e, con maggiore riferimento agli enti locali e piccole e medie imprese ID., *Open data, la PA delude le Pmi: che fare*, *ibidem*.

78. Che vede coinvolti Dipartimento della Funzione Pubblica, AGID e Regioni per definire “una strategia finalizzata al riutilizzo dei dati aperti, in coerenza con le regole di implementazione della direttiva PSI, e le iniziative di collaborazione e confronto tra PA e stakeholders nel contesto delle azioni definite nell’ambito dell’*Open Government Partnership (OGP)*”.

79. Una strategia fondata richiederà una ricognizione dei dati, la stima dei costi, l’ascolto delle istanze e l’interlocuzione, l’adozione di piani di apertura dei dati (non previsto, anche se è prevista la valutazione dirigenziale). Sinora non c’è stato grande interesse per il monitoraggio, per la rilevazione della pubblicazione e utilizzazione dei dati. Sull’importanza delle strategie sugli *open data*, si v. B. VAN LOENEN, *Towards a User-Oriented Open Data Strategy*, *Open Data Exposed, Information Technology and Law*, cit.

un ambiente bene disposto, è fondamentale il ruolo degli entusiasti delle tecnologie, degli sviluppatori indipendenti e delle imprese tecnologiche⁸⁰.

Perché poi l'apertura possa realmente tradursi in una politica efficace, bisognerebbe risolvere alcuni problemi di fondo: servirebbero governi stabili, in grado di investire nell'amministrazione e di portare avanti in maniera continuativa politiche di *e-government*: queste ultime, inoltre, dovrebbero essere inserite in maniera organica in un quadro di riforme. Sul punto, si è osservato che la politica non investe sull'apertura dei dati anche per un problema di ignoranza sui benefici che essa causa⁸¹.

In secondo luogo, servirebbe una *governance* adeguata, dal momento che l'Agid (l'Agenzia per il digitale) non riesce ad essere una efficace autorità di regolazione e vigilanza, sia in virtù dei numerosi e ravvicinati mutamenti organizzativi (che sono anche effetto dell'instabilità politica), sia in quanto dispone di risorse limitate, sia per la istituzione del Commissario straordinario del Governo per l'attuazione dell'Agenda digitale, figura che si sovrappone all'Agid.

In terzo luogo, servirebbe una adeguata formazione nelle amministrazioni, le quali non percepiscono i dati come dei beni di valore⁸²; peraltro, anche laddove ci sia esperienza di commercializzazione dei dati, questa non risulta utile quando si tratti di finanziare la pubblicazione gratuita di dati, che deve essere affrontata dalle amministrazioni come un costo (nonostante si sia dimostrato che nel medio periodo l'apertura dei dati, in virtù della maggiore efficienza conseguibile, comporti risparmi alla spesa)⁸³. La difficoltà di trovare soluzioni al finanziamento dell'apertura ha così l'effetto di condurre le amministrazioni a guardare con sospetto e fastidio alla prospettiva della gratuità, che però è caratteristica fondamentale e imprescindibile degli *open data*⁸⁴. A ciò va aggiunto l'ulteriore problema per cui, nell'attuare le politiche di apertura dei dati, si registra un'incapacità delle amministrazioni a coordinarsi tra loro⁸⁵.

80. M. KASSEN, *Understanding transparency of government from a Nordic perspective: open government and open data movement as a multidimensional collaborative phenomenon in Sweden*, in *Journal of Global Information Technology Management*, 2017, 236-275, 248.

81. *Barriers in working with Open Data*, cit., 10.

82. *Barriers in working with Open Data*, cit., 22.

83. *Barriers in working with Open Data*, cit., 23.

84. *Barriers in working with Open Data*, cit., 23.

85. *Barriers in working with Open Data*, cit., 12.

7. Gli ostacoli nazionali di natura legale alla diffusione degli open data

Il tema dell'apertura dei dati è anche frenato dai vincoli posti dall'ordinamento⁸⁶. Il nostro Paese ha dal punto di vista documentale molte volte mostrato attenzione verso il tema⁸⁷, fino alla recente approvazione, nel giugno del 2019⁸⁸, del *Quarto Piano d'azione nazionale per l'open government 2019-2021*, che impegna oltre 50 amministrazioni pubbliche (centrali e locali)⁸⁹.

La normazione ha il difetto che sia le regole che le linee guida sono le stesse per tutte le amministrazioni e si concentrano solo sugli aspetti tecnici dell'apertura⁹⁰.

In aggiunta, le restrizioni che possono insistere su questi dati derivano per lo più da diritti di proprietà intellettuale e dai vincoli legati alla tutela della riservatezza e di segretezza (di Stato, d'ufficio, industriale)⁹¹: la legge

86. Il già citato art. 50 del CAD, *Disponibilità dei dati delle pubbliche amministrazioni*, prevede che (...) *restano salvi i limiti alla conoscibilità dei dati previsti dalle leggi e dai regolamenti, le norme in materia di protezione dei dati personali ed il rispetto della normativa comunitaria in materia di riutilizzo delle informazioni del settore pubblico*. Sugli ostacoli legali all'apertura, si v. *Barriers in working with Open Data*, cit., 16-17.

87. In aggiunta ai già citati piani adottati a seguito della partecipazione all'*Open Government Partnership*, si v. l'*Agenda nazionale per la valorizzazione del patrimonio informativo pubblico*, a cura dell'Agid (2014), la *Strategia per la crescita digitale 2014-2020*, a cura della Presidenza del Consiglio dei Ministri (2014), che contiene una sezione dedicata agli *open data*.

88. Nell'aprile del 2012 è stato adottato il primo piano italiano, l'*Open Government Partnership Italian Action Plan*. La competenza è del Dipartimento per la Funzione Pubblica della Presidenza del Consiglio dei Ministri, che coordina la partecipazione italiana all'*Open Government Partnership*.

89. Il piano prevede 10 azioni, che includono 170 impegni: in particolare, al di là dell'apertura dei dati (con particolare riferimento ai dati relativi alle politiche di coesione europee e al riutilizzo di beni confiscati), ci si occupa di trasparenza, supporto alla partecipazione, prevenzione della corruzione, la semplificazione della *performance*, i servizi digitali, le competenze digitali. Si v. sul punto S. PIZZICANNELLA, *Open government, ecco l'attuazione del Quarto Piano d'azione nazionale*, 2019, *agendadigitale.eu*.

90. Nel 2017 è stata pubblicata la seconda edizione delle Linee guida: si prevede che un *data manager* si occupi della ricognizione dei dati e strategia di apertura (1.9.1.1.2.); che un ufficio dirigenziale generale sia responsabile per la transizione alla modalità operativa digitale e difensore civico per il digitale (1.9.1.1.1), che sul punto richiama la normativa; che vi sia collaborazione con il responsabile della trasparenza (1.9.1.1.2).

91. Anche a livello europeo: si v. L. DALLA CORTE, *The European Right to Data Protection in Relation to Open Data, Open Data Exposed*, cit., 127. Sulla protezione intellettuale, si v. C. PASQUINELLI, *Open Data e Proprietà Intellettuale. Il Dataset come banca di dati e la questione delle licenze*, CNR Technical Report, 2013, *it.cnr.it*.

sul diritto d'autore italiana (la legge 633/1941) non fa riferimento ai dati e documenti prodotti e pubblicati dalle pubbliche amministrazioni; l'art. 5 fa salvi "i testi degli atti ufficiali dello stato e delle pubbliche amministrazioni" (leggi, regolamenti, provvedimenti, sentenze). Non parrebbero quindi rientrare in senso stretto i contenuti in forma non testuale, come alcuni *dataset* (relativi al traffico o ai trasporti), tabelle (piani), immagini (planimetrie catastali), mappe (geografiche)⁹².

Ulteriori problemi sono posti dalla tutela della riservatezza e dalla segretezza, secondo schemi già conosciuti in relazione ai limiti al diritto di accesso.

Altri problemi invece sono peculiari della disciplina del settore: se da un lato potrebbe sembrare che il criterio dell'*open by default* (per cui, se un'amministrazione pubblica dei dati e non indica una licenza d'uso, i cittadini possono sentirsi liberi di utilizzare quei dati senza dover chiedere il permesso⁹³) agevoli l'apertura dei dati⁹⁴, dall'altro lato, un ostacolo all'apertura dei dati è dato dalla necessità che una licenza d'uso non ci sia.

92. In questo quadro appare particolarmente emblematica la vicenda, che dovrà ancora essere affrontata dall'autorità giudiziaria, *Trenitalia vs Trenit*. *Trenit* è un'applicazione che, facendo ricorso agli *open data*, evidenzia i ritardi nei trasporti. Trenitalia ha ritenuto illegittimo l'uso dei dati che riguardano i propri treni, in quanto non autorizzato, e ha chiesto e ottenuto la sospensione del servizio. Trenitalia ha infatti sviluppato un'altra *app*, *Viaggiatreno*, che utilizza i dati. Trenitalia è un'impresa pubblica (partecipata al 100% dal Gruppo FS – Ferrovie dello Stato italiano, partecipato al 100% dal Ministero dell'Economia e delle Finanze) gestore di servizio pubblico. La vicenda riguarda anche la problematica applicabilità del principio *open by default*, per cui, nel silenzio, si dovrebbe ritenere che i dati siano utilizzabili. Nel 2016 ha inoltre avuto avvio l'iniziativa "*Open Trasporti*", confluita nell'*Action Plan* nazionale. Se il quadro giuridico, come abbiamo visto, non impone ancora che imprese pubbliche che gestiscono un pubblico servizio siano obbligate a condividere i dati, al tempo stesso il fatto che i dati del traffico costituiscano una delle tre voci più significative degli *open data* rende poco giustificabile la presa di posizione contraria all'apertura dei dati relativa al traffico ferroviario.

93. Art. 52 del CAD: *Accesso telematico e riutilizzo dei dati*. 2. *I dati e i documenti che i soggetti di cui all'articolo 2, comma 2, pubblicano, con qualsiasi modalità, senza l'espressa adozione di una licenza di cui all'articolo 2, comma 1, lettera h), del decreto legislativo 24 gennaio 2006, n. 36, si intendono rilasciati come dati di tipo aperto ai sensi all'articolo 1, comma 1, lettere l-bis) e l-ter, del presente Codice. L'eventuale adozione di una licenza di cui al citato articolo 2, comma 1, lettera h), è motivata ai sensi delle linee guida nazionali di cui al comma 7. Sull'importanza dell'adozione di linee guida per un corretto funzionamento degli open data si v. L. DALLA CORTE, *Towards Open Data Across the Pond, Open Data Exposed*, cit., 11.*

94. Art. 52 CAD: "*I dati e i documenti che le amministrazioni titolari pubblicano, con qualsiasi modalità, senza l'espressa adozione di una licenza, si intendono rilasciati come dati di tipo aperto*".

Dal momento che in teoria una licenza ci potrebbe essere, anche nascosta, e ciò impedirebbe l'*open by default*, gli interessati prudenti sono costretti a chiedere se i dati sono liberi, con un aggravio procedurale dall'altro, e con la possibile adozione di contromisure da parte dell'amministrazione, che essa altrimenti non avrebbe neppure pensato di opporre, se non fosse stata contattata⁹⁵.

Sul tema bisognerà verificare come verrà recepita la direttiva europea, ma la vicenda risulta affine a quella della Segnalazione Certificata di Inizio Attività, in cui sono stati infine stilati degli elenchi per identificare le attività certamente oggetto di Scia, certamente più chiarificatori di regole autorizzatorie generali dalla portata incerta⁹⁶.

8. Il ruolo degli intermediari

In teoria la partecipazione, grazie all'uso degli *open data*, è semplificata e aperta a tutti, in quanto i cittadini hanno in mano gli stessi dati dell'amministrazione.

Tuttavia, se l'uso dei dati ha permesso ai giornalisti, professionisti che istituzionalmente hanno interesse ad esaminare i dati, di investigare e raccontare vicende di pubblico interesse⁹⁷, è verosimile che i cittadini non abbiano, nella maggioranza dei casi, competenze e tempo per estrarre i dati, gestirli, rielaborarli. Non si può (e non è neppure auspicabile) pretendere che essi siano revisori, analisti, traduttori, programmatori⁹⁸. Questi sono problemi che si aggiungono a quelli classici della partecipazione, per cui non è detto che i cittadini sappiano fornire argomentazioni informate ai poteri pubblici o che sappiano come partecipare, in quanto può non essere agevole la partecipazione con strumenti aperti. In presenza

95. S. ALIPRANDI, *Open data: serve una norma più chiara. L'open by default non funziona*, 2018, *techeconomy.it*.

96. Sul punto, ci si permette di richiamare le osservazioni effettuate in punto di semplificazione in *Semplificazione e lotta alla corruzione nella legge 241 del 1990*, *Dir. amm.*, 2016, 623.

97. Sul ruolo dei giornalisti, oltre al già citato B.S. NOVECK, *Rights-Based and Tech-Driven: Open Data, Freedom of Information, and the Future of Government Transparency*, cit., si v. anche D. BERLINER - A. INGRAMS - S.J. PIOTROWSKI, *The Future of FOIA in an Open Government Agenda for Freedom of Information Policy and Implementation*, *Vill. L. Rev.*, 2018, 867.

98. C. BIRCHALL, *Interrupting Transparency, Transparency, Society and Subjectivity. Critical Perspectives*, a cura di E. Alloa e D. Thomä, Basingstoke, 2018, 343.

di questi ostacoli può essere complesso rendere effettivo il valore sociale degli *open data*⁹⁹.

Ciò pone inoltre un problema di rappresentatività: se la partecipazione, anche in ragione di questi ostacoli tecnici, non è alla portata di tutti, per cui ne affrontano i costi (anche solo in termini di tempo e impegno) solo coloro che sono davvero interessati¹⁰⁰, con la conseguenza che la platea potenziale dei partecipanti non risulta ampliata in maniera significativa (forse anche giustamente)¹⁰¹; l'aspetto rilevante è che diventa centrale il ruolo degli intermediari, che forniscono servizi per aiutare i cittadini a conoscere, partecipare o collaborare, o ai quali questa possibilità di partecipazione e collaborazione viene esternalizzata.

Di fatto, rientra dalla finestra la figura degli intermediari, che in teoria l'apertura dei dati avrebbe potuto superare. Se l'importanza di questi soggetti è poi abbastanza naturale nel caso di vigilanza (come appunto nel caso dei giornalisti), più problematico è il ruolo di questi soggetti quando i processi di deliberazione e partecipazione richiedono comunque interazione e collaborazione con i cittadini: si è così denun-

99. M. JANSSEN - N. HELBIG, *Innovating and changing the policy-cycle: policy makers be prepared!*, *Gov. Inf. Quart.*, 2015, 99.

100. Già lo si era notato qualche anno fa con riferimento alle procedure di *notice and comment online* rispetto a quelle classiche cartacee: si v. S.M. BENJAMIN, *Evaluating E-Rulemaking: Public Participation and Political Institutions*, in *Duke L.J.*, 2005-2006, 893; C. COGLIANESE, *Citizen Participation in Rulemaking: Past, Present, and Future*, *Duke L. J.*, 55, 2005-2006, 943; C. COGLIANESE - H. KILMARTIN - E. MENDELSON, *Transparency and Public Participation in the Federal Rulemaking Process: Recommendations for the New Administration*, in *Geo. Wash. L. Rev.*, 2008-2009, 924; J.M. DE FIGUEIREDOT, *E-Rulemaking: Bringing Data to Theory at the Federal Communications Commission*, *Duke L.J.*, 55, 2005-2006, 969; C.R. FARINA, *Achieving the potential: the future of federal e-rulemaking. Report of the committee on the status and future of federal e-rulemaking*, *Admin. L. Rev.*, 2010, 279; R. FARINA - P. MILLER *et al.*, *Rulemaking in 140 characters or less: social networking and public participation in rulemaking*, *Pace Law Rev.*, 2011, 382; C.R. FARINA - M.J. NEWHART *et al.*, *Rulemaking 2.0*, *U. Miami L. Rev.*, 2010-2011, 395.

101. Altri temi, connessi alla partecipazione nell'ottica dell'*open government*, che non riguardano il tema in oggetto, sono quelli relativi: alla capacità dei cittadini di rivestire un ruolo davvero centrale rispetto alla complessità attuale; alla possibilità di partecipare a procedimenti diversi da quelli in cui siano coinvolti soggetti determinati; al numero effettivo dei partecipanti e all'interesse effettivo a partecipare; all'impatto della partecipazione *online*; alle modalità e ai tempi di partecipazione; ai profili ideologici coinvolti nella partecipazione; ai costi, alla capacità dell'amministrazione di tenerne conto. Su tutti questi profili sia permesso di rinviare alla voce *Open government*.

ciato che esperti selezionati e non rappresentativi incidono sulla capacità di prendere decisioni¹⁰².

Va tenuto inoltre presente l'atteggiamento delle amministrazioni nei confronti di questi intermediari: allorché la collaborazione implichi un trasferimento di diritti di decisione a esterni, la reazione probabile è di riluttanza da parte degli attori pubblici verso questo tipo di iniziative, il quale atteggiamento fa il paio con la riscontrata incapacità dell'amministrazione di dialogare e recepire le istanze degli intermediari¹⁰³.

Va comunque osservato che forme di intermediazione sono presenti in tutte le attività (del resto, anche in giudizio si ricorre per mezzo di intermediari) e, del resto, essa risponde ad una necessità: non stupisce perciò che i pubblici poteri stessi promuovano il ruolo degli intermediari, affidatari del compito di rendere i dati grezzi accessibili al grande pubblico¹⁰⁴.

9. *La difficoltà di valutazione dell'impatto degli open data*

Mentre è possibile misurare il livello di apertura dei dati, anche solo enumerando i *dataset* disponibili, è ben più difficile misurare l'impatto che l'apertura ha sulla democraticità del sistema e la legittimazione delle istituzioni.

L'assenza di parametri di valutazione certi, poi, sembra la causa dei molto variegati giudizi¹⁰⁵, così che appare che questi ultimi dipendano

102. H. RICHTER - P.R. SLOWINSKI, *The Data Sharing Economy: On the Emergence of New Intermediaries*, IIC, 2019, 4.

103. *Barriers in working with Open Data*, cit., 23.

104. S. BAACK, *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, *Big Data and society*, 2015, 1.

105. J. BATES, *The strategic importance of information policy for the contemporary neoliberal state: the case of Open Government Data in the United Kingdom*, *Gov. Inf. Quart.*, 2014; 388 ritiene che la *Open Government Data Agenda* britannica abbia accresciuto la trasparenza, il coinvolgimento dei cittadini, e consentito di affrontare una serie di politiche anche più ampie e controverse, indirizzate a una forma neoliberale di stato; A. JAAKOLA - H. KEKKONEN - T. LAHTI - A. MANNINEN, *Open data, open cities: experiences from the Helsinki Metropolitan Area. Case Helsinki Region infoshare* www.hri.fi. *Statistical Journal of the IAOS*, 2015, 117, affermano che, più modestamente, con gli *open data* si possono migliorare i processi interni di ricorso ai dati; P. PARYCEK - J. HOCHTL - M. GINER, *Open Government Data implementation evaluation*, *Journ. Theor. Appl. Elect. Comm. Res.*, 2014; 80 ritengono gli *open data* utili a motivare i funzionari, vista la maggiore interazione con gli operatori; S. DAWES - L. VIDIASOVA - O. PARKHIMOVICH, *Planning and desiging in open government data programs: an ecosystem approach*, *Gov. Inf. Quart.*, 2016, 15, e B. WORTHY,

in buona parte dalle aspettative. Si è osservato ad esempio che l'idea che ha l'amministrazione della partecipazione è di consultazione, laddove l'istanza (spesso anche ingenuamente formulata) è di progettazione delle politiche e dei servizi¹⁰⁶.

Uno studio ha mostrato come la metà delle ricerche sul tema non abbia riscontrato un impatto significativo dell'apertura dei dati sul potenziamento della democrazia¹⁰⁷. Così come non sono stati ritenuti significativi i risultati dell'apertura dei dati rispetto all'aspirazione al cambiamento democratico e alla partecipazione pubblica in Cile o al coinvolgimento civile in Ghana¹⁰⁸.

Tuttavia, si deve tenere conto che da un lato la ricerca sugli effetti delle riforme relative agli *open data* è ancora all'inizio, anche perché le iniziative di *open data* sono recenti; dall'altro che gli effetti sono molto complicati da misurare, peraltro nel tempo; ancora, il perseguimento dell'interesse pubblico non si può certo arguire dal numero di utenti che hanno scaricato dati e dalle visualizzazioni degli stessi.

Si è osservato che in Italia (ma non solo) le pratiche partecipative e collaborative poste in atto siano di fatto degli esperimenti e i risultati siano limitati¹⁰⁹: forme di iniziative deliberative *online* sono state adottate sporadicamente; la più importante esperienza, per quanto poco rilevante a livello di diffusione, è costituita dai cd. bilanci partecipativi (Roma, Milano, Bologna), e si è osservato che il buon riscontro di queste iniziative fosse frutto del fatto che il modello, il metodo, le fasi (proposta, progettazione, voto) fossero ben strutturati e definiti e il tutto operasse su un'unica piattaforma tecnologica¹¹⁰. Ad ogni modo, l'impatto di questi esperimenti, che comunque può essere debole o operare su profili margi-

The impact of open data in the UK: complex, unpredictable and political, *Publ. Adm.*, 2015, 788, hanno ritenuto modesto l'impatto in termini di lotta alla corruzione, di controllo dell'uso delle risorse, di miglioramento dell'amministrazione.

106. V. PATRUNO, *Open government: così trasparenza e partecipazione rafforzano la democrazia*, *agendadigitale.eu*, 2018. Sul grande tema della partecipazione all'attività amministrativa generale, per tutti, M. COCCONI, *La partecipazione all'attività amministrativa generale*, Padova, 2010.

107. E.H.J.M. RUIJER - E. MARTINIUS, *Researching the democratic impact of open government data: A systematic literature review*, *Inf. Pol.*, 2017, 233.

108. E.H.J.M. RUIJER - E. MARTINIUS, *Researching the democratic impact of open government data: A systematic literature review*, cit., 2017, 239.

109. V. PATRUNO, *Open government: così trasparenza e partecipazione rafforzano la democrazia*, *agendadigitale.eu*, 2018.

110. *Ibidem*.

nali, quando pure sia misurabile, si scontra poi con dei costi certi¹¹¹, con un effetto disincentivante.

In aggiunta, si tenga presente che l'apertura, quando sia seguita da forme di partecipazione o collaborazione, non conduce necessariamente a un esito irenico: laddove, ad esempio, i cittadini segnalino prontamente le situazioni di emergenza o forniscano informazioni sulla qualità delle strade, tali pratiche, se da un lato contribuiscono alla legittimazione delle istituzioni quando queste ultime reagiscono efficientemente, invece la minano nel caso in ciò non avvenga¹¹². In questo senso, il terzo principio della *collaborazione* si è osservato come andrebbe piuttosto, e in maniera meno conciliante, inteso come binomio collaborazione/conflitto¹¹³ e, del resto, è noto (e già è stato prima citato) il noto binomio *funzione difensiva/collaborativa* in relazione alla partecipazione procedimentale.

10. *Il supposto conflitto tra efficienza e democrazia*

Peraltro, è anche da verificare se l'apertura dei dati sia, nella pratica, intesa e indirizzata ad una maggiore democraticità degli ordinamenti.

In dottrina è evocata una contrapposizione, in merito all'*open government*, tra il profilo della responsabilità democratica da un lato e quello dell'efficienza (come riduzione degli sprechi, miglioramento delle prestazioni e creazione di nuove opportunità economiche) dall'altro¹¹⁴.

Analisi diacroniche hanno mostrato come le politiche di *open government* avrebbero perso nel corso del tempo di precisione riguardo al proprio oggetto, con l'effetto di non promuovere più solo la responsabilità politica, come nei primi interventi, ma anche l'efficienza (sia sotto il versante dei servizi che dell'economia e dell'innovazione) e si è giunti così

111. *Ibidem*.

112. Con *Missing Maps* i dati aperti sono generati dai cittadini e aiutino le organizzazioni pubbliche, attraverso la fornitura di informazioni dettagliate su aree remote in caso di disastri. Con *Promise Tracker* i cittadini brasiliani forniscono informazioni al governo locale, ad esempio sulla qualità delle strade, monitorando così la risposta alle loro promesse. Con *Concerned Citizens of Abra for Good Government*, fin dal 2008 viene monitorata la responsabilità del governo attraverso i dati generati dai cittadini.

113. A. MEIJERA - S. POTJERB, *Citizen-generated open data: An explorative analysis of 25 cases*, *Gov. Inf. Quart.*, 2018, 613.

114. Tra gli altri, nel dibattito D. BERLINER - A. INGRAMS - S.J. PIOTROWSKI, *The Future of FOIA in an Open Government Agenda for Freedom of Information Policy and Implementation*, cit.

ad affermare che, rispetto agli ideali iniziali, stia prevalendo la seconda prospettiva¹¹⁵.

In questo senso si colloca anche il dibattito sul rapporto tra FOIA e *Open data*: *supra* si è fatto riferimento a come il FOIA venga generalmente trattato come una voce dell'*Open Government*. Tuttavia, si sottolinea anche come l'efficienza, che è diventata una parola d'ordine per gli *Open data*, non sia importante nel FOIA, e che anzi quest'ultima disciplina sia più proiettata verso una tutela dei valori democratici, in quanto comunque finalizzata al controllo sulle pubbliche amministrazioni¹¹⁶, laddove invece, e quasi paradossalmente, la disciplina sui dati aperti mirerebbe piuttosto alla valorizzazione del patrimonio informativo in termini di impatto sulla crescita economica¹¹⁷.

Più nello specifico, quanto alla selezione dei dati, ci si sarebbe preoccupati più di aprire i dati per favorirne l'impiego commerciale o il miglioramento dei servizi che per il controllo sull'operato dei poteri pubblici: ciò anche in virtù del fatto che le amministrazioni, nel decidere quali *dataset* proporre, escludono quelli che possono tradursi in una valutazione del proprio operato¹¹⁸.

115. H. YU - G.D. ROBINSON, *The New Ambiguity of Open Government*, cit., 199.

116. D. BERLINER - A. INGRAMS - S.J. PIOTROWSKI, *The Future of FOIA in an Open Government Agenda for Freedom of Information Policy and Implementation*, cit., 876 osservano come imprenditori ed accademici siano focalizzati sull'*Open Government*, mentre avvocati e specialisti di *media* sul FOIA, con i primi interessati ai profili tecnologici e i secondi concentrati su diritti.

117. Dal punto di vista del nostro ordinamento, comunque, i dati pubblicati in virtù del d.lgs. 33 del 2013 non sono dati aperti: si vv. le *Linee guida Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico* dell'Agid (punto 1.5): "le amministrazioni terranno conto delle differenze specifiche tra Open Data, Trasparenza e Condivisione dei dati tra pubbliche amministrazioni per finalità istituzionali. Queste tre azioni mirano a soddisfare esigenze diverse e anche se su alcuni aspetti convergono, fanno sempre riferimento a obiettivi specifici senza mai veramente confluire in un "corpus" organico. Ad esempio, in termini di trasparenza, alcuni documenti resi pubblici a seguito dell'applicazione del D.lgs 33/2013 e s.m.i.23 nella sezione "Amministrazione Trasparente" del sito web istituzionale di una amministrazione devono essere rimossi dopo aver svolto la loro funzione (di solito dopo 3 anni - cfr. art. 14 e 15). In questo senso, essi non possono essere propriamente considerati Open Data, per i quali tali restrizioni temporali non si applicano. Esistono poi dati delle pubbliche amministrazioni che assumono un ruolo importante nell'ecosistema degli Open Data e nella creazione di nuove forme di partecipazione (e.g. edifici, farmacie, musei, turismo, etc.) ma che non risultano nell'elenco dei dati obbligatori da pubblicare ai sensi del D.lgs n. 33/2013 e s.m.i."

118. J. SHKABATUR, *op. cit.*, 117.

Del resto, a livello europeo, come si è visto anche in merito alla nuova direttiva, si tende a sottolineare l'inquadramento degli interventi relativi agli *open data* (cd. trasparenza attiva) nelle politiche economiche di sviluppo della società dell'informazione, mentre gli interventi di trasparenza previsti ad esempio nel d. lgs. 33 del 2013 (cd. passiva) sono inquadrati nelle politiche miranti all'ampliamento della trasparenza governativa: proprio nella prospettiva economica infatti si pone la maggiore attenzione alla qualità dei dati, per impedire che i pubblici poteri agiscano sul mercato in maniera distorsiva¹¹⁹. Anche nella realtà statunitense, tra l'altro, si registra di recente, in particolare con riferimento ai dati di natura ambientale, una insofferenza nei confronti dell'apertura dei dati, qualora e in quanto essa non sia funzionale allo sviluppo economico¹²⁰.

Seguendo questo filone, i *dataset* costituirebbero lo strumento adeguato per migliorare la qualità della vita e dei servizi, non la responsabilità politica e, dalla prospettiva opposta, l'uso dei dati aperti non più solo per ragioni di trasparenza e responsabilità politica ma anche per politiche di mercato spiegherebbe l'impatto limitato sui processi democratici¹²¹.

Queste osservazioni si intrecciano con la sopra indicata difficoltà di valutare l'impatto delle politiche di apertura dei dati.

Molto spesso vengono eccessivamente semplificati i benefici dei dati aperti, viene vista una connessione diretta tra informazioni rese pubbliche e miglioramento della democrazia. Al tempo stesso, quando invece si sostiene che non vi siano stati progressi, come in Cile o Ghana, ciò dimostra che va esaminato anche il contesto, il cd. ecosistema, e adottata una prospettiva che tenga conto di più fattori¹²².

Se inoltre la crisi della legittimazione, così come si è richiamato, è il frutto della insoddisfazione dei cittadini nei confronti delle istituzioni, un miglioramento dell'efficienza, della qualità dei servizi, dell'economia

119. L. DALLA CORTE, *Towards Open Data Across the Pond*, in *Open Data Exposed*, cit., 11.

120. B. HYLAND WOOD, *Open Government Data in an Age of Growing Hostility Towards Science* osserva come le informazioni sui cambiamenti climatici riportate su importanti siti governativi fossero state rimosse. Dubbi sull'atteggiamento della presidenza nei confronti della trasparenza sono anche in C. BIRCHALL, *Interrupting Transparency, Transparency, Society and Subjectivity. Critical Perspectives*, a cura di E. Alloa e D. Thomä, Basingstoke, 2018, 343.

121. S. DAWES - L. VIDIASOVA - O. PARKHIMOVICH, *Planning and designing open government data programs: an ecosystem approach*, cit.

122. Anche se poi J. TAYLOR, *The information polity: towards a two speed future*, *Inf. Polity*, 2012, 227, pur sulla base di una prospettiva olistica, comunque non ha notato nell'ordinamento britannico alcun impatto significativo in ordine alla partecipazione.

è verosimile abbia un effetto favorevole sulla fiducia e quindi sulla legittimazione. Ci si concentra, detto altrimenti, troppo sull'impatto diretto sulla democrazia e non sull'impatto indiretto, riguardo a se le idee e proposte abbiano un effetto sull'operato del governo.

Da questo punto di vista, se ci si aspetta che l'apertura dei dati comporti una immediata riconfigurazione del rapporto tra cittadini e poteri pubblici nel senso della compartecipazione all'assunzione di decisioni strategiche e politiche, verosimilmente è sia da rimeditare bene il fine, quanto l'adeguatezza dello strumento.

11. Conclusioni

In ultima analisi, il tema è affascinante, ma la ricerca sembra lasciare molti punti aperti. Gli *open data* possono essere *uno* strumento di legittimazione delle istituzioni: è molto dubbio che siano *lo* strumento.

Il fenomeno e i movimenti per gli *open data* hanno obbligato le amministrazioni a comprendere che i dati sono dei beni pubblici di importante valore. Le amministrazioni hanno compreso che essi vanno adeguatamente gestiti, per la propria organizzazione o svolgimento delle proprie attività¹²³, e quindi che c'è bisogno di competenze e capacità specifiche. Ha indotto le amministrazioni a ragionare in termini di condivisione dei dati, e non di gelosa conservazione degli stessi¹²⁴.

Come conseguenza, le amministrazioni hanno implicitamente rifiutato la privatizzazione dei dati nei confronti di fornitori e terze parti che sarebbe altrimenti (magari anche inconsapevolmente) avvenuta.

La disciplina sulla trasparenza, in particolare sul versante pubblicazione, si è preoccupata del processo di rafforzamento dell'*accountability* dei

123. Nel presente contributo non ci si sofferma sui benefici specifici che le pubbliche amministrazioni conseguono per la propria attività e organizzazione dall'uso dei dati aperti. Sebbene incentrato sui *big data*, ma con considerazioni traducibili, ci si permette di rinviare, sul punto, a *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei Big Data*, *Dir. pub.*, 2019, 43. In merito ai benefici per le pubbliche amministrazioni, si v. S. BAACK, *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, *Big Data and society*, 2015, 1, il quale osserva che in primo luogo la disponibilità di dati aperti crea più opportunità di sviluppare tecnologie "civiche"; in secondo luogo, le amministrazioni spesso sottopongono al processo di uso dei dati le proprie attività e creano nuovi dati.

124. M. NAPOLITANO, *Il primo decennio di dati aperti è stato un successo, ma non per le ragioni che pensiamo*, 2019, *forumpa.it*.

poteri pubblici, nel senso della messa in questione del loro operato e di democratizzazione attraverso il controllo. Si è proceduto diversamente alla messa a disposizione di dati utili per i cittadini, per consentire maggiore coinvolgimento ed interazione.

Appare illusorio pretendere da parte dell'amministrazione un atteggiamento eccessivamente attivo nell'apertura dei dati, anche perché essa costringe l'amministrazione a condividere il proprio patrimonio, anche mostrando i propri punti deboli; il legislatore però, espressione più diretta della politica, in affanno di consenso, deve rispondere alle istanze di apertura. Si pensi all'esperienza statunitense, nella quale, alla luce del combinato disposto del recentemente adottato *OPEN Government Data Act*¹²⁵ con il *Foundations for Evidence-Based Policymaking Act* (entrato in vigore nel gennaio del 2019, che stabilisce che le agenzie federali, quando elaborano politiche pubbliche, devono provare la loro conoscenza della realtà, anche sotto il profilo della verifica delle politiche in passato attuate), agenzie e ministeri devono aprire i dati relativi alle *politiche pubbliche*, comunicarne i dati in loro possesso, costruire *database* pubblici all'interno del portale unico per l'accesso ai dati¹²⁶.

La vera logica dell'apertura prevale, purché, anche al di là delle intenzioni, essa non si riduca ad una logica di controllo¹²⁷ o di fornitura di dati per consentire lo svolgimento di attività economiche o una migliore efficienza, ma promuova anche forme di dialogo e collaborazione¹²⁸.

125. Sul quale si v. *supra*, nota 49.

126. D. AZZOLINI - M. SISTI, *Evidence-Based Policy e Attività Legislativa. Cosa c'è di nuovo?, Federalismi*, 2019 segnalano come ci sia sempre maggiore attenzione al fenomeno della valutazione delle politiche pubbliche, anche grazie al ruolo del movimento internazionale *Evidence-Based Policy* (EBP). La legge ora citata ha come obiettivi: "il rafforzamento della protezione dei dati e della privacy, il miglioramento dell'accesso ai dati, l'incentivazione delle capacità del governo federale di produrre e utilizzare evidenza valutativa".

127. In questo senso, la concentrazione degli sforzi sull'apertura di basi dati quali SIOPE (Sistema Informativo sulle Operazioni degli Enti pubblici), previsto dall'art. 14 della l. 196 del 31 dicembre 2009, *Open Expo e ItaliaSicura* si pone nella logica del controllo e non della partecipazione o della collaborazione.

128. O. PEREZ, *op. cit.*, 64. Si è in questo senso osservato che le istituzioni pubbliche dovrebbero insistere per includere attivamente i cittadini nei processi decisionali, per creare una più aperta e flessibile forma di democrazia rappresentativa (S. БААКК, *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, in *Big Data and society*, 2015,1).

La funzione conoscitiva nella rivoluzione dei dati

ABSTRACT: L'articolo guarda in modo evolutivo alla funzione conoscitiva, la cui trasformazione dipende dall'aumentata disponibilità di dati (*big data*) a disposizione della pubblica amministrazione. A partire da alcune esperienze già in atto presso alcune amministrazioni (es. INPS e ISTAT), l'articolo dimostra che i *big data* incidono sensibilmente sulla capacità conoscitiva delle amministrazioni, non solo a fini decisionali, ma anche a fini di mera conoscenza e di eventuale e successiva programmazione e pianificazione, per esempio, delle strategie di intervento e del territorio.

1. Introduzione

Nel mio intervento cercherò di porre l'attenzione sugli aspetti problematici che si sono aperti relativamente alla funzione conoscitiva delle amministrazioni con l'emersione di una nuova rivoluzione dei dati. Un tema molto ampio, quasi di coordinamento tra alcuni temi che sono già stati trattati in questa sessione mattutina, e quindi inevitabilmente è un tema che sarò costretto a trattare per punti e spero che mi scuserete se, per questo motivo, ci saranno alcune semplificazioni.

Con il termine rivoluzione dei dati sintetizzo e raggruppo una serie di fenomeni variamente denominati (*internet of things*, *big data*, *data analytics*, intelligenza artificiale), che sono stati descritti nelle loro caratteristiche già nella giornata di ieri, ma che sono tutti accomunati da due caratteri fondamentali:

- 1) la centralità dei dati, della loro raccolta massiva e della loro conservazione scrupolosa.
- 2) l'utilizzo di tecnologie con una grandissima capacità computazionale in grado, attraverso algoritmi matematici, di elaborare questa enorme quantità di dati che si è raccolta e conservata.

2. *Gli algoritmi nella pubblica amministrazione*

Gli algoritmi nelle amministrazioni pubbliche vengono utilizzati, oltre che a fini decisionali¹, anche a fini puramente conoscitivi. È un fenomeno che a livello europeo e a livello internazionale è evidente da tempo² e che sta cominciando a svilupparsi anche qui in Italia. Pensiamo a due grandi enti pubblici come l'Istituto nazionale di statistica (Istat) e l'Istituto nazionale della previdenza sociale (Inps), che utilizzano oramai da tempo, in via sperimentale, i *big data* per le loro elaborazioni.

L'Istat dal 2016 inserisce costantemente nel Programma statistico nazionale l'utilizzo dei c.d. *call detail records*, cioè i numeri che ogni compagnia telefonica assegna all'utente nel momento in cui effettua una chiamata, attraverso i quali l'Istat, elaborando l'ora e l'area da dove è stata effettuata la chiamata, stima, a livello aggregato, i flussi di mobilità delle persone (residenti statici, residenti dinamici, pendolari e visitatori) e li mette a disposizione delle amministrazioni³.

L'Inps, invece, è riuscita a stimare, con i dati dell'ultimo trimestre del 2012, il ricorso alla cassa integrazione da parte delle aziende assicurate nel 2013 attraverso una analisi algoritmica predittiva. L'algoritmo ha previsto il ricorso alla Cassa Integrazione Guadagni di 24.236 aziende. La richiesta effettiva osservata a posteriori è stata di 21.667 richieste. Il risultato della previsione ha avuto un livello di correttezza al 93,1 per cento⁴.

Anche le amministrazioni tradizionali cominciano ad utilizzare questi dati. Pensiamo ai progetti di *smart city* che impegnano i nostri comuni⁵. Il progetto *CityMap* del Comune di Torino, per esempio, ha calcolato

1. Per tutti si veda diffusamente G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, Napoli, Editoriale Scientifica, 2019 e il contributo di Augusti Cerrillo Martinez nella giornata di ieri.

2. Si v. per tutti G. MISURACA - F. MAREDDU - D. OSIMO, *Policy Making 2.0: Unleashing the power of big data for public governance*, Heidelberg Dordrecht, Springer, 2014; M. MACIEJEWSKI, *To do more, better, faster and cheaper: using big data in public administration*, in 83(1S) *International Review of Administrative Sciences* (2017), 120-135.

3. Per la consultazione del *Programma statistico nazionale 2014-2016. Aggiornamento 2016* a cui si fa riferimento si v. <https://www.sistan.it/index.php?id=513>.

4. Un esempio riportato allora anche dalla stampa italiana, precisamente in <https://st.ilsole24ore.com/art/tecnologie/2013-03-17/data-inps-stima-ricorso-082106.shtml?uuid=AbOFFteH>.

5. Per un quadro più generale sulle *smart cities* italiane si v. E. CARLONI, *Città intelligenti e agenda urbana: le città del futuro, il futuro delle città*, in *Munus*, 2016, 2, 235-269 e i contributi presenti nel numero n.4/2015 *Smart cities e amministrazioni intelligenti* della rivista *Istituzioni del federalismo*.

l'«indice di vivibilità» dei quartieri torinesi attraverso l'elaborazione algoritmica di una serie di banche dati pubbliche. L'indice, composto da una serie di indicatori, valuta i quartieri sotto vari aspetti tra cui la loro vitalità, la loro attrattività, la loro peculiarità, le loro criticità. È uno strumento che, al netto di qualsiasi considerazione relativa alla sua efficacia, esemplifica bene il contributo dei dati e della loro elaborazione massiva e algoritmica in termini di programmazione e di pianificazione territoriale⁶.

3. *La conoscenza algoritmica*

Gli algoritmi vengono utilizzati per elaborare una notevole quantità di dati, prevalentemente integrando banche dati diverse, con dati che provengono da fonti differenti, al fine di ricavare nuove informazioni e, soprattutto, una nuova conoscenza sui fenomeni naturali e sociali che si vuole indagare. La conoscenza algoritmica è una conoscenza diversa e originale rispetto al passato perché non è una conoscenza che mira a spiegare le relazioni di causa-effetto dei fenomeni che indaga ma è una conoscenza, invece, che fa emergere delle tendenze, delle correlazioni tra le variabili relative al fenomeno analizzato, fa luce su fenomeni che sono presenti, ma che sono sotto traccia nella società. È una conoscenza che, proprio grazie alla quantità di dati utilizzati, può essere addirittura una conoscenza predittiva, in termini probabilistici ovviamente, con una percentuale di fallimento più o meno alta⁷. La conoscenza algoritmica decodifica la complessità della realtà che le amministrazioni si trovano a dovere amministrare. La conoscenza algoritmica, in generale, rende visibile la “conoscenza invisibile” contenuta nella marea di dati che le amministrazioni posseggono e che senza algoritmi, forse, non sarebbero in grado di elaborare⁸.

6. Per approfondire il progetto è utile visitare il sito internet <http://icxt.di.unito.it/progetti/city-map/>.

7. Sulla conoscenza algoritmica si v. diffusamente V. MAYER SCHOMBERGER - K. CUKIER, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, trad. it. R. Merlini, Milano, Garzanti, 2013. Per i suoi limiti si v. F. SYLOS LABINI, *Big data, complessità e metodo scientifico*, in *Aspenia*, 63, Aspen Institute Italia, dicembre 2013; K.C. DE SOUZA - B. JACOB, *Big Data in the public sector: lessons for practitioners and scholars*, in *Administration & Society* (2014), 1-22.

8. Si v. Task force sull'Intelligenza Artificiale dell'Agenzia per l'Italia Digitale, *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino*, marzo 2018, 32 ss. Un elemento che emerge anche in dottrina. Per tutti si v. R. CAVALLO PERIN, *L'organizzazione delle*

4. *Le sfide poste dalla conoscenza algoritmica alla funzione conoscitiva pubblica*

Utilizzare le evidenti potenzialità conoscitive dell'algoritmo è una grande opportunità per le amministrazioni. Farlo, però, significa affrontare una volta per tutte le questioni problematiche che investono la funzione conoscitiva pubblica, innanzitutto la dimensione organizzativa della funzione conoscitiva, cioè quel complesso di attività (e direi anche l'insieme di strutture che quelle attività svolgono) volte ad organizzare e a produrre la conoscenza pubblica, come le attività di raccolta, di conservazione e di elaborazione dei dati in possesso delle amministrazioni, o come, ad un livello più alto, le attività di coordinamento informativo⁹. Questioni problematiche che sono nuove, ma anche antiche, perché le problematichità che pone la rivoluzione dei dati cascano sulle annose criticità relative alla funzione conoscitiva.

1) Il primo aspetto problematico, inedito, è relativo alla qualità del patrimonio informativo pubblico. I *big data* sono dati che hanno caratteristiche differenti rispetto ai dati tradizionali¹⁰. Il loro utilizzo pone, da un lato, il problema della capacità dei principi di qualità contenuti sia negli artt. 53-54 del Codice dell'amministrazione digitale, sia nelle regole tecniche adottate in questi anni di interpretare questo nuovo fenomeno. Dall'altro, pone il problema della funzione stessa delle banche dati, le quali sembrano tramutarsi da strumenti di garanzia di qualità dei dati pubblici a strumenti di elaborazione massiva dei dati¹¹.

2) Il secondo aspetto problematico, questo annoso per la verità, è relativo al rinnovamento tecnologico delle amministrazioni pubbliche. I

pubbliche amministrazioni e l'integrazione europea, in L. FERRARA - D. SORACE (a cura di), *A 150 anni dall'unificazione amministrativa europea*, vol. I, Firenze University Press, 2016, 19.

9. Si è utilizzata una definizione di funzione conoscitiva che prende spunto da M.P. GUERRA, *Funzione conoscitiva e pubblici poteri*, Milano, Giuffrè, 1996; F. MERLONI, *Sull'emergere della funzione di informazione nelle pubbliche amministrazioni*, in ID. (a cura di), *L'informazione delle pubbliche amministrazioni*, Rimini, Maggioli, 2002. Sulla dimensione organizzativa della funzione conoscitiva si v. E. CARLONI, *Le verità amministrative. L'attività conoscitiva tra procedimento e processo*, Milano, Giuffrè, 2011.

10. Sui caratteri dei *big data* si v. V. MAYER SCHOMBERGER - K. CUKIER, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, cit.

11. Si veda l'intervento di Enrico Carloni in questa stessa sessione mattutina.

big data necessitano di strumenti informatici con capacità di raccolta, di conservazione e di elaborazione del dato diversa rispetto al passato¹². Un elemento che rischia di rendere obsolete le norme sul riutilizzo dei software (art. 68-69 Codice dell'amministrazione digitale) e del tutto inefficaci le politiche di digitalizzazione perseguite con la clausola di invarianza finanziaria. Il rinnovamento tecnologico delle amministrazioni per la gestione dei *big data* ha bisogno di risorse economiche adeguate e costanti¹³.

3) Il terzo aspetto problematico è legato alla “cultura” del dato all'interno delle amministrazioni. La rivoluzione dei dati richiede non solo nuove competenze e nuove professionalità all'interno dei sistemi informativi delle amministrazioni, per utilizzare concretamente questi nuovi strumenti e per aggiornare e verificare il corretto funzionamento dell'algoritmo, ma soprattutto una nuova “cultura” del dato diffusa tra tutti i funzionari pubblici, in modo particolare tra le figure dirigenziali¹⁴. Utilizzare gli algoritmi a fini conoscitivi – ma a ben vedere anche a fini decisionali, utilizzare gli algoritmi *tout court* – significa principalmente avere chiaro qual è il bisogno conoscitivo della propria amministrazione e, conseguentemente, sapere scegliere quali banche dati l'algoritmo deve analizzare e per quali scopi; significa sapere costruire i criteri di elaborazione che l'algoritmo deve utilizzare nell'elaborazione dei dati e, infine, avere le competenze minime per potere comprendere ed interpretare correttamente i risultati dell'analisi algoritmica. Capacità che devono essere appannaggio del dirigente e non del tecnico informatico perché, a ben vedere, è nel momento della scelta di quali dati utilizzare, a fronte di un determinato bisogno conoscitivo, e della delineazione dei criteri di funzionamento dell'algoritmo che risiede la discrezionalità amministrativa nella decisione algoritmica¹⁵.

Un elemento problematico che si inserisce in un contesto in cui l'investimento organizzativo del legislatore negli ultimi anni, relativamente alla gestione dei dati in possesso dell'amministrazione, è rilevante. Basti

12. Per le difficoltà che gli strumenti di gestione dei dati tradizionali hanno nella gestione dei *big data* si v. per tutti A. REZZANI, *Big data. Architettura, tecnologie e metodi per l'utilizzo di grandi basi di dati*, Rimini, Maggioli, 2013.

13. Sullo stato della digitalizzazione, anche relativamente alle risorse impiegate nelle politiche di digitalizzazione delle amministrazioni pubbliche si v. *Relazione della Commissione parlamentare d'inchiesta sulla digitalizzazione e l'innovazione della PA*, 17 dicembre 2018, 25 ss.

14. Una necessità emersa anche nelle sedi istituzionali. Si v. diffusamente Dipartimento della funzione pubblica, *Syllabus “Competenze digitali per la PA”*, ottobre 2018.

15. Parzialmente in questo senso anche la sentenza Cons. St., sez. VI, 08 aprile 2019, n. 2270 relativa alla decisione algoritmica del Miur sull'assegnazione delle cattedre di ruolo.

pensare ai numerosi principi organizzativi a tutela di diritti soggettivi introdotti negli ultimi anni, come il principio *open data by default*¹⁶, oppure i principi di *privacy by design* e *privacy by default*¹⁷. Principi di organizzazione con cui il legislatore ha tentato di affrontare la gestione dei dati pubblici, in particolare nelle sue dinamiche più delicate come il rapporto tra la condivisione dei dati e la tutela della riservatezza¹⁸.

4) Il quarto elemento problematico, inedito, è legato al coordinamento tra la funzione conoscitiva e la funzione informativa, cioè quel complesso di attività predeterminate dalla legge che permettono la diffusione dei dati pubblici. All'interno di un ordinamento come il nostro che ha sposato il paradigma degli *open data*, si pone il problema di quali *big data*, devono essere anche *open data* e quali no¹⁹.

5) Il quinto aspetto problematico è il governo della funzione conoscitiva. Presupponendo che il coordinamento informatico e informativo dei dati è di competenza legislativa esclusiva dello Stato e quindi appannaggio del livello nazionale di governo (art. 117, comma 2, lett. r))²⁰, bisogna decidere definitivamente quale assetto deve avere la funzione conoscitiva nel sistema amministrativo italiano, qual è il ruolo delle varie tipologie di amministrazioni nella composizione della conoscenza pubblica e come si relazionano tra di loro (cioè scegliere quali amministrazioni debbano essere semplici fornitrici di dati, quali debbano essere in parte elaboratrici e in parte fornitrici di dati, quali soltanto elaboratrici della conoscenza). Un problema che influisce sull'intensità di tutti gli altri 4 problemi che abbiamo evidenziato, i quali con una funzione conoscitiva nazionale forte graverebbero di meno sulla parte più debole del sistema amministrativo: le amministrazioni comunali²¹.

16. Per tutti si v. B. PONTI, *Open Data and Transparency: A Paradigm Shift*, in *Informatica e diritto*, 2011, 1-2, 305-320.

17. Per tutti si v. F. PIZZETTI, *Privacy e il diritto europeo alla protezione dei dati personali. Dalla Direttiva 95/46 al nuovo Regolamento europeo*, Torino, Giappichelli, 2016.

18. Sull'equilibrio tra trasparenza e riservatezza si v. l'intervento nella sessione pomeridiana di Benedetto Ponti.

19. Sul rapporto tra *open data* e *big data* si v. J. GURIN, *Open data now*, New York, McGraw-Hill, 2014.

20. Sul coordinamento informativo si v. B. PONTI, *I dati di fonte pubblica: coordinamento, qualità e riutilizzo*, in F. MERLONI (a cura di), *La trasparenza amministrativa*, Giuffrè, Milano, 2008, 405-442

21. In questo senso e in modo più approfondito l'intervento in questa stessa sessione mattutina di Francesco Merloni.

A me pare che queste problematiche debbano essere affrontate seriamente soprattutto perché la rivoluzione dei dati, attraverso le sue potenzialità conoscitive, è capace di contribuire, se correttamente utilizzata, all'oggettivazione della funzione conoscitiva delle amministrazioni pubbliche²². La rivoluzione dei dati potrebbe permettere alle amministrazioni di produrre una maggiore, una migliore e una più oggettiva conoscenza della realtà, di cogliere l'essenza dei fenomeni economici e sociali che le amministrazioni dovrebbero governare.

Una oggettivazione che avrebbe un effetto positivo anche sulla dimensione dinamica della funzione conoscitiva, cioè sui momenti conoscitivi all'interno dei singoli procedimenti. Una funzione conoscitiva oggettivata garantirebbe anche una migliore completezza dell'istruttoria, un migliore accertamento e acclaramento dei fatti, e allo stesso tempo anche una tempistica procedimentale che rispetti il principio di non aggravamento. Il momento conoscitivo istruttorio potrebbe acquistare un nuovo spazio all'interno della decisione amministrativa, la quale oramai sembra appiattita sulla valutazione degli interessi e sul momento decisorio²³.

Utilizzare la conoscenza algoritmica però non è solo foriera di opportunità, ma anche di rischi. La conoscenza algoritmica è una conoscenza che può essere invasiva e penetrante, capace di ledere di per sé, durante l'elaborazione, i diritti di libertà delle persone²⁴.

Sviluppare e regolare gli algoritmi, anche semplicemente a fini conoscitivi, dunque, non è solo necessario in termini di tutela, per mitigare gli aspetti più spinosi del loro utilizzo, ma è anche doveroso ai fini del pieno rispetto dell'art. 97 della Costituzione.

Il Consiglio di Stato nella sentenza sez. VI, 08 aprile 2019, n. 2270 specifica bene questa doverosità, ma la individua solo in riferimento al principio di buon andamento. Personalmente allargherei questa prospettiva perché di fronte ad un'evoluzione tecnologica che, da un lato, allarga i confini della conoscenza possibile dei fenomeni e, potenzialmente, migliora l'oggettività delle rilevazioni, e dall'altro, aumenta la funzionalità dell'azione amministrativa, permettendo un maggiore equilibrio tra principio

22. Si permetta di rimandare a M. FALCONE, "Big data" e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica, in *Rivista trimestrale di diritto pubblico*, 2017, 3, 601-639.

23. *Ibidem*.

24. Per tutti si v. A. MANTELERO (2017), *Regulating Big Data. The guidelines of the Council of Europe in the context of the European data protection framework*, in *Computer Law & Security Report*, 584-602.

di completezza e di non aggravamento della decisione amministrativa, le amministrazioni non solo hanno il dovere di raggiungere quei livelli di maggiore funzionalità dell'azione amministrativa, che il ricorso a queste tecnologie garantisce loro, come suggerisce il giudice amministrativo nella sentenza appena citata. Esse hanno anche il dovere di raggiungere i nuovi confini della conoscenza che le tecnologie garantiscono loro, di raggiungere maggiori livelli di oggettivazione della loro capacità conoscitiva. Elementi, questi ultimi, che sono alla base di qualsiasi decisione amministrativa che voglia definirsi imparziale (nota 25). Solo in questi termini, dunque, rispettosi sia del principio del buon andamento sia del principio di imparzialità della decisione amministrativa, è possibile rispettare pienamente la doverosità che emerge dall'art. 97 della Costituzione²⁵.

25. Sulla doverosità costituzionale del raggiungimento dell'oggettivazione della conoscenza in termini di imparzialità si v. G. PASTORI, *Principi costituzionali sull'amministrazione e principio inquisitorio nel procedimento*, in M. CAMMELLI - M.P. GUERRA (a cura di), *Informazione e funzione amministrativa*, Rimini, Maggioli, 1996, 19 ss.

Martedì 21 maggio - III sessione
Interoperabilità delle banche dati
e funzione amministrativa

Algoritmi e organizzazione dell'amministrazione locale: come declinare il principio di adeguatezza affrontando la complessità*

ABSTRACT: L'articolo indaga le ricadute dell'utilizzo degli algoritmi sull'esercizio della funzione amministrativa esercitata a livello di enti locali. L'articolo evidenzia come gli strumenti di *data analysis* si prestino ad una migliore gestione delle problematiche a cui un'amministrazione deve far fronte e perciò rappresenta uno strumento di valorizzazione degli enti locali. Questi ultimi possono beneficiare degli strumenti tecnologici, che offrono una profonda conoscenza della realtà locale, permettendo una migliore e razionale allocazione delle funzioni amministrative.

1. *L'impatto degli algoritmi sulle categorie del diritto pubblico: considerazioni introduttive*

Un algoritmo è una procedura di calcolo ben definita che, a partire da determinati dati in ingresso, fornisce, attraverso una sequenza finita di passaggi, determinati dati in uscita. Questo, perlomeno, nello scenario tradizionale di utilizzo delle ICTs in cui, con riferimento alle specifiche procedure che si desidera informatizzare, viene progettato e sviluppato un algoritmo che permette l'esecuzione automatica e precisa di tutti i passaggi e di tutte le elaborazioni necessari per portare a termine la sequenza¹.

Lo scopo principale dell'uso degli algoritmi è la risoluzione di problemi, dai più semplici, come l'ordinamento di una lista di numeri o nomi, fino a quelli più complessi, come la gestione delle missioni spaziali e la mappatura del genoma umano. Anche in considerazione del fatto che sono alimentati da enormi quantità di dati (*big data*), gli algoritmi sono ora utilizzati nei settori più disparati e, per quanto rileva in questa sede,

* E. D'ORLANDO, *Algoritmi e organizzazione dell'amministrazione locale: come declinare il principio di adeguatezza affrontando la complessità*, in *Smart City. L'evoluzione di un'idea*, a cura di G.F. Ferrari, Milano, Mimesis, 2020, 529-552.

1. T.H. CORMEN - C.E. LEISERSON - L.R. RIVEST - C. STEIN, *Introduction to Algorithms*, MIT Press, Cambridge, 2010.

pure nell'elaborazione delle politiche pubbliche, negli atti di normazione, nell'azione e nell'organizzazione dell'amministrazione.

Le loro potenzialità, sempre più sofisticate, non sono infatti sfuggite ai decisori pubblici, che ne fanno un uso sempre più frequente. Un ambito nel quale tali potenzialità applicative possono avere in impatto rilevante è quello dell'organizzazione della pubblica amministrazione e, in particolare, di quella locale, di fronte al dilemma del riordino: perseguire la capacità sistemica dell'amministrazione, dando la priorità all'efficienza e alla performance nell'esercizio delle competenze e nell'erogazione dei servizi essenziali, oppure ritenere prevalenti le esigenze di prossimità dell'amministrazione, connessa alla necessaria corrispondenza tra domanda dei cittadini e offerta delle istituzioni.

Stabilire un punto di equilibrio è complicato, anche perché dietro alle due alternative si sviluppano questioni ancora più ampie, che attengono ai principi di buona amministrazione e di democrazia. Le esperienze concrete mostrano però come sia importante cercare di conciliare il più possibile i due aspetti, pur nella consapevolezza, da un lato, di non poter ottenere una perfetta sintesi tra loro e, dall'altro, di non poter attendersi un equilibrio stabile nel tempo.

Gli algoritmi possono rivelarsi uno strumento efficace per fronteggiare le problematiche in cui si dibattono gli enti locali, non solo da oggi e non solo in Italia, ma le questioni che attualmente si pongono presentano un livello di complessità e, al tempo stesso, di urgenza, forse non paragonabili al passato.

La ragione per cui le *ICTs* in generale si rivelano di particolare utilità deriva dal fatto che per affrontare e auspicabilmente risolvere le criticità del governo locale è necessario partire da una conoscenza analitica della realtà sulla quale si opera, per effettuare poi, attraverso l'elaborazione delle informazioni, le scelte più idonee al caso specifico.

Ciò premesso, allorché si discorre di istituzioni, è possibile contare su un accumulo nel tempo di enormi quantità di informazioni di ogni tipo – fiscali, sociali, immobiliari, geologiche, meteorologiche, ecc. – e tale accumulo è in costante aumento. Tuttavia i detentori dei dati non possono più fare affidamento esclusivamente sulle tradizionali dinamiche decisionali di analisi, misurazione e azione, poiché esse si scontrano sempre più spesso con la complessità della realtà.

Gli algoritmi sono un modo per controllare questa complessità.

Di conseguenza, la possibilità di utilizzare le enormi quantità di dati (*big data*), di cui le PA sono tradizionalmente in possesso, per alimentare gli algoritmi, rende questi ultimi molto appetibili nell'elaborazione delle politiche pubbliche e nell'organizzazione dell'amministrazione.

Va da sé che, a fronte delle rilevanti utilità che possono derivare dall'utilizzo dei *big data* tramite algoritmi, è necessario sempre tenere presente che si può così concretizzare un "livello normativo clandestino"². In particolare, quando l'attuazione della legislazione è assicurata da un dispositivo guidato da algoritmi, una sorta di quasi-normazione complementare è stata necessariamente elaborata nella stessa concezione dell'algoritmo³. In altre parole, la *governance* algoritmica potrebbe creare una sorta di potere regolatore nascosto, in ogni caso opaco a tutti coloro che non abbiano accesso all'algoritmo o non siano in grado di interpretarlo.

Da considerare, inoltre, che la concezione e lo sviluppo di algoritmi innovativi sono al giorno d'oggi legati a doppio filo alle ricerche nel settore dell'intelligenza artificiale, ovvero dell'elaborazione di sistemi tecnologici in grado di risolvere problemi o svolgere compiti e attività tipici della mente e dell'abilità umane. In effetti il tentativo di infondere intelligenza ai sistemi artificiali mediante algoritmi di vario tipo è connaturato alla disciplina stessa, dal momento che il sempre maggiore utilizzo degli algoritmi dipende dal fatto che gli studiosi hanno capito che i computer possono imparare per conto proprio se vengono loro fornite poche e semplici istruzioni⁴.

L'intelligenza artificiale offre dunque nuove possibilità ma pone anche nuove questioni, che devono indurre il giurista a riflettere. La posta in gioco è alta perché, in definitiva, si rischia di mettere in discussione il ruolo dell'elemento umano in quei processi decisionali che, da un lato, sono la sede naturale in cui prendono corpo le diverse manifestazioni della democrazia e, dall'altro, sono lo strumento attraverso il quale si concretizzano i diritti dei cittadini in termini di fruizione di servizi.

È indubitabile che affidare agli algoritmi le decisioni ha il vantaggio di poter risolvere questioni complesse con strumenti massimamente razionali, efficienti e neutrali: in quanto equazioni matematiche, gli algoritmi non sono di per sé positivi o negativi, ma dipendono dalle intenzioni e dalle modalità con le quali si utilizzano. Vero è, però, che nell'intelligenza artificiale l'oggetto dell'informatizzazione non sono più proce-

2. J.-B. AUBY, *Le droit administratif face aux défis du numérique*, in *Actualité Juridique Droit Administratif*, 23 aprile 2018, 835 ss.

3. D.W. SCHATUM, *Law and Algorithms in the Public Domain, Etikk i Praksis*, in *Nordic Journal of Applied Ethics*, 10 (1), 2016, 15 ss, reperibile in <http://dx.doi.org/10.5324/eip.v10i1.1973>.

4. S.F. DE ANGELIS, *Artificial Intelligence: How Algorithms make Systems Smart?*, in www.wired.com, 2014.

dure ben definite, secondo modelli precisi sviluppati nell'ambito delle relative discipline tecnico-scientifiche, cosicché l'aspetto deterministico che caratterizza le applicazioni *ICTs* tradizionali sfuma. In altri termini, poiché modellare processi cognitivi complessi non può essere fatto, allo stato attuale delle conoscenze, in modo completo, preciso e affidabile, è possibile un margine di errore⁵.

Tali criticità si sommano, infine, ad altre che non riguardano gli algoritmi in quanto tali, ma il modo in cui la società è attrezzata riguardo all'uso dei dati, alla protezione dei dati e all'utilizzo dei modelli predittivi⁶.

L'uso degli algoritmi oggi pone, dunque, problematiche tecniche che si riflettono in questioni etiche, sintetizzabili nel rischio di deresponsabilizzazione con riferimento a un contesto valoriale⁷; questioni politiche, in termini di imputabilità delle decisioni in relazione al circuito rappresentanza-responsabilità politica⁸; questioni giuridiche, perché è innegabile che gli algoritmi oggi stanno condizionando il diritto e il diritto, dal canto suo, cerca di appropriarsene sul piano dell'analisi scientifica, se non altro per renderli un po' meno opachi e sindacabili (*recte*, revisionabili) in sede giudiziaria, ma con tutti i limiti connessi al perimetro del proprio campo di indagine⁹.

Tenendo presenti i profili problematici così sintetizzati, è possibile analizzare l'ambito applicativo sopra indicato, strutturando l'analisi come segue: perché è necessario riflettere sul tema del riordino della PA locale; come si sono tradizionalmente affrontate le problematiche dell'adeguatezza degli enti locali, sul piano teorico-generale e operativo; come si potrebbero affrontare queste problematiche *pro futuro*, parimenti tanto sul piano dei modelli di riferimento quanto sul versante degli strumenti operativi di cui avvalersi per attuarli.

5. D.-U. GALETTA - J.G. CORVALÁN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in www.federalismi.it, 6 febbraio 2019.

6. E. SIEGEL, *Predictive Analytics: the power to predict who will click, buy, lie or die*, Wiley, New Jersey, 2016.

7. IEEE, *Ethically Aligned Design. A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, II Version*, 2018, spec. 146 ss., reperibile in <https://www.ieee.org/publications/index.html>.

8. L.M. LOPUCKI, *Algorithmic Entities*, in *Washington University Law Review*, 95, 2018, 887 ss.

9. J.A. KROLL - J. HUEY - S. BAROCAS - E.W. FELTEN - J.R. REIDENBERG - D.G. ROBINSON - H. YU, *Accountable Algorithms*, in *University of Pennsylvania Law Review*, 165, 2017, 633 ss.

2. *La complessità come fattore connaturato alla dimensione delle istituzioni e della pubblica amministrazione*

Le esigenze del riordino del governo locale, pur diverse per ogni ordinamento e periodo storico, prendono le mosse da fattori strutturali comuni.

Innanzitutto dal processo di integrazione europea che ha inciso, valorizzandolo, sul ruolo degli enti locali nel *policy making*, innescando un processo di decentramento che ha investito tutti gli ordinamenti europei, erodendo dall'interno il carattere monolitico dell'amministrazione¹⁰.

In secondo luogo, le tendenze globalizzatrici, se in un primo momento sembrava avessero cancellato il territorio e imposto il primato del singolo, hanno poi indotto un processo inverso: la riscoperta della dimensione locale per la sua infungibile valenza di supporto ai processi di sviluppo, quale taratura strategica nell'ottica della competitività¹¹.

Un terzo fattore strutturale che induce politiche di riordino è quello che riguarda gli approcci alla gestione della cosa pubblica. Al tempo in cui si è affermato il paradigma del decentramento, sul piano della gestione amministrativa dominava l'approccio del *New Public Management*, con la conseguente frammentazione, la dipartimentalizzazione nella gestione dei servizi pubblici¹².

Senza contare, infine, un fattore di natura relazionale, tipico dello Stato democratico-sociale, che riguarda i rapporti che intercorrono tra cittadini e istituzioni in termini di qualità e quantità dei servizi erogati dalle amministrazioni.

A questi fattori strutturali, si aggiungono, negli ultimi due o tre lustri, fattori più o meno contingenti, che esasperano certe tendenze o, al contrario, assumono un andamento a esse contrapposto.

Si pensi, da un lato, alla crisi di sovraccarico delle democrazie, con un aumento sempre crescente delle aspettative nei confronti delle istituzioni e della pubblica amministrazione¹³. Dall'altro, si considerino, sotto il profilo dei paradigmi di gestione, le tendenze all'aggregazione e alla cooperazio-

10. P. CARROZZA, *Governo e Amministrazione*, in *Diritto costituzionale comparato*, a cura di P. Carrozza, A. Di Giovine e G.F. Ferrari, Laterza, Roma-Bari, 2017, 1017 ss.

11. M. MARCANTONI - M. BALDI, *Regioni a geometria variabile*, Donzelli, Roma, 2013, 9.

12. C. HOOD, *A Public Management for All Seasons?*, in *Public Administration*, 69, 1, 1991, 4 s.

13. M. FLINDERS, *Defending Politics: Why Democracy Matters in the 21st Century*, Oxford University Press, Oxford, 2012.

ne orizzontale del *Post-New Public Management*¹⁴ e, quanto al contesto economico-finanziario, la crisi globale che ha invertito il paradigma del decentramento a causa della riduzione delle risorse¹⁵.

Tutto ciò rende necessario un ripensamento di quella complessa rete di governo che caratterizza il costituzionalismo europeo contemporaneo, peraltro non solo all'interno degli Stati nazionali, ma anche a livello sovrastatale¹⁶.

Tale complessità fa emergere, in particolare, i profili di inadeguatezza che oggi la tradizionale teoria dell'ordinamento giuridico manifesta allorché applicata alle istituzioni territoriali, sul presupposto che si possa ancora sostenere, con riferimento a ciascuna di esse, la perfetta corrispondenza popolo, territorio, governo, secondo la nota teoria istituzionistica¹⁷. Da qualche tempo fattori come la globalizzazione dell'economia e della società, la competizione tra i territori, la mobilità sempre più facile e veloce, le nuove tecnologie dell'informazione e della comunicazione, fanno sì che quasi nessuno dei confini che le istituzioni territoriali hanno oggi pare idoneo a consentire l'adeguato esercizio delle funzioni tipiche di quel determinato livello istituzionale. È sempre più difficile che gli interessi di un data comunità possano essere adeguatamente soddisfatti unicamente mediante l'esercizio delle funzioni di governo delle istituzioni che direttamente insistono su quel territorio: secondo un processo tipico dell'era della globalizzazione, la risoluzione di buona parte delle problematiche riguardanti un comune, una provincia o una regione, dipende non solo dalle scelte della corrispondente istituzione, secondo la logica classica della rappresentanza che esaurisce

14. T. CHRISTENSEN - P. LÆGREID, *Post New Public Management Reforms: Whole of Government Approaches as a New Trend*, in S. GROENEVELD, S. VAN DE WALLE (a cura di), *New Steering Concepts in Public Management (Research in Public Policy Analysis and Management, Volume 21)*, Emerald Group Publishing Limited, Bingley, 2011, 11 ss.

15. S. MANGIAMELLI, *Crisi economica e distribuzione territoriale del potere politico - Relazione al XVIII Convegno annuale dell'AIC*, in *Rivista AIC*, www.rivistaaic.it, 4, 2013. Per una panoramica comparata, P. VIPIANA (a cura di), *Tendenze centripete e centrifughe negli ordinamenti statali dell'Europa in crisi*, Giappichelli, Torino, 2014.

16. S. CASSESE, *La rete come figura organizzativa della collaborazione*, in A. PREDIERI - M. MORISI (a cura di), *L'Europa delle reti*, Giappichelli, Torino, 2001, 43 ss. Fra i teorici della ripartizione territoriale del potere politico come rete di interrelazioni che si sviluppano in maniera tendenzialmente orizzontale e collaborativa, superando la visione gerarchica e piramidale del potere propria dello Stato liberale, si ricorda altresì D.J. ELAZAR, *Exploring Federalism*, Tuscaloosa, University of Alabama Press, 1987, trad. it. *Idee e forme del federalismo*, Mondadori, Milano, 1998.

17. S. ROMANO, *L'ordinamento giuridico*, Sansoni, Firenze, 1918.

in sé ogni decisione e relativi effetti, bensì anche dall'interazione con altre istituzioni e con le loro decisioni, sia in senso orizzontale che in senso verticale¹⁸. I processi decisionali divengono sempre più complessi e articolati e fanno dipendere la realizzabilità e l'efficacia di una politica pubblica dal coordinamento e della coerenza delle risoluzioni assunte su tale politica a ciascun livello di governo, da quello europeo a quello nazionale a quelli regionali e locali.

È questo oggi il volto della *governance*, ovvero un modo di dispiegarsi dei processi decisionali alternativo o, comunque, più complesso rispetto a quello riferibile al governo istituzionalmente inteso (il *government*), che si manifesta attraverso la deliberazione a contenuto regolativo dell'organo di rappresentanza¹⁹.

Le ripercussioni sui profili organizzativi e funzionali delle istituzioni territoriali si percepiscono chiaramente nella continua ricerca di nuovi assetti. Si pensi al recente accorpamento delle Regioni in Francia²⁰ o, in Italia, al dibattito sulla costituzione delle macroregioni, che ciclicamente si ripropone sia sul piano politico che scientifico. Si considerino, ancora, le tormentate vicende del livello di governo provinciale, che vanno dallo svuotamento dell'ente a Costituzione invariata e con decreto-legge, censurato dalla Corte costituzionale²¹, al proposito dell'eliminazione della provincia dal testo costituzionale contenuto nel ddl costituzionale n. 2613-D della XVII legislatura, fino al più recente ridimensionamento dell'ente operato dalla legge n. 56/2014. Si aggiungano le alterne fortune delle forme di cooperazione intercomunale, di fronte all'irrisolta opzione tra capacità e prossimità²². In una prospettiva sovranazionale, si rifletta sulla ricerca sempre più diffusa di efficaci modalità di cooperazione territoriale e transfrontaliera, concepita tanto come uno strumento di più efficace gestione di problemi che vanno al di là del territorio statale (ambiente, energia,

18. M. CAMMELLI, *Istituzioni pubbliche e sviluppo economico locale*, in R. BIN - L. COEN (a cura di), *Poteri pubblici e sviluppo economico locale*, Cleup, Padova, 2009, 26 ss.

19. J.N. ROSENEAU - E.-O. CZEMPIEL, *Governance without Government: Order and Change in World Politics*, Cambridge University Press, Cambridge Mass., 1992; M.R. FERRARESE, *La governance fra politica e diritto*, il Mulino, Bologna, 2010, e ID., *Governance. Sugli effetti politici e giuridici di una "soft revolution"*, in *Pol. Dir.*, 2, 2014, 161 ss.

20. La *Loi n° 2015-29 du 16 janvier 2015 relative à la délimitation des régions, aux élections régionales et départementales et modifiant le calendrier électoral* ha ridotto, con effetto dal 1° gennaio 2016, il numero delle regioni del territorio metropolitano da 22 a 14.

21. Sentenza n. 220 del 2013.

22. S. BOLGHERINI - M. CASULA - M. MAROTTA, *Il dilemma del riordino*, il Mulino, Bologna, 2018.

reti di trasporto, ecc.), quanto come una modalità di integrazione e di coesione a livello europeo²³.

Queste ripercussioni, per quanto concerne la configurazione degli enti territoriali substatali, giungono a lambire la legittimazione della loro autonomia, ponendo in definitiva in discussione le ragioni stesse della loro esistenza. Un esempio sono le affermazioni estreme del recente passato per cui l'autonomia si giustificerebbe solo se rispondente a criteri di economicità.

3. *L'approccio tradizionale al tema dell'adeguatezza: premesse teorico-generalì e soluzioni organizzative*

È dunque necessario chiedersi come fare fronte alla complessità, che è ormai da ritenere un fattore connaturato alla dimensione delle istituzioni e delle amministrazioni.

Per rispondere alla domanda è necessario fare una breve digressione sull'approccio tradizionale al tema dell'adeguatezza degli enti territoriali, partendo dai modelli teorico-generalì di riferimento.

Utilizzando alcune metafore per sintetizzare i possibili modi di configurare un ente autonomo territoriale²⁴, la prima è quella riferibile all'originaria figura dell' "ente ordinamento": un organismo chiuso in se stesso, che si caratterizza per una rigida perimetrazione delle competenze, che considera la legittimità il modo di presidiare le funzioni, che fa dell'attività regolativa il suo principale modo di essere e di agire. Oggi, però, con l'evoluzione e l'incremento dei compiti dello Stato e dell'amministrazione, accanto alla funzione regolatoria vi sono altre funzioni decisive, poiché le politiche pubbliche di sviluppo e di *welfare* solo in parte si esauriscono nella produzione di regole, ma richiedono l'adozione di scelte per l'allocatione delle risorse, la produzione ed erogazione di servizi, la programmazione.

La medesima prospettiva di chiusura caratterizza il modello dell' "ente azienda", che fa perno sull'idea per cui ogni ente non è solo una forma di rappresentanza politica che decide in forma democratica come esercitare le proprie funzioni, ma è anche un apparato burocratico complesso, che va organizzato e fatto funzionare in modo economico, come un'azienda,

23. L. D'ETTORRE, *Le strategie macroregionali dell'Unione europea: tra cooperazione territoriale europea e multi-level governance*, in www.federalismi.it, 7 novembre 2018, 1 ss.

24. P. CARROZZA, *Le Province della post-modernità: la città territoriale*, in www.federalismi.it, 31 gennaio 2018, 10 s.

così da massimizzare l'impiego e l'utilità delle risorse pubbliche²⁵. Tuttavia anche il modello aziendalistico, per quanto potenzialmente in grado di ottimizzare, dal punto di vista dell'impiego delle risorse, l'attività di un'istituzione pubblica orientandola non solo alla regola ma anche al risultato, risulta inappagante se appiattito sull'idea dell'ente-gestore industriale, tipica dei primi anni Ottanta, del *New Public Management*. La realtà è da allora infatti profondamente cambiata, evolvendosi verso sistemi pubblici plurali, sia dal punto di vista dell'offerta di servizi e interventi che dal punto di vista dell'organizzazione politica della domanda.

Sul piano organizzativo, gli strumenti operativi corrispondenti a questi modelli teorici sono stati tradizionalmente cercati, dunque, partendo dal presupposto che la soluzione a tutti i problemi di disallineamento tra governo e territorio consistesse nell'adeguamento dei confini delle istituzioni: si è pensato che modificando, in genere accorpendo, gli enti territoriali esistenti, essi diventassero più efficienti ed efficaci grazie al principio delle economie di scala, o, semplicemente, alle maggiori risorse disponibili. Ampliando la visuale al panorama comparato, tra gli orientamenti tradizionalmente condivisi per fare fronte a tali criticità vi sono, infatti, i processi di fusione e le forme di associazionismo intercomunale (queste ultime utilizzate anche in funzione alternativa all'ente di area vasta), strumenti che hanno avuto storicamente una resa diversa a seconda del contesto in cui sono stati utilizzati²⁶. In Francia come in Italia, per esempio, l'istituto della fusione non ha dato risultati soddisfacenti, mentre in Germania è stato avviato negli anni Settanta un lungo e articolato processo di fusione durato sin dopo la riunificazione, che è stato decisamente proficuo. Le forme associative intercomunali, quantunque diffuse, in Germania sono considerate un'eccezione alla regola dell'adempimento in proprio dei rispettivi compiti da parte del Comune, la cui autonomia è garantita dal *Grundgesetz* come elemento essenziale del principio democratico, mentre l'ente di area vasta ha il carattere della territorialità ed è rappresentativo della comunità, che ne elegge l'organo titolare dell'indirizzo amministrativo. In Francia, invece, la legislazione di settore prevede una fitta rete di intermunicipalità, che trovano riconoscimento anche a livello costituzionale; le forme più strutturate, ovvero con una fiscalità propria, il potere

25. E. BORGONOVÌ, *Principi e sistemi aziendali per le amministrazioni pubbliche*, Egea, Milano, 2002; per un'analisi comparata, C. POLLITT, G. BOUCKAERT, *La riforma del management pubblico*, Egea, Milano, 2002.

26. E. D'ORLANDO, *Territori e potere. Un nuovo ruolo per regioni ed enti locali*, Forum Edizioni, 2019.

di elaborare politiche integrate e, quindi, a elezione diretta, sono volte a favorire l'integrazione politica tra i Comuni, in vista di future fusioni e di un progressivo ridimensionamento dell'ente intermedio. In Spagna, infine e in controtendenza, l'ente di area vasta, sin dall'origine concepito come raggruppamento di Comuni, è stato rafforzato dalla legislazione degli ultimi anni e convive con diverse forme associative intercomunali, meramente strumentali all'esercizio coordinato di funzioni e servizi o al compimento di opere determinate.

Nell'esperienza italiana, come è noto, parcellizzazione istituzionale e uniformità di ordinamento si annoverano tra i caratteri originari e salienti del governo locale, ma che si sono nel tempo rivelati inadeguati alle esigenze delle comunità amministrate e con essi il legislatore si è misurato in diversi momenti della storia repubblicana, con esiti più o meno sempre insoddisfacenti. Si pensi alla deludente prassi applicativa della legge n. 142/1990; alla fragilità dimostrata dalla riforma Bassanini del 1997, nonostante l'obiettivo di dotare le istituzioni di un assetto più moderno e di migliorare la qualità dei servizi ai cittadini; alla legislazione emergenziale anticrisi, basata sulla convinzione per cui la riduzione della spesa si debba necessariamente tradurre in una riduzione dell'autonomia; alla più recente stagione riformatrice, la cui attuazione procede in modo incerto e comunque ancora condizionato dalla crisi economico-finanziaria²⁷.

4. *Un nuovo paradigma: gli enti autonomi territoriali come attori di un processo dinamico di integrazione*

La complessità, fattore che fa ormai parte integrante della dimensione post-moderna delle istituzioni, fa quindi emergere oggi l'inadeguatezza della tradizionale teoria dell'ordinamento giuridico applicata alle istituzioni territoriali: il governo locale non è più basato sull'unità tra territorio e autorità locali rappresentative, c'è la necessità di un'articolazione delle funzioni di governo e di una flessibilità delle stesse nell'ambito di un modello elastico che deve trovare, d'altro canto, una sintesi nella pianificazione strategica del livello di governo primariamente coinvolto.

Di qui il seguente interrogativo: rincorrere ancora il mito della dimensione ottimale pensando a ristrutturazioni della configurazione territoriale e personale degli enti secondo una logica funzionalista o pensare a una

27. Per un quadro complessivo, L. VANDELLI, *Il sistema delle autonomie locali*, il Mulino, Bologna, 2018, 28 ss.

diversa modalità di traduzione del principio autonomistico nella dimensione istituzionale e amministrativa?

Ripartendo dal piano teorico-generale, il modello che meglio potrebbe rappresentare la condizione post-moderna dell'autonomia può essere offerto da un'ulteriore evoluzione dei precedenti ideal-tipi. Si tratta dell' "ente-sistema" che, preso atto che le funzioni e le attività decisive per il benessere dei cittadini non si possono circoscrivere regolativamente entro i suoi confini, fonda il proprio agire partendo dal presupposto del venire meno della corrispondenza tra rappresentanza politica e territorio e dal fatto che buona parte delle proprie funzioni possono trovare adeguate soluzioni solo interrelandosi e facendo sistema con altri soggetti. L'ente sistema appare perciò prevalentemente relazionale, piuttosto che regolativo; i suoi vertici politici e amministrativi sono tesi soprattutto a costruire relazioni di sistema con gli altri enti omologhi (dunque su un piano orizzontale), con il governo multilivello (stabilendo relazioni politico-istituzionali di tipo verticale) e con la società civile (quindi in senso sussidiario, valorizzando gli apporti della comunità sociale, il partenariato pubblico-privato, l'apporto del terzo settore).

L'attuale fisionomia degli enti locali già oggi si configura come sintesi delle descritte metafore – posto che si tratta di modelli che possono essere definiti con nettezza di contorni e purezza di contenuti solo sul piano speculativo – con una tendenziale prevalenza dei tratti dell'ente sistema, che verosimilmente sono destinati ad accentuarsi in futuro. Ora si tratta di adeguare gli assetti e gli strumenti di azione delle Amministrazioni alle nuove esigenze di relazionalità, perché stabilire relazioni istituzionali, orizzontali e verticali, diventa decisivo per la soddisfazione degli interessi delle comunità.

Si comprende quindi il salto di qualità che può recare con sé questa impostazione: le problematiche poste dalla necessità del riordino non si fronteggiano inseguendo il mito funzionalista, mutando i confini territoriali (fondendo o accorpendo) per adeguarli a una supposta esigenza di sinergie economiche, di efficacia regolativa, ma dotando gli enti territoriali di strumenti che favoriscano la loro relazionalità.

Ed è a questo punto che si delinea chiaramente la convergenza tra l'idea di ente-sistema e l'utilizzo delle *ICTs* come strumento operativo per la realizzazione del paradigma *smart*, le cui potenzialità applicative possono ben svilupparsi anche oltre il contesto urbano in senso stretto e coniugarsi con una nuova visione di governo locale, che oggi rinviene il maggiore punto di criticità, nel contesto italiano ma non solo, nella frammentazione e polverizzazione dei comuni e nelle opzioni organiz-

zative relative al governo di area vasta. Se infatti il concetto di *smart city* non pare applicabile in tutti i contesti, per oggettive differenze di scala (derivanti dalla ridotta concentrazione di abitanti e di imprese), tuttavia la prospettiva *smart* può essere non solo un fattore strategico di sostenibilità, ma anche uno strumento che orienta nella risoluzione delle questioni poste dalla complessità.

La prospettiva da adottare, dunque, è quella della *smart land*, che si basa su (e si sostanzia di) un processo di conoscenza e di costruzione del potenziale proprio di ciascun territorio, fatto naturalmente di realtà eterogenee, complesso e articolato, nel quale si possa costruire un sistema di relazioni che tendano a redistribuire servizi ed opportunità annullando gli effetti storicizzati di contrapposizioni quali città/campagna, montagna/pianura, centro/periferia e rendendo esigibili anche in forme nuove i diritti di cittadinanza.

Una riforma razionale delle autonomie locali – specie per quanto riguarda il problema di fatto irrisolto dei micro-comuni e del governo di area vasta – dovrebbe partire proprio da questo dato, dal momento che il paradigma *smart land*, proprio per come configurato, acquisisce una valenza non solo tecnica, ma anche politica²⁸, poiché costringe a (o induce la necessità di) trovare configurazioni territoriali omogenee e disposte a unificarsi e quindi ad aggregarsi dal basso per volontà di auto-trasformazione. Quindi la tecnologia non unicamente al servizio di una maggiore efficienza dell'apparato amministrativo senza modificare i processi sottostanti e con scarse ricadute esterne e “di sistema” (come nel modello del *New Public Management*), ma strumento utile proprio per permettere una ricomposizione delle comunità intermedie.

Ciò posto, si pone il problema delle competenze necessarie a realizzare le linee di intervento per coniugare *governance* e *smartness*, la cui risoluzione va cercata scegliendo la prospettiva di indagine dalla quale si vuole affrontare la questione.

Se il punto di partenza dell'analisi è quello delle *smart cities*, cioè il necessario (ma non sufficiente) requisito della connettività e dell'ampia fruibilità digitale di comunicazioni e servizi, la legislazione italiana, non fornendo una definizione precisa di *smart cities* e preferendo evocare concetti come “comunità intelligenti”, pone la loro realizzazione nell'ambito delle iniziative dell'Agenda Nazionale Digitale e tratteggia una suddivisione di competenze che non facilita il coordinamento, che è invece

28. A. BONOMI - R. MASIERO, *Dalla smart city alla smart land*, Marsilio, Venezia, 2014, 122.

essenziale, tra i diversi attori istituzionali coinvolti. D'altro canto, l'art. 117, c. 2, lett. r Cost., che menziona tra le materie esclusive dello Stato il "coordinamento statistico e informatico dei dati delle amministrazioni statali, regionali e locali", delinea un riparto della potestà legislativa piuttosto accentratato.

Se il punto di partenza è (quantomeno, anche) la disciplina del governo locale, forse le Regioni possono ritagliarsi maggiori spazi di intervento. Circoscritti, ma in via di espansione, le Regioni ordinarie, considerando non solo l'evoluzione della giurisprudenza costituzionale sui test di bilanciamento tra coordinamento statistico e informatico da un lato e autonomia organizzativa delle amministrazioni dall'altro²⁹, ma anche il potenziale che potrebbe dischiudere l'ampliamento dell'autonomia regionale a seguito dell'attuazione dell'art. 116, c. 3 Cost.: esso si rifletterebbe, infatti, anche sui rapporti regioni-enti locali, dal momento che potenzerebbe anche le funzioni amministrative, il cui svolgimento nel territorio dipende direttamente dalle scelte regionali sulla loro allocazione in capo agli enti di prossimità o di area vasta e sulla determinazione degli ambiti ottimali³⁰, coerentemente con la visione di forma di amministrazione che la Regione vuole concretizzare. Di margini più consistenti, invece, possono già fruire le Regioni a statuto speciale, in particolare quelle che hanno acquisito la titolarità della potestà legislativa esclusiva non solo in relazione agli aspetti ordinamentali del governo locale, ma anche su ulteriori ambiti strategici per realizzare un sistema realmente integrato tra la Regione e gli enti locali del proprio territorio (come, per esempio, la finanza locale, il trasporto pubblico locale, ecc.)³¹.

5. *Gli strumenti operativi per la ricomposizione delle comunità territoriali: la possibile funzione degli algoritmi nei processi di riordino*

A titolo esemplificativo, si può riportare come esempio il metodo seguito per il varo dell'ultima riforma degli enti locali in Friuli Venezia

29. M. CAPORALE, *L'attuazione delle smart cities. Competenze e coordinamento tra livelli di governo*, in *Istituzioni del Federalismo*, 4, 2015, 953 ss.

30. S. MANGIAMELI, *Appunti a margine dell'art. 116, comma 3, della Costituzione*, in *Le Regioni*, 4, 2017 684 s.

31. E. D'ORLANDO - F.E. GRISOSTOLO, *La disciplina degli enti locali tra uniformità e differenziazione*, in F. PALERMO - S. PAROLARI (a cura di), *Le variabili della specialità. Evidenze e riscontri tra soluzioni istituzionali e politiche settoriali*, Esi, Napoli, 2018, 99 ss.

Giulia, che tra gli strumenti operativi utilizzati per risolvere il dilemma del riordino valorizza proprio le ICTs³².

In particolare, in essa viene individuata una dimensione territoriale e istituzionale delle comunità intelligenti non legata all'idea di città, di grosso centro urbano, ma basata sui comuni e le loro forme collaborative, dissociando così l'impostazione da quella adottata dal legislatore nazionale che, abbinando la *smartness* all'attuazione dell'Agenda digitale e non facendo alcun riferimento al livello comunale (art. 20 d.l. n. 179/2012), pare perseguire una logica *top-down* nella realizzazione del paradigma *smart*, senza considerare che qualsiasi processo *smart-oriented* deve necessariamente essere *bottom-up*.

La riforma è stata preceduta da un protocollo di intesa tra la Regione FVG e l'ANCI FVG, nel quale le parti convenivano: di non focalizzare l'attenzione sui contenitori istituzionali, da valutare e plasmare alla fine di un percorso, ma su assetti che consentissero di valorizzare il principio di sussidiarietà nella sua duplice dimensione, orizzontale e verticale e, a quest'ultimo riguardo in particolare, le istituzioni più prossime ai cittadini purché "adeguate"; di ricondurre alle aggregazioni tra comuni non solo una finalità di gestione coordinata dei servizi (aspetto comunque importante per l'efficienza, l'efficacia e l'economicità dell'azione amministrativa), ma anche una funzione di incentivo per uno sviluppo *smart* degli enti locali, secondo quanto previsto dalla strategia Europa 2020³³. Di qui un'ipotesi di riordino del governo locale che prevedeva prioritariamente di allineare i servizi dei Comuni, della sanità e assistenza territoriali e della Regione, realizzando la continuità operativa dei processi, resa possibile dalle nuove tecnologie (come il *cloud computing*³⁴) e da un ambiente di servizi in rete consolidato e omogeneo quale quello dei distretti sanitari.

32. Trattasi della l.r. FVG n. 26/2014 e s.m.i., la cui attuazione non è stata peraltro a oggi completata in previsione di ulteriori interventi riformatori da parte del legislatore regionale. Ai fini che qui rilevano, si ricorda che in FVG sono state abolite le province con la l.cost. n. 1/2016 e che la l.r. n. 26/2014 ha costituito le unioni territoriali intercomunali, inizialmente obbligatorie per tutti i comuni sotto una certa soglia demografica, non elettive, con funzioni proprie nonché gestrici di funzioni di cui i comuni restano titolari e con possibilità di delega di funzioni comunali ulteriori da parte dei comuni: E. D'ORLANDO, *La disciplina del governo locale in Friuli Venezia Giulia*, in *Giornale di diritto amministrativo*, 3, 2018, 297 ss.

33. M.L. MADDALENA, *La digitalizzazione della vita dell'amministrazione e del processo*, in *Foro amm.*, 2016, 2535 ss.

34. Paradigma di erogazione di servizi offerti *on demand* da un fornitore ad un cliente finale attraverso la rete internet – come l'archiviazione, l'elaborazione e la trasmissione di dati –, a partire da un insieme di risorse preesistenti, configurabili e disponibili in remoto

Queste premesse hanno portato a incorporare i principi di sviluppo digitale dei territori e delle comunità nella riforma degli enti locali del FVG, con alcune significative conseguenze: la previsione di un rilevante investimento nelle nuove tecnologie; la formulazione in modo partecipativo dell'agenda digitale; la creazione di un unico sistema informativo per gli enti locali, per rendere disponibili servizi unificati per la PA e i cittadini; l'obiettivo di mirare alla completa interoperabilità fra gli enti locali; la particolare attenzione alla valorizzazione del dato aperto (*open data*).

Il punto cruciale del nuovo assetto riguardava, dunque, l'individuazione del livello di governo in capo al quale allocare la titolarità delle funzioni amministrative e la gestione dei servizi, sulla base di una valutazione di adeguatezza del singolo ente³⁵, da effettuarsi tenendo conto della complessità del contesto in cui esso opera e garantendo ai cittadini uguali servizi con meno risorse, competenze professionali obsolete, vincoli assunzionali e tecnologici. Di qui la necessità di valutare le misure da prendere sulla base dell'analisi di dati oggettivi e di elaborare un sistema complesso di misurazione idoneo a cogliere molteplici dimensioni, posto che la popolazione residente, di per sé, non è una misura di sintesi adeguata.

La valutazione delle dimensioni adeguate è stata effettuata sulla base di un algoritmo, operante secondo il seguente schema³⁶.

- (a) Il concetto da valutare (adeguatezza del comune) è stato scomposto in tre dimensioni da analizzare: contesto territoriale, sociale ed economico; performance economico/finanziaria dell'ente; performance del singolo servizio.
- (b) Ciascuna dimensione è stata a sua volta scomposta in variabili, rispettivamente: ambiente e territorio, società, economia e lavoro, redditi popolazione, fattori funzionali; autonomia finanziaria, equilibrio economico, spesa del personale, gestione dei residui, gestione degli investimenti, gestione dell'indebitamento; variabili del servizio.

sotto forma di architettura distribuita: P. MELL - T. GRANCE, *The NIST Definition of Cloud Computing*, National Institute of Standards and Technology, Gaithersburg, 2011, 2 s.

35. La legge prevedeva, in sintesi, una soglia demografica, variamente compensata da fattori legati all'appartenenza a Comunità montane o alla presenza di minoranze linguistiche, come criterio per stabilire l'obbligatorietà o meno delle gestioni in forma associata o tramite unioni territoriali intercomunali, salva la possibilità di derogare tale soglia nei casi di particolare adeguatezza organizzativa. La proposta dei criteri idonei a determinare i presupposti di adeguatezza organizzativa era demandata all'ANCI FVG.

36. *Sistema di definizione dei parametri per la valutazione dell'adeguatezza dei comuni del Friuli Venezia Giulia*, ANCI FVG, febbraio 2018, 1 ss.

- (c) Ogni variabile è stata collegata a numerosi indicatori (alimentati dai *big data* a disposizione di ogni PA).
- (d) Gli indicatori sono stati raggruppati in indici sintetici nel seguente modo.

Innanzitutto per ogni indice sintetico è stato definito da quali indicatori è composto; un indicatore può essere incluso in più di un indice sintetico e i diversi indici possono essere formati da un numero diverso di indicatori. Per ogni indicatore che compone l'indice si è proceduto dapprima alla standardizzazione statistica dei valori calcolati su tutti i comuni appartenenti al campione di riferimento e, quindi, si è verificato se l'indicatore è positivo (un valore maggiore dell'indicatore indica che il comune ha una migliore prestazione nel campo di riferimento dell'indice – per esempio una maggior quantità di servizio erogato è positivo rispetto all'indice di efficacia del servizio) o negativo (un valore maggiore dell'indicatore indica che il comune ha una peggiore prestazione nel campo di riferimento dell'indice – per esempio una maggior quantità di ore/uomo necessarie ad erogare un'unità del servizio è negativo rispetto all'indice di efficienza del servizio). Per tutti gli indicatori negativi si è proceduto al cambio di segno del valore standardizzato. Infine si è calcolata per ogni comune la media dei valori così ottenuti di tutti gli indicatori che compongono l'indice.

Gli indici sintetici calcolati per ogni dimensione analizzata sono i seguenti. Per il territorio: indice di BES (I_{bes})³⁷; indice di fabbisogno di risorse (I_{fr}); indice di valore del contesto (I_{vc})³⁸. Per ciascuna delle funzioni: indice di efficacia (I_{ec}); indice di efficienza (I_{ez}); indice di economicità (I_{et}). Per la sostenibilità finanziaria: indice di rigidità strutturale (I_{rs}); indice di investimenti (I_i); indice di efficacia (tempestività nei pagamenti) (I_p).

(e) Una volta composti tutti gli indici sintetici si è proceduto alla definizione dell'adeguatezza del singolo comune per la gestione separata delle diverse funzioni e dei servizi, attraverso una combinazione ponderata degli indici sintetici. I valori sono stati ponderati secondo i seguenti pesi: contesto 50%; sostenibilità finanziaria 20%; gestione delle funzioni 30%.

37. Benessere equo e sostenibile, che indica se le caratteristiche del territorio hanno effetto positivo o negativo sulla qualità della vita dei suoi cittadini.

38. Indice che sarà utilizzato dall' algoritmo e che rappresenta una sintesi di BES e di Fabbisogno di Risorse e che viene calcolato considerando un sotto insieme selezionato di indicatori di contesto.

Per ogni dimensione è stato calcolato un indice di sintesi corrispondente alla media degli indici sopra indicati. In particolare:

$$I_t = I_{vc}$$

$$I_{fn} = \text{media di } (I_{ec}), (I_{ez}), (I_{et})$$

$$I_{bil} = \text{media di } (I_{rs}), (I_i), (I_{tp})$$

Per ogni funzione si ha adeguatezza se:

$$I_{bil} \geq \text{soglia e } I_{fn} \geq \text{soglia oppure } [0,2 \times I_{bil} + 0,3 \times I_{fn} + 0,5 \times I_{vc}] \geq \text{soglia}$$

6. ... segue. Prime riflessioni in ordine alle ricadute sulle modalità di azione e di organizzazione della PA

Un primo aspetto che viene in rilievo con riferimento all'adozione del sistema descritto riguarda i riflessi che la digitalizzazione produce sulla relazione dell'amministrazione con il tempo. Essa apre la strada a processi di costante adattamento degli standard e alimenta costantemente il processo decisionale con informazioni in tempo reale sulla situazione che la decisione amministrativa vuole governare. Questo a differenza del tradizionale modo di agire dell'amministrazione, che incide sulla situazione giuridica sottostante in un dato momento dell'evoluzione delle regole applicabili³⁹.

L'algoritmo in oggetto consente infatti di valutare ciascun ente come adeguato, inadeguato, inadeguato ma con un margine di miglioramento subordinato all'attivazione di determinate azioni in un certo intervallo temporale. Il sistema consente pertanto di monitorare come si evolve l'adeguatezza degli enti sul loro territorio: un ente inizialmente adeguato potrebbe non esserlo più in quanto cambiano le condizioni "interne", in termini di gestione delle funzioni, ovvero di gestione del bilancio; oppure si potrebbe verificare una variazione del contesto in cui l'ente agisce, con spostamento del suo livello di adeguatezza in senso positivo o negativo. A differenza di un parametro tendenzialmente "statico", come la popolazione residente, tale sistema consente quindi a tutti gli enti di poter incidere in modo effettivo sulla modificazione della loro posizione nel tempo.

Un ulteriore aspetto da considerare riguarda il modo in cui l'algoritmo è stato strutturato e, più in particolare, l'individuazione degli spazi

39. J.-B. AUBY, *Le droit administratif face aux défis du numérique*, cit., 842 s.

disponibili per valutazioni discrezionali. La scelta della misura della ponderazione delle dimensioni da analizzare, evidentemente, non dipende dal funzionamento dell'algoritmo, ma ne costituisce un necessario elemento operativo, derivante da una decisione discrezionale. Parimenti, la determinazione delle soglie è un dato esterno al funzionamento dell'algoritmo, con il quale il risultato/output dell'operazione algoritmica va raffrontato. Nel caso in oggetto, il peso ponderato delle tre dimensioni, così come l'individuazione dei *benchmarks*, sono derivati da un accordo politico Regione FVG-ANCI FVG, sulla base del principio di leale collaborazione. È chiaro, quindi, che vi è una parte del contenuto dei comandi strutturati nell'algoritmo e poi tradotti nel *software* che è frutto di una scelta discrezionale, politica.

La circostanza è dovuta alla caratteristica della fattispecie di cui si discute (il livello di adeguatezza), che è costruita come un concetto giuridico non analitico, cosicché offre all'amministrazione lo spazio per individuare un'opzione ermeneutica che determina poi il concreto effetto normativo della disposizione sull'assetto delle posizioni giuridiche soggettive sottostanti⁴⁰. D'altro canto, non potrebbe essere diversamente: la decisione di avvalersi di un algoritmo per determinare l'adeguatezza di un ente non deve risolversi in un arretramento delle tutele e delle garanzie di quest'ultimo e, nel caso di specie, non deve determinare il venir meno di quelle valutazioni "adattive" che sono imprescindibili per il funzionamento delle organizzazioni complesse⁴¹.

In sintesi, nel caso di specie la definizione dell'adeguatezza è demandata sì alla *governance* algoritmica, come ormai accade sempre più spesso allorché si tratti di adottare decisioni in contesti multifattoriali, ma l'opacità della metodica, con tutti i rischi connessi⁴², pare temperata adeguatamente dalla procedura negoziale seguita per definire due aspetti fondamentali della struttura della logica del sistema.

Vi sono poi ulteriori aspetti sui quali riflettere, rilevanti sul piano dell'organizzazione della pa e, in definitiva, del buon esito dell'applicazione del sistema.

40. P. OTRANTO, *Decisione amministrativa e digitalizzazione della p.a.*, in *www.federalismi.it*, 17 gennaio 2018, 17 s.

41. Più in generale, sui "comportamenti adattivi" posti in essere dal funzionario che deve assumere decisioni riferite a situazioni non predeterminate dalla norma, S. CIVITARESE MATTEUCCI - L. TORCHIA, *La tecnificazione dell'amministrazione*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione*, Firenze University Press, Firenze, 2016, 34 s.

42. V. *supra*, sub par. 1.

Affinché quest'ultima sia sostenibile è necessario, infatti, che l'alimentazione dei dati per il calcolo degli indicatori avvenga quanto più possibile automaticamente. A tale fine nell'identificazione e selezione degli indicatori di contesto si è fatto riferimento soprattutto ai dati disponibili in banche dati regionali e, dove non disponibili, a banche dati *open* regionali e nazionali (Istat, Arpa, ecc.). Per quelli di bilancio ci si è basati sui dati reperibili in banche dati regionali e nazionali (Banca Dati Amministrazioni Pubbliche). Per gli indicatori relativi alla funzione si è cercato di limitare i dati oggetto di raccolta "manuale" da parte degli enti. Questo presuppone che il sistema informatico che elabora i dati per il calcolo degli indicatori abbia un accesso automatico (quantomeno) alle banche dati regionali e dei comuni.

Di qui i temi fondamentali dell'interoperabilità tra i sistemi e degli *open data*, come necessari presupposti per il corretto funzionamento dell'algoritmo.

Quanto al primo aspetto, da tempo si è ormai affermata l'importanza dell'interoperabilità come strumento di promozione dell'innovazione, in molteplici ambiti⁴³. Con particolare riferimento alle PA, l'art. 1, lett. dd d.lgs. n. 82/2005 (Codice dell'Amministrazione Digitale – CAD) la definisce come la "caratteristica di un sistema informativo, le cui interfacce sono pubbliche e aperte, di interagire in maniera automatica con altri sistemi informativi per lo scambio di informazioni e l'erogazione di servizi" e l'art. 12, c. 2 CAD obbliga le PA a utilizzare, nei rapporti interni, in quelli con altre Amministrazioni e con i privati, le tecnologie dell'informazione e della comunicazione, garantendo l'interoperabilità dei sistemi e l'integrazione dei processi di servizio fra le diverse amministrazioni. A ciò si aggiunge l'importanza che assume l'interoperabilità nel Sistema Pubblico di Connettività (art. 73, co. 1 CAD), quale insieme di infrastrutture tecnologiche e di regole tecniche che, per definizione, assicura l'interoperabilità tra i sistemi informativi delle PA, permette il coordinamento informativo e informatico dei dati tra le amministrazioni dei diversi livelli e tra queste e i sistemi dell'UE ed è aperto all'adesione da parte dei gestori di servizi pubblici e dei soggetti privati. Ciò premesso, è chiaro che, onde garantire un funzionamento fluido dell'algoritmo nel caso di specie, è necessario che l'interoperabilità si realizzi in tutte le sue dimensioni⁴⁴: tecnica, concer-

43. G. CARULLO, *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa*, Giappichelli, Torino, 2017, 129 ss.

44. Delineate nella Comunicazione della Commissione europea del 6 maggio 2015 sulla "Strategia per il mercato unico digitale in Europa", COM (2015) 192, 16 s., reperibile in

nente le modalità di interconnessione dei sistemi informatici e, quindi, la definizione di interfacce, formati dei dati e protocolli; semantica, onde garantire che il significato delle informazioni sia comprensibile da qualsiasi applicazione interconnessa; organizzativa, relativa alla modellazione dei processi, allineando le strutture informative con gli obiettivi organizzativi e agevolando l'integrazione dei processi produttivi. L'assetto delle competenze e delle dotazioni tecniche, umane e strumentali, degli enti dovrà essere configurato dunque in modo conseguente.

Quanto agli *open data*⁴⁵, il discorso nel nostro ordinamento è tradizionalmente legato, se non addirittura confuso, con il più ampio tema della trasparenza amministrativa: si consideri, infatti, che l'art. 53, c. 1 *bis* CAD obbliga le Amministrazioni alla pubblicazione dei dati coerentemente con la normativa sulla trasparenza⁴⁶. Tuttavia è stata altresì sottolineata la valenza ulteriore dei dati aperti, con specifico riferimento alla loro funzione amministrativo-conoscitiva ed economica⁴⁷, che risulta particolarmente rilevante allorché ci si pone dal punto di vista della razionalizzazione

<https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52015DC0192&from=EN>.

45. M. OREFICE, *Gli open data tra principio e azione: lo stato di avanzamento*, in *www.forumcostituzionale.it*, 25 maggio 2015.

46. L'art. 53, c. 1-*bis*, introdotto con la riforma del 2016, dispone, ai sensi dell'art. 9 d.lgs. n. 33/2013 (recante le disposizioni sul "Riordino della disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni"), la pubblicazione anche "(de)l catalogo dei dati e dei metadati, nonché delle relative banche dati in loro possesso e i regolamenti che disciplinano l'esercizio della facoltà di accesso telematico e il riutilizzo di tali dati e metadati".

47. D.-U. GALETTA, *Open-Government, open-data e azione amministrativa*, relazione al convegno "Nuove prospettive dell'amministrazione digitale: open data e algoritmi / *Nouvelles perspectives de l'administration numérique: open data et algorithmes / New Perspectives of Digital Administration: Open Data and Algorithms*", del Network "Smart Cities and Digital Administration. Law and Governance", Venezia, 12 ottobre 2018. L'Autrice evince il carattere polifunzionale degli *open data* dall'art. 1, lett. 1-*ter* CAD, il quale definisce "dati di tipo aperto" quelli aventi le seguenti caratteristiche: "1) sono disponibili secondo i termini di una licenza o di una previsione normativa che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato; 2) sono accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti ai sensi della lettera 1-*bis*), sono adatti all'utilizzo automatico da parte di programmi per elaboratori e sono provvisti dei relativi metadati; 3) sono resi disponibili gratuitamente attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, oppure sono resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione (salvo quanto previsto dall'articolo 7 del decreto legislativo 24 gennaio 2006, n. 36)".

delle risorse disponibili. Ne consegue che la modalità di funzionamento del sistema sopra descritto si rivela utile non solo nella prospettiva della riorganizzazione del governo locale, ma anche per il miglioramento della performance dell'azione amministrativa e del monitoraggio su di essa, tanto da parte del decisore pubblico o degli organi di controllo, quanto da parte dei cittadini.

7. Considerazioni di sintesi

Gli effetti delle ICTs sull'organizzazione della PA possono dunque rivelarsi molto utili, specie per affrontare le problematiche derivanti dalla complessità che caratterizza la dimensione post-moderna delle istituzioni, la quale richiede soluzioni flessibili e strumenti che consentano agli enti di fare fronte alle sempre più pressanti esigenze di relazionalità.

Più in particolare, l'utilizzo della *governance* algoritmica può risultare di ausilio innanzitutto sul piano tecnico: per la revisione dei modelli organizzativi dei comuni basati sulla generazione, gestione e analisi dei dati; per la diffusione di una cultura di *open data*; per consentire la possibilità di un confronto fra enti sotto il profilo della resa delle politiche pubbliche, specie sotto il profilo dell'utilizzo delle risorse.

Da non sottovalutare, inoltre, i risultati politici che essa può determinare. Innanzitutto può dare un impulso alla trasformazione digitale dei territori e promuovere così lo sviluppo dell'*open government*, nella prospettiva del rapporto PA-cittadini. In secondo luogo, può favorire la ricomposizione dei livelli di governo sulla base di logiche di auto-trasformazione (*smart lands*), consentendo l'attivazione di processi di riforma basati su dati oggettivi. In terzo luogo può incentivare la prassi della negoziazione tra regioni ed enti locali e divenire uno strumento operativo per la definizione della forma di amministrazione di cui la regione vuole dotarsi, dando concreta attuazione e fattiva praticabilità ai principi di sussidiarietà, differenziazione e adeguatezza. Infine, ma non in ordine di importanza, può valorizzare l'autonomia e la responsabilità degli amministratori locali: nella scelta delle funzioni e dei servizi che intendono assumere e nella disponibilità a consentire un monitoraggio obiettivo della resa effettiva delle proprie scelte in termini di miglioramento della qualità della vita dei cittadini.

Il Fascicolo sanitario elettronico: opportunità e rischi dell'interoperabilità dei dati sanitari

ABSTRACT: L'articolo esamina il tema dell'interoperabilità delle banche dati, prendendo ad esempio il Fascicolo Sanitario Elettronico. La riflessione muove dagli ostacoli tecnici che impediscono la comunicabilità tra sistemi informativi che gestiscono e detengono dati sanitari per rilevare come gli effetti si ripercuotano non solo sulla gestione dei dati e sull'erogazione delle prestazioni sanitarie, ma anche sull'accesso alle cure e sul diritto fondamentale alla salute. L'articolo si sofferma altresì sugli effetti che il processo di digitalizzazione può avere sulla salute individuale, poiché, per esempio, la telemedicina garantirebbe cure continue, a distanza e personalizzate, ma anche sull'organizzazione della struttura sanitaria, di cui il fascicolo elettronico è solo uno degli effetti della dematerializzazione e di una differente modalità di gestione dei dati privati.

1. *Premessa*

In un ambito come quello sanitario non c'è dubbio che la rivoluzione digitale possa avere effetti di grande portata.

In primo luogo, sulla salute individuale. Solo per fare un esempio l'Unione Europea stima che un telemonitoraggio a casa dei malati di cuore potrebbe migliorare i tassi di sopravvivenza del 15%. Un ulteriore effetto è la riduzione del numero di giorni passati in ospedale, che si aggira intorno al 26%, e, considerando che i decessi per infezioni intraospedaliere sono più che raddoppiati negli ultimi cinque anni, senza contare i contagi per Covid-19 nell'ultimo anno, anche questo è un elemento centrale in termini di salute.

Ma la digitalizzazione della sanità incide anche sull'accesso alle cure migliorandolo ed estendendolo a categorie di soggetti svantaggiati da diversi punti di vista: disagio economico, localizzazione territoriale, ma anche disabilità e non autosufficienza. Sotto quest'ultimo profilo, occorre tenere conto che l'Italia è il Paese «più vecchio» d'Europa con il 21,4% dei cittadini over 65 e il 6,4% over 80, ed è secondo al mondo, preceduto solo dal Giappone. Nel caso delle persone anziane, soprattutto grandi anziani, non autosufficienza e disabilità sono condizioni molto frequenti.

Una diversa conseguenza da non sottovalutare è quella che riguarda i risparmi di spesa. Nel caso del telemonitoraggio dei malati di cuore si parla di un risparmio potenziale del 10% dei costi sanitari.

Ma le valutazioni riguardano anche aspetti di sistema. Ad esempio si è stimato che la dematerializzazione completa delle cartelle cliniche consentirebbe risparmi complessivi di oltre un miliardo e mezzo di euro l'anno e risparmi altrettanto consistenti deriveranno anche dai certificati telematici e dalle ricette elettroniche.

Una importante sfida si apre quindi per la riorganizzazione della sanità, alla quale le risorse tecnologiche offrono finalmente la possibilità di ripensarsi in termini di sempre maggiore personalizzazione delle cure superando parte dei limiti strutturali e organizzativi che ancora oggi ostacolano questo processo. Non c'è dubbio, infatti, che la separazione professionale fra le diverse specialità, la articolazione macro e microorganizzativa delle strutture, il vincolo dei tempi di attesa, la relazione con una pluralità di soggetti rendano ancora oggi molto complicato per la persona il percorso di cura, soprattutto quando riguarda patologie complesse che richiedono interventi diversi e ripetuti. I tempi necessari per la prescrizione, la prenotazione, l'attesa dilatano il tempo della cura e finiscono per assorbire quasi tutto il tempo di vita che diventa tempo di malattia. L'uso degli strumenti digitali può cambiare in meglio tutto questo solo se governato con impegno e attenzione, altrimenti il rischio è quello della dispersione delle energie senza risultati.

2. *La situazione italiana e il Fascicolo Sanitario Elettronico*

Al centro dei processi di digitalizzazione ci sono i dati, prodotti dal cittadino direttamente, ma soprattutto prodotti dai diversi interlocutori della persona nella gestione della propria salute.

Si tratta di dati particolari, che sono massimamente delicati, ma di altissimo interesse per finalità varie, che vanno dalla cura del singolo, al governo del sistema, alla sicurezza sanitaria, alla ricerca, allo sviluppo di economie dedicate: farmaci, prestazioni mediche, assicurazioni, lavoro ecc.

Naturalmente l'interoperabilità delle basi di dati e la loro gestione secondo tecniche di *big data analysis* amplifica a dismisura modi e finalità dell'impiego dei dati stessi.

In Italia l'architettura di fondo di raccolta e organizzazione dei dati relativi al sistema sanitario è rappresentata dal Nuovo Sistema Informa-

tivo Sanitario (NSIS) che è nato all'inizio di questo secolo con l'obiettivo di rendere disponibile, a livello nazionale e regionale, un patrimonio di conoscenze fatto non solo di dati, ma anche di regole e metodologie per misurare qualità, efficienza, appropriatezza e costi. L'analisi integrata di questi dati e informazioni è pensata per essere a supporto del governo del Servizio Sanitario Nazionale, del monitoraggio dei livelli essenziali delle prestazioni sanitarie (LEA) e della spesa sanitaria, condiviso fra i vari livelli istituzionali e centrato sul cittadino.

Il NSIS si articola in due ambiti: il primo nel quale convergono i sistemi informativi orientati al governo del sistema sanitario (come il Sistema informativo dell'assistenza domiciliare, il Sistema Informativo per il Monitoraggio degli Errori in *Sanità*, il Sistema Informativo Nazionale Dipendenze, ma anche l'Osservatorio investimenti pubblici in sanità, ecc.) e il secondo orientato al cittadino, dove troviamo i Centri Unificati di Prenotazione (CUP), le prescrizioni on line, i certificati telematici, la telemedicina e, come vera "archittrave" del sistema, il Fascicolo sanitario elettronico.

Nel Fascicolo sanitario elettronico sono (o meglio: saranno) depositate le nostre informazioni sui ricoveri in ospedale, le visite specialistiche, le prestazioni farmaceutiche, l'assistenza residenziale, l'assistenza domiciliare e gli accessi al Pronto soccorso.

Non si tratta di una cartella clinica elettronica, né tanto meno di un documento medico, ma del 'luogo' del web che il cittadino può attivare e nel quale vengono archiviati e resi estraibili e consultabili i documenti della sua storia clinica.

I contenuti del FSE riguardano: i dati identificativi dell'anagrafica dell'assistito, i dati amministrativi relativi all'assistenza, i documenti sanitari e socio-sanitari, il Patient Summary o Profilo Sanitario Sintetico, il taccuino personale del cittadino e, infine, la eventuale dichiarazione di volontà alla donazione di organi e tessuti.

Una serie di dati chiave per la cura della persona è contenuta nel Patient Summary che riassume la storia clinica del paziente e la sua situazione corrente ed è creato ed aggiornato dal medico di medicina generale che segue il paziente ogni qualvolta intervengono cambiamenti ritenuti rilevanti ai fini della storia clinica del paziente. Il Patient Summary include, infine, anche una serie di dati clinici significativi utili in caso di emergenza.

Lo scopo principale di questo strumento è quello di favorire la continuità di cura, permettendo un rapido inquadramento del paziente al momento di un contatto non predeterminato come ad esempio in situazioni di emergenza o di pronto soccorso.

Gia con il Piano sanitario nazionale del 2003 si mostrava di avere ben chiaro come il patrimonio costituito dalle informazioni sulle condizioni sanitarie del cittadino fosse un elemento strategico al fine di supportare la modernizzazione del Sistema Sanitario Nazionale. In questa prospettiva nel 2004 la Presidenza del Consiglio dei Ministri attivava il “Tavolo di lavoro permanente per la Sanità Elettronica (TSE)”, che, nel 2005 ha prodotto il documento «Una politica condivisa per la sanità elettronica» con il quale si è compiuto un passo avanti importante per la definizione di un percorso programmatico volto all’innovazione digitale nei sistemi socio-sanitari.

Grazie a questa acquisita consapevolezza, il «Piano Industriale per l’Innovazione della Pubblica Amministrazione» del 2008 ha individuato cinque grandi aree di intervento in ambito sanitario e fra esse è stata prevista come centrale la realizzazione del FSE entro giugno 2009.

Si trattava di un obiettivo decisamente complesso che ha richiesto e richiede ancora più tempo di quello originariamente programmato ed è così che nel 2012, oltre al suo inserimento nel Piano di e-Government 2012 del Ministero per la Pubblica Amministrazione, si è arrivati alla prima definizione legislativa di questo strumento.

Ai sensi dell’art. 12, c.1, del d.l. 179 del 2012, il fascicolo sanitario elettronico (FSE) è “l’insieme dei dati e documenti digitali di tipo sanitario e sociosanitario generati da eventi clinici presenti e trascorsi, riguardanti l’assistito”. Tale previsione è stata successivamente attuata dal Decreto del Presidente del Consiglio dei Ministri, 29 settembre 2015, recante il Regolamento in materia di Fascicolo Sanitario Elettronico, che ha prescritto, fra l’altro, contenuti uniformi per i FSE realizzati nell’ambito dei servizi sanitari delle diverse regioni e ha previsto che tali contenuti siano costituiti da un nucleo minimo di dati e documenti obbligatori, nonché da dati e documenti integrativi, uguali per tutti.

Uno degli aspetti più significativi della disciplina normativa del FSE è stata la declinazione delle finalità che si perseguono con tale strumento: 1) una finalità di cura (prevenzione, diagnosi, cura e riabilitazione); 2) una finalità di ricerca (studio e ricerca scientifica in campo medico, biomedico ed epidemiologico); 3) una finalità di governo (programmazione sanitaria, verifica della qualità delle cure e valutazione dell’assistenza sanitaria).

3. Il FSE e l'interoperabilità dei dati

Gli obiettivi perseguiti con la creazione del FSE e la loro funzione anche “di sistema” hanno così reso evidente la necessità di un meccanismo di interoperabilità dedicato.

In breve questo ha dato luogo ad un intervento normativo *ad hoc*: la legge 11 dicembre 2016, n. 232, che ha previsto, attraverso l'introduzione del comma 15 *ter* all'articolo 12 del d.l. 179 del 2012 appena ricordato, la progettazione da parte dell'Agenzia per l'Italia Digitale (AGID), in accordo con il Ministero della salute, il Ministero dell'economia e delle finanze e le regioni, di una infrastruttura nazionale necessaria a garantire l'interoperabilità dei FSE.

Si è avviata così la realizzazione dell'INI, l'Infrastruttura nazionale per l'interoperabilità fra i FSE, curata dal Ministero dell'Economia e delle Finanze attraverso l'utilizzo dell'infrastruttura del Sistema Tessera Sanitaria già realizzata in attuazione dell'art. 50 del D.L. n. 269/2003.

In particolare tale Infrastruttura nazionale ha il compito di garantire l'interoperabilità dei FSE regionali, l'identificazione dell'assistito attraverso l'allineamento con l'Anagrafe Nazionale degli Assistiti (ANA), l'interconnessione dei soggetti previsti per la trasmissione telematica dei dati per le regioni che ne hanno fatto richiesta entro il 31 marzo 2017 e la gestione delle codifiche nazionali e regionali stabilite e rese disponibili dalle Amministrazioni e dagli enti che le detengono.

La realizzazione dell'INI è fondamentale anche per poter partecipare al sistema di interoperabilità europeo dei dati sanitari.

A questo proposito, già nel 2011 con la direttiva n. 2011/24/UE sull'assistenza sanitaria transfrontaliera, è stato istituito l'Health Network, che ha adottato nel 2013 le Linee Guida Europee (LGE) per il Patient Summary, che definiscono in dettaglio i contenuti informativi (ovvero un insieme minimo di informazioni che descrivono le condizioni clinico-sanitarie di ciascun cittadino-paziente), individuano gli standard semantici, sintattici e tecnologici che possono essere adottati, le architetture e i modelli di riferimento.

Attraverso questo ed altri strumenti inseriti nel Piano di Azione e-Health 2012-2020 la Commissione Europea intende rimuovere la gran parte degli ostacoli all'interoperabilità per garantire l'assistenza sanitaria transfrontaliera.

In questo quadro si inserisce anche il programma europeo *Connecting Europe Facility* (CEF), nell'ambito del quale saranno attivati per primi i servizi a sostegno dell'interoperabilità comunitaria del Patient Summary

e dell'ePrescription/eDispensation mediante la realizzazione di una infrastruttura dedicata e connessa alla rete nazionale di interoperabilità dei fascicoli sanitari regionali. Il progetto per la sua realizzazione: "Deployment of generic cross border eHealth services in Italy" è iniziato ufficialmente nel gennaio 2017.

Come già accennato prima, senza un adeguato sistema di interoperabilità le funzioni del FSE non possono realizzarsi pienamente e i tre scopi che la legge ha previsto per esso non possono essere effettivamente raggiunti.

La prima finalità, certamente la più importante, è quella individuale di prevenzione, diagnosi e cura. L'interoperabilità in questa prospettiva è essenziale per poter consentire alla persona di rivolgersi a diversi professionisti, diverse strutture, ubicate anche in regioni diverse e persino fuori dal territorio nazionale.

Ma anche al di là degli aspetti "logistici", non c'è dubbio che l'interoperabilità favorisca anche un approccio multiprofessionale alle cure, così importante in una fase come questa, in cui la iperspecializzazione delle professionalità mediche fa correre il rischio della perdita di visione complessiva della salute, con l'abbassamento, in certi casi, della capacità diagnostica soprattutto delle malattie metaboliche e rare.

Ma all'approccio multiprofessionale deve affiancarsi anche quello multidisciplinare, con particolare riferimento alla integrazione fra prestazioni di salute e prestazioni sociali. Sempre di più, infatti, il benessere individuale in cui si sostanzia lo stato di buona salute dipende da fattori diversi: sanitari, ma anche sociali. Di qui l'importanza di una condivisione e interoperabilità nel popolamento e nella consultazione dei dati fra diversi ambiti di servizio alla persona.

Infine l'interoperabilità appare essenziale anche per la prevenzione e la diagnosi precoce. Gli screening generalizzati riguardano solo alcune patologie. Altre altrettanto gravi restano non indagate per la scarsa disponibilità delle risorse impiegabili a questo fine.

Non c'è dubbio che poter selezionare le persone che presentano un rischio maggiore o mostrano segnali che potrebbero preludere allo sviluppo di una malattia consentirebbe di dirigere gli screening in maniera più mirata e sostenibile.

La funzione essenziale dell'interoperabilità dei dati (in questo caso da impiegarsi senza l'utilizzo di quelli identificativi degli assistiti) a fini di governo del sistema sanitario (programmazione, monitoraggio, valutazione, ecc.) a livello regionale e nazionale non ha bisogno di essere dimostrata e lo stesso vale per le finalità di ricerca. Piuttosto per queste ultime si deve se mai segnalare il limitato numero di soggetti che possono perseguirle

e che corrispondono a quelli che possono trattare i dati a fini di governo del sistema: Ministero della salute, Ministero del lavoro e delle politiche sociali, regioni e province autonome.

L'impiego dei dati a fini di ricerca può avvenire unicamente “nei limiti delle rispettive competenze” e, fatto salvo il caso del Ministero della sanità, che ha organismi ed enti da esso dipendenti che si occupano di ricerca scientifica, negli altri casi tale finalità rischia di non trovare alcuno spazio.

La sensibilità per la centralità dei dati sanitari contenuti nel FSE è del resto piuttosto recente, se solo si pensa che nelle linee guida del 2009 del Garante per la privacy si prevedeva che “a garanzia dell’interessato, le finalità perseguite” dovessero essere ricondotte solo alla prevenzione, diagnosi, cura e riabilitazione dell’interessato medesimo, con esclusione di ogni altra finalità”.

4. *Rischi e opportunità dell’interoperabilità in materia di FSE*

L’interoperabilità, una volta realizzata, potrà rendere possibile la collaborazione tra Pubbliche amministrazioni, ma anche tra queste e soggetti terzi, per mezzo di soluzioni tecnologiche sempre più avanzate.

Naturalmente il primo rischio che si prospetta in questo quadro è quello per la privacy delle persone, tema ben presente sin dalla prima progettazione del FSE: tant’è che, ben prima che il legislatore intervenisse con norme ad hoc, il Garante per la privacy aveva già adottato linee guida apposite per tutelare gli utenti e garantire la sicurezza nella conservazione dei dati.

Oggi la disciplina normativa del FSE prevede particolari garanzie centrate sull’autodeterminazione della persona. Il FSE deve essere attivato su richiesta della persona interessata e non è quindi strumento del sistema sanitario, ma dell’individuo. Per popolarlo di dati è necessario un ulteriore specifico consenso. A questo proposito il comma 3-bis dell’articolo 12, prevede che: “Il FSE può essere alimentato esclusivamente sulla base del consenso libero e informato da parte dell’assistito, il quale può decidere se e quali dati relativi alla propria salute non devono essere inseriti nel fascicolo medesimo”.

Dal consenso all’alimentazione dei dati, si distingue il successivo consenso alla consultazione che deve essere specificamente autorizzata per finalità di prevenzione, diagnosi e cura. Mentre per le finalità di studio e ricerca e programmazione sanitaria i dati possono essere impiegati senza identificazione del paziente (art. 12, commi 3-bis e 5, del d.l. n.179 del 2012).

Altro aspetto delicato riguarda l'informazione che precede il consenso.

Il regolamento di attuazione del FSE del 2015 prevede che l'informativa deve indicare tutti gli elementi richiesti dall'art. 13 del Codice Privacy.

L'articolo 78 del Codice in materia di protezione dei dati personali dispone che tali informazioni debbano essere fornite in forma chiara e tale da rendere agevolmente comprensibili gli elementi indicati. Alla base vi è la consapevolezza che, soprattutto in campo medico sanitario, l'informazione, da sola, non è spesso sufficiente, poiché deve essere elaborata dalla persona che la riceve, e chi la fornisce deve accertarsi che l'interessato o il suo rappresentante l'abbia compresa: l'elaborazione e la trasmissione dell'informazione sono importanti tanto quanto l'informazione in sé, premessa imprescindibile dell'esercizio dell'autodeterminazione terapeutica come dell'autodeterminazione informativa.

La delicatezza dello strumento FSE emerge anche nelle disposizioni da ultimo inserite nell'articolo 78 del Codice che, al comma 5, lettera c-bis, prevede oggi che l'impiego dei dati personali nel quadro del FSE rappresenti una forma di trattamento che presenta rischi specifici per i diritti e le libertà fondamentali, nonché per la dignità dell'interessato. Questo richiede un supplemento informativo che sia tale da evidenziare questi rischi.

Infine all'assistito è garantito il diritto, come in precedenza previsto dalle Linee Guida Nazionali del 2011, all'oscuramento dei dati e documenti sanitari e socio-sanitari che lo riguardano a tutti i soggetti abilitati all'accesso senza che questi ultimi vengano a conoscenza del fatto che l'assistito abbia effettuato tale scelta (oscuramento dell'oscuramento).

Un altro rischio, oltre a quelli per la privacy, riguarda la qualità dei dati e la loro sicurezza. A questo proposito il regolamento attuativo del 2015 dettaglia minuziosamente le misure di sicurezza che gli operatori debbono adottare per la protezione dei dati trattati mediante il FSE.

Mentre, per la consultazione in sicurezza dei dati, gli accessi al FSE da parte degli operatori del FSN, dovranno essere tracciabili e limitati al tempo strettamente necessario alla erogazione delle cure.

Ancora può essere segnalato come rischio quello della autovulnerabilità delle persone attraverso la disponibilità dei propri dati clinici. L'utente può avere accesso a dati che una volta erano generalmente mediati dal medico, mentre oggi si possono ottenere direttamente. Questo consente alla persona di rivolgersi liberamente a servizi medici o sedicenti tali che vengono erogati, in forma di consulenze ed eventuale vendita di medicinali, in rete.

L'esigenza di garantire la sicurezza dei pazienti rende necessario che il sistema pubblico (a livello non solo nazionale) si attrezzi per validare

anche servizi e professionisti che operano unicamente in rete. Si tratta di una sfida di grande portata che richiede misure non semplici da prevedere e a cui in questa sede non si può che riservare solo questo breve accenno.

Da ultimo merita porre attenzione alle opportunità che la disponibilità di dati offerta dal FSE e dalla interoperabilità possono aprire, soprattutto in termini di *Big Data Analysis*.

Gli ambiti sono diversi. Ne cito solo alcuni e in maniera rapidissima. Il primo riguarda le politiche sanitarie, regionali, nazionali, ma anche sovranazionali: si va qui dalla possibilità di conoscere in maniera molto più accurata il quadro epidemiologico, fino alla opportunità di prevenzione su scala mondiale della diffusione di certe patologie, che può essere incrementata dalla possibilità di monitorare fattori di rischio o dalla capacità predittiva che hanno dati marginali, quali l'impennata nella richiesta di alcuni accertamenti diagnostici o l'improvviso aumento nella prescrizione di certi farmaci, ecc.

Un secondo ambito riguarda la sostenibilità nell'impiego delle risorse, che può essere straordinariamente migliorata attraverso la conoscenza che la disponibilità di grandi quantità di dati sanitari personali possono offrire sotto il profilo del rapporto fra invecchiamento della popolazione e patologie emergenti, ma anche attraverso l'identificazione dei fattori indiretti che migliorano la salute e consentono un più adeguato impiego delle risorse.

L'ultimo aspetto che voglio segnalare riguarda la possibilità dei sistemi sanitari nazionali di interloquire da una posizione di minore debolezza con le grandi forze del mercato, e penso qui soprattutto a Big Pharma. Soltanto sistemi pubblici che conoscono se stessi e la propria utenza effettiva e potenziale possono sperare di rapportarsi con le multinazionali farmaceutiche senza esserne sopraffatti.

Vicende di questi ultimi anni hanno mostrato come la capacità di governo dei dati da parte degli Stati sia un elemento chiave per potersi relazionare con le imprese farmaceutiche, soprattutto in occasione della commercializzazione di farmaci innovativi, il cui prezzo per i sistemi sanitari dipende in maniera significativa anche dalla quantità del prodotto acquistato. In frangenti come questi è essenziale poter programmare con un sufficiente grado di certezza il fabbisogno complessivo per poter contrattare con il produttore il costo del farmaco.

Tenuto conto del fatto che l'approccio farmacologico è quello prevalente nel trattamento delle malattie più diffuse e che il costo dei farmaci è divenuta una delle voci di spesa più importanti dei sistemi sanitari, la gestione dei dati, anche attraverso adeguati sistemi di interoperabilità, si configura oggi senza dubbio come una delle più importanti leve di governo della sanità.

L'infrastruttura necessaria per creare interoperabilità tra pubbliche amministrazioni

ABSTRACT: L'articolo affronta il tema dell'interoperabilità dal punto di vista informatico, ponendo l'accento sulle infrastrutture necessarie affinché la comunicazione tra sistemi informatici pubblici sia possibile. La struttura a silos su cui si basa il sistema informativo della pubblica amministrazione italiana risulta inadeguato all'approccio della *big data analysis* che, a contrario, richiede la piena comunicabilità tra sistemi informativi affinché il reperimento dei dati su cui condurre sperimentazioni sia quanto più facile e mirato.

1. Introduzione

Con piacere intervengo in questo convegno sulla pubblica amministrazione con i BigData, la cui natura multidisciplinare arriva ad abbracciare la scienza dell'informazione e dei dati che è la materia della mia ricerca. Un convegno con apertura coraggiosa e necessaria, visto la crescente interdipendenza degli aspetti etici, giuridici e tecnologici nella gestione dei dati.

In particolare vorrei parlare delle banche dati della Pubblica Amministrazione e della loro integrazione. Le banche dati sono tradizionalmente organizzate secondo un modello detto "datasilos": insiemi di dati contenuti in diversi database ognuno dei quali organizzato secondo uno *schema*. Lo schema deve essere definito al momento della progettazione del database prima di iniziare a memorizzare i dati e definisce l'organizzazione logica dei dati in esso contenuti. Immaginiamo come semplice caso uno schema "scuola" con due tabelle: alunni e classi; i campi di ogni tabella hanno degli attributi, per esempio nome, cognome, matricola per gli alunni e anno e sezione per le classi. Lo schema definisce anche le relazioni fra alunni e classi, per esempio il fatto che ogni alunno deve appartenere a una sola classe ma che ogni classe deve avere almeno un certo numero di alunni. In definitiva lo schema descrive la conoscenza *ex-ante* che il progettista del database ha dei dati e delle loro correlazioni.

Uno dei fattori di complessità nell'incrociare dati di database diversi è proprio l'assenza di un unico schema. Incrociare i dati significa sostan-

zionalmente creare un nuovo database e quindi un nuovo schema che li unisca. Un'operazione che deve essere ripetuta ad ogni nuovo incrocio di dati con una crescita esponenziale della complessità e del costo. Gli aspetti tecnici non sono l'unico problema. Giuridicamente, in accordo alla General Data Protection Regulation (GDPR), i dati possono avere vincoli di localizzazione, vincoli temporali, vincoli di accesso. La copia dei dati è un'operazione che può facilmente indurre violazioni di questi vincoli; una copia dei dati potrebbe rimanere accessibile anche dopo la data ultima di validità delle informazioni. Ritorno su questo aspetto.

Un altro motivo per cui l'approccio *datasilos* è poco adatto alla ricerca di nuove informazioni – come correlazioni fra alcune categorie di fornitori e di bandi pubblici – è il fatto che i database esprimono in modo efficiente solo le relazioni definite dallo schema al momento della progettazione e quindi sono intrinsecamente poco adeguati alla ricerca di nuove relazioni fra i dati.

Oggi in Italia sono presenti oltre diecimila datacenter con diversi database della pubblica amministrazione: i dati comunali (fino ad arrivare al singolo PC nel comune di montagna), regionali, dell'anagrafe, degli ospedali, delle targhe automobilistiche, delle imposte, etc. Incrociare i dati è praticamente impossibile.

2. *L'approccio BigData*

Allora come si va oltre? Un paradigma di successo per incrociare i dati è il cosiddetto approccio *BigData*, che si basa sul modello "data lake". Un data lake, figurativamente un lago di dati, è un insieme di dati eterogenei (testo, immagini, video, etc.) che si distingue da un *datasilos* proprio per l'assenza di uno schema definito *ex-ante*; un lago che riceve dati da diversi immissari, cioè sorgenti di dati come dati dell'anagrafe, dati provenienti dai social network, dall'agenzia delle entrate. In un data lake è possibile cercare correlazioni non originariamente previste o prevedibili. La ricerca di queste correlazioni è dominio della cosiddetta "data science", che è oggi supportata da strumenti software per "BigData analytics" molto maturi e disponibili su diverse piattaforme on-premise e cloud commerciali (Microsoft Azure, Amazon AWS) e opensource (OpenStack). Esattamente in questo punto l'Intelligenza Artificiale entra in gioco. L'insieme delle tecniche legate all'Intelligenza Artificiale altro non sono che metodi per analizzare dati al fine di segmentarli, ridurli in cluster, classificarli, etc. Da questo punto di vista l'unica differenza fra i metodi statistici tradizionali e

quelli dell'Intelligenza Artificiale è che questi ultimi “imparano” a filtrare i dati guardando molti esempi, senza regole precostituite. Così come i bambini imparano a riconoscere gli oggetti o gli odori.

Per tornare all'analogia del lago, i “data scientists” sono proprio i professionisti che “pescano” informazioni dal lago di dati. Ovviamente la qualità del pescato dipende fortemente dalla qualità dei dati che arrivano nel lago in termini di pulizia, completezza, copertura, veracità e anche dalla abilità dei data scientists nell'individuare strumenti di pesca (cioè di analisi) sufficientemente selettivi per uno specifico obiettivo. Sono professionisti con un forte background di statistica e informatica in grado di valutare per ogni singolo caso se e quali metodi utilizzare scegliendo in un insieme di strumenti che va da quelli statistici tradizionali (Business Intelligence) a quelli dell'Intelligenza Artificiale. Chiaramente non esiste un metodo buono per tutti gli scopi, la scelta è complessa e per questo i data scientists sono oggi fra i professionisti più richiesti e pagati del mercato professionale tecnologico. L'università di Torino ha da qualche anno attivato un corso di dottorato innovativo in data science supportato da diversi dipartimenti, fra cui il dipartimento di informatica, matematica e economia.

3. *Datasilos e Datalake nella pubblica amministrazione*

Parrebbe quindi che il paradigma datalake sia la soluzione adatta per l'organizzazione dei dati della Pubblica Amministrazione. Purtroppo la risposta non è così semplice. Un datalake è tecnicamente un sistema OLAP (On Line Analytics Processing) che è adatto a cercare nuove relazioni fra i dati ma è totalmente inadeguato a mantenere la consistenza e la correttezza dei dati nel tempo per cui i tradizionali datasilos (tecnicamente detti OLTP - On Line Transaction Processing) basati su schemi sono ancora l'unica soluzione. Di fatto oggi è necessario utilizzare insieme datasilos e datalake. Per questo il problema relativo alla costruzione di sistemi di immagazzinamento e analisi che siano veramente GDPR compliant va oltre il singolo database o datacenter ma richiede una visione che abbraccia una rete di sistemi connessi e cooperanti. Da questo punto di vista, le attuali procedure di certificazione di sicurezza o di rispetto della GDPR penso assolvano più alla necessità di creare opportunità di business per le società di consulenza che altro. Faccio un paio di esempi relativi alla GDPR.

Il primo esempio riguarda lo scopo per cui i dati vengono raccolti, che in alcuni casi deve essere esplicitamente dichiarato (come per i dati

sanitari). La frontiera della ricerca medica si muove su un territorio sconosciuto e cerca relazioni ignote cercando correlazioni fra diversi fonti di conoscenza: cartelle cliniche, immagini TAC, DNA/RNA. Il possibile uso della scoperte è spesso non definito a priori. Vogliamo contare i prodotti che oggi funzionano bene o rispondono a uno scopo completamente diverso da quello per cui erano stati pensati?

Alla fine dell'Ottocento, ad Atlanta, il farmacista John Stith Pemberton, riprendendo la formula del "Vin Mariani", una miscela di vino e foglie di coca creata dal farmacista corso Angelo Mariani, inventò la Coca-Cola: un elisir contro mal di testa e stanchezza. Sempre alla fine dell'Ottocento, all'Ospedale Saint Mary di Londra, il chimico Alder Wright, ottiene, da una sintesi chimica della morfina, l'eroina. Tuttavia la scoperta non è ritenuta interessante. Una ventina di anni dopo l'eroina è nuovamente sintetizzata dal dottor Felix Hoffmann, ricercatore della casa farmaceutica Bayer. La Bayer commercializza da subito l'eroina come medicinale per il trattamento della tosse, dei problemi respiratori e per combattere la dipendenza dalla morfina.

Gli esempi sono deliberatamente riferiti a un'epoca in cui la conoscenza non veniva estratta dai dati, come invece avviene per i BigData. Oggi la situazione è completamente diversa e la frontiera della ricerca medicina si usa tecniche di BigData e Intelligenza Artificiale e quindi dipende direttamente dalla GDPR e dalle leggi che disciplinano l'uso dei dati. A questo riguardo a Torino abbiamo da poco attivato un nuovo trial clinico che si svolge presso l'Ospedale San Luigi come collaborazione fra dipartimento di radiomica e informatica. È uno studio sul polmone finalizzato ad eliminare molte delle biopsie oggi in uso nella pratica clinica utilizzando tecniche di Intelligenza Artificiale per prevedere le varianti genetiche a partire da immagini TAC dei pazienti osservati. Uno studio assolutamente nuovo di cui neanche noi riusciamo a immaginare esattamente tutti i contorni, che potrebbero essere anche fuori dall'alveo dello scopo per cui i dati del trial clinico sono stati originariamente raccolti.

Un secondo esempio riguarda le reti di sistemi, che abbiamo visto essere davvero necessarie a mantenere e analizzare i dati. Per funzionare le reti devono copiare i dati sistemi datasilos/OLTP, finalizzati alla archiviazione (come quello di ANAC o le cartelle cliniche dell'ospedale) a sistemi datalake/OLAP, finalizzati all'analisi. Secondo la GDPR i dati possono essere vincolati all'uso all'interno di una finestra temporale e di un luogo geografico specifici. Quindi da un lato la copia dei dati da sistemi OLTP a OLAP è necessaria per l'efficacia dell'analisi, dall'altro ogni copia è un pericolo per l'integrità dei vincoli perché è tecnicamente difficilissimo se

non impossibile tracciare la vita di tutte le copie dei dati e renderle inaccessibili quando il dato originale arriva alla sua scadenza. E inoltre, cosa ne sarebbe dei processi di analisi che dipendono da quei dati? La catena di conseguenze è ad oggi ignota e personalmente non vedo nessun modo serio per utilizzare i dati per estrarre informazioni e rispettare alcune delle prescrizioni GDPR. In linea di principio l'insieme delle tecniche conosciute come "distributed ledger" (blockchain) potrebbero essere una soluzione, ma questo argomento è fuori dagli scopi di questa relazione. A questo proposito mi fa piacere ricordare che grazie alla Professoressa Racca, l'Università di Torino ha appena stipulato un accordo con ANAC e Team Digitale della Presidenza del Consiglio per poter accedere al database dei bandi pubblici e studiare possibili irregolarità all'interno degli stessi. A questo fine copieremo i dati ANAC nel sistema HPC4AI/OLAP.

4. *Le prospettive future*

Credo quindi di poter sostenere che il futuro prossimo sia ricco di sfide per gli informatici, per i giuristi, ma anche e soprattutto per i team multidisciplinari. Ricordo una scuola estiva, nel 2003, in cui ero giovane dottorando, sul tema del fraud detection mediante tecniche di data mining. Era già allora tecnicamente possibile, dal punto di vista strettamente algoritmico, cercare correlazioni per esempio fra stile di vita e dichiarazione dei redditi ma non c'era la capacità tecnica di organizzare i dati in datalake per poterli incrociare facilmente e non c'era neanche forse la volontà di farlo. Volontà che può arrivare solo attraverso la profonda comprensione degli aspetti giuridici ma anche organizzativi e tecnici da parte dei team a supporto dei legislatori e dei manager della Pubblica Amministrazione. Purtroppo questa è solo una condizione necessaria ma non sufficiente alla modernizzazione del sistema paese che rimane sempre appeso alle buone intenzioni senza riuscire ad andare oltre.

Negli ultimi anni tuttavia questa volontà pare emergere. Guardiamo con interesse al Team Digitale della Presidenza del Consiglio che aveva lo scopo dichiarato di ridurre il numero di datacenter della Pubblica Amministrazione da molte migliaia a una decina di datacenter di importanza nazionale, centralizzando alcuni importanti database, come per esempio l'anagrafe. E proprio l'anagrafe e l'esperienza della città di Torino nella integrazione dalla anagrafe locale (gestita da CSI) a nazionale necessaria per la transizione alla carta di identità elettronica ci suggerisce che i processi di consolidamento sono necessari ma complessi sia tecnicamente

che politicamente. Il consolidamento dei datacenter di fatto aggredisce piccole o grandi rendite di posizione che negli anni si sono incrostate e l'energia e la competenza necessaria per rimuoverle è spesso molto più grande di quanto possa inizialmente apparire.

E a questo proposito, secondo me, un solo team radicato a Roma è un approccio al problema "donchisciottesco". Specialmente perché il lavoro necessita specifiche competenze e diffuse collaborazioni su tutto il territorio nazionale.

Sulle competenze l'America della Silicon Valley ha costruito il proprio monopolio. Per citare Mario Tchou che a metà degli anni cinquanta era il giovane ingegnere responsabile dei progetti più innovativi di Olivetti: "Perché le cose nuove si fanno solo con i giovani. Solo i giovani ci si buttano dentro con entusiasmo, e collaborano in armonia senza personalismi e senza gli ostacoli derivanti da una mentalità consuetudinaria". Tchou aveva guidato il gruppo di lavoro per progettare il primo calcolatore interamente italiano: la Calcolatrice Elettronica Pisana (CEP). Adriano Olivetti fu fra i primi al mondo a investire sulle competenze informatiche; lo fece a Pisa perché l'Università aveva accumulato un ingente capitale per sviluppare un acceleratore di particelle e fu Enrico Fermi a suggerire al rettore dell'Università di Pisa di usare i soldi per fare qualcosa di nuovo: una calcolatrice elettronica. Un'esperienza da cui poi nacque nel 1969 a Pisa il primo dipartimento di Informatica italiano (dove anche io ho studiato). Queste stesse competenze portarono il gruppo di Tchou a sviluppare la macchina "Elea 9003", il primo computer della storia interamente realizzato con componenti a stato solido (senza valvole) e che di fatto ha aperto la frontiera della miniaturizzazione e della velocità dei calcolatori elettronici.

In ambito scientifico, sui temi come BigData e Intelligenza Artificiale noi informatici dobbiamo essere estremamente attenti a non diventare esclusivamente dei selezionatori di giovani talenti da spedire in America. Per questo voglio ricordare HPC4AI, un nuovo centro di competenza sui temi del calcolo ad alte prestazioni, del BigData e l'Intelligenza Artificiale che l'Università e il Politecnico di Torino (sotto il mio coordinamento). HPC4AI è stato recentemente finanziato con 4.5M€ (50% Regione Piemonte, 50% Università e Politecnico) con l'obiettivo di creare un grande laboratorio federato dove poter sperimentare nuovi sistemi e tecniche per l'analisi dei dati in diversi ambiti, dalla biologia alla medicina alla giurisprudenza. Il recente accordo fra Università di Torino e ANAC per l'analisi dei bandi pubblici si appoggia a questa infrastruttura di ricerca.

Da questo scopo primario derivano diversi altri obiettivi, fra i quali:

- Sviluppare e mantenere le competenze su questi temi. Perché le competenze sono la base della catena del valore del trasferimento tecnologico e dell'innovazione.
- Sviluppare gli strumenti necessari alla transizione digitale e l'analisi dei dati. Perché se anche si compisse la transizione digitale della Pubblica Amministrazione ma questa avvenisse utilizzando esclusivamente tecnologia di proprietà di multinazionali ci sarebbe poco da gioire. Il funzionamento del nostro paese, dall'anagrafe alla raccolta delle tasse, dipenderebbe in modo esclusivo da società estere. Un lock-in tecnologico che potrebbe essere molto costoso e lesivo della sovranità nazionale (Italiana e Europea).
- Supportare la transizione al digitale della Pubblica Amministrazione e il consolidamento delle banche dati sono nuovi processi, sviluppati e testati localmente e poi integrati su scala nazionale. Perché l'idea che un solo team, a Roma, da solo e contro tutti possa davvero digitalizzare tutta l'Italia è piuttosto ingenua. Anche senza considerare gli aspetti politici, questo team avrebbe bisogno di centinaia o migliaia di persone (ne ha una trentina adesso) che dovrebbero integrare sistemi regionali e comunali sparsi in tutto il territorio, generare una quantità di processi e proprietà intellettuale che oggi solo i centri partecipati delle università possono garantire. In Piemonte, oltre HPC4AI che è specificamente caratterizzato sulla ricerca scientifica, si può menzionare il CSI che oggi gestisce la gran parte dei dati regionali e comunali e che - almeno secondo me - dovrebbe essere una componente di un sistema federato nazionale per la Pubblica Amministrazione.

L'amministrazione come fornitore e come fruitore di dati personali pubblici: sono praticabili soluzioni basate sulla *Big Data Analytics/Machine Learning*?

ABSTRACT: L'articolo affronta il tema del regime giuridico dei dati pubblici, alla luce delle regole poste dal Regolamento generale per la protezione dei dati personali alla pubblica amministrazione: il principio di finalità cristallizzato dal Regolamento si traduce per l'amministrazione pubblica in un rafforzamento del principio di legalità a cui la stessa deve sottostare. L'articolo evidenzia che dalla convergenza di questi principi deriva un regime di gestione dei dati pubblici più rigoroso e stringente rispetto alla gestione privata, tale da rappresentare un potenziale ostacolo alla sperimentazione del machine learning nel settore pubblico. In particolare, lo specifico regime di gestione dei dati pubblici incide sulle modalità in cui i dati devono essere resi disponibili verso i terzi e quindi sugli obblighi che la pubblica amministrazione ha in quanto "fornitore" di dati e nondimeno sulle modalità di raccolta degli stessi poiché l'amministrazione -in quanto "fruitrice" di dati- deve procedere alla raccolta in un modo funzionalizzato all'interesse pubblico.

1. *Prima premessa: il framework asimmetrico*

L'amministrazione raccoglie, conserva e gestisce una mole enorme di dati personali, le cui potenzialità sono ben lungi dall'essere pienamente sfruttate secondo le logiche e le tecniche di elaborazione della *Big Data analytics* (BDA) e del *Machine Learning* (ML), per la significativa, fondante ragione che l'uso di questi dati raccolti e detenuti dal settore pubblico è soggetto ad un regime che – a tutela della libertà, della dignità, della riservatezza, etc. – ne regola in modo stringente l'utilizzo.

Tale regime – si intende, quello relativo all'uso dei dati personali da parte del settore pubblico – ha specifiche caratteristiche che, sotto l'apparente omogeneità delle regole poste a tutela dei dati personali così come disegnate dal GDPR¹, vedono nettamente distinti i presupposti di raccolta e di utilizzo dei dati personali da parte dei soggetti che esercitano funzioni un potere o una funzione pubblica, rispetto a tutti gli altri soggetti dell'or-

1. Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016 - Regolamento generale sulla protezione dei dati.

dinamento. Infatti, mentre questi ultimi hanno come fine il perseguimento di un “legittimo interesse” autonomamente individuato² e come presupposto di liceità all’uso innanzitutto il consenso dell’interessato, nel caso dei titolari (del trattamento) che esercitano un potere o una funzione pubblica il fine è integralmente etero-determinato³, mentre la raccolta e l’uso dei dati generalmente (salvo alcune categorie di dati particolari) prescinde dal consenso dell’interessato. La raccolta e l’uso risultano in questo caso leciti solo nella misura in cui il loro trattamento sia connesso al perseguimento della finalità/compito di interesse pubblico così come definito dal legislatore ed affidato o “attribuito” all’amministrazione (o al privato che eserciti una funzione pubblica conferita dalla legge).

Il principio di finalità, connesso al presupposto di raccolta e utilizzo dei dati personali, proietta in modo forte il principio di legalità sul trattamento dei dati personali da parte della pubblica amministrazione, alla quale risulta precluso l’uso di dati personali al di fuori del perimetro di liceità disegnato dalla interazione tra principio di legalità (attribuzione di un potere/di un compito di interesse pubblico da parte della legge) e principio di finalità (che è proprio del regime del GDPR).

In verità, il principio di finalità si proietta ben oltre il perimetro dell’esercizio delle funzioni pubbliche, costituendo in effetti una sorta di *grundnorm* che permea ed informa di sé tutta la disciplina eurounitaria di tutela/circolazione dei dati personali. In effetti, tutti i (differenti) presupposti di liceità dei dati personali sono connotati in termini di *finalità* del trattamento (esecuzione di un contratto⁴, adempimento di un obbligo legale⁵, salvaguardia degli interessi vitali dell’interessato o di un’altra persona fisica⁶), ivi compreso il meccanismo del previo consenso, che

2. Cfr. l’art. 6, comma 1, lett. f) del GDPR, ai sensi del quale è lecito “il trattamento [quando] è necessario per il perseguimento del legittimo interesse del titolare del trattamento o di terzi, a condizione che non prevalgano gli interessi o i diritti e le libertà fondamentali dell’interessato che richiedono la protezione dei dati personali, in particolare se l’interessato è un minore”.

3. Infatti, ai sensi del combinato disposto tra l’art. 6, comma 1, lett. f), l’art. 6, comma 3, lett. b) e l’art. 2-ter, comma 1 del Codice privacy, la base giuridica del trattamento di dati necessario per l’esecuzione di un compito di interesse pubblico o connesso all’esercizio di pubblici poteri di cui è investito il titolare del trattamento”, è costituita esclusivamente da una norma di legge o, nei casi previsti dalla legge, da una norma di regolamento.

4. Cfr. l’art. 6, comma 1, lett. b) del GDPR.

5. Cfr. l’art. 6, comma 1, lett. c) del GDPR.

6. Cfr. l’art. 6, comma 1, lett. d) del GDPR.

infatti non costituisce una delega in bianco, ma va espresso “per una o più specifiche *finalità*”⁷, che condizionano la liceità di ogni successivo uso.

Tuttavia – e qui sta la differenza, sulle cui conseguenze vorremmo riflettere – il soggetto di diritto comune dispone della possibilità (e dell’autonomia) di scegliere e determinare le finalità rispetto alle quali raccogliere il consenso presso dell’interessato. Diversamente, al soggetto che raccolga dati personali in vista dell’esercizio di un compito di interesse pubblico, questo spazio di manovra è del tutto precluso, sì che per esso (diversamente che per i soggetti di diritto comune) il principio/vincolo di finalità opera in modo più stringente, poiché – in modo del tutto fisiologico rispetto allo statuto della funzione amministrativa – la finalità risulta integralmente etero-determinata.

Il framework asimmetrico (tra titolari che esercitano *funzioni pubbliche* e titolari che si determinano *autonomamente*) così disegnato dal GDPR comporta (insieme ad altri fattori, di cui si dirà in seguito) una serie di effetti sulla concreta praticabilità della BDA/ML nell’ambito del settore pubblico, di cui si forniranno alcuni esempi concreti.

Per comodità espositiva, tratteremo il tema distinguendo il caso del soggetto pubblico (*rectius*, del titolare del trattamento “necessario per l’esecuzione di un compito di interesse pubblico o connesso all’esercizio di pubblici poteri”⁸) come *fornitore* di dati personali, dal caso del soggetto pubblico quale *fruitore* di dati personali, in ottica di BDA/ML.

2. *Seconda premessa: principio di finalità e logica della BDA/ML*

Prima di procedere è però necessario introdurre un secondo ordine di considerazioni, utile ad evidenziare la tendenziale *incompatibilità* tra le logiche sottese alla BDA/ML e il principio di finalità (cardine del modello di tutela dei dati personali, in particolare nel contesto eurounitario). Il principio di finalità, infatti, è funzionale a garantire l’autonomia e/o il controllo da parte dell’interessato sull’uso dei dati da parte dei titolari del trattamento. Ciò che presuppone che l’interessato (o l’autorità di controllo) sia in condizioni di verificare che il trattamento sia funzionalmente connesso con la specifica finalità per la quale è stato fornito il consenso (o che sorregge la funzione pubblica, etc.), e che tale connessione sia mantenuta anche in seguito (in caso di trattamenti ulteriori). Tuttavia, le logiche del

7. Cfr. l’art. 6, comma 1, lett. a) del GDPR.

8. Cfr. l’art. 6, comma 1, lett. e) del GDPR.

BDA/ML entrano in conflitto con questo modello di tutela, dal momento che loro caratteristica precipua è quella di consentire l'emersione di conoscenza nuova ed *inattesa*, a partire dalla rilevazione di correlazioni/interrelazioni non immediatamente evidenti. Ciò comporta la radicale imprevedibilità dell'uso dei dati, sia con riferimento alla tipologia di trattamento, sia con riferimento alla finalità cui tale uso può in seguito risultare utile/funzionale⁹. L'avvento di queste tecniche di elaborazione dei dati mette in crisi il principio di finalità, che infatti da più parti è ritenuto non più idoneo quale istituto a tutela dell'interessato, in ragione – essenzialmente – della sua ineffettività¹⁰. Questa osservazione offre molte ragioni di perplessità circa l'impianto della disciplina eurounitaria (utili ed opportune *de jure condendo*), che in questa sede è doveroso richiamare¹¹, ma che non costituiscono l'oggetto delle nostre riflessioni. A noi interessa piuttosto evidenziare – *de jure condito* – come questa incompatibilità *di fondo* tra principio di finalità e logiche della BDA/ML interagisca con il framework asimmetrico di cui si è detto sopra. Come vedremo, la maggiore cogenza del principio di finalità nell'ambito dei trattamenti finalizzati all'esercizio di funzioni pubbliche determina alcuni effetti (enfaticizzati anche da alcune scelte in sede di applicazione/completamento del quadro normativo eurounitario in sede nazionale) che amplificano la distanza (in termini di concreta

9. “If the reframing of consent in data protection rules has been instrumental in ensuring the continuous enhancement of the expression of user autonomy and control, new technologies are challenging its limits. There is growing skepticism over the efficiency of consent as a pervasive legal ground for legitimate personal data processing. The design of algorithmic data processing makes the unpredictable and even unimaginable use of data a feature, not a bug, which is directly at odds with the rights and obligations depicted in data protection rights and obligations such as the purpose specification obligation. How can explicit (or even informed) consent be given for specified data processing purposes when the process itself is not transparent or when the purpose is impossible to predict, specify, and explain *ex ante*?”; così A. GIANNOPOULOU, *Algorithmic systems: the consent is in the detail?*, in *Internet Policy Review*, 2000, 9(1), 3; adde M.L. JONES - E. EDENBERG - E. KAUFMAN, *AI and the Ethics of Automating Consent*, in *IEEE Security & Privacy*, 2018, 16(3), 64 ss.

10. Cfr. L. EDWARDS - M. VEALE, *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?*, in *IEEE Security & Privacy*, 16(3), 2018, 46 ss.; D. KAMARINOU - C. MILLARD - J. SINGH, *Machine learning with personal data*, in *Queen Mary School of Law Legal Studies*, Research Paper n. 247/2016; M.S. GAL, *Algorithmic Challenges to Autonomous Choice*, in *Michigan Telecommunications and Technology Law Review*, 2018, 25(1), 59 ss.

11. Cfr. A. MANTELERO, *Regulating big data. The guidelines of the Council of Europe in the context of the European data protection framework*, in *Computer Law & Security Review*, 2017, 5, 584 ss.

praticabilità) tra settore pubblico e settore privato. In sintesi, tali effetti paiono essere anche la conseguenza del diverso *tasso di ineffettività* del principio di finalità (per contrasto con le logiche della BDA/ML) che connota il settore pubblico rispetto a quello privato. Il primo, infatti, è oggetto di procedure di controllo più sistematiche e pervasive, ciò che determina un *maggior grado di effettività* del principio di finalità, con la conseguenza (come vedremo) che la realizzazione di servizi, procedure, e funzioni basati sul BDA/ML incontra ostacoli molto severi.

3. *L'amministrazione come fornitore di dati personali*

Il principio di finalità osta, in tutta evidenza, a che un'amministrazione possa – di propria iniziativa – rendere disponibili a soggetti terzi i dati personali che ha raccolto o che conserva per fini istituzionali, a meno che tale trattamento (la comunicazione, ovvero la diffusione) non sia espressamente prevista dalla legge. Simili previsioni di legge sono ricorrenti, in effetti, per quanto concerne la comunicazione di dati personali *tra amministrazioni*, e di questo aspetto ci occuperemo nel prossimo paragrafo. Per quanto concerne invece la comunicazione all'esterno del settore pubblico (o, *rectius*, al di fuori del perimetro dell'esercizio delle funzioni pubbliche), viene in rilievo una specifica categoria di dati personali che, proprio per effetto di specifiche previsioni legislative, possono essere rese conoscibili a chiunque. Si tratta di quel particolare sottoinsieme di dati personali che l'ordinamento ritiene di rendere conoscibili a chiunque, a fini di *trasparenza*; ossia, per assicurare l'accesso a quelle informazioni mediante la cui conoscenza i cittadini sono posti nella condizione di *comprendere* l'organizzazione, le azioni e le scelte effettuate dei poteri pubblici, al fine di operare quel controllo diffuso che è risolto essenziale del principio democratico¹². Lo stesso GDPR contempla questa categoria di dati, sebbene rimetta la determinazione della ampiezza e della consistenza, ed il relativo regime di accessibilità generalizzata, alla competenza del legislatore nazionale¹³. In questo modo, mentre il regime di tutela dei dati personali è

12. Cfr. F. MERLONI, *Trasparenza delle istituzioni e principio democratico*, in F. MERLONI (a cura di), *La trasparenza amministrativa*, Milano, 2008, 9 ss.; E. CARLONI, *La "casa di vetro" e le riforme. Modelli e paradossi della trasparenza amministrativa*, in *Dir. Pubbl.*, 2009, 779 ss.

13. Cfr. l'art. 86 del GDPR, che recita "I dati personali contenuti in documenti ufficiali in possesso di un'autorità pubblica o di un organismo pubblico o privato per l'esecuzione

(sarebbe?) uniforme a livello UE, vi sono invece tanti (differenti) regimi di accesso ai dati personali quante sono le legislazioni nazionali che operano in concreto il bilanciamento autorizzato dall'art. 86 del GDPR¹⁴.

Nell'ordinamento nazionale italiano, vi sono due principali istituti giuridici differenti che realizzano tale bilanciamento, e che pertanto concorrono a determinare il sottoinsieme di dati personali detenuti dai soggetti pubblici (o da chi eserciti una funzione pubblica) che possono essere fruiti all'esterno, da parte di soggetti terzi. Si tratta in particolare dei dati personali che debbono essere obbligatoriamente pubblicati sui siti delle amministrazioni pubbliche per effetto delle previsioni del d.lgs. n. 33/2013 ("Riordino della disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni"), e che pertanto come tali sono fruibili (o, quantomeno, accessibili) da parte di chiunque; si tratta poi dei dati personali la cui *disclosure* consegua ad una richiesta di accesso civico generalizzato di cui all'art. 5, comma 2 del medesimo decreto legislativo (richiesta potenzialmente attivabile da chiunque).

Questo sottoinsieme di dati personali può essere oggetto di trattamenti (da parte di terzi che vi accedono o ne fruiscono) secondo la logica della BDA/ML? Una prima risposta affermativa potrebbe dedursi dalla chiara statuizione di cui all'art. 3 del decreto "trasparenza", che sancisce (anche) il diritto a riutilizzare i dati resi pubblici mediante pubblicazione o mediante accesso generalizzato¹⁵. Come vedremo di qui a breve, però, le cose non stanno esattamente in questi termini.

di un compito svolto nell'interesse pubblico possono essere comunicati da tale autorità o organismo *conformemente al diritto* dell'Unione o *degli Stati membri* cui l'autorità pubblica o l'organismo pubblico sono soggetti, al fine di conciliare l'accesso del pubblico ai documenti ufficiali e il diritto alla protezione dei dati personali ai sensi del presente regolamento" (corsivo aggiunto).

14. Per quanto concerne bilanciamento asimmetrico tra tutela dei dati personali e diritto alla trasparenza, che caratterizza l'assetto normativo all'incrocio tra ordinamento UE e ordinamenti nazionali degli stati membri, sia consentito rinviare a B. PONTI, *Il luogo adatto dove bilanciare. Il "posizionamento" del diritto alla riservatezza e alla tutela dei dati personali vs il diritto alla trasparenza nella sentenza n. 20/2019*, in *Istituzioni del federalismo*, 2019, 2, 525 ss.

15. Cfr. l'art. 3, comma 1 del d.lgs. 33/2013, che recita "Tutti i documenti, le informazioni e i dati oggetto di accesso civico, ivi compresi quelli oggetto di pubblicazione obbligatoria ai sensi della normativa vigente sono pubblici e chiunque ha diritto di conoscerli, di fruirne gratuitamente, e di utilizzarli e riutilizzarli ai sensi dell'articolo 7".

3.1 *I dati personali pubblici resi disponibili secondo il meccanismo della pubblicazione/diffusione*

Cominciando dai dati personali oggetto di pubblicazione obbligatoria, occorre considerare che le risposte al quesito circa la effettiva riutilizzabilità dei dati personali pubblicati e diffusi ai sensi del decreto trasparenza, alla luce della disciplina di tutela dei dati personali, sono mutate nel tempo, al mutare del quadro normativo di riferimento. In vigenza del precedente quadro normativo (costituito dalla direttiva 95/46/CE e dalla disciplina nazionale di recepimento, raccolta nel d.lgs. n. 196/2003, il cd. “Codice privacy”), le indicazioni formulate dall’autorità di settore miravano a contenere il problema “a valle” (ossia, in una fase procedurale successiva alla fuoriuscita dei dati personali determinata dall’adempimento dell’obbligo di pubblicazione). In particolare, l’approccio delle linee guida varate nel 2014 per fornire indicazioni alle amministrazioni¹⁶ su come contemperare obblighi di pubblicazione e regime di tutela dei dati personali tendeva a scaricare sulle amministrazioni stesse (e non sui fruitori esterni) il compito di dare effettività al principio di finalità, al fine di contemperare il “diritto al riutilizzo” sancito dall’art. 3 del d.lgs. n. 33/2013 (e dal successivo art. 7) con i principi a tutela dei dati personali. Era dunque compito delle amministrazioni limitare (o addirittura proibire) la possibilità di (ri)utilizzo dei dati personali (pure) resi obbligatoriamente disponibili, a seguito di una “rigorosa valutazione di impatto in materia di protezione dei dati, al fine di ridurre il rischio di perdere il controllo sulle medesime informazioni o di dover far fronte a richieste di risarcimento del danno da parte degli interessati”¹⁷. Le linee guida costruivano quindi un meccanismo che tendeva a collocare il peso delle responsabilità per eventuali danni derivanti dalle attività di riutilizzo poste in essere dai terzi (anche quelle consistenti nell’applicazione della logica BDA/LM) anche sulle amministrazioni. Una soluzione di massima precauzione, anche se di dubbia efficacia, dal momento che le licenze in questione non avevano altro valore che ribadire obblighi e presupposti di liceità (gravanti sui terzi riutilizzatori) già imposti dall’ordinamento, e sembravano onerare le amministrazioni di

16. Cfr. le “Linee guida in materia di trattamento di dati personali, contenuti anche in atti e documenti amministrativi, effettuato per finalità di pubblicità e trasparenza sul web da soggetti pubblici e da altri enti obbligati”, adottate con provvedimento del 15 maggio 2014.

17. Cfr. *ivi*, parte I, par. 6. “Limiti al ‘riutilizzo’ di dati personali (artt. 4 e 7 del d. lgs. n. 33/2013)”.

un compito impossibile da perseguire (quello di non perdere il controllo delle informazioni).

Qualcosa cambia con l'avvento del GDPR, che all'art. 5 ("Principi applicabili al trattamento dei dati personali") introduce il cd principio di responsabilizzazione, in base al quale è "il titolare del trattamento [che] è competente per il rispetto del paragrafo 1 e in grado di provarlo («responsabilizzazione»)"¹⁸. In base a questo principio, la responsabilità per il mancato rispetto dei requisiti di liceità (ossia, il "il rispetto del paragrafo 1") ricade sul titolare del trattamento medesimo (ossia, sul soggetto che "determina le finalità e i mezzi del trattamento di dati personali"), così che lo schema disegnato dalle linee guida del 2014 pare perdere di mordente, risultando difficile chiamare a rispondere le amministrazioni per trattamenti di cui non risultano titolari (le p.a. essendo titolari del trattamento consistente nella pubblicazione, mentre i titolari dei trattamenti consistenti in eventuali riutilizzi di tali dati da parte di terzi sono appunto questi ultimi).

Forse non è casuale, dunque, che – una volta indebolita la possibilità di inibire "a valle" la possibilità di riutilizzo dei dati (personali) oggetto di pubblicazione obbligatoria – l'attenzione si sia spostata "a monte", ovvero sul perimetro dei dati oggetto di pubblicazione obbligatoria (oggetto, cioè, di fuoriuscita). Come noto, il *casus belli* è stato rappresentato dall'estensione (nel 2016¹⁹) a tutti i titolari di incarico dirigenziale degli obblighi di pubblicità reddituale e patrimoniale già previsti nel 2013 in capo ai titolari degli organi di indirizzo politico. La nota sentenza della Corte costituzionale, chiamata dal T.A.R. Lazio a verificare la compatibilità di tale disciplina con plurimi principi costituzionali e euorunitari, è rilevante sotto molteplici aspetti²⁰, non ultimo il fatto che il modello di pubblicità/trasparenza

18. Cfr. l'art. 5, comma 2 del GDPR.

19. L'ampliamento dell'ambito soggettivo di applicazione degli obblighi di pubblicità relativa ai dati personali patrimoniali e reddituali, con l'inclusione di tutti i titolari di incarico dirigenziale avviene ad opera del d.lgs. n. 97/2016, adottato in attuazione della delega contenuta nella legge 124/2015, cd. legge "Madia".

20. Nella sentenza 19/2020 la Corte, infatti, ritorna sulla svolta operata con la precedente sentenza n. 269/2017, e relativa al trattamento da riservare all'incidente di costituzionalità nel caso in cui un medesimo diritto riceva protezione (o copertura) sia da parte del diritto costituzionale interno, sia da parte della Carta dei diritti fondamentali dell'Unione europea, ciò che ha comprensibilmente attirato l'attenzione dei commentatori, interessati a verificare se ed in che misura il Giudice delle leggi avrebbe confermato gli approdi di quella epocale pronuncia del 2017; cfr. *ex multis* A. RUGGERI, *La Consulta rimette a punto i rapporti tra diritto euorunitario e diritto intero con una pronuncia in chiaroscuro (a prima lettura di Corte cost. n. 20 del 2019)*, in *Consulta Online*, 23 febbraio 2019; O. POLLICINO - G. REPETTO, *Not to be Pushed Aside: the Italian Constitutional Court and the European Court*

adottato con il d.lgs. n. 33/2013 non ne è uscito stravolto sotto il profilo oggettivo, ma solo limitato quanto alla sua applicazione soggettiva²¹. Ciò che qui interessa sottolineare, però, è che l'esistenza e l'affermazione delle tecniche e le logiche della BDA/ML hanno giocato un ruolo non secondario nell'indirizzare tanto il giudice remittente, quanto la Consulta, in quella parte del giudizio in cui il regime delle pubblicità dei dati reddituali e patrimoniali è stato valutato come sproporzionato. Già il T.A.R. (in sede di rimessione) aveva evidenziato che la mole enorme di dati resi disponibili fosse inintelligibile (e quindi, scarsamente utile proprio a fini di trasparenza) dalla generalità dei cittadini, il cui gran numero risulta privo di quella capacità di lettura che è propria invece di quei "soggetti complessi a vario titolo operanti nell'ordinamento vigente, che, essendo in possesso di strumenti idonei a decrittare importanti masse di informazioni, risultano, a legislazione vigente [...] allo stato i soli in grado di trarre dalle stesse conclusioni coerenti con quanto complessivamente reso disponibile e con gli obiettivi propri della legislazione di cui trattasi"²². Argomentazione ripresa e confermata dalla Consulta, che ha notato come "la pubblicazione di quantità così massicce di dati, infatti, non agevola affatto la ricerca di quelli più significativi a determinati fini (nel nostro caso particolare, ai fini di informazione veritiera, anche a scopi anticorrittivi) se non siano utilizzati efficaci strumenti di elaborazione, che non è ragionevole supporre siano a disposizione dei singoli cittadini.", soggiungendo poco oltre il rischio "di generare 'opacità per confusione', proprio per l'irragionevole mancata selezione, a monte, delle informazioni più idonee al perseguimento dei legittimi obiettivi perseguiti"²³. Come si vede, proprio l'esistenza di strumenti e tecniche idonei a estrarre conoscenza da una massa di informazione altrimenti illeggibile (e confusa), strumenti però indisponibili alla generalità dei cittadini, ma concentrati in pochi soggetti "complessi" comporta una

of Justice, in *VerfBlog*, 27 febbraio 2019, G. BRONZINI, *La sentenza n. 20/2019 della Corte costituzionale italiana: verso un riavvicinamento all'orientamento della Corte di giustizia?*, in *Questione Giustizia*, 4 marzo 2019.

21. Cfr. B. PONTI, *Il luogo adatto dove bilanciare*, cit., 542-543. Per altro, il legislatore (d.l. n. 162/2019, art. 1, comma 7) ha risposto all'appello della Corte a rivedere in modo organico i termini del bilanciamento, ed ha formulato a questo fine alcuni criteri direttivi da tradurre in un regolamento governativo i cui termini di approvazioni e dapprima fissati al 31.12.2020 sono stati in seguito prorogati al 30.04.2021).

22. Così T.A.R. Lazio (sezione I quater), ordinanza del 19 settembre 2017, n. 9828 (corsivo mio).

23. Corte Cost., sentenza 27 febbraio 2019, n. 20, p.to 5.3.1. *Cons. dir.*

censura nei confronti del meccanismo di “fuoriuscita” dei dati personali così congegnato. V’è, in queste considerazioni, una prefigurazione del potere conoscitivo di “pochi soggetti complessi” (quella che in letteratura è definita anche *digital dominance*²⁴) che pare alimentarsi proprio della rottura/superamento del principio di finalità. Non a caso, il giudice pare prediligere invece un meccanismo che selezioni “a monte” le informazioni utili e finalizzate a realizzare la trasparenza: un approccio questo che appare pienamente coerente con il principio di finalità.

3.2 *I dati personali pubblici resi disponibili secondo il meccanismo dell’accesso generalizzato*

Ancora più significativo è il caso dell’altro meccanismo di fuoriuscita di dati personali, quello che consegue ad un accesso Foia. La disciplina legislativa contempla esplicitamente la circostanza che i dati personali possa essere oggetto di *disclosure*, sia perché l’accesso civico generalizzato di cui all’art. 5, comma 2 del d.lgs. n. 33/2013 (come risultante dalle modifiche apportate dal d.lgs. n. 97/2016) si applica “a tutti dati e i documenti detenuti dalla pubblica amministrazione”, sia perché l’art. 5-bis, comma 2 contempla la possibilità di rifiutare l’accesso qualora il diniego sia necessario per evitare un pregiudizio concreto all’interesse privato alla “protezione dei dati personali, in conformità con la disciplina legislativa in materia”: ciò comporta che in assenza di tale pregiudizio concreto il dato personale è (invece) accessibile. Questo quantomeno *in teoria*. Nella prassi, invece, si è affermato un indirizzo interpretativo che ha condotto, nella stragrande maggioranza dei casi, ad un diniego di accesso a dati personali, anche quando si è trattato di dati personali particolarmente connessi all’esercizio di una funzione pubblica (e quindi, più funzionali di altri alla finalità della trasparenza). Ciò che qui interessa sottolineare, tuttavia, non è tanto la proporzione dei dinieghi a fronte degli accoglimenti (comunque significativa), quanto le ragioni addotte per motivare tali dinieghi. Va notato che tali motivazioni sono elaborate essenzialmente dal Garante privacy, che deve essere sempre sentito nelle procedure di ricorso in sede amministrativa che abbiano ad oggetto dati personali. Orbene, nella più larga maggioranza dei casi il parere del Garante (a supporto dei dinieghi di accesso) delinea il seguente percorso argomentativo: quando un dato personale è oggetto di *disclosure* a seguito di una richiesta di accesso civico generalizzato, quel dato acquista il regime

24. Cfr. M. MOORE - D. TAMBINI (a cura di), *Digital dominance: the power of Google, Amazon, Facebook, and Apple*, New York, Oxford University Press, 2018.

del dato pubblico, in quanto conoscibile da chiunque; in conseguenza di questo *status*, il dato personale può essere riutilizzato (ai sensi della disciplina pertinente); poiché al momento della richiesta di accesso non può essere apprezzato in quali contesti tale riutilizzo potrebbe essere effettuato (in seguito) né a quali effetti esso potrebbe dare luogo, e poiché tali effetti potrebbero essere anche di carattere pregiudizievole per l'interessato, allora per evitare questo potenziale pregiudizio l'accesso viene rifiutato²⁵. Dunque, la non prevedibilità degli effetti del riutilizzo (ossia, il carattere distintivo della logica sottesa alla BDA/ML e che si pone in antitesi al principio di finalità) opera (nell'interpretazione del Garante) come un rubinetto che tende a "chiudere" il flusso di dati personali all'esterno del soggetto pubblico. Anche in questo caso, dunque, emerge l'incompatibilità tra il meccanismo di tutela dei dati personali (prevedibilità della finalità/tipologia d'uso, così da poterne apprezzare la "pericolosità") e logica della BDA/ML, con la prima che opera al fine di evitare la concreta praticabilità della seconda.

Va sottolineato che in entrambi i casi qui sopra rappresentati, la limitazione o l'esclusione della fuoriuscita di dati personali finisce per frustrare anche la praticabilità di soluzioni basate su BDA/ML che abbiano come obiettivo proprio la trasparenza, e che pertanto risulterebbero pienamente in linea con il principio di finalità²⁶. La minimizzazione del rischio (potenzialmente derivante da usi non preventivabili e pregiudizievoli) comporta anche la preclusione alle potenziali opportunità. Tale approccio – pur comprensibile, soprattutto quando proposto dall'autorità che ha come (unica) missione istituzionale quella di tutelare i dati personali – appare però non pienamente allineato al quadro normativo eurounitario, che mediante l'esplicitazione del principio di responsabilizzazione sembra prediligere

25. È ricorrente, nei pareri del Garante sulle richieste di accesso civico "Foia", la seguente formula: "si ritiene che l'ostensione dei dati e delle informazioni personali oggetto dell'istanza di accesso civico – considerando che «Tutti i documenti, le informazioni e i dati oggetto di accesso civico [...] sono pubblici e chiunque ha diritto di conoscerli, di fruirne gratuitamente, e di utilizzarli e riutilizzarli ai sensi dell'articolo 7» (art. 3, comma 1, del d. lgs. n. 33/2013), sebbene il loro ulteriore trattamento vada in ogni caso effettuato nel rispetto dei limiti derivanti dalla normativa in materia di trattamento dei dati personali – sia suscettibile di determinare, a seconda dei casi e del contesto in cui possono essere utilizzati da terzi, proprio quel pregiudizio concreto alla tutela della protezione dei dati personali previsto dall'art. 5-bis, comma 2, lett. a), del d. lgs. n. 33/2013" (*ex multis* cfr. Garante privacy, parere del 10 aprile 2017, n. 190).

26. Quanto all'esercizio del diritto al riutilizzo dei dati pubblici, come strumentale alla realizzazione della trasparenza, sia consentito rinviare a B. PONTI, *La mediazione informativa nel regime giuridico della trasparenza: spunti ricostruttivi*, in *Diritto dell'informazione e dell'informatica*, 2019, 383 ss.

modelli regolatori diversi da quelli integralmente fondati sull'esigenza di prevenire ogni possibile rischio, anche a costo di impedire la concreta praticabilità di soluzioni che – quantomeno in certi casi – potrebbero risultare non solo utili, ma anche coerenti con il principio di finalità.

4. *L'amministrazione come fruitore di dati personali*

Il secondo profilo che vorremmo rapidamente indagare è quello della praticabilità di soluzioni basate sulla logica BDA/LM applicata (anche) a dati personali raccolti e detenuti dall'amministrazione, questa volta da parte della stessa amministrazione. Si tratta di un profilo di enorme rilievo, anche in considerazione della mole di dati personali detenuti dal settore pubblico, complessivamente inteso. Il vincolo principale – come già sottolineato – è quello del principio di finalità come connesso a quello di legalità/funzionizzazione dell'azione amministrativa. I dati personali (già detenuti) possono essere (ri)utilizzati solo per finalità che siano le medesime che ne avevano giustificato la raccolta, ovvero per ulteriori finalità che siano valutate come “compatibili” con le prime²⁷. Le finalità sono quelle definite (in termini eteonomi) dal quadro normativo che attribuisce poteri e compiti di interesse pubblico. Pertanto, l'amministrazione subisce (fisiologicamente) questo doppio vincolo: riceve le finalità dall'ordinamento (principio di legalità/attribuzione) e può (ri)utilizzare i dati solo per trattamenti che siano necessari per conseguire tali finalità (o finalità con queste compatibili). In via di prima approssimazione, dunque, si potrebbe ritenere che una soluzione BDA/LM (che contempra anche il trattamento di dati personali) che sia progettata per realizzare finalità identiche o compatibili con quelle affidate all'amministrazione costituisca una strada astrattamente percorribile.

L'approfondimento della normativa (eurolunitaria e nazionale), nonché della prassi maturata negli ultimi anni, ci consente però di apprezzare un quadro regolatorio affatto complesso, nel quale le condizioni effettive di praticabilità di soluzioni BDA/ML risultano particolarmente sacrificate.

Vediamone le ragioni.

4.1 *Il quadro normativo*

In primo luogo, occorre considerare che il patrimonio informativo è distribuito tra differenti amministrazioni ed enti pubblici; con riferimento

27. Cfr. l'art. 6, comma 4 del GDPR.

a quella parte costituita da dati personali (largamente maggioritaria), tale (accentuato) decentramento non costituisce affatto un accidente della storia, ma è piuttosto il frutto di una precisa scelta di politica del diritto, maturata nel corso degli anni settanta del secolo scorso, come risposta “di sistema” ai pericoli derivanti dalla concentrazione di potere conoscitivo connessa all’emergere delle tecnologie di memorizzazione ed elaborazione dei dati. Quella scelta era il frutto di riflessioni consapevoli e particolarmente avanzate²⁸, e non è nostra intenzione metterla in discussione in termini retrospettivi (tutt’altro). Tuttavia, occorre pure fare i conti con alcuni effetti a lungo termine di quella scelta, e che hanno come punto di caduta il framework asimmetrico di cui s’è detto in apertura. Mentre le banche dati di dati personali sono strutturalmente *distribuite* all’interno del settore pubblico, gli attori privati (*rectius*, quei soggetti “complessi” e dominanti, cui faceva riferimento il T.A.R. Lazio nel 2017) possono procedere alla loro illimitata *concentrazione*, e proprio sulla base del principio del consenso (e delle sue ormai evidenti debolezze). Cosa comporti questo in termini di differente praticabilità (nei due contesti) di soluzioni basate su BDA/ML non è difficile immaginarlo.

In termini concreti, poiché i dati personali sono “distribuiti” può accadere (anzi, è l’occorrenza fisiologicamente largamente prevalente) che lo sviluppo di una soluzione di BDA/ML necessiti della integrazione tra banche dati differenti e diversamente dislocate, ossia di un trattamento preliminare: la comunicazione/duplicazione della banca dati. Su questo punto, il legislatore nazionale ha in parte utilizzato gli spazi regolatori aperti dal GDPR, che consente di mantenere o introdurre requisiti e misure ulteriori a garanzia della liceità e della correttezza dei trattamenti²⁹. In particolare, il Codice della privacy, come da ultimo modificato per adeguarsi al GDPR, prevede che se il trattamento è effettuato da una amministrazione diversa da quella che ne è titolare, la comunicazione dei dati da parte dell’amministrazione titolare dei dati è ammessa solo se prevista dalla legge. Se manca tale previsione nella legge, la comunicazione è ammessa quando è comunque necessaria per lo svolgimento di compiti di interesse pubblico e lo svolgimento di funzioni istituzionali, e può essere iniziata se è decorso il termine di quarantacinque giorni dalla

28. Su tutti, lo studio precursore di S. ΡΟΔΟΥΤΑ, *Elaboratori elettronici e controllo sociale*, Bologna, 1973.

29. Cfr. l’art. 6, comma 2 del GDPR, che consente di introdurre requisiti e condizioni ulteriori proprio con riferimento ai trattamenti necessari per lo svolgimento di compiti di interesse pubblico.

relativa comunicazione al Garante, senza che lo stesso abbia adottato una diversa determinazione delle misure da adottarsi a garanzia degli interessati³⁰. Consideriamo che l'autorizzazione legislativa sarebbe comunque necessaria nel caso di trattamenti (ivi compresa la comunicazione tra amministrazioni di dati personali) con finalità non compatibili con quella originaria. Ciò comporta che il legislatore nazionale ha (in effetti) introdotto un obbligo di segnalazione al Garante per tutti i casi in cui la comunicazione/il trasferimento dei dati dall'amministrazione che ne è titolare a quella che intende effettuare il riutilizzo sia funzionale ad un trattamento che è stato valutato (dalle stesse amministrazioni coinvolte) come compatibile con la finalità originari. L'obbligo di segnalazione consente al Garante di indicare eventuali misure a garanzia degli interessati, ma costituisce anche occasione per una valutazione di secondo grado circa la compatibilità tra il nuovo trattamento e quello originario. Anche in questo caso, dunque, viene messa in opera una soluzione volta ad accrescere il tasso di effettività di principio di finalità, una misura che lascia intendere che il principio di responsabilizzazione ha minore rilevanza, quando si tratta di trattamenti posti in essere da soggetti che esercitano poteri pubblici o funzioni/compiti di interesse pubblico. L'asimmetria del framework regolatorio ne risulta così ulteriormente rafforzata.

4.2 *La prassi*

Un secondo elemento che occorre considerare è l'interpretazione dei principi rilevanti in materia elaborata dall'autorità di garanzia, proprio con riferimento alle tecniche di trattamento dei dati che fanno riferimento alle logiche della BDA/ML. Si è già notato che il Garante privacy ha una considerevole voce in capitolo (da giocarsi generalmente in via preventiva, a dispetto del principio di responsabilizzazione), pertanto questo fattore ha un peso decisivo nel definire i margini di concreta praticabilità di queste tipologie di trattamento da parte dei soggetti pubblici.

Le decisioni maturate nel corso degli ultimi anni sono molteplici, e di notevole rilievo, sebbene non tutte comparabili tra loro. Ad esempio, il caso del sistema "SAVIO" – elaborato dall'INPS al fine di concentrare le visite fiscali sui certificati di malattia che presentassero un profilo di più elevato rischio di comportamento anomalo e quindi più probabilmente scorretto o fraudolento (che certamente si basava sul procedure di BDA/ML quali

30. Cfr. l'art. 2-ter, comma 2 del d.lgs. n. 196/2006 (Codice della privacy) così come introdotto dal d.lgs. n. 101/2018.

il *data-mining*), sistema che è stato valutato alla luce del quadro normativo allora vigente (e quindi precedente all'entrata in vigore del GDPR) – presentava significative peculiarità, configurandosi come trattamento integralmente automatizzato, volto alla profilazione (art. 22 GDPR) e che utilizzava dati idonei a rivelare lo stato di salute (art. 9 GDPR). Un insieme di elementi che rendono indispensabile non solo una specifica autorizzazione di rango legislativo (che mancava, e manca tutt'ora), ma anche la previsione di misure appropriate e specifiche per tutelare i diritti fondamentali dell'interessato. In questo senso, i provvedimenti assunti da Garante³¹, e rivolti a interrompere l'uso del sistema (ed anche a sanzionare l'amministrazione) appaiono del tutto aderenti/vincolati dal quadro normativo.

Vi sono tuttavia altri casi in cui – anche in assenza di alcuni di questi elementi – il Garante segnala la tendenza ad interpretare le soluzioni basate sul *machine learning* come comunque incompatibili con il principio di finalità, dal momento la loro applicazione non consente di apprezzare in anticipo i criteri, le modalità e gli esiti del trattamento. Un caso di scuola è rappresentato dalla prolungata interlocuzione tra il Garante e l'Agenzia delle Entrate con riferimento all'uso dei dati integrati nell'Anagrafe rapporti finanziari al fine profilare il rischio di evasione fiscale, così da indirizzare i controlli in modo più efficace. Una interlocuzione affatto significativa, poiché per un verso testimonia in concreto come le logiche del principio di finalità (così come interpretate da Garante) risultino incompatibili con quelle di correlazione statistica che sono proprie del ML³²; dall'altro, indicano come in questo caso (ma analogo è il caso

31. Cfr. Garante privacy, ordinanza d'ingiunzione 29 novembre 2018, n. 492.

32. “Prendiamo ad esempio la vicenda dell'anagrafe dei rapporti e dei conti finanziari dove, dal 2011, accedono i dati relativi ai saldi iniziali, alla giacenza media e ai saldi finali di ogni conto, con i relativi dati anagrafici. In un provvedimento del 17 aprile 2012 il Garante ha stabilito che, per consentire l'utilizzo dei dati, l'Agenzia dovesse sottoporre al Garante stesso preliminarmente i criteri per l'elaborazione delle liste di contribuenti a rischio di evasione. I criteri di profilazione, tuttavia, sono l'esito e non il presupposto dell'analisi dei dati. Se i dati rivelano correlazioni significative tra alcune caratteristiche personali (ad esempio, uno scostamento significativo tra le giacenze medie sul conto e i dati della dichiarazione, magari combinati in modo non lineare con tutte le altre caratteristiche individuali) e i comportamenti a rischio (ad esempio la presenza di un'evasione accertata e definitiva), allora quelle caratteristiche personali diventano un criterio di rischio. Ma queste correlazioni sono rivelate dai dati, non decidibili a priori. E, soprattutto, a posteriori queste correlazioni possono essere difficilmente spiegabili o comprensibili se non, appunto, come regolarità statistiche. Venendo alla sperimentazione, la questione diventa quindi la seguente: il Garante accetterebbe di trovarsi di fronte un insieme di criteri di individuazione dei diversi gradi di rischio che non sono spiegabili e valutabili se non sulla base degli esiti

INPS/SAVIO) l'interlocuzione tra Amministrazione e Garante ha come esito lo scaricare sul legislatore il compito di sciogliere i nodi controversi (liceità del trattamento, e della profilazione, tipologia dei dati da trattare, logiche del trattamento, etc.)³³. Ciò che – per un verso – conferma (e, semmai, consolida ulteriormente) il principio di stretta legalità che regge il trattamento dei dati personali da parte dei soggetti pubblici; per altro verso, testimonia della difficile praticabilità tanto della sperimentazione, quanto della messa a regime di queste soluzioni, costrette a fare i conti con quadro regolatorio che, quando non apertamente non ostile³⁴, si rivela comunque estremamente rigido (in ragione del ruolo eminente svolto proprio dal legislatore).

A contrario, è significativo richiamare la recente esperienza della Regione Veneto che, per monitorare e tracciare la diffusione del virus così ed intervenire in modo mirato per contenere i focolai e circoscrivere il

puramente o almeno prevalentemente statistici dell'incrocio dei dati avvenuto a monte? La risposta affermativa implica la rinuncia del Garante a porre in questione la logicità astratta di quei criteri (al limite andrebbe valutata la metodologia statistica utilizzata, ma è opportuno che questa funzione sia svolta dal Garante della privacy?). La risposta negativa implica l'impossibilità di arrivare ai criteri attraverso un'analisi approfondita dei dati", così A. SANTORO, *Lotta all'evasione e privacy dei cittadini. Replica al Presidente del Garante privacy*, in *lavoce.info*, 4.10.2019, testo disponibile al sito www.lavoce.info.

33. La disciplina legislativa istitutiva dell'Anagrafe dei rapporti presso l'Agenzia delle Entrate, finalizzata a integrare una serie di informazioni provenienti da fonti differenti, così da consentire un'analisi del rischio di evasione ed elusione fiscale, risale al 2011 (art. 11 del d.l. n. 201/2011, il cosiddetto "salva Italia"), è stata oggetto di successive modifiche in quattro occasioni, anche al fine di adeguarsi alle indicazioni formulate nei pareri espressi dal Garante privacy, che in varie occasioni (a cominciare dal parere n. 145 del 17 aprile 2012) aveva frustrato i tentativi di dare seguito alla realizzazione di soluzioni volte a profilare il rischio di comportamenti fiscalmente scorretti. Da ultimo, anche la legge di bilancio per il 2020 è intervenuta nel tentativo di rimuovere gli ostacoli (connessi al regime di tutela dei dati personali, ed emersi nel corso della prolungata interlocuzione tra Agenzia e Garante) che ancora si frapponevano alla realizzazione di soluzioni di analisi dei dati basate sul ML, e volte a profilare il rischio di evasione (cfr. l. n. 160/2029, art. 1, commi 681-686).

34. Un'ulteriore caso degno di nota è rappresentato dalle notevoli difficoltà incontrate in sede di elaborazione ed adozione del Piano statistico nazionale (PSN) 2017-2019 ed aggiornamento 2019, riconducibili in larga misura all'implementazione da parte dell'Istat di metodologie di estrazione ed aggregazione dei dati volte a semplificare le metodologie di raccolta, evitare duplicazioni, favorire la realizzazione di lavori statistici anche a partire dalla raccolta e l'integrazione di microdati da fonti amministrative, ciò che ha suscitato le perplessità del Garante, che ha più volte protratto l'istruttoria sui punti qualificanti del piano. Per farne un'idea, si vedano i provvedimenti del Garante del 9 maggio 2018 e del 13 febbraio 2019.

contagio, ha predisposto a tempo di record una soluzione basata sull'incrocio di tre diverse banche dati (l'anagrafe sanitaria, che fornisce gli indirizzi dei contagiati e dei conviventi; quella dei dipendenti del sistema sanitario; nonché il database di Veneto Lavoro, l'agenzia regionale che raccoglie i dati di tutti i dipendenti delle aziende e dei datori di lavoro). Una soluzione che ha dimostrato una notevole efficacia, che ha supportato la politica una politica di prevenzione e contenimento del contagio di successo³⁵. Come testimoniato dagli stessi protagonisti della vicenda, la predisposizione del “cruscotto” a supporto delle attività di biosorveglianza è stata possibile solo “forzando” la disciplina a tutela dei dati personali³⁶, quantomeno finché non è intervenuto il legislatore dell'emergenza ad allentarne le maglie³⁷.

Non è quindi un caso che la recente indagine conoscitiva congiunta da parte di tre autorità indipendenti sul tema dei *Big Data*³⁸ si occupi solo in termini estremamente marginali del settore pubblico (per altro, proprio a carico della parte curata da Garante privacy), considerato un ambito di scarso rilievo per la conoscenza del fenomeno. Si può inoltre notare che solo a marzo del 2018 il governo ha lanciato un'iniziativa per diffondere la consapevolezza delle opportunità della cd. Intelligenza Artificiale nel settore pubblico, iniziativa che (al di là della stesura di un sintetico *Libro Bianco*) non ha prodotto effetti significativi.

35. Per una illustrazione della piattaforma si veda l'articolo di *Engeneering* (la azienda produttrice del software) #FASE 2: *Engineering e l'analisi integrata dei dati per la protezione dei cittadini da Sars-Covid-2*, in www.eng.it, 2 maggio 2020. Vedi anche R. LUNA, *I dati del virus, il Veneto e la piattaforma di cui abbiamo bisogno*, disponibile al sito www.repubblica.it, 2 maggio 2020.

36. «Inutile negarlo – riconosce Gubian [responsabile dell'unità operativa complessa dei sistemi informativi di Azienda Zero, l'ente sanitario cui fanno capo tutte le Asl del Veneto, *N.d.R.*] – in tempi normali non si sarebbero potute incrociare queste banche dati. L'abbiamo fatto nell'interesse superiore della salute pubblica, partendo dall'idea che al sistema di contrasto non debba sfuggire neppure un caso positivo perché potrebbe essere generatore di morti. È chiaro che finita l'emergenza tutto dovrà rientrare», cfr. A. PASUALETTO, *Coronavirus, Crisanti: «Così ho violato le regole sui tamponi e ho fatto bene»*, 1 giugno 2020, disponibile al sito www.corriere.it.

37. Per effetto delle disposizioni contenute nell'art. 14 del d.l. n. 14/2020, convertito in l. n. 6/2020.

38. Cfr. AGCM, Agcom, Garante privacy, *Indagine conoscitiva sui Big data*, 2019, disponibile al sito www.agcom.it.

5. *Concludendo*

La rapida analisi che precede ha inteso evidenziare alcuni effetti prodotti dal framework asimmetrico che delinea la praticabilità di soluzioni di trattamento dei dati personali, ed in particolare quelle basate sulla logica BDA/LM. Gli effetti di questa asimmetria pongono il settore pubblico (ciascuno dei suoi attori: le amministrazioni, i gestori di servizi pubblici, tutti i soggetti che esercitano compiti di interesse pubblico) in una condizione di oggettivo svantaggio competitivo rispetto agli attori del settore privato. Sotto questo profilo, tale asimmetria costituisce anche una manifestazione della condizione di *digital dominance* che i “soggetti complessi” esercitano nell’attuale assetto socio-economico, ed anche politico, a livello internazionale. La scarsa praticabilità della BDA/ML nel settore pubblico è dunque anche il sintomo del fatto che il potere conoscitivo si colloca oramai stabilmente fuori dal perimetro dei poteri pubblici: un assetto complessivo di cui la disciplina a tutela dei dati personali è solo un tassello, sebbene particolarmente importante.

Interoperabilità dei dati e riflessi organizzativi: il caso della conservazione digitale

ABSTRACT: L'articolo ricostruisce la nozione di interoperabilità dei sistemi informatici, a partire dalla sua evoluzione normativa e dai tentativi di interoperabilità che si sono susseguiti, per approdare alla definizione finale offerta dal d.lgs. 82 del 2005 (Codice dell'amministrazione digitale), che valorizza l'utilizzo delle interfacce pubbliche e aperte come strumento di lavoro e di gestione dei rapporti interorganici, intersoggettivi e con i privati. A partire da un caso pratico rappresentato dallo standard tecnico SInCRO per la conservazione dei documenti, l'articolo evidenzia le contraddizioni del sistema di conservazione digitale, ove la promozione dell'interoperabilità dei sistemi informatici affinché essi siano aperti e accessibili, si scontra con il carattere proprietario degli standard tecnici (SInCRO) richiesti dal legislatore.

1. Introduzione

Il presente contributo mira ad analizzare il concetto di interoperabilità – a partire dalla nozione normativa contenuta nel Codice dell'amministrazione digitale (CAD) – per verificare se ed in che misura tale caratteristica sia rinvenibile nella disciplina sulla conservazione digitale, e quindi nelle disposizioni tecniche che regolano le modalità attraverso cui le pubbliche amministrazioni sono tenute a custodire nel tempo i propri dati e documenti.

In questo senso, il tema dell'interoperabilità dei dati può essere inquadrato nell'ambito della definizione di informatica tradizionale che ci ha fornito il Prof. Tasso¹ quale «*esecuzione automatica, precisa e deterministica di tutti i passaggi*». Lo scambio e l'allineamento dei dati e dei formati funzionale all'interfacciamento di diversi sistemi informatici è infatti una caratteristica essenziale delle tecnologie dell'informazione e della comunicazione.

1. In questo Volume, C. Tasso, Attori, processi, meriti e responsabilità nell'utilizzo di algoritmi di Intelligenza Artificiale: il caso del Machine Learning, 19 e ss.

Che il tema dell'interoperabilità non sia nuovo, del resto, ce lo ha confermato il Prof. Cavallo Perin² quando ha richiamato l'art. 1, c. 3, lett. a), del d.lgs. 39/1993 che già prevedeva che «*lo sviluppo dei sistemi informativi automatizzati [...] risponde ai [...] criteri [di] integrazione ed interconnessione dei sistemi medesimi*». Senonché lo stesso Prof. Cavallo Perin ci ha altresì ricordato che tale d.lgs. è rimasto completamente inattuato.

L'obiettivo perseguito dal citato d.lgs. 39/1993 non è tuttavia stato abbandonato. Viceversa, i successivi interventi legislativi hanno progressivamente intensificato gli sforzi verso l'interoperabilità dei sistemi. Da ultimo, la rilevanza dell'interoperabilità è confermata dall'attenzione che vi presta il legislatore nel d.lgs. 82/2005 (CAD). Sotto un profilo quantitativo, tale nozione viene attualmente in rilievo ben 28 volte nel CAD. Sotto un profilo qualitativo, tale caratteristica dei sistemi informatici ha un importante ruolo in molteplici parti del testo normativo, come ad esempio nel c.d. Sistema pubblico di connettività di cui all'art. 73 CAD, che consiste in un «*insieme di infrastrutture tecnologiche e di regole tecniche che assicura l'interoperabilità tra i sistemi informativi delle pubbliche amministrazioni, permette il coordinamento informativo e informatico dei dati tra le amministrazioni centrali, regionali e locali e tra queste e i sistemi dell'Unione europea ed è aperto all'adesione da parte dei gestori di servizi pubblici e dei soggetti privati*».

Considerato dunque che il legislatore insiste da ormai tre decenni sulla necessità di rendere i sistemi delle pubbliche amministrazioni interoperabili, conviene capire perché tale caratteristica sia ritenuta tanto importante, partendo dalla nozione stessa di interoperabilità.

2. La nozione di interoperabilità

La definizione di interoperabilità ci è fornita dal d.lgs. 82/2005 (CAD), ai sensi del quale la stessa è la «*caratteristica di un sistema informativo, le cui interfacce sono pubbliche e aperte, di interagire in maniera automatica con altri sistemi informativi per lo scambio di informazioni e l'erogazione di servizi*»³. Tra i vari aspetti interessanti della norma, preme in questa sede sottolineare che tale nozione richiede espressamente che le interfacce di un sistema interoperabile siano «*pubbliche e aperte*».

2. In questo Volume, R. CAVALLO PERIN, *Pubblica amministrazione e data analysis*, 11 e ss.

3. Art. 1, c. 1, lett. dd), CAD.

Tali qualità delle interfacce sono particolarmente importanti in quanto lo sviluppo di un sistema informatico che si vuole rendere interoperabile con altri deve essere conforme a dette interfacce. Solo in tal modo detto sistema informatico, una volta ultimato, può essere effettivamente in grado di comunicare automaticamente con altri dispositivi conformi alle medesime interfacce. Ne discende quindi che, ove le specifiche tecniche delle interfacce non siano conosciute, diventa estremamente difficile – se non impossibile – poter realizzare sistemi tra loro interoperabili. Sicché si può certamente condividere la scelta del legislatore di aver previsto che, per poter parlare di interoperabilità, siano necessarie interfacce «*pubbliche e aperte*».

Non viene tuttavia chiarito dalla norma cosa si intenda con tali due aggettivi in tale contesto. In mancanza di una definizione specifica, possiamo evincerne il significato ricavandolo per analogia da altre disposizioni relative a sistemi informatici che presuppongono tali caratteristiche, fornendone al contempo una definizione.

Quanto al carattere pubblico delle interfacce, si può fare riferimento alla definizione di «*dato pubblico*» fornita dall'art. 2, c. 1, d.lgs. 36/2006 ai sensi del quale è tale «*il dato conoscibile da chiunque*». In base a tale disposizione si può perciò ritenere che le interfacce siano qualificabili come «*pubbliche*» laddove siano «*conoscibili da chiunque*». Il che appare coerente con quanto si è esposto poc'anzi in merito alla necessità di conoscere le interfacce per poter realizzare sistemi tra loro interoperabili.

Quanto al carattere «*aperto*», può essere utile la definizione di «*dati di tipo aperto*» di cui al CAD⁴. Questi sono «*i dati che presentano le seguenti caratteristiche: 1) sono disponibili secondo i termini di una licenza o di una previsione normativa che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato; 2) sono accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti ai sensi della lettera l-bis), sono adatti all'utilizzo automatico da parte di programmi per elaboratori e sono provvisti dei relativi metadati; 3) sono resi disponibili gratuitamente attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, oppure sono resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione salvo quanto previsto dall'articolo 7 del decreto legislativo 24 gennaio 2006, n. 36*». In questo caso, le interfacce, per essere aperte, devono quindi rispondere a tre requisiti: essere sottoposti a licenze senza restrizioni di

4. V. art. 1, c. 1, lett. l-bis, CAD.

utilizzo, essere facilmente accessibili online ed essere tendenzialmente gratuite, salvo circoscritte eccezioni.

Quanto alla funzione assoluta dall'interoperabilità, si può guardare alla definizione fornita dalla Commissione europea, secondo cui tale caratteristica dei sistemi informatici è il «*mezzo grazie al quale ottenere un intercollegamento tra sistemi, informazioni e metodi di lavoro: all'interno di una stessa amministrazione o tra amministrazioni diverse; a livello nazionale o in tutta Europa, oppure con le imprese*»⁵. In altri termini, grazie all'interoperabilità è possibile scambiare dati in modo automatizzato tra un sistema informatico ed un altro.

Per le amministrazioni ciò significa che l'interoperabilità è una caratteristica particolarmente rilevante affinché si possano realizzare sistemi in grado di parlare tra loro, sia nell'ambito dei rapporti interorganici, sia in quelli intersoggettivi, nonché nelle comunicazioni con i privati. È infatti opportuno che i sistemi di cui si dotino le amministrazioni siano realizzati in modo tale da consentire sia che diversi uffici possano scambiarsi dati, sia che ciò sia possibile in rapporto a tutte le amministrazioni che siano coinvolte in un dato procedimento.

Se ne può dunque agevolmente dedurre che l'interoperabilità è una caratteristica dei sistemi informatici particolarmente importante laddove vi sia una frammentazione della fattispecie.

3. *Frammentazione della fattispecie ed interoperabilità dei sistemi*

Ove più enti siano dotati di autonomia organizzativa in relazione ai sistemi informatici da adottare, l'eventuale mancanza di un coordinamento tra i diversi centri decisionali può determinare il rischio che vengano implementati molteplici sistemi informatici tra loro non interoperabili. È dunque ragionevole immaginare che, in un sistema multilivello e improntato alla tutela delle autonomie locali e regionali, e dove si inseriscono, anche a livello statale, molteplici soggetti dotati di forte autonomia ed indipendenza, il processo di digitalizzazione delle amministrazioni possa portare ad un contesto nel quale i sistemi adottati dai diversi enti non sono tra loro interoperabili in mancanza di misure di armonizzazione degli stessi.

Per ovviare a tale problema il legislatore ha adottato diverse strategie. In proposito si può anzitutto ricordare l'art. 69 CAD ai sensi del quale le amministrazioni, ove possibile, dovrebbero rendere riutilizzabili, e quindi

5. Comunicazione della Commissione del 26 settembre 2003, (COM(2003) 567), p. 21.

riutilizzare, i propri sistemi informatici, così che più enti, anche se tra loro autonomi, siano dotati dei medesimi mezzi digitali. Il che, oltre a poter contenere i costi facilitando economie di scala, può anche ridurre la frammentazione dei sistemi informatici.

Un altro strumento attraverso cui il legislatore ha inteso garantire l'interoperabilità dei sistemi delle pubbliche amministrazioni è stato attraverso il conferimento all'AgID, quale soggetto dotato di elevata competenza tecnica, del compito di individuare standard tecnici a cui i sistemi adottati dalle amministrazioni devono uniformarsi. La definizione di standard tecnici è difatti una tecnica ampiamente diffusa per garantire l'interoperabilità. La conformazione allo standard comporta la necessità di adeguare il funzionamento di una o più parti di un sistema ad un dato modello – ad esempio un formato di file – di modo che qualsiasi altro software realizzato conformemente a tale standard sia parimenti in grado di interfacciarsi con tale modello.

Poiché la definizione di standard costituisce un mezzo per garantire l'interoperabilità tra sistemi informatici, e poiché il concetto di interoperabilità – come declinato dal CAD – presuppone che le interfacce siano «*pubbliche e aperte*», per poter considerare uno standard effettivamente idoneo a garantire l'interoperabilità è necessario che lo stesso sia anch'esso pubblico e aperto. Viceversa, la realizzazione di un sistema interoperabile può divenire finanche impossibile in mancanza di una precisa definizione del modello tecnico a cui è necessario conformarsi.

4. *La necessità di standard aperti: il caso dello standard SInCRO*

Il carattere pubblico ed aperto richiesto dal legislatore affinché le interfacce siano effettivamente qualificabili come interoperabili non appare rinvenibile in uno dei più fondamentali aspetti della vita digitale delle amministrazioni, ed anche dei privati, ossia la conservazione digitale.

La conservazione digitale è in sostanza il metodo attraverso cui si garantisce che i documenti informatici siano fruibili anche decorso un certo tempo dalla loro creazione. Anche nella nostra esperienza quotidiana possiamo scontrarci con i problemi posti dalla necessità di accedere ad un file conservato su un dispositivo non più diffuso (floppy disk, cd rom, ecc.), ovvero in un formato obsoleto (ad es. WordStar). Il che ha spinto il legislatore a dettare una specifica disciplina volta a prevenire tale situazione.

L'art. 43, c. 1, del CAD prevede in proposito che «*gli obblighi di conservazione e di esibizione di documenti si intendono soddisfatti a tutti gli effetti di legge a mezzo di documenti informatici, se le relative procedure*

sono effettuate in modo tale da garantire la conformità ai documenti originali e sono conformi alle Linee guida». Le Linee Guida sulla formazione, gestione e conservazione dei documenti informatici⁶ dell'AgID (di seguito, Linee Guida) dell'ottobre 2020, che hanno sostituito quelle previgenti del 2015⁷, offrono anzitutto una definizione tecnica di conservazione, identificandola come «insieme delle attività finalizzate a definire ed attuare le politiche complessive del sistema di conservazione e a governarne la gestione in relazione al modello organizzativo adottato, garantendo nel tempo le caratteristiche di autenticità, integrità, leggibilità, reperibilità dei documenti»⁸.

Per quanto qui di interesse, viene chiarito che «l'interoperabilità tra i sistemi di conservazione dei soggetti che svolgono attività di conservazione è garantita dall'applicazione delle specifiche tecniche del pacchetto di archiviazione definite dalla norma UNI 11386 - Standard SInCRO - Supporto all'Interoperabilità nella Conservazione e nel Recupero degli Oggetti digitali»⁹.

L'obiettivo di garantire l'interoperabilità tra sistemi informatici attraverso la definizione di standard comuni è quindi in concreto qui perseguita identificando lo standard SInCRO quale normativa tecnica di riferimento. In base dunque alla citata nozione di interoperabilità di cui al CAD, peraltro espressamente richiamata dalle Linee Guida¹⁰, ci si dovrebbe aspettare che detta normativa tecnica sia liberamente accessibile e consultabile, onde integrare il concetto di interfacce «pubbliche e aperte»¹¹.

L'Allegato 4 delle Linee Guida, relativo a «standard e specifiche tecniche», non fornisce tuttavia una qualche ulteriore indicazione sui contenuti di tale standard, né tantomeno le norme tecniche da rispettare per conformarsi a questo. In altri termini, né le Linee Guida né i relativi Allegati

6. Versione del 9 settembre 2020, che hanno sostituito quelle precedenti del dicembre 2015, disponibile sul sito istituzionale dell'AgID all'indirizzo https://trasparenza.agid.gov.it/archivio19_regolamenti_0_5385.html.

7. V. 1.0 del dicembre 2015, denominate *Linee guida sulla conservazione dei documenti informatici*.

8. V. Allegato 1, Glossario dei termini e degli acronimi, alle *Linee Guida*, pagg. 5-6.

9. V. para. 4.2, pag. 33 delle *Linee Guida* del 2020. Le previgenti *Linee guida sulla conservazione dei documenti informatici* dell'AgID del 2015 rimandavano in proposito a quanto stabilito dal DPCM 3 dicembre 2013. Anche l'Allegato 4 di dette previgenti Linee Guida, relativo alle specifiche tecniche del pacchetto di archiviazione, del pari indicava che i file oggetto di conservazione dovessero essere inseriti all'interno di un pacchetto di dati, con indice strutturato secondo lo standard SInCRO.

10. V. Allegato 1, Glossario dei termini e degli acronimi, alle *Linee Guida*, pag. 9.

11. Espressione di cui alla definizione di interoperabilità *ex art. 1, c. 1, lett. dd)*, CAD e Allegato 1, Glossario dei termini e degli acronimi, alle *Linee Guida*, pag. 9.

contengono una descrizione degli elementi da implementare per rendere un sistema interoperabile secondo lo standard SInCRO. Per conoscere la struttura vera e propria dell'indice da realizzare è necessario fare riferimento alle norme tecniche di cui al catalogo UNI.

Secondo quanto risultante dal catalogo dell'Ente nazionale italiano di unificazione (UNI), la prima versione dello standard SInCRO è del 2010 (codice UNI 11386:2010)¹². Tale versione è stata aggiornata nel 2020 con l'adozione dello standard UNI 11386:2020, rilasciato il 7 maggio 2020, proprio qualche mese prima dell'adozione delle nuove Linee Guida. A scanso di equivoci, viene espressamente indicato nella scheda tecnica di questa versione che *«la norma definisce la struttura dell'insieme di dati a supporto del processo di conservazione e recupero degli oggetti digitali, individuando gli elementi informativi necessari alla creazione dell'indice di conservazione e descrivendone sia la semantica sia l'articolazione per mezzo del linguaggio formale XML»*¹³

Per poter realizzare un sistema di conservazione conforme alle Linee Guida, e dunque alla previsione di cui all'art. 43, c. 1, CAD che ad esse rinvia, è perciò necessario scaricare le specifiche tecniche dello standard SInCRO UNI 11386:2020 dal catalogo UNI. Senonché l'accesso al relativo file – così come a quello di cui alla versione del 2010 – è protetto da un sistema di *Digital Rights Management* (DRM), ed è possibile solo pagando i relativi diritti, che conferiscono una licenza di utilizzo limitata.

Ne risulta dunque che, a ben vedere, il sistema di conservazione digitale di cui al citato art. 43 CAD, così come declinato nella disciplina tecnica – tanto in quella attuale così come quella di cui alle previgenti Linee Guida del 2015 –, non appare in linea con la citata nozione di interoperabilità in quanto uno degli standard su cui tale sistema si basa è, a ben vedere, di tipo proprietario.

5. *Conclusion*

L'adozione di specifiche tecniche basate su standard proprietari quali quelli adottati per la disciplina della conservazione digitale appare in antitesi con i criteri e principi dettati dal legislatore per l'individuazione

12. V. la scheda tecnica dello standard UNI 11386:2010 all'indirizzo <http://store.uni.com/catalogo/uni-11386-2010>.

13. V. la scheda tecnica dello standard UNI 11386:2020 all'indirizzo <http://store.uni.com/catalogo/uni-11386-2020>.

delle specifiche tecniche. Il che, come nel caso della conservazione, può determinare conseguenze negative per una vasta gamma di soggetti: per le amministrazioni, per le imprese, i professionisti, e per tutti coloro che siano tenuti ad interagire con i sistemi basati su detto standard, ossia in questo caso, con i sistemi volti alla conservazione digitale.

La difficoltà ad accedere alle specifiche tecniche limita le opportunità di sviluppo di nuove soluzioni, riducendo quindi le opzioni a valle, ed aumentando il rischio di *lock-in*. Il che, per quanto riguarda la conservazione digitale, è facilmente verificabile attraverso una ricerca online. Le soluzioni disponibili sono poche, costose e quasi esclusivamente fornite da operatori italiani.

Anche il legislatore si è del resto dimostrato consapevole dell'eccessiva complessità dell'attuale sistema di conservazione avendo dettato, all'art. 25 del d.l. 76/2020 (recante «*misure urgenti per la semplificazione e l'innovazione digitale*»), norme precisamente volte a determinare una «*semplificazione in materia di conservazione dei documenti informatici [...]*». L'intervento, tuttavia, si è limitato a disporre il passaggio da un sistema di previo accreditamento da parte di AgID dei soggetti conservatori – ossia di coloro che mirano ad offrire sul mercato servizi di conservazione – ad un sistema basato sulla mera sussistenza di determinati requisiti soggettivi¹⁴. Si è in altri termini rimosso il requisito del previo accreditamento per l'offerta di servizi di conservazione. Non si è invece in alcun modo intervenuto sugli aspetti tecnici qui evidenziati, che invece appaiono essenziali nella determinazione del livello di complessità del sistema.

Il risultato, in altre parole, è che laddove le interfacce adottate da un'amministrazione non siano effettivamente aperte e pubbliche, quella che può apparire come una misura volta alla promozione dell'interoperabilità – quale l'adozione a livello normativo di un dato standard – può invece produrre l'effetto opposto, riducendo le opzioni a disposizione degli utenti.

Non solo. Nella prospettiva di garantire la maggior circolazione possibile delle tecnologie, e favorire quindi una più ampia scelta ad amministra-

14. Più precisamente, si è previsto che chi voglia offrire il servizio di conservazione digitale debba possedere «*requisiti di qualità, di sicurezza e organizzazione individuati, nel rispetto della disciplina europea, nelle Linee guida di cui all'art. 71 relative alla formazione, gestione e conservazione dei documenti informatici nonché in un regolamento sui criteri per la fornitura dei servizi di conservazione dei documenti informatici emanato da AgID*». Da sottolineare che, ai sensi del medesimo art. 25, fino all'adozione delle Linee Guida e del Regolamento resta in vigore il precedente sistema.

zioni e privati nella selezione delle soluzioni conformi ad un determinato standard, appare preferibile fare affidamento a standard aventi rilievo internazionale. In tal modo le soluzioni adottate dal nostro legislatore possono avere maggiori *chance* di attrarre fornitori stranieri, a tutto beneficio dell'offerta ai clienti finali, aumentando al contempo la visibilità delle iniziative adottate nel nostro ordinamento e, quindi, le possibilità che queste siano esportate altrove.

Infine, poiché la standardizzazione comporta un'attività piuttosto complessa, e non è priva di rischi in ordine al risultato, come l'esempio qui illustrato ben evidenzia, non ci si può esimere anche dal sottolineare che tale sforzo dovrebbe essere limitato a quei soli casi in cui vi sia un'effettiva esigenza di garantire l'interoperabilità tra sistemi informatici. In altri termini, l'interoperabilità dovrebbe essere intesa quale mezzo per conseguire un qualche risultato utile, e non quale semplice fine. Il che, per concludere, non appare immediatamente rinvenibile nel caso qui esaminato relativo alla conservazione digitale. Se, infatti, sino ad oggi la conservazione cartacea era lasciata alla libera disponibilità dei privati, salvo l'onere di esibire la documentazione ove prescritto, non si vede perché, nel mondo digitale, si sia invece ritenuto di dover imporre "dall'alto" le specifiche modalità operative di conservazione nei più minimi dettagli tecnici, invece di lasciare anche in questo caso la scelta sulle modalità di conservazione di dati e documenti ai titolari degli stessi.

Il che, naturalmente, non avrebbe superato l'onere di esibire la documentazione (informatica) nei casi e nei modi previsti dalla legge, né avrebbe impedito di prevedere norme tecniche volte a garantire l'interoperabilità dei dati nella sola fase di esibizione di dati e documenti, senza dover normare anche l'attività di vera e propria conservazione. In altre parole, si sarebbe potuto evitare di regolare ogni singola fase della conservazione digitale, limitando la definizione di norme tecniche alla sola fase di esibizione della documentazione – ossia quella fase che davvero viene resa più efficiente dall'esistenza di sistemi interoperabili – lasciando per il resto la responsabilità di una corretta conservazione a ciascun soggetto.

Dalla smart citizenship alla cittadinanza digitale

ABSTRACT: Il contributo propone una analisi sul perimetro giuridico della cittadinanza digitale in Italia. Appare chiara la spinta, anche dell'Unione europea, verso una digitalizzazione delle amministrazioni, nonché delle loro attività e servizi, che implica una relazione partecipativa con i cittadini, realizzatasi in primo luogo attraverso una serie di iniziative maturate a livello locale e accorpate nel concetto di "smart cities". Meno chiaro il complessivo quadro normativo e i riferimenti costituzionali attraverso cui realizzare questa dimensione nella relazione tra amministrazioni e amministrati. Ci si interroga quindi su quanto sia possibile configurare quella digitale come una ulteriore declinazione dei sistemi di diritti e di doveri di cittadinanza più noti alla dottrina e alla giurisprudenza, in particolare con riferimento alla cittadinanza amministrativa.

1. *Premessa*

Il tema dei diritti dei singoli in internet riguarda diversi aspetti del rapporto tra gli utenti e la rete. Se un significativo dibattito si è svolto sui diritti umani cosiddetti di quarta generazione e sui diritti che, in generale, devono essere tutelati nell'uso della rete e verso i fornitori di servizi internet, meno ci si è soffermati sulle implicazioni per i diritti dei singoli nell'ambito del rapporto con i pubblici poteri, in virtù dell'uso che le pubbliche amministrazioni fanno della rete e delle possibilità offerte dalle nuove tecnologie. Le riflessioni più recenti riguardano la tutela dei dati personali e, più in generale, le questioni di sicurezza legate alla gestione dei dati dei singoli, anche nel rapporto tra diversi ordinamenti, ma il tema ha una rilevanza più ampia e trasversale, pensando alla natura e intensità di scambi e di servizi che la pubblica amministrazione ha con i suoi particolari utenti, anche quando opera *online*.

Il tema emerge in qualche modo già con le prime politiche in materia di *e-government*, promosse a livello internazionale, europeo ed anche nazionale, fin dagli anni '90, ma in questa fase prevaleva un'altra impostazione, un altro approccio nell'applicazione della telematica nelle amministra-

zioni pubbliche, in cui prevaleva un'impostazione volta a informatizzare atti e procedimenti, a passare dalla carta al bit, ma non si riteneva ancora pienamente che ciò implicasse una trasformazione significativa dell'organizzazione e dell'attività delle PA in ragione delle potenzialità innescate dall'evoluzione tecnologica. Anche la tecnologia era a un diverso livello di avanzamento, per cui non era possibile il livello di interazione attualmente praticabile tra le amministrazioni e tra queste e gli utenti¹.

Le politiche di *e-government* hanno, da allora, mutato il loro approccio, passando dalla definizione di *e-government*/governo elettronico², a quella di *digitization*/digitalizzazione delle amministrazioni a quella più recente di *digital government*/governo digitale. Parallelamente è mutato il modo

1. Non è possibile in questa sede ricostruire il percorso di affermazione dell'*e-government* tra livello internazionale, europeo e nazionale. Per quanto utile a contestualizzare l'affermazione e l'evoluzione del fenomeno: F. MERLONI (a cura di), *Introduzione all'e-Government*, Torino, 2005. Può essere utile richiamare le *Linee guida del Governo per lo sviluppo della Società dell'Informazione*, Roma, giugno 2002, secondo cui "Il modello è una Pubblica Amministrazione orientata all'utente, cittadino ed impresa, fornitrice di moderni servizi, creatrice di valore "pubblico", con cui sia facile operare. Una Pubblica Amministrazione efficiente e trasparente nei suoi compiti e nel suo grande patrimonio informativo, fattore di innovazione e di competitività per il Paese. La realizzazione di un tale modello di *e-government* poggia su moderne infrastrutture "abilitanti" che ne assicurano in modo efficiente e sicuro alcune funzionalità di base. Un sistema di *e-government* nei suoi sviluppi più avanzati rappresenterà anche un potente strumento di coinvolgimento e partecipazione dei cittadini ai processi decisionali, evolvendo verso modelli innovativi di *e-democracy*". Nonostante questi auspici la critica più ricorrente alle politiche di *e-government* perseguite in questa fase era che fosse stata dedicata maggiore attenzione agli aspetti tecnologici che a quelli del servizio al cittadino e della radicale riorganizzazione delle attività amministrative. Si veda il position paper del gruppo di lavoro Astrid coordinato da Bruno Dente: *Federalismo informatico e rinnovamento delle istituzioni: dieci tesi sull'E-government*, www.astrid-online.it, 2002. All'evoluzione del concetto e delle politiche di *e-government* e in particolare a un accresciuta considerazione della partecipazione dei cittadini si è affiancato inoltre il percorso verso la trasparenza della PA, che in cui l'uso delle tecnologie ha assunto un peso estremamente rilevante. Si veda almeno E. CARLONI, *L'amministrazione aperta*, Rimini, 2014, 193 e ss. Sulla democrazia elettronica/digitale vedi G. FIORIGLIO, *Democrazia elettronica: presupposti e strumenti*, Padova, 2017, e P. COSTANZO, *La "democrazia digitale". Precauzioni per l'uso*, in *Diritto pubblico*, 1, 2019, 77 ss.

2. Nella definizione datane per la prima volta da Gartner Group, Western Europe Government Sector: IT Solution opportunities, 2000. Ad oggi non esiste un'unica definizione di *e-government*. Con riferimento all'Italia si veda F. MERLONI (a cura di), *Introduzione all'e-government*, cit. Per una contestualizzazione delle politiche di *e-government* e i principali temi rilevanti per la PA. S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione* (online), 2016, Firenze University Press, in particolare, a firma dei due curatori, *La tecnificazione della Pubblica Amministrazione*, 7 ss.

con cui si guarda agli utenti dei servizi online delle PA³ attraverso queste politiche e gli strumenti normativi adottati all'interno dei vari Paesi. Da un approccio *government centred*, incentrato sul ruolo dell'amministrazione si è passati a un sistema più orientato ai bisogni dell'utente/cittadino (*user/citizen centred*) e la prossima tappa, che in alcuni paesi o in singole iniziative è già realtà, è quella di una dimensione *people-citizen driven/user voices*, che parta dalle esigenze delle persone, dalla loro voce come motore delle attività della PA⁴.

Questo percorso è anche quello seguito dall'Unione Europea, dai primi anni '90 e fino ad oggi, con le proprie politiche di *e-government*, che hanno segnato tappe importanti con la Dichiarazione di Malmö del 2009 e, più di recente, nel 2017, con la Dichiarazione di Tallinn, nonché la Dichiarazione di Berlino del 2020⁵. Oggi la digitalizzazione delle pubbliche

3. Sulla definizione di servizi pubblici *online* G. CAMMAROTA, *Erogazione on line dei servizi pubblici e teleprocedure amministrative*, in *Dir. Pubbl.*, 3, 2001, 991 ss. e ID., *Digitalizzazione dell'amministrazione e servizi pubblici on line. Lineamenti del disegno normativo*, in *Dir. Pubbl.*, 2019, 1, 117 e ss.; G. CAMMAROTA, *Servizi in rete della pubblica amministrazione*, in *Digesto delle discipline pubblicistiche*, V Aggiornamento, 2012, Torino, 616 ss. In particolare sulla prospettiva partecipativa dei servizi pubblici *online*: G. CAMMAROTA, *Servizi pubblici online e partecipazione migliorativa*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La Tecnificazione*, cit., 113 e ss.

4. Si veda OECD, *Recommendation of the Council on Digital Government Strategies*, 15 luglio 2014, in cui si definisce il *Digital Government*: "...[it] refers to the use of digital technologies, as an integrated part of governments' modernisation strategies, to create public value. It relies on a digital government ecosystem comprised of government actors, non-governmental organisations, businesses, citizens' associations and individuals which supports the production of and access to data, services and content through interactions with the government." Più di recente, sempre OECD, *Digital Government Strategies for Transforming Public Services in the Welfare Areas, Comparative Study*, 2016, in cui si legge: "The challenge is not to introduce digital technologies into public administrations (digitisation); it is more transformative. The challenge is to integrate the use of digital technologies into public sector modernisation efforts (digital government)". Si veda in questo senso anche il recente documento dell'OECD *Strengthening Digital Government*, <https://www.oecd.org/going-digital/strengthening-digital-government.pdf>, del marzo 2019, che si fonda sul precedente "*Recommendation of the Council on Digital Government Strategies*".

5. Il riferimento attuale delle politiche UE in materia è rappresentato dal recente Piano di azione dell'UE per l'*e-government* 2016-2020, *Accelerare la trasformazione digitale della pubblica amministrazione*, Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni, *Piano d'azione dell'UE per l'eGovernment 2016-2020*, Bruxelles, 19 aprile 2016 COM(2016) 179 final. *Ministerial Declaration on eGovernment approved unanimously in Malmö, Sweden, on 18 November 2009* e Tallinn, *Declaration on eGovernment at the ministerial meeting during Estonian Presidency of the Council of the EU on 6 October 2017*. Berlin Declaration

amministrazioni in Europa, e quindi anche le politiche di *e-government*, sono perseguite principalmente nell'ambito della realizzazione di un Mercato Unico Digitale Europeo⁶, in cui si riconosce il ruolo essenziale delle amministrazioni pubbliche nel complessivo percorso verso il mercato digitale. L'Unione europea interviene disciplinandone aspetti rilevanti, malgrado la perdurante mancanza di competenze sull'assetto amministrativo interno degli stati membri⁷, secondo il principio detto di "indifferenza". Tale principio appare sempre più messo in discussione dalle evoluzioni tecnologiche e dagli scenari che si prefigurano per le pubbliche amministrazioni europee. In questo senso può essere letta anche la Dichiarazione di Tallinn, da cui in particolare emerge l'impegno ad orientare la digitalizzazione delle PA su alcuni obiettivi prioritari, nel rispetto della centralità dell'utente e delle sue aspettative quando si relaziona con la PA⁸.

È dunque sempre più chiaro, nell'Unione Europea ma anche nei singoli ordinamenti, che non può esserci amministrazione digitale senza un'adeguata considerazione dei suoi utenti. Questa relazione dinamica riporta ad alcuni temi classici della relazione amministrazione-utenti ma pone evidentemente anche nuove e diverse questioni che non hanno trovato ancora un compiuto inquadramento giuridico.

Questo risulta evidente anche dall'incerto riferimento costituzionale che, in Italia, può essere invocato a copertura del rapporto tra PA digitale e utenti e da una incompiuta individuazione dei diritti in gioco. In questa sede si può fare solo un cenno alla straordinaria valenza che ha il diritto alla tutela dei dati personali⁹. E quantomeno un cenno merita anche il

on Digital Society and Value-Based Digital Government at the ministerial meeting during the German Presidency of the Council of the European Union on 8 December 2020.

6. *Communication From the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions a Digital Single Market Strategy for Europe*, COM/2015/0192 final. Il più recente piano di *e-government* UE (*E-Government Action Plan 2016-2020*) non per caso è parte integrante della strategia per il *Digital Single Market*.

7. Nonostante l'UE teoricamente non interferisca, nel rispetto del principio di attribuzione di cui all'art. 5 T.U.E. e all'art. 4 T.U.E., par. 2, sull'organizzazione istituzionale interna dei Paesi membri. Sul principio di attribuzione vedi almeno D.U. GALETTA, *Le fonti del diritto amministrativo europeo*, 91, in M.P. CHITI (a cura di), *Diritto amministrativo europeo*, Milano, 2013.

8. Le cinque priorità individuate sono: *Digital-by-default*; *Inclusiveness and accessibility*; *Once only*; *Trustworthiness and security*; *Openness and transparency*; *Interoperability by default*.

9. A fronte della sconfinata produzione scientifica sulla tutela dei dati personali, anche in rapporto con la PA, specie a seguito dell'adozione del Reg. Ue 679/2016 e del d.lgs. 101

fondamentale tema del *digital divide*, nell'ambito del quale l'accesso ad internet deve essere considerato in un'accezione multidimensionale (di accesso alla rete, di infrastrutture materiali, di accessibilità, di condizioni di utilizzo della rete, di competenze digitali, di accessibilità dei servizi e dei contenuti...), tanto per i singoli quanto per le amministrazioni. Per quest'ultime la sfida delle tecnologie si pone in modo pressante ma con notevoli ostacoli, in parte di tipo normativo ma anche in ordine alla capacità di spesa e di competenze all'interno delle stesse PA¹⁰.

2. *Diritti digitali e pubblica amministrazione. Una lettura della cittadinanza in chiave "smart"*

In tempi recenti il tema dei diritti dei singoli nei confronti delle PA in chiave digitale è stato posto a livello internazionale, con alcuni tratti di novità, all'interno del dibattito sulle "smart cities" e sull'idea di un'amministrazione (locale) intelligente, volta a perseguire la sostenibilità in ambito urbano, da realizzare in larga parte attraverso l'uso delle nuove tecnologie¹¹.

del 10 agosto 2018, sia consentito di proporre solo pochi riferimenti bibliografici. Per un quadro generale e una contestualizzazione delle riforme succedutesi sulla tutela dei dati personali in una prospettiva europea e nazionale, F. PIZZETTI, *Privacy e il diritto europeo alla protezione dei dati personali*, Torino, 2016; V. CUFFARO - R. D'ORAZIO - V. RICCIUTO (a cura di), *I dati personali nel diritto europeo*, Torino, 2019, in cui, se si vuole, M. CAPORALE, *Aspetti particolari del trattamento di dati personali in ambito pubblico: accesso ai documenti amministrativi e sistemi di identificazione personale*, 495 ss.; G. FINOCCHIARO, *La protezione dei dati personali in Italia. Regolamento UE n. 2016/679 e d.lgs. 10 agosto 2018, n. 101*, Bologna, 2019. Per i rapporti tra PA e tutela dei dati personali E. CARLONI, *L'amministrazione aperta*, Rimini, 2014.

10. Come emerge dai lavori della Commissione Parlamentare di Inchiesta sul Livello di Digitalizzazione e Innovazione delle Pubbliche Amministrazioni e Sugli Investimenti Complessivi Riguardanti il Settore delle Tecnologie e della Comunicazione, istituita con Deliberazione della Camera dei Deputati del 14 giugno 2016, relazione conclusiva presentata il 26 ottobre 2017. E come accertano le rilevazioni periodiche svolte dalla Commissione europea e tradotte nel DESI Digital Economy and Society Index (DESI). Secondo i dati DESI sull'Italia (2019) il livello delle competenze digitali di base e avanzate degli italiani è al di sotto della media UE. Solo il 44 % degli individui tra i 16 e i 74 anni possiede competenze digitali di base (57 % nell'UE). Con questi dati l'Italia si colloca al ventiseiesimo posto fra gli Stati membri dell'Ue e si trova quindi al di sotto della media Ue.

11. Non si intende in questa sede riproporre il dibattito sulle smart cities. Per un inquadramento del fenomeno, in particolare in Italia e Spagna, si veda M. CAPORALE, *L'attua-*

Le *smart cities* nascono infatti con il proposito di superare, principalmente attraverso la tecnologia, le sfide che il progressivo affollamento delle città pone in tutti i paesi del mondo. L'obiettivo è quindi di rendere sostenibili città in cui convergono enormi masse di popolazione e di garantire servizi adeguati. Le sfide innescate dalla crisi pandemica hanno solo cambiato alcune delle variabili considerate, ma le città intelligenti rimangono comunque un tema sviluppo nelle politiche di digitalizzazione, che probabilmente includeranno anche ulteriori problematiche e dimensioni territoriali diverse.

La definizione che risulta maggiormente diffusa e condivisa sul piano internazionale è quella proposta dall'Università di Vienna², che individua sei assi principali attorno ai quali si svilupperebbero le *smart cities*: *smart economy*; *smart mobility*; *smart environment*; *smart people*; *smart living*; *smart governance*. La definizione che ne scaturisce considera quindi che «A Smart City is a city well performing in a forward-looking way in these six characteristics, built on the “smart” combination of endowments and activities of self-decisive, independent and aware citizens»¹².

zione delle smart cities. Competenze e coordinamento tra livelli di governo, in Istituzioni del Federalismo, 4, 2015, 949 ss.; AA.Vv., *Ville intelligente, ville démocratique?*, Actes du colloque organisé le 13 février 2013 par Berger-Levrault et la Chaire Mutations de l'action publique et du droit public (MADP) de Sciences Po, Berger-Levrault, Paris, 2014; G.F. FERRARI (a cura di), *La prossima città*, Mimesi, Milano-Udine, 2017. Con riferimento al sistema spagnolo F. GARCÍA RUBIO (a cura di), *Las nuevas perspectivas de la ordenación urbanística y del paisaje: smart cities y rehabilitación. Una perspectiva hispano-italiana*, Madrid, Fundación Democracia y Gobierno Local, 2017, in cui se si vuole, M. CAPORALE, *El régimen de las smart cities en Italia*.

12. R. GIFFINGER - C. FERTNER - H. KRAMAR - R. KALASEK - N. PICHLER-MILANOVIĆ - E. MEIJERS, *Smart cities. Ranking of European medium-sized cities*, University of Technology, Vienna, 2007, http://www.smart-cities.eu/download/smart_cities_final_report.pdf. Non è inutile precisare che la ricerca si fonda sulle città di medie dimensioni e che la definizione proposta in chiave “operazionale”. Alcune ricognizioni delle numerose definizioni delle smart cities si trovano in A. COCCHIA, *Smart and Digital City: A Systematic Literature Review*, in R.P. DAMERI - C. ROSENTHAL-SABROUX (eds.), *Smart City. How to Create Public and Economic Value with High Technology in Urban Space*, Springer International Publishing, Switzerland, 2014; T. PARDO - N. TAEWOO, *Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th Annual International Conference on Digital Government Research*, ACM, New York, 2011, 282 ss.; A. CARAGLIU - C. DEL BO - P. NIJKAMP, *Smart cities in Europe*, *Journal of Urban Technology*, 18(2), 2011, 65 ss. con riferimento al concetto di *smart people* e *smart citizenship* W. CASTELNUOVO - G. MISURACA - A. SAVOLDELLI, *Citizen's Engagement and Value Co-production in Smart and Sustainable Cities*, *Conference Paper*, 2015,

Secondo questa definizione aperta di *smart city*, e comunque secondo tutti i principali studi in materia, una componente essenziale del concetto di *smart city* è la “smart people”. Come per le *smart cities* anche della *smart people* non c’è una definizione univoca. Può essere utile utilizzare quella fornita nel Rapporto *Mapping Smart Cities in the Eu*¹³, in cui per *smart people* si intende “e-skills, working in ICT-enabled working, having access to education and training, human resources and capacity management, within an inclusive society that improves creativity and fosters innovation. As a characteristic, it can also enable people and communities to themselves input, use, manipulate and personalise data, for example through appropriate data analytic tools and dashboards, to make decisions and create products and services”.

Nella letteratura sulle *smart cities* il passaggio dal concetto di “people” a quello di “citizens” e di “citizenship” si è consumato senza un dibattito troppo accurato, che pure sarebbe stato opportuno in considerazione dei concetti e delle problematiche interpretative che il termine “cittadinanza” evoca.

In questa assimilazione tra *smart people* e *smart citizens* è opportuno evidenziare che il rapporto dei singoli con la città e con i suoi servizi “smart” non è comunque qualificato da speciali e particolari vincoli, in particolare non dal rapporto di “cittadinanza” che lega gli enti territoriali che abitualmente amministrano le città, e cioè i comuni, con i propri cittadini. Infatti si potrebbero considerare, quali citizen della *smart city* in una prospettiva di *users/consumers* sia residenti che, ad esempio, turisti, o in generale le persone che, anche indipendentemente da un vincolo di stabilità e continuità, vivono la città e fruiscono dei servizi offerti in quell’area urbana. Occorre tuttavia quanto meno domandarsi quale diritto questi soggetti hanno di fruire dei servizi digitali delle PA, in questo caso locali, anche considerando che il concetto di *smart city* e *smart people* si sviluppano intorno ai principi di inclusività e di partecipazione, che dovrebbero essere due degli assi fondamentali su cui fondare l’uso delle nuove tecnologie¹⁴.

International Conference on Public Policies Proceedings. L. SARTORI, *Alla ricerca della smart citizenship*, in *Istituzioni del Federalismo*, 4, 2015, 927 ss.

13. Studio commissionato dal Parlamento europeo e pubblicato nel 2014, [http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET\(2014\)507480_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf).

14. In questo senso sia consentito rinviare ancora a M. CAPORALE, *Administracion local inclusiva y ciudadanía digital*, in M.L. GOMÉZ JIMENÉZ (a cura di), *Estudios sobre el Estatuto Jurídico de las Ciudades Inteligentes*, Atelier, Barcelona, 63-78. Si veda anche, per

Oltre al tema dell'inclusione e della partecipazione e quindi, di nuovo, del superamento del *digital divide*, sono emersi altri elementi problematici su cui è opportuno interrogarsi. Non è inutile ricordare che gli *smart citizens* non sono solo “consumatori” di informazioni e servizi ma assumono un ruolo essenziale anche nell'alimentazione e nell'uso dei dati, senza i quali sarebbe impossibile governare le città in modo “smart” e questo pone una serie considerevole di problemi in ordine alla corretta gestione dei dati personali, da parte di quali soggetti e alle responsabilità per il loro trattamento¹⁵.

Al di là, dunque, di una certa retorica della partecipazione e nella consapevolezza che i dati degli utenti giocano un ruolo essenziale nelle dinamiche di realizzazione della *smart citizenship*, qualunque sia la forma e il livello di governo territoriale competente nell'amministrazione delle città o dei territori “smart” (in Italia principalmente comuni e città metropolitane), le città dovranno fronteggiare la sfida di erogare servizi per numero crescente di persone, presenti in alta concentrazione nel sistema urbano, così da rendere indispensabile garantire servizi più efficienti, più economici e che favoriscano una coesione sociale indispensabile in contesti densamente abitati. Ma gli strumenti che adotteranno toccano evidentemente, anche solo nei due casi citati (partecipazione e tutela dei dati personali) sfere di diritti che non possono essere contenuti solo nella dimensione locale.

Questi temi, che si sono posti nelle esperienze “smart” sviluppate principalmente dalle autonomie territoriali, si pongono dunque con pari forza anche nel rapporto tra cittadini e amministrazioni diverse da quelle locali.

2.1 *Smart citizenship e cittadinanza digitale in Italia*

In Italia il legislatore, con qualche originalità, con il d.l. 179/2012 (convertito nella l. 221/2012) ha compiuto una scelta molto forte, prevedendo in una legge dello stato alcuni aspetti legati all'attuazione delle

uno inquadramento del tema, F. GASPARI, *Smart city, agenda urbana multilivello e nuova cittadinanza amministrativa*, Napoli, 2018.

15. La complessità dei temi non consente in questa sede di evidenziare il ruolo delle compagnie ITC e dei loro interessi nelle *smart cities*. Appare utile almeno un richiamo al volume di S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Luiss University Press, Roma, 2019. Su questi aspetti vedi il recente A. VENANZONI, *Smart cities e capitalismo della sorveglianza*, in *Forum di quaderni costituzionali*, 20 ottobre 2019. L'uso dei dati dei cittadini è talora considerato, criticamente, come il principale obiettivo per la realizzazione delle *smart cities*, che caratterizzerebbe l'interessamento e coinvolgimento delle *tech companies* in modo non propriamente democratico e comunque non con il principale obiettivo di erogare servizi più efficienti da parte delle amministrazioni.

smart city, dall'altro si è distaccato da questa definizione valorizzando la componente umana e sociale¹⁶. La legge parla infatti di "comunità intelligenti" non di città intelligenti, scelta probabilmente giustificata dal desiderio di valorizzare la componente umana, il capitale sociale delle *smart cities* ma anche di non ancorare la definizione di *smart city* a un preciso livello amministrativo di governo locale. In quello stesso periodo era in corso la riforma che avrebbe portato alla definizione, nel 2014, delle città metropolitane come nuovo livello di autonomia territoriale, che supera la concezione del comune come tradizionale livello istituzionale di governo della "città"¹⁷. Il modello delle comunità intelligenti disegnato dal legislatore è stato ripreso nei documenti attuativi dell'Agenda digitale europea e da quelli approvati dall'Agenzia per l'Italia Digitale, AgID, nell'esercizio delle competenze affidatele, principalmente attraverso le linee guida da queste adottate ma anche attraverso lo specifico strumento del Piano Triennale per l'Informatica nelle Pubbliche Amministrazioni, anche se non nella sua prima stesura¹⁸. Occorre infatti sottolineare che, dopo un

16. Decreto legge 179/2012, convertito, con modifiche, in legge 17 dicembre 2012, n. 221. L'art. 20, «Comunità intelligenti», è stato modificato dal decreto legislativo n. 10 del 22 gennaio 2016 e di nuovo dal decreto legislativo 26 agosto 2016, n. 179, che interviene sul Codice di amministrazione digitale CAD.

17. Sul nuovo assetto delineato dalla legge Delrio e in particolare sulle province e sulle città metropolitane si veda almeno: L. VANDELLI, *Città metropolitane, province, unioni e fusioni di comuni*, con la collaborazione di P. BARRERA - T. TESSARO - C. TUBERTINI, Rimini, 2014; A. STERPA (a cura di), *Il nuovo governo dell'area vasta. Commento alla legge 7 aprile 2014 n. 56*, Jovene, Napoli, 2014; F. PIZZETTI, *La riforma degli enti territoriali Città metropolitane, nuove province e unioni di comuni Legge 7 aprile 2014, n. 56 (Legge "Delrio")*, Giuffrè, Milano, 2015. Il concetto stesso di città necessiterebbe di profondi distinguo e analisi, anche in relazione alla sua sovrapposizione o meno con il comune, che qui non è possibile nemmeno accennare. Sia sufficiente menzionare che il Testo unico degli enti locali, Tuel, d.lgs. 267/2000, non esplicita il nesso tra un preciso livello di governo locale e la città.

18. L'AgID è stata fondata dall'art. 19, decreto legge 22 giugno 2012, n. 83, convertito con modifiche alla legge 7 agosto 2012, n. 134, nota come "Crescita-Italia". È stato modificato dal decreto legislativo 179/2016, che ha aggiunto un nuovo art. 14-bis del CAD, d.lgs. 82/2005, che ora costituisce il riferimento per dare un perimetro regolamentare ad AgID. AgID ha le sue competenze nello sviluppo di città intelligenti e, in generale, nella digitalizzazione delle pubbliche amministrazioni italiane. L'articolo 20 del D.L. 179/ 2012 attribuisce infatti all'AgID il compito di: "definire strategie e obiettivi, coordinare il processo di attuazione, predisporre gli strumenti tecnologici ed economici per il progresso delle comunità intelligenti". In particolare, AGID oltre a predisporre il Piano nazionale delle comunità intelligenti e a monitorarne l'attuazione, è tenuta ad "emanare le linee guida recanti definizione di standard tecnici, compresa la determinazione delle ontologie dei servizi e dei dati delle comunità intelligenti, e procedurali nonché di strumenti finanziari innovativi

inizio promettente e addirittura con una legge statale, il tema delle *smart cities*/comunità intelligenti ha avuto alterne fortune. Si consideri che nel primo Piano triennale per l'informatica delle pubbliche amministrazioni, 2017-2019, il tema non è nemmeno menzionato e non è stato quindi considerato una priorità nell'attuazione della digitalizzazione delle PA. Nel secondo Piano triennale, 2019-2021 compare ma con riferimento a una dimensione più ampia, si parla infatti di *smart landscape* più che di *smart cities* e, per quanto qui di interesse, emerge con chiarezza una critica alle iniziative sviluppate fino a questo momento, in quanto l'approccio complessivo delle amministrazioni italiane sarebbe stato eccessivamente circoscritto al contesto urbano e, inoltre, quasi tutte le PA avrebbero preso in maggior considerazione gli aspetti correlati al "cittadino" e ai suoi diritti, trascurando, in qualche modo colpevolmente, le esigenze delle imprese¹⁹. Negli stessi documenti, peraltro, non si fa nemmeno riferimento ad altri profili e dimensioni di "cittadinanza digitale". Solo nel Piano Triennale 2019-2021 il tema affiora in termini diversi, sulla base delle riforme del Codice dell'Amministrazione Digitale su cui ci soffermeremo tra poco.

Da quanto rilevato emerge che la definizione della *smart citizenship* è avvenuta in una dimensione fondamentalmente urbana o di area metropolitana, in larga parte in virtù delle competenze delle autonomie territoriali operanti in questi ambiti. Tuttavia la necessità di accelerare la digitalizzazione delle pubbliche amministrazioni, che, storicamente ma soprattutto dopo la crisi economica, è stata vista come una via necessaria per rendere più efficienti i costi delle amministrazioni e di semplificare i sistemi amministrativi per favorire gli investimenti e lo sviluppo delle imprese, ha fatto sì che in Italia si affermassero in modo crescente esigenze di coordinamento delle politiche di digitalizzazione da realizzare a livello nazionale, con riferimento a tutte le amministrazioni e non solo a quelle delle autonomie territoriali. Si è quindi assistito in qualche modo a una centralizzazione nell'esercizio delle competenze legislative e nell'adozione di regole tecniche²⁰ adottate a livello nazio-

per lo sviluppo delle comunità intelligenti; istituire e gestire la Piattaforma nazionale delle comunità intelligenti". Questo compito è stato assolto anche attraverso l'adozione dei due Piani triennali per l'informatica nelle pubbliche amministrazioni, rispettivamente per gli anni 2017-2019 e 2019-2021, entrambi reperibili al sito www.agid.gov.it.

19. Piano triennale per l'informatica nelle pubbliche amministrazioni, 2019-2010, 159 ss.

20. Con riferimento alle regole tecniche quale strumento normativo idoneo alla digitalizzazione delle pubbliche amministrazioni, da adottarsi normalmente con decreto ministeriale, questo oggi sono state sostituite da "Linee guida" adottate dall'AgID, con

nale, come emerge dalle competenze affidate all'AgID anche in materia di comunità intelligenti cui si è fatto cenno e, in generale, in materia di digitalizzazione delle pubbliche amministrazioni. Nonostante le forme di collaborazione e condivisione con il sistema delle autonomie territoriali, appare evidente che si sia cercato di ricondurre a una omogeneità di modelli e strumenti, sul piano nazionale, le esperienze condotte a livello locale nell'ambito delle varie iniziative "smart".

3. *La cittadinanza digitale oltre la dimensione locale*

Per quanto riguarda più tipicamente i diritti degli utenti nei rapporti con la PA attraverso strumenti digitali e la rete è evidente che questi non vivono solo in una dimensione locale. La caratterizzazione in chiave "smart" e locale di questi diritti si innesta su un concetto più ampio dei rapporti digitali con la PA. Ecco dunque che il tema della smart citizenship confluisce in quello qualitativamente diverso della "cittadinanza digitale", secondo la formulazione introdotta nel nostro ordinamento nel Codice dell'Amministrazione Digitale (da ora in poi CAD)²¹.

Per comprendere come si possa qualificare la cittadinanza digitale rispetto a tutte le PA, locali e centrali, occorre guardare in primo luogo al sistema di competenze legislative tra i diversi livelli di governo, statale e regionale, in materia di digitalizzazione, da ricondurre al "coordinamento

conseguenti riflessioni in ordine alla collocazione, nel sistema delle fonti, di aspetti a volte anche molto rilevanti. Art. 71 CAD come modificato dall'art. 66, comma 1, D.Lgs. 13 dicembre 2017, n. 217. Sulla natura delle Linee guida vedi almeno G. MORBIDELLI, *Linee guida ANAC: comandi o consigli?*, in *Dir. Amm.*, 2016, 3, 273 ss. F. CINTIOLI, *Il sindacato del giudice amministrativo sulle linee guida, sui pareri del c.d. precontenzioso e sulle raccomandazioni di ANAC*, in *Dir. Proc. Amm.*, 2017, S. MORETTINI, *Il soft law nelle autorità indipendenti: procedure oscure e assenza di garanzie?*, in *Osservatorio sull'Air*, www.osservatorioair.it.

21. D.Lgs. 82 del 7 marzo 2005. Per un primo inquadramento sul Codice, E. CARLONI, *Codice dell'amministrazione digitale: commento al D.Lgs. 7 marzo 2005*, Rimini, 2005. A seguito delle numerosissime riforme, per una analisi dell'attuale testo B. CAROTTI, *L'amministrazione digitale: le sfide culturali e politiche del nuovo codice*, in *Giornale di Diritto Amministrativo*, 2017, 1, 7 ss. Sul contesto normativo generale in cui questo inarrestabile flusso di riforme del CAD si colloca E. CARLONI, *Tendenze recenti...*, cit. Tra i pochi interventi della dottrina sugli aspetti che, fin dalla prima stesura del CAD, implicavano una prospettiva centrata sugli utenti e consideravano i diritti di questi ultimi in una prospettiva di cittadinanza, C. FLICK - V. AMBRIOLA, *La cittadinanza amministrativa telematica fra previsioni normative ed effettività*, in *Dir. Informatica*, 6, 2006, 825 ss.

informativo e informatico”, per come definito dalla Costituzione italiana all’art. 117, comma 2, lett. r)²².

A fronte del cospicuo ritardo tecnologico dell’Europa in generale e dell’Italia in particolare, si stanno affermando le ragioni di un coordinamento informativo e informatico sempre più intenso da parte dello stato. In particolare, la Corte costituzionale, in arresti ormai piuttosto risalenti, aveva affermato che il perimetro di tale coordinamento dovesse essere limitato ad aspetti meramente tecnici. Con la più recente sentenza 251/2016 la Corte sembra invece riconoscere un vigore maggiore a questa competenza statale, ravvisando la necessità di offrire garanzie uniformi su tutto il territorio nell’accesso ai servizi della PA implementati, con normativa nazionale, per attuare l’Agenda digitale europea nel nostro paese realizzando alcuni interventi per la digitalizzazione delle PA in ambiti di cui si assumeva la violazione della competenza legislativa regionale. In questa prospettiva la Corte evoca, anche per la digitalizzazione delle PA, la determinazione dei livelli essenziali delle prestazioni, di competenza esclusiva statale, che l’ordinamento italiano applica in altri contesti, ma non spinge oltre questa assimilazione²³.

22. Che si è espresso prima con l’attuazione dell’Agenda digitale europea con l’Agenda digitale italiana e anche di alcune declinazioni a livello locale. Il Dipartimento è stato istituito con Decreto del Presidente del Consiglio dei Ministri 19 giugno 2019, Modifiche al decreto del Presidente del Consiglio dei ministri 1° ottobre 2012 recante «Ordinamento delle strutture generali della Presidenza del Consiglio dei ministri, in G.U. n.199 del 26 agosto 2019). L’attuale Governo italiano ha inoltre istituito il Ministero (senza portafoglio) per l’Innovazione Tecnologica e la Digitalizzazione.

23. Corte costituzionale, sentt. 16 gennaio 2004, n. 17 e 26 gennaio 2005, n. 31 e 25 novembre 2016, n. 251. In generale, M. PIETRANGELO, *Brevi note sul coordinamento informativo informatico e statistico dei dati dell’amministrazione statale, regionale e locale*, in *Informatica e diritto*, 2004, 1-2, 35 ss.; E. D’ORLANDO, *Profili costituzionali dell’amministrazione digitale*, in *Dir. inf. Inf.*, 2011, 213 ss.; F.M. LAZZARO, *Coordinamento informativo e pubbliche amministrazioni*, in *Ist. fed.*, 2011, 23 ss.: E. CARLONI, *Tendenze recenti e nuovi principi della digitalizzazione pubblica*, in *Giornale Dir. Amm.*, 2015, 2, 148 ss. A. MEALE, *Organizzazione della P.A. - il principio di leale collaborazione tra competenze statali e regionali*, in *Giur. It.*, 2017, 3, 733; E. CARLONI, *Digitalizzazione pubblica e differenziazione regionale*, in *Giornale Dir. Amm.*, 2018, 6, 698. La sentenza 251/2016 recita: “Esse [le competenze di cui all’art. 117, comma 2, lett. r), n.d.a.] ... Assolvono, inoltre, all’esigenza primaria di offrire ai cittadini garanzie uniformi su tutto il territorio nazionale, nell’accesso ai dati personali, come pure ai servizi, esigenza che confina anche con la determinazione di livelli essenziali delle prestazioni. Tanto basta per confermare la piena competenza dello Stato, coerente con l’impegno, dallo stesso assunto, di uniformarsi alle indicazioni provenienti dall’Unione europea”. Sui livelli essenziali delle prestazioni esiste una bibliografia considerevole. Per un quadro recente vedi almeno, M. BELLETTI, *I “livelli essenziali delle prestazioni concernenti*

Il bilanciamento tra competenze statali e regionali, l'equilibrio nelle relazioni tra centro e autonomie territoriali sono stati oggetto delle riflessioni e delle politiche nazionali che hanno trovato composizione nel CAD, anche con riferimento alla cittadinanza digitale.

Anche nel testo attuale del CAD è innegabile un rinnovato centralismo, che emerge in primo luogo nel ruolo riconosciuto all'AgID e che può ritenersi confermato anche dalla recente istituzione del Dipartimento per la Trasformazione Digitale.

In particolare le più recenti riforme al CAD che confermano questo centralismo e che hanno introdotto la cittadinanza digitale nel nostro ordinamento in una prospettiva statale, sono state adottate nell'ambito di una vasta legge delega, la legge 124/2015, che ha fatto delle riforme all'amministrazione digitale un perno per la complessiva semplificazione della PA²⁴. Tra i criteri direttivi della delega troviamo infatti l'individuazione di strumenti per definire il livello minimo di sicurezza, qualità, fruibilità, accessibilità e tempestività dei servizi on line delle amministrazioni pubbliche nonché la previsione di speciali regimi sanzionatori e premiali per le amministrazioni stesse (art. 1, c. 1, lett. a) l. 124/2015). Come evidenziato negli studi parlamentari a corredo della legge delega con questa formulazione si sarebbe inteso evocare il regime dei livelli essenziali delle prestazioni concernenti i diritti civili e sociali che devono essere garantiti su tutto il territorio nazionale (art. 117, c. 2, lett. m) Cost.). In tale lettura la delega avrebbe dovuto creare i presupposti per applicare anche ai servizi on line della PA il principio della competenza statale nella determinazione dei livelli essenziali, al fine di estendere una serie di obblighi di digitalizzazione a tutte le amministrazioni regionali. Nulla di tutto ciò è oggi entrato nel CAD in cui sopravvive un art. 7 che rinvia di nuovo all'AgID, per i servizi oggetto del CAD, per la fissazione di standard e livelli di qualità, validi per tutte le amministrazioni e gli enti, nazionali e locali, tenuti ad applicare il CAD²⁵.

i diritti civili e sociali...” alla prova della giurisprudenza costituzionale. Alla ricerca del parametro plausibile...”, in Istituzioni del Federalismo, 3/4, 2003, 614 e ss., con riferimento alla sanità C. BOTTARI (a cura di), I livelli essenziali delle prestazioni sociali e sanitarie, Santarcangelo di Romagna, 2014.

24. L. 7 agosto 2015, n. 124, Deleghe al Governo in materia di riorganizzazione delle amministrazioni pubbliche, in G.U. n. 187 del 13 agosto 2015.

25. Dossier del Servizio Studi del Senato sull'A.S. n. 1577-A, “Riorganizzazione delle amministrazioni pubbliche”, n. 203, aprile 2015; Camera dei Deputati, Scheda di lettura l. 7 agosto 2015, n. 124 del 15 ottobre 2015, n. 303/3. Il testo dell'art. 7 attualmente in vigore, rivolto alle PA tenute ad applicare il CAD come da art. 2, c. 2 CAD recita: “01.

Rispetto dunque a questo scenario, in cui si imporrebbero standard e livelli di qualità uniformi su tutto il territorio nazionale e per tutte le amministrazioni, e in cui si dovrebbe raggiungere l'obiettivo di realizzare una PA che presti servizi omogenei a tutti i cittadini, si pone quindi il tema dei diritti esigibili, da parte dei cittadini rispetto ai servizi stessi, diritti che sono stati posti, appunto, sotto la definizione di una "Carta della cittadinanza digitale".

Il concetto non è definito espressamente nel CAD che ne fornisce tuttavia per la prima volta un perimetro giuridico.

4. *La cittadinanza digitale nell'orizzonte normativo del CAD*

La Legge 124/2015 dedica alla Carta della cittadinanza digitale l'art. 1, scelta che evidenzia quanto prioritario fosse il tema nell'agenda delle riforme per la PA complessivamente perseguite da questa legge. La scelta di dedicare al Codice e ai diritti di cittadinanza digitale l'articolo 1 della legge è sintomatica della centralità che si è inteso riconoscere alle tecnologie dell'informazione e della comunicazione nei rapporti tra cittadini, imprese e pubbliche amministrazioni, con l'ambizione di spostare l'attenzione dal processo di digitalizzazione ai diritti digitali di cittadini e imprese²⁶.

È anche opportuno ricordare che la Legge delega n. 124/2015 nasce sotto il segno della "Accelerazione e semplificazione nei servizi per i cittadini e le imprese"²⁷. Si conferma così ancora una volta che per il legislatore italiano la digitalizzazione della PA è comunque un tutt'uno con la semplificazione e modernizzazione dell'attività amministrativa e questa precisa prospettiva delimita anche il perimetro della cittadinanza digitale. I decreti attuativi della legge 124/2015 (d.lgs. 179/2016 e il successivo d.lgs.

Chiunque ha diritto di fruire dei servizi erogati dai soggetti di cui all'articolo 2, comma 2, in forma digitale e in modo integrato, tramite gli strumenti telematici messi a disposizione dalle pubbliche amministrazioni e il punto di accesso di cui all'articolo 64-bis, anche attraverso dispositivi mobili. 1. I soggetti di cui all'articolo 2, comma 2, provvedono alla riorganizzazione e all'aggiornamento dei servizi resi, sulla base di una preventiva analisi delle reali esigenze degli utenti e rendono disponibili on-line i propri servizi nel rispetto delle disposizioni del presente Codice e degli standard e dei livelli di qualità individuati e periodicamente aggiornati dall'AgID con proprie Linee guida tenuto anche conto dell'evoluzione tecnologica..."

26. Così la relazione di accompagnamento allo schema di decreto legislativo di riforma del CAD del 2017 - D. Lgs. 13 dicembre 2017, n. 217.

27. Art. 1, ddl 1577/2014, XVII Legislatura.

217/2017)²⁸ interpretano la delega rinominando la seconda sezione del CAD “Carta della cittadinanza digitale” e collocando al suo interno alcune disposizioni che ne costituiscono quindi la struttura essenziale. Come dicevamo nessuna definizione viene data della cittadinanza digitale in sé, che deve essere quindi desunta dal complesso delle disposizioni contenute in questa sezione, dagli artt. 3 a 11 del CAD.

Per quanto risulta dai lavori preparatori della legge delega e dei decreti legislativi di attuazione, tutta la prima parte del CAD è stata concepita per essere investita di questa funzione di Carta della cittadinanza digitale. Attraverso la nuova intestazione della seconda sezione del CAD, si è inteso realizzare un’enfaticizzazione non solo formale della cittadinanza digitale in modo che questa permei il disegno complessivo della riforma del CAD in una prospettiva *citizen centred*²⁹. Le modalità individuate passano però attraverso una serie di servizi specifici, principalmente attraverso l’affermarsi di una serie di disposizioni sul diritto ad una identità e ad un domicilio digitale, alla fruizione di servizi pubblici on-line in modo accessibile ed integrato, anche con dispositivi mobili, sul diritto di partecipare effettivamente al procedimento amministrativo per via elettronica così come ad effettuare pagamenti on-line che non definiscono, non esauriscono il possibile orizzonte della cittadinanza digitale ma ne costituiscono una prima materiale attuazione³⁰.

In assenza di una definizione esplicita di cittadinanza digitale, questa viene comunemente riferita all’art. 3 del CAD che sancisce il diritto di *chiunque* di usare, in modo accessibile ed efficace, le soluzioni e gli strumenti predisposti dallo stesso Codice, anche ai fini dell’esercizio dei diritti di accesso e della partecipazione al procedimento amministrativo,

28. D.Lgs. 26 agosto 2016, n. 179, Modifiche ed integrazioni al Codice dell’amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, ai sensi dell’articolo 1 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche e D.Lgs. 13 dicembre 2017, n. 217, Disposizioni integrative e correttive al decreto legislativo 26 agosto 2016, n. 179, concernente modifiche ed integrazioni al Codice dell’amministrazione digitale, di cui al decreto legislativo 7 marzo 2005, n. 82, ai sensi dell’articolo 1 della legge 7 agosto 2015, n. 124, in materia di riorganizzazione delle amministrazioni pubbliche.

29. Così ancora la Relazione illustrativa dello schema di decreto legislativo del 2017 “Le linee portanti di tale nuovo intervento legislativo sono: ... (b) sottolineare con maggior forza la natura di carta della cittadinanza digitale della prima parte del Codice...”.

30. Per la valenza in qualche modo “retorica” della cittadinanza digitale nel CAD e per l’approccio per obiettivi di breve, medio e lungo periodo del CAD, individuati nel parere del Consiglio di Stato in sede consultiva con il parere della Commissione speciale del 17 marzo 2016 v. S. CIVITARESE - L. TORCHIA, *La tecnificazione...*, cit., in particolare 12 e 13.

nei rapporti le Pubbliche Amministrazioni e con gli altri enti individuati nell'art. 2, c. 2 CAD.

Questo riconoscimento va letto in modo complementare all'obbligo espresso nell'art 2, c. 1, CAD, che sancisce il cosiddetto principio di priorità digitale, o *digital first* o anche *digital by default*, secondo cui lo Stato, le Regioni e le autonomie locali assicurano la disponibilità, la gestione, l'accesso, la trasmissione, la conservazione e la fruibilità dell'informazione in modalità digitale e si organizzano ed agiscono a tale fine utilizzando con le modalità più appropriate e nel modo più adeguato al soddisfacimento degli interessi degli utenti le tecnologie dell'informazione e della comunicazione. Completa questa prospettiva l'art. 7, che prevede il diritto di chiunque di fruire dei servizi erogati dalle PA in forma digitale e in modo integrato, tramite gli strumenti telematici messi a disposizione dalle pubbliche amministrazioni e che tali servizi devono essere organizzati e aggiornati sulla base di una preventiva analisi delle reali esigenze degli utenti.

Il CAD quindi esprime un rafforzamento della centralità dell'utente, in una complementarietà tra doveri delle amministrazioni e diritti dei cittadini, ma in una prospettiva per lo più applicativa che non esprime appieno il potenziale della cittadinanza digitale.

Prima del CAD, in una legge che, sbagliando, si potrebbe definire di settore, era contenuta una prospettiva analoga e forse anche più ampia. La legge 9 gennaio 2004, n. 4, "Disposizioni per favorire l'accesso dei soggetti disabili agli strumenti informatici", cosiddetta Legge Stanca, si caratterizza per esordire, già nel suo testo originario, all'art. 1, c. 1, con il riconoscimento del diritto *di ogni persona* ad accedere a tutte le fonti di informazione e ai relativi servizi, ivi compresi quelli che si articolano attraverso gli strumenti informatici e telematici e prosegue, al comma 2, tutelando e garantendo in particolare alle persone con disabilità il diritto ad accedere a servizi informatici e telematici della PA. In questo senso sembra che, da un lato, la legge faccia "troppo": afferma la valenza generale, apparentemente anche al di là dell'oggetto specifico della legge, di un diritto, quello all'informazione e ai relativi servizi, anche in forma telematica, per ogni persona³¹. I destinatari sono quindi tutti, chiunque, diremmo oggi secondo la terminologia utilizzata dal CAD. L'oggetto sono i diritti di informazione e relativi servizi, quindi l'ancoraggio costituzionale

31. "La Repubblica riconosce e tutela il diritto di ogni persona ad accedere a tutte le fonti di informazione e ai relativi servizi, ivi compresi quelli che si articolano attraverso gli strumenti informatici e telematici", art. 1, c. 1, l. 4/2004.

sembra essere inequivocabilmente quello di cui all'art. 21 Cost. In una legge dedicata essenzialmente a strumenti e servizi informatici della PA anche questo passaggio merita di essere sottolineato, anche se il diritto così richiamato non è riconosciuto esplicitamente nei confronti della PA, ma è affermato nella sua valenza più ampia e generale³².

Il primo comma dell'art. 1 avrebbe quindi avuto, già al momento dell'approvazione originaria della legge, una portata rivoluzionaria, a prenderlo sul serio, in quanto affermava (e afferma) il diritto a godere dei servizi dell'informazione offerti dalle nuove tecnologie in termini del tutto generali, anche se poi precisati con riferimento ai diritti di accessibilità delle persone con disabilità.

Nel CAD invece non c'è una riconduzione dei diritti ivi riconosciuti a principi e diritti costituzionali a differenza di quanto evidenziato nella legge Stanca. Eppure sembra ineludibile, anche qui, il richiamo tanto all'art. 97 Cost., vista, se non altro, la collocazione della riforma del CAD nell'alveo della legge delega per la semplificazione dell'amministrazione, ma anche dell'art. 21 Cost. Per quanto sia problematica l'affermazione dei diritti di informazione amministrativa³³, in quel "chiunque" contenuto tanto nell'art. 3 quanto nell'art. 7 CAD non si può non leggere un rinvio allo stesso "chiunque" riconosciuto titolare della libertà di accesso civico ai sensi del d.lgs. 33/3013 (art. 2, c. 1)³⁴, libertà che si attua, anche

32. L. n. 4/2004, art. 1, commi 1 e 2: "1. La Repubblica riconosce e tutela il diritto di ogni persona ad accedere a tutte le fonti di informazione e ai relativi servizi, ivi compresi quelli che si articolano attraverso gli strumenti informatici e telematici. 2. È tutelato e garantito, in particolare, il diritto di accesso ai servizi informatici e telematici della pubblica amministrazione e ai servizi di pubblica utilità da parte delle persone con disabilità, in ottemperanza al principio di uguaglianza ai sensi dell'articolo 3 della Costituzione". Nel senso indicato sembrano deporre anche i lavori preparatori, v. XIV Legislatura, Scheda lavori preparatori, Atto parlamentare n. 3486, "Norme per il diritto di accesso ai servizi e alle risorse telematiche pubbliche e di pubblica utilità da parte dei cittadini diversamente abili". Sulla portata generale di questo riconoscimento: P. COSTANZO, *L'accesso ad internet in cerca d'autore*, in *Dir. Internet*, 3, 2005, 247 ss. In particolare, anche se non del tutto convincente, v. la prospettiva di M. PIETRANGELO, *Il diritto all'uso delle tecnologie nei rapporti con la pubblica amministrazione: luci ed ombre*, in *Inf. e dir.*, 1-2, 2005, 73 ss. Dall'altro lato tale diritto, così fortemente declamato, sembra poi precisato, in termini soggettivi e oggettivi, nella tutela e garanzia del diritto di accesso ai servizi informatici e telematici della pubblica amministrazione e ai servizi di pubblica utilità per le persone con disabilità, sotto la copertura espressa dell'art. 3 Cost., del principio di uguaglianza.

33. G. GARDINI, *Il codice della trasparenza: un primo passo verso il diritto all'informazione amministrativa?*, in *Giornale dir. amm.*, 2014, 878 ss.

34. Art. 1, c. 2, d.lgs. 33 del 14 marzo 2013, come modificato, su delega della stessa legge "Madia" che ha dato luogo alle modifiche del CAD, dal d.lgs. 97 del 25 maggio 2016. Nella

nel disegno del Codice della Trasparenza, principalmente attraverso la pubblicazione di dati, documenti e informazioni nei siti web delle PA e comunque secondo modalità che in via prioritaria si avvalgono della rete internet. Questa lettura è anche sorretta dal richiamo alla partecipazione al processo democratico (art. 9, CAD) e alla promozione da parte delle PA dell'uso delle nuove tecnologie per promuovere una maggiore partecipazione dei cittadini, per facilitare l'esercizio dei diritti politici e civili e per migliorare la qualità degli atti delle PA³⁵.

5. *Rilievi e prospettive per l'affermazione di uno statuto della cittadinanza digitale*

Nell'analisi condotta si è proposta una lettura della cittadinanza digitale che partisse dal basso, dalle esperienze svolte a livello locale, facendo solo riverberare, in queste riflessioni, l'accezione di cittadinanza come status che definisce i rapporti tra cittadino e urbs, amministrata dall'ente territoriale competente, e quella che vede il cittadino che si rapporta con una dimensione non solo locale dei poteri pubblici.

A questo riguardo abbiamo volutamente lasciato sullo sfondo il tema della definizione del concetto di cittadinanza *tout court*, seguendo in qualche modo quella che sembra essere stata la scelta del Legislatore, che non ha inteso portare l'uso del termine "cittadinanza" alla piena assunzione delle conseguenze giuridiche che implicherebbe. Tuttavia indagare la tenuta delle ordinarie categorie del diritto con cui si fa riferimento a questo termine appare un passaggio assolutamente necessario. Questo in particolare nella declinazione della cittadinanza amministrativa, che sembra potere accogliere tutte le istanze connesse allo statuto della cittadinanza digitale, almeno nel suo stato di attuazione e definizione e soprattutto nella sua attuazione a livello locale³⁶.

vasta bibliografia che ha approfondito la portata delle riforme del Codice della Trasparenza, v. almeno B. PONTI (a cura di), *Nuova trasparenza amministrativa e libertà di accesso alle informazioni*, Maggioli, Rimini, 2016.

35. Oltre al possibile utilizzo, di forme di consultazione preventiva per via telematica sugli schemi di atto da adottare (art. 9 CAD).

36. Sulla cittadinanza amministrativa vedi almeno C.E. GALLO, *La pluralità delle cittadinanze e la cittadinanza amministrativa*, in *Dir. amm.*, 3, 2002, 481 e ss.; R. CAVALLO PERIN, *La configurazione della cittadinanza amministrativa*, in *Dir. amm.*, 2004, 201 e ss.; A. BARTOLINI - A. PIOGGIA (a cura di), *Cittadinanze amministrative*, Firenze, 2016. Per le prime riflessioni su questi aspetti a breve distanza dall'adozione del CAD v. C. FLICK

Se si considera infatti la cittadinanza amministrativa quale “appartenenza ad una comunità diversa da quella sovrana, con legittimazione a posizioni soggettive che non dipendono dallo status di cittadino-sovrano... come legittimazione, che nei servizi pubblici tuttavia è sempre stata riconosciuta a ciascun amministrato”³⁷ sembra che la cittadinanza digitale possa esser attratta nell’orbita di questa definizione di cui occorrerebbe tuttavia ponderare con rigore le implicazioni in relazione alle fonti nazionali e sovranazionali, in particolare di quelle europee.

Come abbiamo evidenziato, nella dimensione locale e nell’esperienza “smart” condotta da tante amministrazioni sembrano effettivamente prevalere gli accenti che valorizzano la componente umana, il capitale sociale delle città, in un’ottica più attenta a inclusione e partecipazione, mentre nella dimensione di cittadinanza digitale espressa dal CAD questa sembra assumere più una configurazione limitata, per ora, al diritto/dovere di fruire di una serie di specifici servizi digitali nei confronti della PA.

L’apparato di norme predisposto colloca la cittadinanza digitale in un contesto di più ampio riconoscimento di diritti, sia pure senza un esplicito ancoraggio costituzionale, ma questi prendono corpo solo attraverso il diritto a godere di un primo nucleo di servizi digitali, fondante il rapporto tra PA e cittadini.

L’obiettivo del CAD è quindi, oggi, quello di individuare un nucleo di servizi che rispondano a standard omogenei su tutto il territorio nazionale (art. 7 CAD), e da affidare principalmente a competenze esercitate a livello centrale. Sempre a norma del CAD in caso di violazione degli obblighi di fruire dei servizi in forma digitale e in modo integrato (art. 7 CAD) è competente il neo-istituito difensore civico digitale (art. 17 CAD), ufficio originariamente previsto presso ogni amministrazione ma che oggi è previsto solo presso l’AgID.

La unicità di questa figura a livello nazionale lascia intendere che lo sguardo sull’effettività del rispetto delle previsioni del CAD debba essere unico sul tutto il territorio nazionale e per tutte le amministrazioni.

Inoltre per quando riguarda l’implementazione della cittadinanza digitale, a norma del CAD, l’AgID, è competente per la redazione e pubblicazione sul proprio sito di una “guida di riepilogo” dei diritti di citta-

- V. AMBRIOLA, *La cittadinanza amministrativa telematica fra previsioni normative ed effettività*, in *Dir. informatica*, 6, 2006, 825 ss. Sul rapporto tra cittadinanza ed era digitale vedi il brillante contributo di P. TÜRK, *La citoyenneté à l’ère numérique*, in *Revue de droit publique*, 3, 2018, 623 ss.

37. R. CAVALLO PERIN, *op. ult. cit.*

dinanza digitali previsti dal CAD (art. 17, c. 1-quinquies), ad oggi non adottata. Ancora una volta si tratta di un atto adottato a livello centrale, sia pure con le forme di consultazione e coordinamento previste nel CAD tra Agid e sistema delle autonomie.

Se questo è l'orizzonte entro cui sembra muoversi il legislatore qualche spunto di riflessione ulteriore muove effettivamente dall'AgID che, nel più recente Piano triennale per l'Informatica nella Pubblica Amministrazione (2019-2021), individua l'obiettivo di supportare l'attuazione della cittadinanza digitale nella tutela della sua natura di "diritto/dovere": diritto del cittadino digitale alla fruizione dei servizi; dovere del cittadino ad adeguarsi alle modalità richieste dall'*e-government* per poter esercitare il suo diritto. A parte l'evidenza della conferma che il CAD si muove strettamente in un orizzonte di *e-government*, questa prospettiva appare assolutamente interessante sul piano delle riflessioni per un percorso verso uno statuto della cittadinanza digitale costruito anche sulla base delle tradizionali definizioni di cittadinanza, che implicano il binomio diritti/doveri, e tuttavia rimane un riferimento isolato e scisso dal contesto in cui si è ritenuto di collocare, allo stato attuale, la Carta della cittadinanza digitale.

Se è giusto declinare la cittadinanza digitale in una prospettiva di diritti ma anche di doveri dei cittadini occorre però anche riflettere anche su come la PA sta applicando effettivamente il principio del *digital first* di cui all'art. 2 CAD: l'impianto della riforma e l'adozione degli strumenti previsti, come sembra, rimangono concentrati su una digitalizzazione della PA con una scarsa attenzione alle competenze digitali del cittadino e a una sua educazione alla cittadinanza digitale in generale, per cui il richiamo al dovere del cittadino di adeguarsi alle modalità richieste dall'*e-government* rischia di tradursi nel gravare eccessivamente gli utenti, senza essersi preoccupati di farne crescere le competenze digitali³⁸.

Da questo complesso di riferimenti, ancora frammentari e imprecisi, non sembra ancora giunto il tempo di parlare di un vero e proprio statuto della cittadinanza digitale, quanto piuttosto di una progressiva definizione, per fasi, di servizi ed elementi essenziali per condividere una prospettiva omogenea sul territorio nazionale, partendo appunto

38. E. CARLONI, *Tendenze recenti...*, cit. L'educazione alla cittadinanza digitale, in termini generali, costituisce di specifiche iniziative del Consiglio d'Europa. Inoltre il tema è stato incluso di recente tra quelli che dovrebbero essere oggetto del nuovo insegnamento obbligatorio previsto dalla Legge 20 agosto 2019, n. 92, *Introduzione dell'insegnamento scolastico dell'educazione civica*.

dai primi pilastri dell'identità digitale, dei pagamenti elettronici e così via, secondo i contenuti del CAD. È dunque comprensibile l'opzione del legislatore, nella corsa verso la digitalizzazione della PA e dei suoi servizi, per un più serrato coordinamento a livello statale e di una attuazione della cittadinanza digitale principalmente attraverso i servizi digitali che il CAD implementa. Ma questo dovrebbe essere una prospettiva da superare presto, lasciandosi alle spalle la mera realizzazione di servizi digitali a favore dell'affermazione di un più pieno e maturo statuto di cittadinanza digitale, che espliciti diritti e doveri dei cittadini nei confronti delle PA e una più limpida esigibilità di questi diritti su tutto il territorio nazionale.

Rimane peraltro trascurata dal CAD e dalla sua applicazione anche la qualificazione giuridica dei soggetti titolari dei diritti ivi contenuti. La formula "chiunque" utilizzata dall'art. 3 e ripresa dall'art. 7 del CAD lascia intendere che qualunque soggetto si interfacci con le PA italiane ha diritto di vantare i diritti riconosciuti dal CAD. Questa scelta evoca peraltro quella fatta propria dal legislatore del Codice della Trasparenza che riconosce a "chiunque" la libertà di accesso ai dati e ai documenti detenuti dalle pubbliche amministrazioni nelle due forme di accesso civico contenute del decreto: tramite l'accesso civico cosiddetto generalizzato e tramite la pubblicazione di documenti, informazioni e dati concernenti l'organizzazione e l'attività delle pubbliche amministrazioni, il cosiddetto accesso civico semplice, e le modalità per la loro realizzazione. La scelta di individuare una medesima legittimazione attiva (da riconoscere appunto a "chiunque") per il godimento dei diritti riconosciuti nella riforma dei due provvedimenti (artt. 3 e 7, CAD e art. 1, c. 2 d.lgs. 33/2013, Codice della Trasparenza) non può essere casuale, sia in considerazione del fatto che le modifiche sono scaturite dalla medesima legge di delegazione e nel medesimo orizzonte politico sia in ragione della rilevanza che la componente tecnologica riveste nel Codice della Trasparenza, che realizza l'accesso civico, nelle sue due principali accezioni, sostanzialmente attraverso modalità telematiche. Questo nesso che si è voluto evidenziare tra i due provvedimenti rafforza la lettura in prospettiva democratica della relazione tra PA e cittadini in chiave digitale, tanto per realizzare la libertà di accesso civico quanto per garantire la possibilità di fruire dei servizi della PA secondo le tecnologie digitali.

Tuttavia l'estensione di questa legittimazione, è stata precisata, in via applicativa, fino ad oggi, in modo diverso: per l'accesso civico è stata riconosciuta indipendentemente dalla titolarità di una situazione giuridica soggettiva connessa e dall'essere cittadino italiano o residente nel

territorio dello Stato³⁹. Per la cittadinanza digitale, negli strumenti che ne costituiscono prima attuazione, si deve constatare che questa accezione è in realtà più ristretta. Ad esempio possono richiedere le credenziali del Sistema Pubblico di Identità Digitale solo i cittadini italiani o i soggetti con permesso di soggiorno e residenti in Italia maggiorenni⁴⁰. Ad oggi non proprio chiunque, quindi.

Oltre agli evidenziati problemi interpretativi sulla legittimazione soggettiva, la scelta del legislatore di definire in modo scarso riferimenti costituzionali e principi generali della cittadinanza digitale – con limitati riferimenti al principio democratico e all’alfabetizzazione digitale – e i blandi riferimenti alla esigibilità dei diritti connessi alla cittadinanza digitale, sembrano avere l’obiettivo di superare l’impasse tecnologico della PA nazionale attraverso servizi e modalità precise e determinate, preoccupandosi solo in un secondo momento dell’eventuale costruzione di un vero e proprio statuto della cittadinanza digitale.

Viceversa sarebbe stato utile e forse più costruttivo partire da una definizione dello statuto giuridico della cittadinanza digitale e da questo fare dipendere estensione, natura ed esigibilità dei servizi offerti. Approfondire il contesto e il rinvio a quali concetti di cittadinanza, la qualificazione o meno dei diritti digitali come diritti sociali, l’eventuale applicazione della competenza statale esclusiva non solo in base all’art. 117, comma 3 lett. r) ma anche lett. m), e cioè la determinazione dei livelli essenziali delle prestazioni concernenti i diritti civili e sociali che devono essere garantiti su tutto il territorio nazionale, evocato ma non pienamente affermato dalla citata giurisprudenza costituzionale. Questo ultimo passaggio, se compiuto fino in fondo, superate le notevoli difficoltà che implica, avrebbe il pregio

39. Così l’ANAC, determinazione n. 1309 del 28 dicembre 2016, *Linee guida recanti indicazioni operative ai fini della definizione delle esclusioni e dei limiti all’accesso civico di cui all’art. 5 co. 2 del d.lgs. 33/2013, Art. 5- bis, comma 6, del d.lgs. n. 33 del 14/03/2013 recante «Riordino della disciplina riguardante il diritto di accesso civico e gli obblighi di pubblicità, trasparenza e diffusione di informazioni da parte delle pubbliche amministrazioni».*

40. Art. 64, CAD, *Sistema pubblico per la gestione delle identità digitali e modalità di accesso ai servizi erogati in rete dalle pubbliche amministrazioni*; D.P.C.M. 24 ottobre 2014, *Definizione delle caratteristiche del sistema pubblico per la gestione dell’identità digitale di cittadini e imprese (SPID), nonché dei tempi e delle modalità di adozione del sistema SPID da parte delle pubbliche amministrazioni e delle imprese*; Determinazione AGID, n. 189/2016, *Modifica ai regolamenti SPID e Regolamento recante le modalità attuative per la realizzazione dello SPID del 22 luglio 2016*. Dal complesso di queste deriva l’obbligatorietà di alcuni documenti, tra cui, ad es., il Codice fiscale. Le specifiche richiamate sono riportate al link <https://www.spid.gov.it/richiedi-spid>.

di rafforzare e rendere più trasparenti gli equilibri tra centro e autonomie, valorizzando il lavoro svolto da tante di queste.

Non resta che auspicare che il percorso, costruito dal legislatore apparentemente in senso inverso, porti comunque a una ponderazione e definizione degli effettivi diritti in gioco e della loro possibile implementazione e tutela, a livello nazionale e a livello locale e senza trascurare il livello europeo.

L'Unione europea rileva infatti non solo per l'intensa attività di promozione della digitalizzazione delle pubbliche amministrazioni e per le norme che, in modo sempre più puntuale, arrivano a disciplinarne specifici aspetti, ma anche come fondamento e cornice di specifici diritti di cittadinanza digitale come reso evidente e urgente, in alcuni casi specifici, dalla crisi pandemica. Laddove si decidesse di portare a conseguenze giuridicamente più pregnanti il concetto di cittadinanza digitale, non si potrebbe prescindere dunque dalla dimensione anche della cittadinanza europea⁴¹.

41. Questa lettura si fonda sui principi del diritto dell'Unione, connessi alla libertà di circolazione e di stabilimento di persone, imprese e servizi, oltre che almeno dagli artt. 41 (Diritto a una buona amministrazione) e 42 (Diritto di accesso ai documenti) della Carta dei diritti fondamentali dell'Unione europea. Più puntualmente legittimano questa prospettiva alcuni atti, relativamente recenti. Si pensi ad esempio al Reg. Ue 2018/1724 del Parlamento Europeo e del Consiglio, del 2 ottobre 2018 che istituisce uno sportello digitale unico per l'accesso a informazioni, procedure e servizi di assistenza e di risoluzione dei problemi e che modifica il regolamento (UE) n. 1024/2012. Tale Regolamento nasce con l'obiettivo principale di agevolare le interazioni tra cittadini e imprese, da un lato, e autorità competenti, dall'altro, fornendo accesso a soluzioni *online* per i principali servizi. Si consideri inoltre che non solo il Piano d'azione dell'Unione per l'*e-government* 2016-2020 indica lo sportello tra le azioni previste per il 2017 ma, ai fini di una considerazione dei diritti digitali in una prospettiva di cittadinanza europea, la relazione della Commissione del 24 gennaio 2017, "Rafforzare i diritti dei cittadini in un'Unione di cambiamento democratico - Relazione sulla cittadinanza dell'UE 2017", considera lo sportello una priorità per i diritti dei cittadini dell'Unione.

La partecipazione procedimentale per legittimare gli algoritmi nel procedimento amministrativo*

ABSTRACT: L'articolo guarda ai principi del procedimento amministrativo e fra questi, in particolare, al diritto di partecipazione per individuare nella normativa esistente l'appiglio legislativo per l'introduzione degli algoritmi nell'attività della pubblica amministrazione. L'articolo concentra la sua analisi sull'uso dell'algoritmo nella fase decisionale del procedimento amministrativo e sulle ricadute che questa scelta ha sul contenuto del provvedimento. Data l'incapacità di cogliere le ragioni di fatto e di diritto dell'agire algoritmico, l'individuazione di forme di controllo preventivo e di partecipazione dell'interessato nella fase antecedente l'adozione del provvedimento potrebbero rappresentare lo strumento per accelerare l'uso dell'intelligenza artificiale nel settore pubblico. A partire da questa constatazione e dalla riconducibilità del software dell'algoritmo al documento informatico, si evidenzia come le garanzie che già esistono per l'esercizio del potere pubblico possano trovare applicazione anche per i nuovi strumenti tecnologici utilizzati dall'amministrazione pubblica.

1. *La tecnologia nell'amministrazione*

Da almeno trent'anni, le riforme che hanno interessato l'amministrazione pubblica hanno cercato di avviare un processo di modernizzazione delle procedure amministrative e dell'organizzazione pubblica, attraverso l'introduzione sistematica delle nuove tecnologie.

Tali riforme sono state ispirate da un duplice intento: l'uno che vede nelle nuove tecnologie gli strumenti per il raggiungimento degli obiettivi di economicità, efficacia ed efficienza attraverso un processo di semplificazione¹; l'altro che vede in esse lo strumento capace di allineare la pubblica amministrazione al cambiamento socio-economico della società.

* Relazione ampiamente rivista a seguito del Convegno "L'amministrazione pubblica con i *big data*", Torino 20 e 21 maggio 2019.

1. Molto spesso, infatti, gli interventi legislativi a favore dell'informatizzazione della pubblica amministrazione, sono stati mossi da un'esigenza di semplificazione, nella consapevolezza che ciò ha un riflesso economico sulla pubblica amministrazione, così

La tecnologia è stata vista come lo “stimolo permanente di riforme sostanziali” che attraverso la “radicale trasformazione del metodo di gestione del lavoro amministrativo”² può superare le disfunzioni insite nell’apparato pubblico, attraverso la semplificazione dei procedimenti amministrativi e la riorganizzazione degli apparati pubblici.

Le prime applicazioni informatiche, tra cui gli applicativi gestionali, di archiviazione di documenti o i programmi di videoscrittura (*word processor*) dei testi, non sono state sintomatiche di un’innovazione sensibile poiché l’ingresso nell’attività amministrativa non ha inciso sulle modalità di esternazione della volontà amministrativa: la manifestazione di volontà continuava a circolare in modalità cartacea.

Soltanto le successive applicazioni informatiche hanno permesso l’avvio della dematerializzazione dei documenti originariamente cartacei e hanno così impresso un’accelerata al processo di digitalizzazione del settore pubblico.

La migrazione dell’attività amministrativa verso la dimensione digitale ha così favorito una sensibile trasformazione dell’organizzazione pubblica³: è stato evidente che per la circolazione in rete di documenti amministrativi informatici fosse necessaria l’interconnessione tra sistemi informativi delle pubbliche amministrazioni, nonché un sistema di connettività pubblico. Tale infrastruttura materiale era essenziale al consolidamento della circolazione in modalità telematica dei documenti⁴ tra amministrazioni pubbliche e nei confronti dei terzi. Alla forma cartacea

G. GUZZARDO, *Semplificazione amministrativa e competitività nel governo del territorio*, Bari, 2008; P. LAZZARA, *Principio di semplificazione e situazioni giuridico-soggettive*, in *Dir. amm.*, 2011, 679, un collegamento stretto tra semplificazione ed interazione telematica con la pubblica amministrazione.

2. G. DUNI, *Il procedimento amministrativo tra l. 7 agosto 1990 n. 241 ed introduzione dell’amministrazione telematica*, in *Foro Amm.*, fasc. 1, 1995, 238.

3. A tal proposito, la lettura in combinato disposto degli artt. 58, 59, 60 del TU 445/2000 valorizza la tecnologia come strumento di supporto che, grazie a procedure applicative e di gestione documentale, permettono alle amministrazioni pubbliche di ricercare e visualizzare documenti detenuti da altre amministrazioni e procedere all’acquisizione diretta delle informazioni.

4. È da notare che la modalità telematica ha permesso la circolazione non solo dei documenti, ma anche dei dati e delle informazioni, indipendentemente dalla forma del supporto, così l’articolo 50 del d.lgs. 82/2005 cit. ha ampliato l’ambito di applicazione dell’accesso documentale; ciò anche grazie all’intervento normativo dell’art. 5, d.lgs. 14 marzo 2013, n. 33 che permettendo l’accesso a informazioni e dati della pubblica amministrazione ha valorizzato “il dato conoscitivo come tale, a prescindere cioè da quelli che sono i vincoli derivanti dalle sue modalità di organizzazione e conservazione”.

dei documenti amministrativi si è così affiancata la forma elettronica, che ha avuto riconoscimento soltanto con l'art. 3 del d.lgs. 12 febbraio 1993, n. 39, il cui primo comma ha riconosciuto la validità generale degli atti predisposti attraverso sistemi informatici⁵, mentre il secondo comma ne ha individuato i requisiti che tali atti devono possedere per essere considerati validi⁶.

Le prime interconnessioni⁷ tra sistemi informativi e tra banche dati pubbliche, nonché il miglioramento delle prestazioni degli strumenti informatici hanno permesso che intorno agli anni Novanta il settore pubblico avviasse una prima automazione dell'attività amministrativa e, in particolare, dell'elaborazione del contenuto del provvedimento.

La teorica del provvedimento amministrativo si è così arricchita dell'atto amministrativo *ad elaborazione elettronica* anche definito come atto amministrativo *automatico*⁸, caratterizzato dalla circostanza che la decisione è assunta dal *software* installato sull'elaboratore elettronico in ragione dei dati inseriti⁹.

così F. MOROLLO, *Documento elettronico fra e-government e artificial intelligence (AI)*, in *Federalismi.it*, n. 2, 2015, 16.

5. Successivamente abrogato dal d.lgs. 26 agosto 2016, n. 179.

6. L'atto emanato attraverso sistemi informatici doveva indicare l'autorità emanante, anche detta *fonte*, il responsabile dell'immissione dei dati, della riproduzione su supporto cartaceo o digitale, della trasmissione anche telematica, dell'emanazione dell'atto, nonché del nominativo del soggetto responsabile, soltanto qualora fosse richiesta la particolare firma autografa dell'atto.

7. A. USAI, *Le prospettive di automazione delle decisioni amministrative in un sistema di teleamministrazione*, in *Dir. Inf. e Inform.*, 1993, 164 ss.; A. SCALA, *L'automazione nella redazione degli atti amministrativi*, in *Nuova rass.*, 1995, 1792 e ss.; G. DUNI, *Il documento informatico: profili amministrativi*, in *Leg. e Giust.*, 1995, 362 e ss.; nonché M. MINERVA, *L'attività amministrativa in forma elettronica*, in *Foro amm.*, 1997, 1301 e s.; U. FANTIGROSSI, *Automazione e pubblica amministrazione. Profili giuridici*, Il Mulino, Bologna, 1993.

8. Usa il termine in modo intercambiabile P. OTRANTO, *Decisione amministrativa e digitalizzazione della p.a.*, in *Federalismi*, n. 2, 2018, 2, 15, nonché la sez. III bis del TAR Lazio che con la sentenza 27 marzo 2016, n. 3769 offre una ricostruzione della teorica dell'atto elaborato elettronicamente.

9. Tale attività è stata definita *screen-level bureaucracy* in ragione del ruolo preminente svolto dall'elaboratore elettronico nello svolgimento dell'attività amministrativa. Sviluppa l'evoluzione dalla fase di *screen-level bureaucracy*, a quella di *screen-level bureaucracy* per giungere infine ad una fase più avanzata e cioè la *system-level bureaucracy*. M. D'ANGELOSANTE, *La consistenza del modello dell'amministrazione "invisibile" nell'età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni*, cit., 157 e s.

La prima fase dell'automazione decisionale è stata limitata alle sole ipotesi¹⁰ riconducibili al ragionamento condizionale della logica informatica basata sullo schema *Se (If)... Allora (Then)*. In tali casi, l'algoritmo utilizzato è di tipo deterministico, ossia è programmato¹¹ per fornire una risposta univoca al verificarsi di una determinata condizione¹², senza che essa apporti nuova conoscenza, ma mera *elaborazione di conoscenza*¹³: il dispositivo del provvedimento amministrativo è dunque il frutto dell'applicazione sillogistica ai fatti delle istruzioni impartite dal programmatore.

Oggi, invece, l'amministrazione pubblica può disporre di strumenti informatici basati su algoritmi di apprendimento automatico, che offrono all'amministrazione una conoscenza che, per le modalità con cui viene raggiunta, può essere talvolta inaspettata, ma sicuramente granulare e approfondita¹⁴.

La pubblica amministrazione che utilizza detti algoritmi predittivi a scopi decisionali evolve verso¹⁵ un modello di amministrazione intel-

10. Sulla necessità che l'applicazione degli strumenti informatici all'attività amministrativa richieda una previa normalizzazione del testo e cioè che il testo sia composto da parole a cui corrispondano "una e una sola funzione" e che rappresentino "enunciati giuridici determinati" si veda A. MASUCCI, *L'atto amministrativo informatico: primi lineamenti di una ricostruzione*, Jovene, Napoli, 1993, 19.

11. Detti sistemi informativi hanno natura meramente strumentale poiché la loro capacità decisionale dipende esclusivamente dalle istruzioni previamente impartite, cosicché la decisione è frutto di una predeterminazione degli obiettivi. Sulla stretta dipendenza della capacità decisionale del computer dalla programmazione e dunque dalla volontà dell'uomo di cui sono una mera riproduzione meccanica e ripetitiva si veda R. BORRUSO, *Computer e diritto*, Milano, 1978, nonché V. FROSINI, *Riflessi sociali dell'informatizzazione*, in *Scritti in onore di M.S. Giannini*, Milano, 1988.

12. Si tratta in questo caso del *formalismo procedurale* ossia di quelle tipologie di ragionamento che rientrano nella categoria dei "Rule-Based Systems (RBS)", così definito da P.L.M. LUCATUORTO, *Teorie e modelli del diritto per il ragionamento giuridico automatico*, Milano, LED, 2009.

13. Così A. USAI, *Le prospettive di automazione delle decisioni amministrative in un sistema di teleamministrazione*, cit., 168.

14. Dette potenzialità oggi disponibili si devono al parallelo aumento delle capacità computazionali dei sistemi informatici e alla parallela diffusione di grandi quantità di dati (*big data*), la cui provenienza è differente e la cui alimentazione è continua sì da offrire un quadro sempre più ampio e completo all'amministrazione pubblica. In questa continua e ampia disponibilità di dati che le nuove tecnologie si differenziano dai sistemi esperti che già potevano offrire soluzioni utili anche se a partire da un numero incompleto ed approssimato di informazioni, in quanto hanno conoscenza dei problemi da affrontare e delle metodologie per decidere.

15. Sull'idea dell'unicità dell'amministrazione, sia essa *tradizionale, digitale o intelligente* si esprime P. PIRAS, *Il tortuoso cammino verso un'amministrazione nativa digitale*, in *Diritto dell'Informazione e dell'Informatica*, fasc. 1, 2020, 44.

ligente (*smart administration*)¹⁶, ove la tecnologia¹⁷ è elevata da “ausilio per le decisioni”¹⁸ a “essenza stessa dell’attività amministrativa”¹⁹. Questo cambiamento ha spinto l’amministrazione verso un modello di amministrazione *ragionevole*, che attraverso l’uso di algoritmi può emanare provvedimenti caratterizzati da *razionalità operativa*²⁰.

Sempre più spesso, infatti, le amministrazioni pubbliche si stanno organizzando per sfruttare il potenziale informativo che può essere estrapolato dai dati a disposizione delle stesse: la combinazione della mole di dati a disposizione della pubblica amministrazione e la conoscenza che da questa si ottiene con l’uso di algoritmi ha reso evidente che le tradizionali modalità di assunzione delle decisioni da parte delle amministrazioni pubbliche richiedano un ripensamento significativo.

Ne deriva che la possibilità di inferire decisioni da grandi quantità di dati in tempi rapidi rappresenta una forma di gestione della complessità e permette di giungere a decisioni maggiormente ragionevoli.²¹

16. Così I.M. DELGADO, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, in *Istituzioni del Federalismo*, 2019, 662 che tuttavia invita a riflettere sull’opportunità che le nuove tecnologie possano delineare un modello di amministrazione *razionale*. Detto “stadio di evoluzione” è successivo all’Amministrazione digitale, che a sua volta rappresenta una trasformazione dell’Amministrazione tradizionale. È stato ricordato che l’amministrazione digitale è stata erroneamente considerata come *altra* tipologia rispetto all’amministrazione tradizionale, così determinando un approccio ostruzionistico

17. Sottolinea il passaggio da una dimensione dell’informatica di tipo documentale ad una metadocumentaria, M. D’ANGELOSANTE, *La consistenza del modello dell’amministrazione “invisibile” nell’età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione*, vol. IV, in L. FERRARA - D. SORACE (a cura di), *A 150 anni dall’unificazione amministrativa italiana. Studi*, Firenze University Press, Firenze, 2016, 155-180.

18. *Ibidem*.

19. Si esprime in questo senso M. MINERVA, *L’attività amministrativa in forma elettronica*, cit., 1301.

20. I.M. DELGADO, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, cit., 662; F. MERUSI, *Ragionevolezza e discrezionalità amministrativa*, Napoli, 2011, 51, afferma che il provvedimento discrezionale non possa essere ricondotto a razionalità, “visto che la ragionevolezza è informata da una logica opaca e che le molte manifestazioni della discrezionalità non rispondono ad una logica unitaria”. Sulla maggior razionalità del provvedimento amministrativo grazie all’uso delle nuove tecnologie si veda anche F. MOROLLO, *Documento elettronico fra e-government e artificial intelligence (AI)*, cit., 17.

21. Sul punto, si veda R. CAVALLO PERIN - I. ALBERTI, *Atti e procedimenti amministrativi digitali*, in *Diritto dell’Amministrazione pubblica digitale*, Giappichelli, 2020, 119-158.

Purtuttavia, ai benefici in termini di realizzazione del canone del buon andamento della pubblica amministrazione, l'utilizzo degli algoritmi nel procedimento decisionale reca seco talune criticità (§ par. 3) a cui, però, il diritto amministrativo può offrire le opportune soluzioni (§ par. 3.1).

2. *Gli algoritmi nel procedimento decisionale ovvero il provvedimento digitale*

Nella categoria delle decisioni amministrative automatizzate ovvero del provvedimento digitale, rientrano quelle decisioni, caratterizzate dalla circostanza che la tecnologia è stata inserita nel procedimento decisionale, in totale o parziale assenza del funzionario.

Le prime tipologie di decisioni amministrative automatizzate sono state rappresentate dai verbali di accertamento emessi a seguito della rilevazione automatica delle infrazioni al codice della strada per superamento del limite di velocità rilevate dall'autovelox oppure dalla creazione delle liste dei componenti le commissioni di abilitazione scientifica nazionale. In dette semplici ipotesi, l'automazione del procedimento decisionale è il risultato della programmazione *ex ante* dei limiti all'esercizio del potere amministrativo. La decisione amministrativa automatizzata dipende dunque dalla predeterminazione dei limiti che il sistema informatico – attraverso cui il potere amministrativo si esternalizza – automaticamente ed in autonomia rispetta.

L'evoluzione tecnologica sta affiancando ipotesi ove il dispositivo della decisione amministrativa è il risultato della capacità predittiva e del calcolo probabilistico che un algoritmo di apprendimento automatico ha ottenuto, a partire dall'elaborazione dei precedenti applicativi oppure dall'individuazione di ricorrenze che esso stesso individua nel *datalake* di riferimento.

In tali ipotesi - ancora più che nelle precedenti applicazioni di algoritmi deterministici- l'accuratezza richiesta nel definire l'algoritmo di apprendimento automatico è operazione ancor più essenziale, poiché maggiore è il grado di autonomia che l'algoritmo può raggiungere.

La sua capacità predittiva e l'accuratezza della decisione non è solo influenzata dalle regole logico-informatiche e dalla scelta delle variabili ossia delle classi di input, ma anche e soprattutto dalla qualità e dalla quantità del bacino di dati su cui l'algoritmo lavora. La precisione del modello è influenzata soprattutto dalla quantità e dalla varietà della casistica messa a disposizione dello stesso, ossia in base alla rilevanza e alla rappresentatività della banca dati e del *datalake* di riferimento. Da questi, gli algoritmi

di apprendimento automatico individuano autonomamente modelli di comportamento che la logica lineare causa-effetto non evidenzerebbe. Proprio il capovolgimento della logica decisionale usata rende talvolta imperscrutabile il meccanismo decisionale, sicché il provvedimento digitale appare talvolta privo di razionalità tra i presupposti di fatto e gli *output*, alimentando così le resistenze nell'introdurre nel settore pubblico dette tecnologie. L'opacità algoritmica è ciò che appare maggiormente in contrasto con i principi di trasparenza e di pubblicità che informano l'azione amministrativa.

2.1 Applicazioni algoritmiche nell'amministrazione pubblica

Il bacino di riferimento su cui può operare un algoritmo predittivo è rappresentato dalla casistica settoriale di riferimento previamente individuata, da tutti i nuovi casi che l'algoritmo ha individuato, nonché dal ricco flusso di dati²² proveniente da differenti fonti che alimentano la base di dati a disposizione (*big data*).

Così, per il settore degli appalti il bacino di riferimento può essere la casistica previamente individuata di tutte le ipotesi in cui il soccorso istruttorio è stato correttamente esercitato oppure no, dalché facendone derivare la lecita o illecita ammissione della domanda di gara.²³

Definito il bacino di analisi dei dati, l'algoritmo predittivo di apprendimento automatico dovrà essere addestrato a riconoscere autonomamente le ipotesi future, nonché a predire – al ricorrere di modelli di comportamento simili – quelle in cui il soccorso istruttorio è stato correttamente esercitato perché finalizzato a sopperire a *mere inesattezze*²⁴ oppure no, così realizzando un'inammissibile *modifica sostanziale della domanda*²⁵.

22. Il riferimento è ai dati che i sensori di rilevamento collegati in rete (*Internet of things*) raccolgono e inviano automaticamente alla *dashboard* operativo, che alimenta il bacino di dati a disposizione dell'amministrazione pubblica.

23. Per ulteriori esempi, si veda R. CAVALLO PERIN - I. ALBERTI, *Atti e procedimenti amministrativi digitali*, cit., 141.

24. Dall'analisi della giurisprudenza è stato evidenziato che rientrano in dette ipotesi la mancata allegazione di documenti quando il contenuto era comunque ricavabile da altri (Cons. St., sez. V, 27 marzo 2020, n. 2146; Cons. St., sez. V, 16 marzo 2020, n. 1881) o per l'inosservanza delle formalità di firma digitale quando cui sopperisce il sistema d'autenticazione delle persone al portale (Cons. St., sez. III, 19 marzo 2020, n. 1963).

25. Rientrano in quest'ipotesi la mancanza anche parziale dei requisiti di partecipazione (Cons. St., sez. V, 9 marzo 2020, n. 1671) o con un'inammissibile presentazione – postuma – del certificato di esecuzione di lavori (Cons. St., sez. V, 28 dicembre 2017, n. 6135).

L'algoritmo diviene capace di classificare autonomamente ogni nuovo caso nella casistica corrispondente (es. mere inesattezze documentali o modifica sostanziale della domanda), ne apprende le qualificazioni normative corrispondenti e le conseguenze giuridiche (es. ammissione od esclusione della domanda di partecipazione alla gara), così potendo elaborare l'idoneo dispositivo del provvedimento (es. ammissione od esclusione del concorrente alla gara).

L'aumento e la varietà dei casi a disposizione dell'algoritmo incrementano l'accuratezza della decisione algoritmica e perfezionano l'approssimazione probabilistica che determina il contenuto del provvedimento finale, così supportando l'amministrazione pubblica ad agire in modo ragionevole, entro i limiti di proporzionalità e ragionevolezza che il sindacato giurisdizionale ha individuato nei casi precedenti.

Le tecniche di data mining cominciano ad essere applicate soprattutto nell'attività di vigilanza ispettiva, tra cui l'Istituto Nazionale della Previdenza Sociale (INPS)²⁶ avviando un progetto che dall'analisi dei flussi assicurativi e contributivi, dei conti individuali e aziendali individua le anomalie nelle posizioni dei contribuenti e può attivare il potere di verifica e di accertamento conseguente.

L'Agenzia delle Entrate ha recentemente²⁷ cominciato ad orientare i controlli sulla situazione reddituale e sulle posizioni contributive, a partire dalle percentuali di rischio di evasione che l'algoritmo ha associato ad ogni singolo contribuente e che vengono assunti dalla stessa come indicatori utilizzabili per indirizzare l'attività di accertamento.

26. Circolare del 16 febbraio 2010 n. 23 in tema di "Funzione di accertamento e verifica amministrativa - Attuazione del nuovo modello organizzativo delle strutture territoriali di produzione previsto dalla circolare n. 102 del 12/08/2009" ha introdotto strumenti di intelligenza artificiale basati su tecniche di *data mining* al fine di orientare la propria attività ispettiva

27. L'art. 1, comma 682 della legge 27 dicembre 2019 n. 160 recante norme in materia di "*Bilancio di previsione dello Stato per l'anno finanziario 2020 e bilancio pluriennale per il triennio 2020-2022*" ha introdotto la possibilità per l'Agenzia delle Entrate di ricorrere agli strumenti informatici per indirizzare il potere di accertamento. Tuttavia, indicazioni operative sullo sfruttamento "razionale, completo e coerente" si ritrovano già nella circolare n. 16/E del 28 aprile 2016 dell'Agenzia delle entrate. In essa, infatti, è riconosciuto il valore conoscitivo che si può ricavare dalle banche dati, con la conseguenza che si riconosce l'importanza di arricchire tali banche dati con dati "qualitativamente corretti", allo scopo di "supportare in modo sempre più efficace l'attività di analisi del rischio e lo sviluppo e implementazione di nuovi percorsi di indagine e selezione" dei contribuenti da sottoporre a controllo.

Tale algoritmo opera su un grande database, chiamato SuperAnagrafe ove sono raccolti dati derivanti dai movimenti bancari in entrata e in uscita, dalle dichiarazioni dei redditi e da altre fonti di reddito come i dati delle fatture, i dati dei pagamenti e dei consumi, i prodotti finanziari, le obbligazioni, nonché di quelli provenienti dalle autorità fiscali estere: nel caso in cui, l'algoritmo dell'amministrazione rilevi un eccessivo discostamento tra le spese sostenute e quanto dichiarato attribuisce un indicatore di "evasione" e l'amministrazione può decidere di avviare i controlli sulla posizione del contribuente. Soltanto al ricorrere di questa circostanza, l'amministrazione può servirsi altresì delle informazioni derivanti dalle fonti aperte²⁸, tra cui *articoli di giornale, social network e siti web*²⁹.

Il supporto algoritmico rappresenta così un indubbio ausilio non solo per l'attività vincolata, ma anche per l'attività discrezionale della pubblica amministrazione, poiché in ragione dell'iniziale opera di classificazione dei fatti rilevanti e del conseguente *addestramento* dell'algoritmo, esso può apprendere le differenti conseguenze normative che ne derivano ed elaborare il provvedimento, a prescindere dalla natura vincolata o discrezionale del potere esercitato.

3. *L'imperscrutabilità della logica algoritmica*

I benefici dell'intelligenza artificiale a favore della pubblica amministrazione richiedono purtuttavia che il ricorso agli algoritmi nel settore pubblico rispetti il punto di equilibrio rappresentato dall'assenza di una eccessiva costrizione dei diritti dei privati.³⁰ Ciò che qui interessa³¹ è verifi-

28. La Circolare 1/2018 "*Manuale operativo in materia di contrasto all'evasione e alle frodi fiscali*" della Guardia di Finanza riconosce il valore dei dati derivanti da fonti aperte quali elementi informativi integranti i dati provenienti da alcune specifiche banche dati, cui è riconosciuto valore giuridico.

29. L'Agenzia delle Entrate, per esempio, sfrutta la conoscenza ricavabile dal web (intesa come *fonte aperta*) per recuperare informazioni sulle caratteristiche degli immobili compravenduti e sulla zona in cui si trovano e la confronta con le quotazioni rilevate dall'Osservatorio del mercato immobiliare, per verificare che non sia stato sottodichiarato il valore del prezzo di acquisto di un immobile; in particolare.

30. Riconosce questa necessità, I. MARTÍN DELGADO, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, in *Istituzioni del Federalismo*, 2019, 649.

31. Date le esigenze di economia del lavoro non si intende trattare delle criticità connesse all'algoritmizzabilità del linguaggio giuridico, caratterizzato da antinomie e da concetti giuridici indeterminati; per alcune riflessioni sulla normalizzazione del linguaggio giuridico

care l'impatto di dette tecnologie nella dinamica procedimentale e quindi sull'esercizio della funzione pubblica, affinché si possa riflettere e superare il determinismo tecnologico rappresentato dalla scarsa conoscenza di come il meccanismo decisionale funzioni³².

L'esplorazione delle ragioni giuridiche di fatto e di diritto e cioè la motivazione del dispositivo del provvedimento amministrativo è uno degli istituti di maggiore garanzia del privato di fronte all'esercizio del potere amministrativo. In ragione di ciò, l'esercizio discrezionale del potere amministrativo è limitato.

L'istituto trova la sua ragion d'essere nella scelta istituzionale compiuta affinché la dichiarazione di volontà che ne deriva sia espressione di un processo intellegibile che mostri la ragionevolezza esistente tra il ricorrere dei presupposti di fatto indicati dalla norma attributiva del potere e la scelta in concreto compiuta che ha contemperato gli interessi privati con l'interesse pubblico.

Si è detto che, in ragione della logica di tipo probabilistica, anziché logico-deduttiva, l'uso dell'algoritmo di apprendimento automatico nel procedimento decisionale pone proprio alcune criticità in termini di *spiegabilità* del procedimento logico-deduttivo sotteso alla decisione.

L'imperscrutabilità dei percorsi argomentativi degli algoritmi di apprendimento automatico che ne deriva contrasta con alcuni fra i principi cardini dell'azione amministrativa, prima fra tutti la trasparenza, nondimeno l'imparzialità dell'azione amministrativa.

Il rischio che la decisione amministrativa automatizzata possa perpetrare delle ineguaglianze di trattamento è possibilità non remota, su cui anche il Consiglio di Stato nella sentenza 13 dicembre 2019, n. 8765 si è espresso nell'affermare che l'uso dell'intelligenza artificiale nel settore pubblico debba sottostare a taluni principi, in particolare la non discriminazione algoritmica e la trasparenza rinforzata³³.

ai fini della sua formalizzazione in linguaggio informatico si veda V. FROSINI, *Telematica e informatica giuridica*, voce in *Enc. dir.*, Milano, 1992 e neanche sul tema dell'imputabilità del provvedimento digitale, poiché si condivide l'idea che esso sia pur sempre imputabile al funzionario secondo le comuni regole della teorica dell'organo amministrativo, così I. MARTÍN DELGADO, *La riforma dell'amministrazione digitale: Un'opportunità per ripensare la pubblica Amministrazione*, in S. CIVITARESE MATTEUCCI - L. TORCHIA, *La tecnificazione*, cit., 146-148; anche G. DUNI, *L'amministrazione digitale. Il diritto amministrativo nella evoluzione telematica*, cit., 77, che infatti non ipotizza l'attribuzione di volontà alla macchina.

32. M. D'ANGELOSANTE, *La consistenza del modello dell'amministrazione "invisibile" nell'età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni*, cit., 157 e s.

33. Da ultimo si veda E. CARLONI, *I principi della legalità algoritmica*, in *Dir. Amm.*, 2/2020.

Si ritiene pertanto che la decisione algoritmica debba essere oggetto di un contraddittorio tra i partecipi al procedimento, così superando le criticità che l'uso di algoritmi di apprendimento automatico solleva.

3.1 *Il connubio tra capacità analitica algoritmica e partecipazione procedimentale*

Del carattere predittivo degli algoritmi si è detto per evidenziare le potenzialità in termini di rilevazione della realtà, tuttavia preme ora sottolineare come la loro introduzione nel procedimento decisionale pubblico ha fatto emergere una differente visione del fatto rilevato. La capacità analitica di cui sono capaci gli algoritmi valorizza il *fatto* non solo come elemento in sé considerato, cui giustapporre una scelta d'interesse pubblico, ma anche come dato organizzato. Quest'ultima accezione permette l'individuazione dei soggetti che sono portatori di interessi (pubblici, privati o diffusi), che permettono di entrare nella dinamica procedimentale in quanto requisiti di legittimazione al procedimento (art. 9, l. 7 agosto 1990, n. 241), sicché il responsabile deve offrire una valutazione preliminare agli stessi (art. 6, c. 1, lett. a), l. 7 agosto 1990, n. 241), utile a consentire la partecipazione sin dalla fase istruttoria.

Sotto questo profilo, la partecipazione non è istituto che mira a erodere il potere pubblico, quanto invece mezzo per apportare nel procedimento interessi pubblici o privati che non sarebbero diversamente conoscibili dal responsabile del procedimento e che impedisce all'amministrazione pubblica di fare scelte unilaterali e in via esclusiva e che anzi, la vincola alla composizione degli interessi coinvolti nel perseguimento dell'interesse pubblico³⁴.

Poiché si ritiene che anche il provvedimento digitale, sia "manifestazione unilaterale di volontà, giudizio, conoscenza di una pubblica amministrazione, nell'esercizio delle sue funzioni, avente rilevanza esterna, vale a dire capacità di produrre effetti giuridici nei confronti dei terzi"³⁵ e atto d'autorità³⁶ al pari di un atto amministrativo tradizionalmente ela-

34. R. CAVALLO PERIN, *Atto autoritativo e diritto amministrativo*, in *Al di là del nesso autorità / libertà: tra legge e amministrazione*, a cura di S. Perongini, Giappichelli Editore, Torino, 2017, 23-24.

35. M.S. GIANNINI, *Atto amministrativo*, voce in *Enc. Dir.*, IV, Milano, 1959, 157, ma anche G. ZANOBINI, *Corso di diritto amministrativo*, Milano, 1954.

36. Sul tema si veda R. CAVALLO PERIN, *Atto autoritativo e diritto amministrativo*, in *Al di là del nesso autorità / libertà: tra legge e amministrazione*, a cura di S. Perongini, Giappichelli Editore, Torino, 2017, 17-30, che evidenzia la duplice finalità istituzionale dello stesso, sia

borato, l'equiparazione permette la diretta applicabilità di tutti gli istituti del diritto amministrativo.

In ragione di ciò, a prescindere dagli strumenti tecnologici utilizzati dall'amministrazione pubblica, il potere amministrativo dovrà pur sempre essere esercitato nel perseguimento dell'interesse pubblico che la legge ha attribuito alla sua cura. La certezza dell'amministrare pubblico e le garanzie previste nei confronti dei privati, pertanto, non subiscono alcuna limitazione.

Gli algoritmi di apprendimento automatico o altri strumenti tecnologici usati dall'amministrazione pubblica nella fase istruttoria sono nient'altro che strumenti d'ausilio per il perseguimento dell'interesse pubblico: è pur sempre questo a rappresentare il limite all'esercizio del potere, nonché la ragione giustificativa della partecipazione degli interessati. Infatti, ricondurre il provvedimento digitale alla teorica dell'atto autoritativo mostra che la *forma* tecnologica non incide sulle forme di tutela che l'ordinamento riconosce – quand'anche in via interpretativa – ai privati. Ciò vuol dire che l'introduzione dell'intelligenza artificiale nell'esercizio del potere amministrativo, non modifica in alcun modo la sostanza dello stesso e cioè la sua *funzionalizzazione* verso l'interesse pubblico, che è e rimane causa del provvedimento digitale.

Pertanto, attraverso il potere riconosciuto al responsabile del procedimento di comunicare le informazioni necessarie ai notiziati dalla comunicazione di avvio del procedimento (artt. 7 e 8, l. 241/1990, cit.), questi possono intervenire. L'oggetto della comunicazione sarà la decisione algoritmica elaborata raggiunta al termine della fase istruttoria che così può essere sottoposta al contraddittorio procedimentale.

Il confronto tra amministrazione pubblica e privati sulla decisione algoritmica permetterà alla prima di conoscere compiutamente le particolarità del caso concreto che le saranno rappresentate dai privati: a seguito del vaglio dell'amministrazione pubblica, essa potrà recepirle e discostarsi dalla decisione algoritmica, oppure rigettarle. Se in detto ultimo caso, è già previsto che il rigetto delle osservazioni debba essere motivato (artt. 3 e 10-*bis*, l. 7 agosto 1990, n. 241), nell'altra ipotesi, invece, si ritiene che sia la natura giuridica di atto generale della decisione algoritmica a comportare la motivazione in caso di discostamento dalla stessa.

come atto che garantisce uniformità di governo, sia come prodotto formale dell'attività giuridica delle amministrazioni pubbliche.

4. *Riflessioni conclusive*

Le nuove tecnologie *dentro* l'amministrazione pubblica hanno avviato un cambiamento dell'organizzazione verso logiche che superano l'autoreferenzialità delle singole unità organizzative e aprono verso nuove logiche³⁷ che fanno della trasparenza e della comunicazione gli "strumenti essenziali in funzione della semplificazione amministrativa, della modernizzazione e della condivisione"³⁸.

Attraverso l'uso delle nuove tecnologie, infatti l'istituto della partecipazione non è solo più diritto legato ad un procedimento in corso, di cui il privato intende esercitare i relativi diritti di accesso agli atti e di presentazione delle memorie (art. 10, l. 7 agosto 1990, n. 241), ma anche diritto di partecipazione ai processi decisionali, per l'esercizio dei diritti politici e civili, attraverso le forme della consultazione preventiva per via telematica sugli schemi di atto da adottare (art. 9, d.lgs. 14 marzo 2005, n. 82).

L'introduzione delle nuove tecnologie nell'organizzazione amministrativa ha accentuato la facilità di far circolare l'informazione e di aumentare la comunicazione tra amministrazioni pubbliche e verso i privati avviando "un vero e proprio interscambio comunicativo"³⁹.

L'aumentato bisogno di trasparenza sempre più preteso dai privati ha perciò avviato un modello organizzativo che ha trasformato l'amministrazione in un'amministrazione aperta (*open government*), basata sui principi della trasparenza e della partecipazione.

Così, è stata riconosciuta la rilevanza pubblicistica dell'algoritmo nel procedimento amministrativo in forza della sua riconducibilità alla nozione di documento amministrativo e dello svolgimento di attività amministrativa sostanziale, in forza della quale le posizioni soggettive sono garantite con il riconoscimento dei diritti di partecipazione e di accesso anche al codice sorgente del software che esegue l'algoritmo (T.A.R. Lazio, sez. III-*bis*, 27 marzo 2016, n. 3769).

Si può concludere che l'introduzione degli algoritmi nel procedimento decisionale pubblico ha avuto una duplice funzione: da un lato rinvigorire

37. L'analisi dei pregressi applicativi potrebbe rivelare "l'opacità di un esercizio del potere amministrativo o giurisdizionale in determinati settori o periodi", come nota R. CAVALLO PERIN, *I Cambiamenti dell'amministrazione pubblica italiana tra rivoluzione liberale, repubblicana e l'avvento dell'intelligenza artificiale*, in corso di pubblicazione.

38. P. PIRAS, *Organizzazione, tecnologie e nuovi diritti*, in *Dir. informatica*, fasc. 3, 2005, 595.

39. *Ibidem*.

il processo di digitalizzazione dell'amministrazione digitale, dall'altro cogliere l'innovazione tecnologica nel diritto amministrativo per individuare nuovi limiti del potere amministrativo.

A tal proposito, la disciplina dell'Unione europea, seppur con riferimento al trattamento dei dati personali, offre un quadro normativo valevole anche per le pubbliche amministrazioni che legittima le decisioni automatizzate e nondimeno quelle *esclusivamente* automatizzate (art. 22, Regolamento 2016/679/UE), al ricorrere di talune condizioni, quali il consenso dell'interessato (art. 6, par. 1, lett. a, Regolamento 2016/679/UE), la necessità di eseguire un contratto (art. 6, par. 1, lett. b, Regolamento 2016/679/UE), l'adempimento da parte del titolare del trattamento di un obbligo legale (art. 6, par. 1, lett. c, Regolamento 2016/679/UE), la salvaguardia di interessi vitali che riguardano l'interessato o altra persona fisica (art. 6, par. 1, lett. d, Regolamento 2016/679/UE), l'esecuzione di un compito per ragioni di interesse pubblico o che sia connesso all'esercizio di pubblici poteri ad esso assegnati (art. 6, par. 1, lett. e, Regolamento 2016/679/UE) o il perseguimento di un legittimo interesse del titolare stesso (art. 6, par. 1, lett. f, Regolamento 2016/679/UE).

L'autorizzazione da parte di una norma speciale del diritto dell'Unione o di uno Stato membro, il previo esplicito consenso dell'interessato o la necessità di concludere o eseguire un contratto sono eccezioni ove l'esclusività della decisione algoritmica è contemperata dal diritto dell'interessato di: *richiedere l'intervento umano, di far valere le proprie ragioni, di contestare la decisione* (art. 22, § 3, Regolamento 2016/679/UE), fatta salva in ogni caso la tutela «dei diritti, delle libertà e dei legittimi interessi dell'interessato» (art. 22, § 2, lett. b, Regolamento 2016/679/UE).

Tali eccezioni sono ipotesi conosciute dall'ordinamento italiano e che ben si conciliano con la disciplina garantistica assicurata dal diritto amministrativo, in particolare dalla disciplina sul procedimento amministrativo (l. 7 agosto 1990, n. 241).

La disciplina europea sul trattamento dei dati rafforza dunque i diritti di partecipazione al procedimento tradizionalmente inteso o svolto per tramite delle nuove tecnologie e che può perciò restituire effettività all'amministrazione pubblica⁴⁰.

40. Sul punto si veda anche D.-U. GALETTA, *Digitalizzazione e diritto ad una buona amministrazione (Il procedimento amministrativo, fra diritto UE e tecnologie ICT)*, in *Diritto dell'Amministrazione pubblica digitale*, Giappichelli, 2020, 85-117.

Postfazione

Riunire a maggio 2019 molti studiosi a Torino per parlare di *big data* e d'intelligenza artificiale nella pubblica amministrazione, sembrava allora un'impresa ambiziosa e fors'anche singolare, certo anche come occasione per festeggiare il mio sessantesimo compleanno. Un convegno all'epoca eccentrico per il luogo, l'oggetto e la natura interdisciplinare del colloquio, che assumeva di per sé un carattere dirompente, già solo per avere partecipato ad un dibattito che – quanto meno all'inizio - ciascuno ha visto come non proprio.

Stupisce in particolare vedere che nei diciotto mesi successivi i contributi sulla digitalizzazione della pubblica amministrazione sono stati numerosi, un profluvio di attenzioni, che hanno perciò oggettivamente omaggiato i primi coraggiosi autori sull'informatizzazione e i *big data* comparsi nella seconda metà del secolo scorso. Ben augurante sembra essere stato pensare a Torino come nuova capitale dell'intelligenza artificiale, che potrebbe accogliere a breve l'Istituto Italiano per l'Intelligenza Artificiale (come deciso dal Consiglio dei ministri il 3 settembre 2020).

Sebbene di rivoluzione informatica, di società dell'informazione e della comunicazione se ne fosse sentito parlare da anni, l'intelligenza artificiale rimaneva perlopiù un tema che restava di dominio esclusivo degli informatici o degli statistici, mentre quello dell'analisi dei *big data* e della qualità dei dati appariva piuttosto raccogliere l'interesse del mondo economico, con una crescente attenzione verso la tutela della concorrenza e della privacy.

Fuori da tali ambiti di collaborazione, il giurista che studia il funzionamento e l'organizzazione della pubblica amministrazione ha dovuto comprendere i nuovi strumenti tecnologici con la necessaria e quanto mai utile collaborazione degli informatici o degli statistici, acquisendo un sapere nuovo e contribuendo così alla realizzazione di prototipi o di veri e propri prodotti commerciali, capaci di contenere in sé – sin dall'origine – le peculiarità di quel diritto che è applicato alla pubblica amministrazione, con i suoi principi, limiti e possibilità.

Ci è stato detto in questi anni che l'uso della tecnologia e delle connessioni internet possono offrire ai cittadini una città sostenibile, intelligente, perché la si vuole più capace di soddisfare i bisogni dei molti e differenti gruppi che fruiscono della città, organizzata a rispondere alle esigenze spazio-temporali e alle scelte che sono per essi preferenziali, cioè di chi la vive periodicamente, selezionando tempi e modi di una sua pacifica e facile fruizione, modo di vivere *smart* che la recente pandemia ha esteso ben oltre la città e reso universale.

Ora emerge che tali comportamenti di persone o gruppi non sono comportamenti casuali, ma disciplinati, in cui forti sono i caratteri del diritto non scritto e dello *jus involontarium*. Una potenza ordinante del diritto che ora giova della potenza conoscitiva dell'intelligenza artificiale, uno strumento che ci rivela il molteplice, che rivela il reale nella sua forza organizzata di persone, che è perciò solo istituzionale, per ciò solo diritto obiettivo. Una potenza conoscitiva e al tempo stesso ordinante che svela nuove prospettive per l'amministrazione pubblica, nuove questioni che debbono essere superate, o pericoli che debbono essere sventati.

L'organizzazione di questo convegno ha voluto raccogliere questa attenzione, cercando di sospingere il dibattito verso considerazioni che cerchino di comprendere ancora una volta se l'innovazione tecnologica rechi con sé, o sia levatrice, di una vera e propria rivoluzione sociale e istituzionale.

Torino 18 dicembre 2020

Il curatore

Bibliografia

- AA.VV., *Smart cities e amministrazioni intelligenti*, *Istituzioni del federalismo*, n. 4, 2015.
- AA.VV., *L'amministrazione di fronte alla sfida digitale*, *Istituzioni del federalismo*, n. 3, 2019.
- AA.VV., *Ville intelligente, ville démocratique?*, Actes du colloque organisé le 13 février 2013 par Berger-Levrault et la Chaire Mutations de l'action publique et du droit public (MADP) de Sciences Po, Berger-Levrault, Paris, 2014.
- M.L. ABATE - K.V. DIEGERT *et al.*, *A Hierarchical Approach to Improving Data Quality*, in *Data Quality*, 1, 1998.
- AGENZIA PER L'ITALIA DIGITALE, *Piano triennale per l'informatica nelle pubbliche amministrazioni*, in agid.gov.it.
- AGENZIA PER L'ITALIA DIGITALE, *Libro Bianco sull'Intelligenza Artificiale al servizio del cittadino - Versione 1.0* (a cura della Task force sull'Intelligenza Artificiale dell'Agenzia per l'Italia Digitale), in agid.gov.it, 2018.
- AGENZIA PER L'ITALIA DIGITALE, *Linee guida nazionali per la valorizzazione del patrimonio informativo pubblico*, in agid.gov.it, 2017.
- AGENZIA PER L'ITALIA DIGITALE, *Strategia per la crescita digitale 2014-2020*, in agid.gov.it.
- AGENZIA PER L'ITALIA DIGITALE, *Verso il sistema nazionale di e-Government, Linee strategiche*, Roma, 2007 (già in innovazione.gov.it).
- C.C. AGGARWAL, *Neural Networks and Deep Learning: A Textbook*, Springer, 2018.
- C.C. AGGARWAL, *Recommender Systems – The Textbook*, Springer, 2016.
- M. J. AHN, *Whither E-Government? Web 2.0 and the Future of E-Government, Web 2.0 Technologies and Democratic Governance*, in *Pub. Adm. Inf. Tech.*, 2012.
- L. AL-HAKIMCHE, *Global E-Government: Theory, Applications and Benchmarking*, Hershey, 2007.
- A. ALEMANNI, *Unpacking the principle of openness in EU Law. Transparency, participation and democracy*, in *Eur. Law Rev.*, 2014.
- S. ALIPRANDI, *Open data: serve una norma più chiara. L'open by default non funziona*, 2018, in techeconomy.it.
- U. ALLEGRETTI, *Democrazia partecipativa*, in *Enc. dir., Annali*, V, Milano, 2011.
- G. AMATO, *Autorità semi-indipendenti e autorità di garanzia*, in *Riv. trim. dir. pub.*, 1997.
- G. ARENA, *Trasparenza amministrativa*, *Diz. dir. pub.*, a cura di S. Cassese, VI, Milano, 2006.

- G. ARENA, *Introduzione all'amministrazione condivisa*, in *Studi parlamentari e di politica costituzionale*, 1997.
- G. ARENA, *La "Legge sul diritto alla informazione" e la pubblicità degli atti dell'Amministrazione negli Stati Uniti*, in *Pol. dir.*, 1978.
- J.-B. AUBY, *Il diritto amministrativo di fronte alle sfide digitali*, in *Istituzioni del Federalismo*, 2019.
- J.-B. AUBY, *Le droit administratif face aux défis du numérique*, in *Actualité Juridique Droit Administratif*, 23 aprile 2018.
- AGCM - AGCOM - GARANTE PRIVACY, *Indagine conoscitiva sui Big data*, 2019, in www.agcom.it.
- AUTORITÀ PER L'INFORMATICA NELLA PUBBLICA AMMINISTRAZIONE - AIPA, *Piano triennale per l'informatica della pubblica amministrazione, 2002-2004*, già in www.cnipa.gov.it.
- G. AVANZINI, *Decisioni amministrative e algoritmi informatici*, Napoli, Editoriale Scientifica, 2019.
- D. AZZOLINI - M. SISTI, *Evidence-Based Policy e Attività Legislativa. Cosa c'è di nuovo?, Federalismi*, 2019.
- S. BAACK, *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, *Big Data and society*, 2015.
- S. BAACK, *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, *Big Data and society*, 2015.
- F. BANNISTER - R. CONNOLLY, *The Trouble with Transparency: A Critical Review of Openness in e-Government, in Policy and Internet*, vol. 3, 2011.
- A. BARTOLINI - A. PIOGGIA (a cura di), *Cittadinanze amministrative*, Firenze, 2016.
- M. BASALDELLA - G. SERRA - C. TASSO, *The Distiller framework: Current state and future challenges*, in *Digital Libraries and Multimedia Archives. 14th Italian Research Conference on Digital Libraries* (a cura di G. Serra, C. Tasso), CCIS, Springer Verlag, 2018.
- J. BATES, *The strategic importance of information policy for the contemporary neoliberal state: the case of Open Government Data in the United Kingdom*, in *Gov. Inf. Quart.*, 2014.
- C. BATINI - M. SCANNAPIECO, *Qualità dei dati. Concetti, metodi, tecniche*, Milano, 2008.
- M. BELLETTI, *I "livelli essenziali delle prestazioni concernenti i diritti civili e sociali..." alla prova della giurisprudenza costituzionale. Alla ricerca del parametro plausibile...*, in *Istituzioni del Federalismo*, n. 3-4, 2003.
- S.M. BENJAMIN, *Evaluating E-Rulemaking: Public Participation and Political Institutions*, in *Duke L. J.*, 2005-2006.
- B. BENNATO, *Il computer come macroscopio. Big data e approccio computazionale per comprendere i cambiamenti sociali e culturali*, Milano, Angeli, 2015.
- D. BERLINER - A. INGRAMS - S.J. PIOTROWSKI, *The Future of FOIA in an Open Government Agenda for Freedom of Information Policy and Implementation*, in *Vill. L. Rev.*, 2018.
- R. BIFULCO, «Democrazia deliberativa», in *Enc. dir., Annali*, IV, Milano, 2011.

- C. BIRCHALL, *Interrupting Trasparency, Transparency, Society and Subjectivity. Critical Perspectives*, a cura di E. Alloa e D. Thomä, Basingstoke, 2018.
- P. J. BIRKINSHAW - A. HICKS, *The law and public information on UK. quality, access and re-use*, in *Dir. Pub.*, 2007.
- L. BLOMGREN BINGHAM, *The next generation of administrative law: building the legal infrastructure for collaborative governance*, in *Wis. L. Rev.*, 2010.
- L. BOBBIO, *Dilemmi della democrazia partecipativa*, *Dem. dir.*, 2007.
- S. BOLGHERINI - M. CASULA - M. MAROTTA, *Il dilemma del riordino*, il Mulino, Bologna, 2018.
- D. BOLLIER, *The promise and peril of Big data*, Washington, 2010.
- A. BONOMI - R. MASIERO, *Dalla smart city alla smart land*, Marsilio, Venezia, 2014.
- A. BONOMO, *Informazione e Pubblica Amministrazione. Dall'accesso ai documenti alla disponibilità delle informazioni*, Bari, 2012.
- E. BORGONOV, *Principi e sistemi aziendali per le amministrazioni pubbliche*, Egea, Milano, 2002.
- P. BORGNA (a cura di), *Manuale di sociologia*, Torino, Utet, 2008
- R. BORRUSO, *Computer e diritto*, Milano, 1978, nonché V. FROSINI, *Riflessi sociali dell'informaticizzazione*, in *Scritti in onore di M.S. Giannini*, Milano, 1988.
- C. BOTTARI (a cura di) *I livelli essenziali delle prestazioni sociali e sanitarie*, Santarcangelo di Romagna, 2014.
- R. BRAUNEIS - E.P. GOODMAN, *Algorithmic transparency for the smart city*, in *Yale Journal of Law & Technology*, n. 20, 2017.
- G. BRONZINI, *La sentenza n. 20/2019 della Corte costituzionale italiana: verso un riavvicinamento all'orientamento della Corte di giustizia?*, in *Questione Giustizia*, 4 marzo 2019.
- B. BUGARIC, *Openness and transparency in public administration: challenges for public law*, in *Wiscon. Intern. law journ.*, 2004.
- S. CALZOLAIO, "Digital (and privacy) by default". *L'identità costituzionale della amministrazione digitale*, in *Giornale di storia Costituzionale*, 2016.
- Q. CAMERLENGO, *Ritratto costituzionale della legittimazione politica*, in *Rivista AIC*, 2016.
- G. CAMMAROTA, *Digitalizzazione dell'amministrazione e servizi pubblici on line. Lineamenti del disegno normativo*, in *Dir. Pubbl.*, n. 3, 2019.
- G. CAMMAROTA, *Servizi in rete della pubblica amministrazione*, in *Digesto delle discipline pubblicistiche*, V Aggiornamento, 2012, Torino.
- G. CAMMAROTA, *Erogazione on line dei servizi pubblici e teleprocedure amministrative*, in *Dir. Pubbl.*, n. 3, 2001.
- G. CAMMAROTA, *Servizi pubblici online e partecipazione migliorativa*, in S. Civitarese Matteucci, L. Torchia (a cura di), *La tecnificazione*, Firenze University Press, Firenze, 2016.
- M. CAMELLI, *Istituzioni pubbliche e sviluppo economico locale*, in R. Bin, L. Coen (a cura di), *Poteri pubblici e sviluppo economico locale*, Padova, 2009.
- M. CAPORALE, *L'attuazione delle smart cities. Competenze e coordinamento tra livelli di governo*, in *Istituzioni del Federalismo*, 4, 2015.

- M. CAPORALE, *El régimen de las smart cities en Italia.*, in F. GARCÍA RUBIO (a cura di), *Las nuevas perspectivas de la ordenación urbanística y del paisaje: smart cities y rehabilitación. Una perspectiva hispano-italiana*, Madrid, Fundación Democracia y Gobierno Local, 2017.
- M. CAPORALE, *Aspetti particolari del trattamento di dati personali in ambito pubblico: accesso ai documenti amministrativi e sistemi di identificazione personale* in V. CUFFARO - R. D'ORAZIO - V. RICCIUTO (a cura di), *Protezione e libera circolazione dei dati personali nel diritto europeo. il regolamento generale 2016/679 (e le direttive 2016/680 e 2016/681 sul trattamento dei dati in ambito penalistico)*, Giappichelli, 2018.
- M. CAPORALE, *Administracion local inclusiva y ciudadanía digital*, in M.L. GOMÉZ JIMENÉZ (a cura di), *Estudios sobre el Estatuto Jurídico de las Ciudades Inteligentes*, Atelier, Barcelona, 2018.
- A. CARAGLIU - C. DEL BO - P. NIJKAMP, *Smart cities in Europe*, in *Journal of Urban Technology*, n. 18, 2011.
- F. CARDARELLI, *Il codice dell'amministrazione digitale*, in *Libro dell'Anno del diritto 2017*, Roma, 2017.
- D. CARDON, *Che cosa sognano gli algoritmi. Le nostre vite al tempo dei Big data*, Milano, Mondadori, 2016.
- E. CARLONI, *I principi della legalità algoritmica*, in *Dir. Amm.*, n. 2, 2020.
- E. CARLONI, *Algoritmi su carta. Politiche di digitalizzazione e trasformazione digitale delle amministrazioni*, in *Dir. Pub.*, 2019.
- E. CARLONI, *Digitalizzazione pubblica e differenziazione regionale*, in *Giornale Dir. Amm.*, n. 6, 2018.
- E. CARLONI, *Città intelligenti e agenda urbana: le città del futuro, il futuro delle città*, in *Munus*, n. 2, 2016.
- E. CARLONI, *Il nuovo diritto di accesso generalizzato e la persistente centralità degli obblighi di pubblicazione*, *Dir. amm.*
- E. CARLONI, *Tendenze recenti e nuovi principi della digitalizzazione pubblica*, in *Giornale Dir. Amm.*, n. 2, 2015.
- E. CARLONI, *L'amministrazione aperta*, Rimini, 2014.
- E. CARLONI, *I principi del codice della trasparenza (artt. 1, commi 1 e 2, 2, 6)*, in B. Ponti (a cura di), *La trasparenza amministrativa dopo il d.lgs. 14 marzo 2013*, n. 33, Rimini, 2013.
- E. CARLONI, *Le verità amministrative. L'attività conoscitiva tra procedimento e processo*, Milano, Giuffrè, 2011.
- E. CARLONI, *La riforma del Codice dell'amministrazione digitale*, in *Giorn. dir. amm.*, n. 5, 2011.
- E. CARLONI, *La "casa di vetro" e le riforme. Modelli e paradossi della trasparenza amministrativa*, *Dir. pub.*, 2009.
- E. CARLONI, *La qualità delle informazioni pubbliche. L'esperienza italiana nella prospettiva comparata*, in *Rivista trimestrale di diritto pubblico*, 2009.
- E. CARLONI, *La qualità dei dati pubblici*, in B. Ponti (a cura di), *Il regime dei dati pubblici. Esperienze europee e ordinamento nazionale*, Rimini, 2008.

- E. CARLONI, *Nuove prospettive della trasparenza amministrativa: dall'accesso ai documenti alla disponibilità delle informazioni*, in *Dir. pub.*, 2005.
- E. CARLONI, *Codice dell'amministrazione digitale: commento al D.Lgs. 7 marzo 2005*, Rimini, 2005.
- E. CARLONI, *La qualità delle informazioni diffuse dalle amministrazioni federali negli Stati Uniti*, in *Giornale di diritto amministrativo*, 2002.
- B. CAROTTI, *L'amministrazione digitale: le sfide culturali e politiche del nuovo codice*, in *Giornale di Diritto Amministrativo*, n. 1, 2017.
- C. CARPINETO, *Open data PA, bilanciare rischi privacy e benefici: un nuovo approccio*, 2018, in *agendadigitale.eu*.
- P. CARROZZA, *Le Province della post-modernità: la città territoriale*, in *www.federalismi.it*, 31 gennaio 2018.
- P. CARROZZA, *Governo e Amministrazione*, in *Diritto costituzionale comparato*, a cura di P. Carrozza, A. Di Giovine e G.F. Ferrari, Laterza, Roma-Bari, 2017.
- G. CARULLO, *Gestione, fruizione e diffusione dei dati dell'amministrazione digitale e funzione amministrativa*, Giappichelli, Torino, 2017.
- L. CASINI, *L'inchiesta pubblica. Analisi comparata*, *Riv. trim. dir. pub.*, 2007.
- W. CASTELNUOVO - G. MISURACA - A. SAVOLDELLI, *Citizen's Engagement and Value Co-production in Smart and Sustainable Cities*, *Conference Paper, International Conference on Public Policies Proceedings*, 2015.
- S. CASSESE, *Che cosa resta dell'amministrazione pubblica?*, in *Riv. trim. dir. pub.*, n. 1, 2019.
- S. CASSESE, *La partecipazione dei privati alle decisioni pubbliche. saggio di diritto comparato*, *Riv. trim. dir. pub.*, 2007.
- S. CASSESE, *La rete come figura organizzativa della collaborazione*, in A. PREDIERI - M. MORISI (a cura di), *L'Europa delle reti*, Giappichelli, Torino, 2001.
- S. CASSESE - C. FRANCHINI (a cura di), *I garanti delle regole. Le autorità amministrative indipendenti*, Bologna, 1996.
- R. CAVALLO PERIN - I. ALBERTI, *Atti e procedimenti amministrativi digitali*, in *Diritto dell'Amministrazione pubblica digitale*, Giappichelli, 2020.
- R. CAVALLO PERIN, *I Cambiamenti dell'amministrazione pubblica italiana tra rivoluzione liberale, repubblicana e l'avvento dell'intelligenza artificiale*, in corso di pubblicazione.
- R. CAVALLO PERIN, *Atto autoritativo e diritto amministrativo*, in *Al di là del nesso autorità / libertà: tra legge e amministrazione*, a cura di S. Perongini, Torino, Giappichelli Editore, 2017.
- R. CAVALLO PERIN, *L'organizzazione delle pubbliche amministrazioni e l'integrazione europea*, in L. FERRARA - D. SORACE (a cura di), *A 150 anni dall'unificazione amministrativa europea*, vol. I, Firenze University Press, 2016.
- R. CAVALLO PERIN, *La configurazione della cittadinanza amministrativa*, in *Dir. amm.*, 2004.
- A. CAUDURO, *Il diritto di accesso a dati e documenti amministrativi come promozione della partecipazione: un'innovazione limitata*, *Dir. Amm.*, 2017.

- CERRILLO I MARTÍNEZ, AGUSTÍ (2019). *How can we open the black box of public administration? Transparency and accountability in the use of algorithms*. *Revista Catalana de Dret Públic*, (58), 13-28. <https://doi.org/10.2436/rcdp.i58.2019.3277>.
- V. CERULLI IRELLI, «Sussidiarietà (dir. amm.)», in *Enc. giur.*, XII, Roma, 2003.
- S. CHINELLATO, *Sperimentazione di una Metodologia per lo Sviluppo di Sistemi Basati su Conoscenza - Un'Applicazione in Campo Zootecnico*, Tesi di Laurea A.A. 1994-95, Università di Udine.
- A. CHADWICK, *Web 2.0: New Challenges for the Study of E-Democracy in an Era of Informational Exuberance*, in *A Journal of Law and Policy*, 2008-2009.
- E. CHELL, *Le autorità amministrative indipendenti nella forma di governo*, Associazione per gli studi e le ricerche parlamentari. *Quaderno n. 11*, Torino, 2000.
- R. CHIEPPA, *La trasparenza come regola della pubblica amministrazione*, *Dir. econ.*, 1994.
- S. CHIGNARD, *A brief history of Open data*, *parisinnovationreview.com*, 2014.
- P. CHIRULLI, *Istituti di partecipazione e semplificazione amministrativa, Italiadecide. Rapporto 2015. Semplificare è possibile: come le pubbliche amministrazioni potrebbero fare pace con le imprese*, a cura di C. Pinelli, 2015.
- T. CHRISTENSEN - P. LÆGREID, *Post New Public Management Reforms: Whole of Government Approaches as a New Trend*, in S. GROENEVELD - S. VAN DE WALLE (a cura di), *New Steering Concepts in Public Management (Research in Public Policy Analysis and Management, Volume 21)*, Emerald Group Publishing Limited, Bingley, 2011.
- F. CINTIOLI, *Il sindacato del giudice amministrativo sulle linee guida, sui pareri del c.d. precontenzioso e sulle raccomandazioni di ANAC*, in *Dir. Proc. Amm.*, 2017.
- S. CIVITARESE MATTEUCCI, «Umano troppo umano». *Decisioni amministrative automatizzate e principio di legalità*, in *Diritto Pubblico*, 2019.
- S. CIVITARESE MATTEUCCI - L. TORCHIA, *La tecnificazione dell'amministrazione*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione*, Firenze University Press, Firenze, 2016.
- A. CLARKE - H. MARGETTS, *Governments and citizens getting to know each: other? Open, closed, and Big Data in public management reform*, in *Policy and Internet*, 2014.
- B. COCCAGNA, *Libero accesso nelle politiche di open data: trasparenza, apertura e autoorganizzazione nel riutilizzo delle informazioni del settore pubblico*, in *Cib. dir.*, 2011.
- M. COCCONI, *La partecipazione all'attività amministrativa generale*, Padova, 2010.
- A. COCCHIA, *Smart and Digital City: A Systematic Literature Review*, in R.P. DAMERI - C. ROSENTHAL-SABROUX (eds.), *Smart City. How to Create Public and Economic Value with High Technology in Urban Space*, Springer International Publishing, Switzerland, 2014.
- C. COGLIANESE, *The Transparency President? The Obama Administration and Open Government, Governance*, 2009.
- C. COGLIANESE, *Citizen Participation in Rulemaking: Past, Present, and Future*, *Duke L. J.*, 55, 2005-2006.
- C. COGLIANESE, *The Transparency President? The Obama Administration and Open Government, in Governance*, 2009.

- C. COGLIANESE - H. KILMARTIN - E. MENDELSON, *Transparency and Public Participation in the Federal Rulemaking Process: Recommendations for the New Administration*, in *Geo. Wash. L. Rev.*, 2008-2009.
- S. COGNETTI, *Quantità e qualità della partecipazione*, Milano, 2000.
- S. COLEMAN, *Making the E-Citizen: A Socio-Technical Approach to Democracy, Connecting democracy: online consultation and the flow of political communication*, a cura di S. Coleman e P.M. Shane, Cambridge (US), 2011.
- D. CORLETTI, *Autorità indipendenti e giudice amministrativo*, in P. CAVALIERI - G. DELLE VEDOVE - P. DURET, *Autorità indipendenti e agenzie*, Padova, 2003.
- T.H. CORMEN - C.E. LEISERSON - L.R. RIVEST - C. STEIN, *Introduction to Algorithms*, MIT Press, Cambridge, 2010.
- F. COSTANTINO, *Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei Big Data*, *Dir. pub.*, 2019.
- F. COSTANTINO, *Lampi. Nuove frontiere delle decisioni amministrative tra "open" e "big data"*, in *Diritto amministrativo*, 2017.
- P. COSTANZO, *La "democrazia digitale". Precauzioni per l'uso*, in *Diritto pubblico*, 1, 2019.
- P. COSTANZO, *L'accesso ad internet in cerca d'autore*, in *Dir. Internet*, 3, 2005.
- M. CLARICH, *Autorità indipendenti. Bilancio e prospettive di un modello*, Bologna, 2005.
- A. CROSETTI - F. FRACCHIA (a cura di), *Procedimento amministrativo e partecipazione. Problemi, prospettive ed esperienze*, Milano, 2002.
- C. CUDIA, *Trasparenza amministrativa e pretesa del cittadino all'informazione*, *Dir. pub.*, 2007.
- V. CUFFARO, R. D'ORAZIO, V. RICCIUTO (a cura di), *I dati personali nel diritto europeo*, Torino, 2019.
- M. D'ANGELOSANTE, *La consistenza del modello dell'amministrazione "invisibile" nell'età della tecnificazione: dalla formazione delle decisioni alla responsabilità per le decisioni*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione*, vol. IV, in L. FERRARA - D. SORACE (a cura di), *A 150 anni dall'unificazione amministrativa italiana. Studi*, Firenze University Press, Firenze, 2016.
- M. D'ALBERTI, *Lezioni di diritto amministrativo*, 4°, Torino, 2019.
- M. D'ALBERTI, *Processi decisionali delle amministrazioni pubbliche. un'analisi comparata (Relazione al Convegno annuale dell'Associazione Italiana dei Professori di Diritto Amministrativo (AIPDA), Bergamo, 5-7 ottobre 2017)*, *Dir. pub.*, 2018.
- L. D'ETTORRE, *Le strategie macroregionali dell'Unione europea: tra cooperazione territoriale europea e multi-level governance*, in www.federalismi.it, 7 novembre 2018.
- E. D'ORLANDO, *Territori e potere. Un nuovo ruolo per regioni ed enti locali*, Forum edizioni, 2019.
- E. D'ORLANDO, *Algoritmi e organizzazione dell'amministrazione locale: come declinare il principio di adeguatezza affrontando la complessità*, in G.F. Ferrari (a cura di), *Smart City. L'evoluzione di un'idea*, Milano, Mimesis, 2020.
- E. D'ORLANDO - F.E. GRISOSTOLO, *La disciplina degli enti locali tra uniformità e differenziazione*, in F. Palermo - S. Parolari (a cura di), *Le variabili della specialità. Evidenze e riscontri tra soluzioni istituzionali e politiche settoriali*, Esi, Napoli, 2018.

- E. D'ORLANDO, *Profili costituzionali dell'amministrazione digitale*, in *Dir. inf. Inf.*, 2011.
- L. DALLA CORTE, *The European right to data protection in relation to open data*, in B. VAN LOENEN - G. VANCAUWENBERGHE - J. CROMPVOETS (eds.), *Open data exposed*, Information Technology and Law Series, Vol. 30, 2018.
- S. DAWES - L. VIDIASOVA - O. PARKHIMOVICH, *Planning and designing open government data programs: an ecosystem approach*, in *Gov. Inf. Quart.*, 2016.
- S.F. DE ANGELIS, *Artificial Intelligence: How Algorithms make Systems Smart?*, in *www.wired.com*, 2014.
- E. DE BLASIO, *Democrazia digitale. Una piccola introduzione*, Roma, 2014.
- J.M. DE FIGUEIREDOT, *E-Rulemaking: Bringing Data to Theory at the Federal Communications Commission*, in *Duke L. J.*, 55, 2005-2006.
- D. DE NART - C. TASSO, *A Personalized Concept-Driven Recommender System for Scientific Libraries*, in *Procedia Computer Science* 38, 2014.
- K.C. DE SOUZA - B. JACOB, *Big Data in the public sector: lessons for practitioners and scholars*, in *Administration & Society*, 2014.
- M.C. DE VIVO - A. POLZONETTI - P. TAPANELLI, *Open data, Business intelligence e Governance nella Pubblica amministrazione*, *Inf. Dir.*, 2011.
- F. DI DONATO, *Lo stato trasparente. Linked open data e cittadinanza attiva*, Pisa, 2010.
- F. DI MASCIO, *Miti e realtà degli "open data" all'italiana*, in *Giornale di diritto amministrativo*, 2017.
- G. DUNI, *Il procedimento amministrativo tra l. 7 agosto 1990 n. 241 ed introduzione dell'amministrazione telematica*, in *Foro Amm.*, fasc. 1, 1995.
- G. DUNI, *Il documento informatico: profili amministrativi*, in *Leg. e Giust.*, 1995.
- G. DUNI, *L'amministrazione digitale. Il diritto amministrativo nella evoluzione telematica*, Giuffrè, 2008.
- P. DUNLEAVY - H. MARGETTS et al., *The second wave of digital era governance*, in *ssrn.com*, 2010.
- P. DUNLEAVY - H. MARGETTS et al., *New Public Management is Dead-Long Live Digital-Era Governance*, in *J. Public Adm. Res. Theory*, 2006.
- P. DUNLEAVY - H. MARGETTS et al., *Digital Era Governance. IT Corporations, the State, and e-Government*, Oxford, 2006.
- L. EDWARDS - M. VEALE, *Enslaving the Algorithm: From a 'Right to an Explanation' to a 'Right to Better Decisions'?*, in *IEEE Security & Privacy*, 16-3, 2018.
- D.J. ELAZAR, *Exploring Federalism*, Tuscaloosa, University of Alabama Press, 1987, trad. it., *Idee e forme del federalismo*, Mondadori, Milano, 1998.
- M. EPPLER, *Managing Information Quality. Increasing the Value of Information in knowledge-intensive Products and Processes*, Berlin, 2003.
- F. FAINI, *La strada maestra dell'open government: presupposti, obiettivi, strumenti*, in *Cib. dir.*, 2013.
- F. FAINI, *Trasparenza, apertura e controllo democratico dell'amministrazione pubblica*, in *Cib. dir.*, 2014.

- M. FALCONE, "Big data" e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica, in *Rivista trimestrale di diritto pubblico*, 2017.
- M. FALCONE, "Big data" e pubbliche amministrazioni: nuove prospettive per la funzione conoscitiva pubblica, in *Rivista trimestrale di diritto pubblico*, n. 3, 2017.
- U. FANTIGROSSI, *Automazione e pubblica amministrazione. Profili giuridici*, Bologna, Il Mulino, 1993.
- R. FARIA - P. MILLER et al., *Rulemaking in 140 characters or less: social networking and public participation in rulemaking*, *Pace Law Rev.*, 2011.
- C.R. FARINA - M.J. NEWHART et al., *Rulemaking 2.0*, *U. Miami L. Rev.*, 2010-2011.
- C.R. FARINA, *Achieving the potential: the future of federal e-rulemaking. Report of the committee on the status and future of federal e-rulemaking*, *Admin. L. Rev.*, 2010.
- R. FERRARA, *La partecipazione al procedimento amministrativo: un profilo critico (Relazione al convegno "La legge generale sul procedimento amministrativo", Torino, 29 maggio 2017)*, *Dir. amm.*, 2017.
- M.R. FERRARESE, *Governance. Sugli effetti politici e giuridici di una "soft revolution"*, in *Pol. Dir.*, 2, 2014.
- M.R. FERRARESE, *La governance fra politica e diritto*, il Mulino, Bologna, 2010.
- G.F. FERRARI (a cura di), *La prossima città*, Milano-Udine Mimesi, 2017.
- G. FIORIGLIO, *Democrazia elettronica: presupposti e strumenti*, Padova, 2017.
- K. FINK, *Opening the government's black boxes: freedom of information and algorithmic accountability in Information, Communication & Society*, 2017.
- G. FINOCCHIARO, *La protezione dei dati personali in Italia. Regolamento UE n. 2016/679 e d.lgs. 10 agosto 2018, n. 101*, Bologna, 2019.
- J. FISHENDEN - M. THOMPSON, *Digital government, open architecture, and innovation: why public sector it will never be the same again*, in *Journ. Pub. Adm. Res. Theory*, 2012.
- C. FLICK - V. AMBRIOLA, *La cittadinanza amministrativa telematica fra previsioni normative ed effettività*, in *Dir. Informatica*, n. 6, 2006.
- M. FLINDERS, *Defending Politics: Why Democracy Matters in the 21st Century*, Oxford University Press, Oxford, 2012.
- S. FOÀ, *La nuova trasparenza amministrativa*, *Dir. amm.*, 2017.
- A. FUNG - M. GRAHAM - D. WEIL, *Full disclosure, The Perils and Promises of Transparency*, Cambridge, 2007.
- M.S. GAL, *Algorithmic Challenges to Autonomous Choice*, in *Michigan Telecommunications and Technology Law Review*, 2018.
- D.-U. GALETTA, *Digitalizzazione e diritto ad una buona amministrazione (Il procedimento amministrativo, fra diritto UE e tecnologie ICT)*, in *Diritto dell'Amministrazione pubblica digitale*, Giappichelli, 2020.
- D.-U. GALETTA - J.G. CORVALÁN, *Intelligenza Artificiale per una Pubblica Amministrazione 4.0? Potenzialità, rischi e sfide della rivoluzione tecnologica in atto*, in *www.federalismi.it*, 6 febbraio 2019.

- D.-U. GALETTA, *Open-Government, open-data e azione amministrativa*, relazione al convegno “Nuove prospettive dell’amministrazione digitale: open data e algoritmi / *Nouvelles perspectives de l’administration numérique: open data et algorithmes / New Perspectives of Digital Administration: Open Data and Algorithms*”, del Network “*Smart Cities and Digital Administration. Law and Governance*”, Venezia, 12 ottobre 2018.
- D.-U. GALETTA, *La trasparenza, per un nuovo rapporto tra cittadino e pubblica amministrazione: un’analisi storico-evolutiva, in una prospettiva di diritto comparato ed europeo (Relazione al 62 Convegno di Studi Amministrativi, Varenna, 22-24 settembre 2016)*, in *Riv. It. Dir. Pub. Com.*, 2016.
- D.-U. GALETTA, *Le fonti del diritto amministrativo europeo*, 91, in M.P. CHITTI (a cura di), *Diritto amministrativo europeo*, Milano, 2013.
- C.E. GALLO, *La pluralità delle cittadinanze e la cittadinanza amministrativa*, in *Dir. amm.*, 3, 2002.
- L. GALLINO, *La società. Perché cambia, come funziona. Un’introduzione sistemica alla sociologia*, Torino, Paravia, 1980.
- F. GARCÍA RUBIO (a cura di), *Las nuevas perspectivas de la ordenación urbanística y del paisaje: smart cities y rehabilitación. Una perspectiva hispano-italiana*, Madrid, Fundación Democracia y Gobierno Local, 2017.
- G. GARDINI, *Il paradosso della trasparenza in Italia: dell’arte di rendere oscure le cose semplici*, *Federalismi.it*, 2017.
- G. GARDINI, *Il codice della trasparenza: un primo passo verso il diritto all’informazione amministrativa?*, in *Giornale dir. amm.*, 2014.
- M. GASCO-HERNANDEZ (a cura di), *Open Government. Opportunities and Challenges for Public Governance*, New York, 2014.
- F. GASPARI, *Smart city, agenda urbana multilivello e nuova cittadinanza amministrativa*, Napoli, 2018.
- U. GASSER, *Framing information Quality Governance Research*, in ID. (a cura di), *Information Quality Regulation. Foundations, Perspectives, Applications*, Baden Baden, 2004.
- K. GAVELIN - S. BURALL - E. WILSON, *Open Government: Beyond Static Measures*, 2009.
- M.S. GIANNINI, *Atto amministrativo*, voce in *Enc. Dir.*, IV, Milano, 1959, 157, ma anche G. ZANOBINI, *Corso di diritto amministrativo*, Milano.
- A. GIANOPOULOU, *Algorithmic systems: the consent is in the detail?*, in *Internet Policy Review*, 2000.
- R. GIFFINGER - C. FERTNER - H. KRAMAR - R. KALASEK - N. PICHLER-MILANOVIC - E. MEIJERS, *Smart cities. Ranking of European medium-sized cities*, Vienna, University of Technology, 2007.
- F. GIGLIONI, *Le soluzioni istituzionali alla qualità dell’informazione statistica*, in *Diritto pubbl.*, 2005.
- F. GIGLIONI, *Il principio di sussidiarietà orizzontale nel diritto amministrativo e la sua applicazione*, in *Foro amm CdS*, 2009.

- M. GIORGIO, *Trasparenza e pubblicità dei dati delle pubbliche amministrazioni (Commento a delib. ANAC 28 dicembre 2016, n. 1310)*, in *Giorn. dir. amm.*, 2018.
- E. GIOVANNINI, *Scegliere il futuro. Conoscenza e politica al tempo dei Big data*, Bologna, Mulino, 2014
- G. GIUFFRIDA - S. GOZZO - F. MAZZEO RINALDI - V. TOMASELLI, *Applyin Network Analysis to Online News Big Data*, in F. GRESELIN - F. MOLA - M. ZENGA (a cura di), *Classification and Data Analysis*, Milano, Universitas Studiorum, 2017.
- G. GIUFFRIDA - S. GOZZO - F. MAZZEO RINALDI - V. TOMASELLI, *Big Data and Network Analysis: a Promising Integration for Decision-Making*, in LAURO C. - AMATURO E. - GRASSIA M.G. - ARAGONA B. - MARINO M. (a cura di), *Data Science and Social Research. Epistemology, Methods, Technology and Applications*, Berlin Heidelberg, Springer, 2017.
- G. GIUFFRIDA - S. GOZZO - F. MAZZEO RINALDI - V. TOMASELLI, *Extracting info from political news through Big Data Network Analysis*, in *Rivista Italiana di Economia Demografia e Statistica*, Vol. LXXI, n. 1, 2017.
- M. GNES, *La decertificazione. Dalle certificazioni amministrative alle dichiarazioni sostitutive*, Rimini, 2014.
- B. GOLDSTEIN - L. DYSON (a cura di), *Beyond Transparency*, S. Francisco, 2013.
- S.J. GOULD, *Legitimacy*, in V. BOGDANOR (ed.), *The Blackwell Encyclopedia of Political Institutions*, Oxford-New York, 1987.
- G. GRASSO, *Le autorità amministrative indipendenti della Repubblica. Tra legittimità costituzionale e legittimazione democratica*, Milano, 2006.
- L. GRIMALDI, *Il principio di sussidiarietà orizzontale tra ordinamento comunitario e ordinamento nazionale*, Bari, 2006.
- R. GRIMALDI - M.A. GALLINA - S.M. CAVAGNERO, *Big data e tracce digitali. Una ricerca internazionale tra studenti universitari*, in *Sociologia e Ricerca Sociale*, n. 109 (numero monografico "Sulle tracce dei Big Data. Questioni di metodo e percorsi di ricerca" a cura di Agnoli M.S. e Parra Saiani P.), 2017.
- R. GRIMALDI - M.A. GALLINA - S.M. CAVAGNERO, *Big data e tracce digitali. Una ricerca internazionale tra studenti universitari*, in Trincherò R. - Parola A. (a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, Angeli, 2016.
- R. GRIMALDI (a cura di), *A scuola con i robot*, Bologna, Il Mulino, 2015.
- R. GRIMALDI (a cura di), *Disuguaglianze digitali nella scuola*, Milano, Angeli, 2006.
- R. GRIMALDI (a cura di), *Metodi formali e risorse di rete*, Milano, Angeli, 2005.
- R. GRIMALDI (a cura di), *Le risorse culturali della Rete*, Milano, Angeli, 2003.
- M.P. GUERRA, *Circolazione dell'informazione e sistema informativo pubblico: profili giuridici dell'accesso interamministrativo telematico. Tra Testo Unico sulla documentazione amministrativa e codice dell'amministrazione digitale*, in *Diritto pubblico*, 2005.
- M.P. GUERRA, *Funzione conoscitiva e pubblici poteri*, Milano, Giuffrè, 1996.
- G. GUIDA - C. TASSO, *Design and Development of Knowledge-Based Systems: from Life Cycle to Development Methodology*, John Wiley and Sons, Chichester, UK 1994.

- J. GURIN, *Open data now*, NewYork, McGraw-Hill, 2014.
- G. GUZZARDO, *Semplificazione amministrativa e competitività nel governo del territorio*, Bari, 2008.
- J. HIRSCH, *The state's new clothes: NGOs and the internationalization of states*, *Rethinking Marxism*, 2003.
- C. HOOD, *A Public Management for All Seasons?*, in *Public Administration*, 69, 1, 1991.
- T. IM - W. CHO - G. PORUMBESCU - J. PARK, *Internet, trust in government, and citizen compliance*, *J. Public Adm. Res. Theory*, 2012.
- A. JAAKOLA - H. KEKKONEN - T. LAHTI - A. MANNINEN, *Open data, open cities: experiences from the Helsinki Metropolitan Area. Case Helsinki Region infoshare www.hri.fi. Statistical Journal of the IAOS*, 2015.
- M. JANSSEN - N. HELBIG, *Innovating and changing the policy-cycle: policy makers be prepared!*, in *Gov. Inf. Quart.*, 2015.
- M.L. JONES - E. EDENBERG - E. KAUFMAN, *AI and the Ethics of Automating Consent*, in *IEEE Security & Privacy*, 2018.
- D. KAMARINOU - C. MILLARD - J. SINGH, *Machine learning with personal data*, in *Queen Mary School of Law Legal Studies*, Research Paper n. 247, 2016.
- M. KASSEN, *Understanding transparency of government from a Nordic perspective: open government and open data movement as a multidimensional collaborative phenomenon in Sweden*, in *Journal of Global Information Technology Management*, 2017, 236-275.
- V. KHATRI - B.M. SAMUEL, *Analytics for managerial work. Communications of the ACM*, 62, 2019.
- B. KOHLER-KOCH - B. FINKE, *The institutional shaping of EU-society relations: a contribution to democracy via participation?*, in *Journ. Civ. Soc.*, 2007.
- J.A. KROLL - J. HUEY - S. BAROCAS - E.W. FELTEN - J.R. REIDENBERG - D.G. ROBINSON - H. YU, *Accountable Algorithms*, in *University of Pennsylvania Law Review*, 165, 2017.
- H. LANDEMORE, *Inclusive Constitution-Making: The Icelandic Experiment*, in *Journ. Pol. Phil.*, 2015.
- D. LATHROP - L. RUMA, *Open Government: Collaboration, Transparency, and Participation in Practice*, Sebastopol, 2010.
- P. LAZZARA, *I procedimenti ad istanza di parte. Dalla disciplina generale sul procedimento (l. 241/90) alla direttiva servizi (2006/123)*, Napoli, 2008.
- P. LAZZARA, *Principio di semplificazione e situazioni giuridico-soggettive*, in *Dir. amm.*, 2011.
- F.M. LAZZARO, *Coordinamento informativo e pubbliche amministrazioni*, in *Ist. fed.*, 2011.
- Y.W. LEE - D.M. STRONG - B.K. KAHN - R.Y. WANG, *AIMQ: A Methodology for Information Quality Assessment*, in *Information & Management*, 2002.
- D. LINDERS, *From e-government to we-government: Defining a typology for citizen coproduction in the age of social media*, *Gov. inform. quart.*, 2012.
- L.M. LOPUCKI, *Algorithmic Entities*, in *Washington University Law Review*, n. 95, 2018.
- N. LONGOBARDI, *Autorità amministrative indipendenti e sistema giuridico-istituzionale*, Torino, 2004.

- M. LUCIANI, *Democrazia rappresentativa e democrazia partecipativa*, in *Rivista AIC*, 2003.
- M. LUCIANI, *Il paradigma della rappresentanza di fronte alla crisi del rappresentato*, Milano, 2001.
- P.L.M. LUCATUORTO, *Teorie e modelli del diritto per il ragionamento giuridico automatico*, Milano, LED, 2009.
- R. LUNA, *I dati del virus, il Veneto e la piattaforma di cui abbiamo bisogno*, disponibile al sito www.repubblica.it, 2 maggio 2020.
- M. MACIEJEWSKI, *To do more, better, faster and cheaper: using big data in public administration*, in *International Review of Administrative Sciences*, 2017.
- M.L. MADDALENA, *La digitalizzazione della vita dell'amministrazione e del processo*, in *Foro amm.*, 2016.
- G. MANCOSU, *Trasparenza amministrativa e open data: un binomio in fase di rodaggio, federalismi.it*, 2012.
- F. MANGANARO, *L'evoluzione del principio di trasparenza amministrativa*, astrid-online.it, 2009.
- S. MANGIAMELI, *Appunti a margine dell'art. 116, comma 3, della Costituzione*, in *Le Regioni*, n. 4, 2017.
- S. MANGIAMELI, *Crisi economica e distribuzione territoriale del potere politico - Relazione al XVIII Convegno annuale dell'AIC*, in *Rivista AIC*, www.rivistaaic.it, 4, 2013.
- A. MANTELERO, *Regulating Big Data. The guidelines of the Council of Europe in the context of the European data protection framework*, in *Computer Law & Security Report*, 2017.
- A. MANZELLA, *Parlamento Europeo e Parlamenti Nazionali come Sistema*, *Rivista AIC*, n. 1, 2015.
- M. MARCANTONI - M. BALDI, *Regioni a geometria variabile*, Donzelli, Roma, 2013.
- F. MARCELLI - P. MARSOCCI - M. PIETRANGELO (a cura di), *La rete Internet come strumento di partecipazione politica. Una prospettiva giuridica*, Napoli, 2015.
- M. MARGOLIS - G. MORENO - RIAHO, *E-Government, customers and citizens, Understanding E-Government in Europe: Issues and challenges*, a cura di G. Nixon, V.N. Koutrakou e R. Rawal, Londra-New York, 2010.
- I. MARINO, *Autonomie e democrazia. Profilo dell'evoluzione dell'autonomia e della sua ricaduta sul sistema giuridico*, in *Nuove autonomie*, 2007.
- D. MARONGIU, *Qualità, durata e modalità di erogazione della pubblicazione*, in B. Ponti (a cura di), *Nuova trasparenza amministrativa e libertà di accesso alle informazioni*, Rimini, 2016.
- D. MARONGIU, *I dati delle pubbliche amministrazioni come patrimonio economico nella società dell'informazione*, in *Inf. dir.*, 2008.
- S. MARTIN, *Engaging with citizens and other stakeholders*, in T. BOVAIRD - E. LOFFLER (a cura di), *Public management and governance*, Londra, 2009.
- I. MARTÍN DELGADO, *Automazione, intelligenza artificiale e pubblica amministrazione: vecchie categorie concettuali per nuovi problemi?*, in *Istituzioni del Federalismo*, 2019.
- I. MARTÍN DELGADO, *La riforma dell'amministrazione digitale: Un'opportunità per ripensare la pubblica Amministrazione*, in S. CIVITARESE MATTEUCCI - L. TORCHIA (a cura di), *La tecnificazione*, Firenze University Press, Firenze, 2016.

- F. MARTINES, *La digitalizzazione della pubblica amministrazione*, in *La rivista di diritto dei media*, n. 2, 2018.
- F. MARZANO, *La trasparenza nella Pubblica Amministrazione passa dall'Open Data o l'Open Data passa dalla trasparenza?*, *Inf. dir.*, 2011.
- A. MASUCCI, *L'atto amministrativo informatico: primi lineamenti di una ricostruzione*, Jovene, Napoli, 1993.
- V. MAYER SCHOMBERGER - K. CUKIER, *Big data. Una rivoluzione che trasformerà il nostro modo di vivere e già minaccia la nostra libertà*, trad. it. R. Merlini, Milano, 2013.
- F. MAZZEO RINALDI - G. GIUFFRIDA - T. NEGRETE, *Real-time monitoring and evaluation - Emerging news as predictive process using Big Data based approach*, in PETERSSON G. - BREUL J.D. (eds), *Cyber Society, Big Data and Evaluation*, vol. 24, New Brunswick, NJ, Transaction, 2017.
- A. MCAFEE - E. BRYNJOLFSSON, *Big data: The management revolution*, in *Harv. Bus. Rev.*, 2012.
- W.S. MCCULLOCH - W.A. PITTS, *A logical calculus of the ideas immanent in nervous activity*, in *Bull. math Biophysics* 5, 1943.
- K. MCNUTT, *Public engagement in the Web 2.0 era: Social collaborative technologies in a public sector context*, in *Can. Publ. Adm.*, 2014.
- A. MEALE, *Organizzazione della P.A. - il principio di leale collaborazione tra competenze statali e regionali*, in *Giur. It.*, 2017, 3.
- A. MEIJERA - S. POTJERB, *Citizen-generated open data: An explorative analysis of 25 cases*, in *Gov. Inf. Quart.*, 2018.
- P. MELL, T. GRANCE, *The NIST Definition of Cloud Computing*, National Institute of Standards and Technology, Gaithersburg, 2011.
- A. MELONCELLI, *L'informazione amministrativa*, Rimini, 1983.
- F. MERLONI - R. CAVALLO PERIN (a cura di), *Al servizio della Nazione. Etica e statuto dei funzionari pubblici*, Milano, 2009.
- F. MERLONI (a cura di), *La trasparenza amministrativa*, Milano, 2008.
- F. MERLONI (a cura di), *Introduzione all'@government. Pubbliche amministrazioni e società dell'informazione*, Torino, 2005.
- F. MERLONI, *Sull'emergere della funzione di informazione nelle pubbliche amministrazioni*, in ID. (a cura di), *L'informazione delle pubbliche amministrazioni*, Rimini, 2002.
- F. MERLONI, *Trasparenza delle istituzioni e principio democratico*, in F. MERLONI (a cura), *La trasparenza amministrativa*, Milano, 2008.
- F. MERUSI, *Ragionevolezza e discrezionalità amministrativa*, Napoli, 2011.
- E. MINARDI (a cura di), *Codici e paradigmi per rileggere lo sviluppo locale*, Teramo, 2011.
- F. MINAZZI, *Il principio dell'open data by default nel codice dell'amministrazione digitale: profili interpretativi e questioni metodologiche*, *Federalismi.it*, 2013.
- M. MINERVA, *L'attività amministrativa in forma elettronica*, in *Foro amm.*, 1997.
- G. MISURACA - F. MAREDDU - D. OSIMO, *Policy Making 2.0: Unleashing the power of big data for public governance*, Heidelberg Dordrecht, Springer, 2014.

- T.M. MITCHELL, *Machine Learning*, McGraw-Hill 1997.
- MIUR (2016), *Big data@Miur, Rapporto del gruppo di ricerca*, www.istruzione.it/allegati/2016/bigdata.pdf.
- A. MOLITERNI, *Il principio di trasparenza: recenti tendenze e prospettive future*, in *Riv. it. sc. giur.*, 2014.
- J. MOOR, *The Dartmouth College Artificial Intelligence Conference: The Next Fifty years*, in *AI Magazine* 27, 2006.
- M. MOORE - D. TAMBINI (a cura di), *Digital dominance: the power of Google, Amazon, Facebook, and Apple*, New York, Oxford University Press, 2018.
- G. MORBIDELLI, *Linee guida ANAC: comandi o consigli?*, in *Dir. Amm.*, n. 3, 2016.
- S. MORETTINI, *Il soft law nelle autorità indipendenti: procedure oscure e assenza di garanzie?*, in Osservatorio sull'AI, www.osservatorioair.it.
- F. MOROLLO, *Documento elettronico fra e-government e artificial intelligence (AI)*, in *Federalismi.it*, n. 2, 2015.
- K. MOSSBERGER - C. TOLBERT, *The effects of E-Government on trust and confidence in government*, in *Pub. Adm. Rev.*, 2003.
- T. NAM, *Suggesting frameworks of citizen-sourcing via Government 2.0*, in *Gov. inf. Quart.*, 2012.
- M. NAPOLITANO, *Il primo decennio di dati aperti è stato un successo, ma non per le ragioni che pensiamo*, 2019, forumpa.it.
- D. NATALE - M.C. PAOLETTI - A. SIMONETTA, *La qualità dei dati e l'informazione statistica*, in *Rivista degli infortuni e delle malattie professionali*, 2012.
- A. NATALINI - G. VESPERINI (a cura di), *Il big bang della trasparenza*, Napoli, 2015.
- P.G. NIXON - V.N. KOUTRAKOU - R. RAWAL (a cura di), *Understanding E-Governmente in Europe: Issues and challenges*, Londra-New York, 2010.
- B.S. NOVECK, *Rights-Based and Tech-Driven: Open Data, Freedom of Information, and the Future of Government Transparency*, in *Yale Hum. Rts. & Dev. L.J.*, 2017.
- B.S. NOVECK, *Wiki Government: how technology can make government better, democracy stronger, and citizens more powerful*, Washington D.C., 2009.
- P. OTRANTO, *Decisione amministrativa e digitalizzazione della p.a.*, in *Federalismi*, n. 2, 2018.
- M. OREFICE, *Gli open data tra principio e azione: lo stato di avanzamento*, in www.forumcostituzionale.it, 25 maggio 2015.
- A. PAJNO, *L'esercizio di attività in forme contenziose*, in S. CASSESE - C. FRANCHINI (a cura di), *I garanti delle regole*, Bologna, 1996.
- V. PANDOLFINI, *Il sociologo e l'algoritmo. L'analisi dei dati testuali al tempo di Internet*, Milano, Angeli, 2017.
- T. PARDO - N. TAEWOO, *Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th Annual International Conference on Digital Government Research*, New York, ACM, 2011.
- W. PARKS, *The Open Government Principle: Applying the Right to Know under the Constitution*, in *George Wash. Law Rev.*, 1957.

- P. PARYCEK - J. HOCHTL - M. GINNER, *Open Government Data implementation evaluation*, in *Journ. Theor. Appl. Elect. Comm. Res.*, 2014.
- C. PASQUINELLI, *Open Data e Proprietà Intellettuale. Il Dataset come banca di dati e la questione delle licenze*, CNR Technical Report, 2013.
- G. PASQUINO, *I problemi della rappresentanza politica*, in *XXI Secolo*, 2009.
- G. PASTORI, *Principi costituzionali sull'amministrazione e principio inquisitorio nel procedimento*, in M. CAMMELLI - M.P. GUERRA (a cura di), *Informazione e funzione amministrativa*, Rimini, Maggioli, 1996.
- P. PATRITO - F. PAVONI, *La disciplina del riutilizzo dei dati pubblici dal punto di vista del diritto amministrativo*, in *Dir. inf.*, 2012.
- F. PATRONI GRIFFI, *Élite politiche e amministrative tra crisi della democrazia e riforme dell'amministrazione*, in *Federalismi*, 2017.
- V. PATRUNO, *Dieci anni di open data: ora servono "dati vivi"*, *forumpa.it*.
- V. PATRUNO, *Open data nel piano triennale ICT PA: la prima strategia unica su riuso e valorizzazione*, *agendadigitale.eu*, 2018.
- V. PATRUNO, *Open government: così trasparenza e partecipazione rafforzano la democrazia*, *agendadigitale.eu*, 2018.
- A. PELED - N. KARINE, *Towards Open Data for Public Accountability: Examining the US and the UK Models*, in *SSRN Electronic Journal*, OAD. doi:10.2139/SSRN.2546017, 2015.
- O. PEREZ, *Open Government, technological innovation, and the politics of democratic disillusionment: (E-)Democracy from Socrates to Obama*, in *Journ. law pol. inf. Soc.*, 2012.
- M. PIETRANGELO, *Brevi note sul coordinamento informativo informatico e statistico dei dati dell'amministrazione statale, regionale e locale*, in *Informatica e diritto*, n. 1-2, 2004.
- M. PIETRANGELO, *Il diritto all'uso delle tecnologie nei rapporti con la pubblica amministrazione: luci ed ombre*, in *Informatica e diritto*, 1-2, 2005.
- C. PINELLI, *Il deficit democratico europeo e le risposte del Trattato di Lisbona*, in *Rassegna Parlamentare*, 2008.
- P. PIRAS, *Servizi pubblici e nuove tecnologie*, in *Dir. inf.*, 2006.
- P. PIRAS, *Il tortuoso cammino verso un'amministrazione nativa digitale*, in *Diritto dell'Informazione e dell'Informatica*, n. 1, 2020.
- P. PIRAS, *Organizzazione, tecnologie e nuovi diritti*, in *Dir. informatica*, n. 3, 2005.
- G. PIZZANELLI, *La partecipazione dei privati alle decisioni pubbliche. Politiche ambientali e realizzazione delle grandi opere infrastrutturali*, Milano, 2010.
- F. PIZZETTI, *Privacy e il diritto europeo alla protezione dei dati personali. Dalla Direttiva 95/46 al nuovo Regolamento europeo*, Torino, Giappichelli, 2016.
- F. PIZZETTI, *Privacy e il diritto europeo alla protezione dei dati personali*, Torino, 2016.
- F. PIZZETTI, *La riforma degli enti territoriali Città metropolitane, nuove province e unione di comuni Legge 7 aprile 2014, n. 56 (Legge "Delrio")*, Milano, Giuffrè, 2015.
- B. PONTI, *La mediazione informativa nel regime giuridico della trasparenza: spunti ricostruttivi*, in *Diritto dell'informazione e dell'informatica*, 2019.

- B. PONTI, *Il luogo adatto dove bilanciare. Il “posizionamento” del diritto alla riservatezza e alla tutela dei dati personali vs il diritto alla trasparenza nella sentenza n. 20/2019, in Istituzioni del federalismo*, n. 2, 2019.
- B. PONTI (a cura di), *Nuova trasparenza amministrativa e libertà di accesso alle informazioni*, Rimini, Maggioli, 2016.
- B. PONTI (a cura di), *La trasparenza amministrativa dopo il d.lgs. 14 marzo 2013*, n. 33, Rimini, 2013.
- B. PONTI, *I dati di fonte pubblica: coordinamento, qualità e riutilizzo*, in F. MERLONI (a cura di), *La trasparenza amministrativa*, Giuffrè, Milano, 2008.
- B. PONTI, *Il patrimonio informativo pubblico come risorsa. i limiti del regime italiano di riutilizzo dei dati delle pubbliche amministrazioni*, in *Dir. pub.*, 2007.
- B. PONTI (a cura di), *Il regime dei dati pubblici. Esperienze europee e ordinamento nazionale*, 2007.
- O. POLLICINO - G. REPETTO, *Not to be Pushed Aside: the Italian Constitutional Court and the European Court of Justice*, in *VerfBlog*, 27 febbraio 2019.
- C. RAIOLA, *Le consultazioni telematiche delle Autorità indipendenti. Gli effetti dell’AIR su tecniche e caratteristiche*, osservatorioair.it, 2012.
- C.G. REDDICK, *Citizen interaction and e-government: evidence for the managerial, consultative, and participatory models*, in *Transforming Government: People, Process and Policy*, 2011.
- R. REINIKKA - J. SVENSSON, *The Power of Information in Public Services: Evidence from Education in Uganda*, *Journal of Public Economics*, 2011, 956.
- A. REZZANI, *Big data. Architettura, tecnologie e metodi per l’utilizzo di grandi basi di dati*, Rimini, Maggioli, 2013.
- S. RICCI, *Note in tema di “riutilizzo dell’informazione pubblica” e diritto alla privacy*, in *federalismi.it*, 2005.
- H. RICHTER - P.R. SLOWINSKI, *The Data Sharing Economy: On the Emergence of New Intermediaries*, IIC, 2019.
- B. RITTMER, *Building Europe’s Parliament. Democratic Representation Beyond the Nation State*, Oxford, 2005.
- A. RIVIEZZO, *Autorità amministrative indipendenti e ordinamento costituzionale*, in *Quad. cost.*, 2005.
- D. ROBINSON - H. YU *et al.*, *Government Data and the Invisible Hand*, in *Yale J.L. & Tech.*, 2009.
- N. RODEAN, *Participatory Democracy: Mechanism of Better Regulation in Europe*, in *Acta U. Danubius Jur.*, 201.
- S. RODOTÀ, *Tecnopolitica*, 2, Roma-Bari, 2004.
- A. ROMANO, *Il Cittadino e la pubblica amministrazione*, in *Il diritto amministrativo degli anni 80, Atti del XXX Convegno di studi della scienza dell’amministrazione*, Milano, 1987.
- S. ROMANO, *L’ordinamento giuridico*, Sansoni, Firenze, 1918.
- F. ROSENBLATT, *The perceptron: A probabilistic model for information storage and organization in the brain*, in *Psychological Review* 65(6), 1958.

- J.N. ROSENEAU - E.-O. CZEMPIEL, *Governance without Government: Order and Change in World Politics*, Cambridge University Press, Cambridge Mass., 1992.
- A. RUGGERI, *La Consulta rimette a punto i rapporti tra diritto europolitico e diritto intero con una pronuncia in chiaroscuro (a prima lettura di Corte cost. n. 20 del 2019)*, in *Consulta Online*, 23 febbraio 2019.
- E.H.J.M. RUIJER - E. MARTINIUS, *Researching the democratic impact of open government data: A systematic literature review*, in *Inf. Pol.*, 2017, 233.
- D. RUSHKOFF, *Open source democracy. How online communication is changing offline politics*, Londra, 2003.
- S. RUSSEL - P. NORVIG, *Artificial Intelligence: A Modern Approach, 3rd Ed.*, Pearson 2010.
- G.C. SALERNO, *Servizi di interesse generale e sussidiarietà orizzontale fra ordinamento costituzionale e ordinamento dell'Unione europea*, Torino, 2010.
- R. SANNA, *Dalla trasparenza amministrativa ai dati aperti. Opportunità e rischi delle autostrade informatiche*, Torino, 2018.
- C.P. SANTACROCE, *Conquiste e arretramenti nella trasparenza amministrativa che cambia*, in *Dir. soc.*, 2017.
- L. SARTORI, *Alla ricerca della smart citizenship*, in *Istituzioni del Federalismo*, n. 4, 2015.
- L. SARTORI, *Open Government: what else?*, in *Istituzioni del Federalismo*, n. 3-4, 2013.
- M. SAVINO, *Il FOIA italiano. la fine della trasparenza di Bertoldo (Commento a d.lg. 25 maggio 2016, n. 97)*, in *Giorn. dir. amm.*, 2016.
- M. SAVINO, *The right to open public administrations in Europe: emerging legal standards*, Sigma Paper n. 46, Oecd, 2010.
- A. SCALA, *L'automazione nella redazione degli atti amministrativi*, in *Nuova rass.*, 1995.
- D.W. SCHARTUM, *Law and Algorithms in the Public Domain, Etik i Praksis*, in *Nordic Journal of Applied Ethics*, 10, 2016.
- F. SCIACCHITANO, *Disciplina e utilizzo degli Open Data in Italia*, in *Medialaws.eu*.
- SHAILY KUMAR, *The importance of analytics in the public sector*, 17th of April, 2018, downloaded at: <https://www.digitalistmag.com/future-of-work/2018/04/17/importance-of-analytics-in-public-secotr-06061094>, editor: 2019 SAP SE.
- J. SHKABATUR, *Transparency with(out) accountability: Open Government in the United States*, in *Yale Law Pol. Rep.*, 2012.
- E. SIEGEL, *Predictive Analytics: the power to predict who will click, buy, lie or die*, Wiley, New Jersey, 2016.
- M.R. SPASIANO, *La partecipazione al procedimento quale fonte di legittimazione dell'esercizio del potere: un'ipotesi ricostruttiva*, in *Dir. amm.*, 2002.
- F. SYLOS LABINI, *Big data, complessità e metodo scientifico*, in *Aspenia*, 63, Aspen Institute Italia, dicembre 2013.
- C.R. SUNSTEIN, *Infotopia. How many minds produce knowledge*, Oxford, 2006.
- A. STERPA (a cura di), *Il nuovo governo dell'area vasta. Commento alla legge 7 aprile 2014 n. 56*, Napoli, Jovene, 2014.
- J. TAYLOR, *The information polity: towards a two speed future*, in *Inf. Polity.*, 2012.

- J. TAUBERER, *Open Government Data: The Book*, in *opengovdata.io*, 2, 2014.
- M. TIMO, *Cittadinanza amministrativa e democrazia partecipativa*, in *Osservatorio AIC*, 2016.
- L. TORCHIA, *Autonomia dei soggetti e funzionalità del sistema: condizioni di qualità dell'informazione statistica*, in *Riv. ital. dir. pubbl. com.*, 1999.
- P. TÜRK, *La citoyenneté à l'ère numérique*, in *Revue de droit publique*, 3, 2018.
- B. UBALDI, *Open Government Data: Towards Empirical Analysis of Open Government Data Initiatives*, *Oecd Working Papers on Public Governance*, 22, 2013.
- A. USAI, *Le prospettive di automazione delle decisioni amministrative in un sistema di teleamministrazione*, in *Dir. Inf. e Inform.*, 1993.
- L. VANDELLI, *Il sistema delle autonomie locali*, il Mulino, Bologna, 2018.
- L. VANDELLI, *Città metropolitane, province, unioni e fusioni di comuni*, con la collaborazione di P. BARRERA, T. TESSARO, C. TUBERTINI, Rimini, 2014.
- B. VALTYSSON, *Democracy in disguise: the use of social media in reviewing the Icelandic Constitution*, *Media, Culture & Society*, 2014.
- A. VENANZONI, *Smart cities e capitalismo della sorveglianza*, in *Forum di quaderni costituzionali*, 2019.
- M.R. VICENTE - A. NOVO, *An empirical analysis of e-participation. The role of social networks and e-government over citizens' online engagement*, in *Gov. Inf. Quart.*, 2014.
- S. VILLAMENA, *Il c.d. FOIA [Freedom Of Information Act] (o accesso civico 2016) ed il suo coordinamento con istituti consimili*, in *Federalismi.it*, 2016.
- L. VIOLANTE (a cura di), *La fiducia nelle società democratiche*, Soveria Mannelli, 2014.
- P. VIPIANA (a cura di), *Tendenze centripete e centrifughe negli ordinamenti statali dell'Europa in crisi*, Giappichelli, 2014.
- R.Y. WANG - D.M. STRONG - L. GUARASCIO, *An Empirical Investigation of Data Quality Dimensions. A Data Consumer's Perspective*, working paper TDQM-94-01, MIT, Cambridge (Ma), 1984.
- A.M. WARREN - A. SULAIMAN - N.I. JAAFAR, *Social media effects on fostering online civic engagement and building citizen trust and trust in institutions*, in *Gov. Inf. Quart.*, 2014.
- WATSON IoT, *Descriptive, predictive, prescriptive: Transforming asset and facilities management with analytics*, downloaded at: <https://www.ibm.com/downloads/cas/3V9AA9Y5>, editor: IBM Corp. 2017.
- M. WEBER, *Economia e società*, trad. it., Milano, 1961.
- J.H. WEILER, *La Costituzione dell'Europa*, Bologna, 2003.
- F. WELLE DONKER, *Funding Open Data*, in *Open Data Exposed, Information Technology and Law*, a cura di B. van Loenen et al., 2018.
- B. WORTHY, *The impact of open data in the UK: complex, unpredictable and political*, in *Publ. Adm.*, 2015.
- J. YAKOWITZ, *Tragedy of the Data Commons*, in *Harv. Journ. Law Tech.*, 2011.
- H. YU - G.D. ROBINSON, *The New Ambiguity of Open Government*, in *Ucla Law Rev. Disc.*, 2012.

- X. ZHAO - H.D. XU, *E-Government and Corruption: A Longitudinal Analysis of Countries*, in *Int. J. Pub. Adm.*, 38, 2015.
- S. ZUBOFF, *Il capitalismo della sorveglianza. Il futuro dell'umanità nell'era dei nuovi poteri*, Luiss University Press, Roma, 2019.

Notizie sugli Autori

ISABELLA ALBERTI, Dottoranda di diritto amministrativo nell'Università degli Studi di Torino.

MARCO ALDINUCCI, Professore ordinario di Informatica nell'Università di Torino.

MARINA CAPORALE, Ricercatrice (RTDA) di diritto amministrativo nell'Università Telematica "Leonardo da Vinci", professoressa incaricata nell'Università di Bologna.

ENRICO CARLONI, Professore ordinario di diritto amministrativo nell'Università degli Studi di Perugia.

GHERARDO CARULLO, Ricercatore (RTDB) di diritto amministrativo nell'Università degli Studi di Milano.

ROBERTO CAVALLO PERIN, Professore ordinario di diritto amministrativo nell'Università degli Studi di Torino.

AGUSTÍ CERRILLO I MARTÍNEZ, Catedrático de Derecho administrativo nell'Universitat Oberta de Catalunya.

FULVIO COSTANTINO, Ricercatore (RTDB) di diritto amministrativo nell'Università degli Studi di Macerata.

ELENA D'ORLANDO, Professoressa ordinaria di Diritto pubblico comparato nell'Università degli Studi di Udine.

FABIANA DI PORTO, Professoressa associata di diritto dell'economia nell'Università degli studi del Salento.

MATTEO FALCONE, Assegnista di ricerca di diritto amministrativo nell'Università degli Studi di Perugia.

RENATO GRIMALDI, Professore ordinario di Metodologia della ricerca nell'Università degli studi di Torino.

FRANCESCO MERLONI, Professore ordinario di diritto amministrativo nell'Università degli Studi di Perugia.

ROSA MEO, Professoressa associata di Informatica nell'Università degli Studi di Torino.

MIRKO LAI, Assegnista di ricerca di Informatica nell'Università degli Studi di Torino.

UGO PAGALLO, Professore ordinario di informatica giuridica nell'Università degli studi di Torino.

PAOLO PASTERIS, Personale tecnico amministrativo Dipartimento di Informatica nell'Università degli Studi di Torino.

ALESSANDRA PIOGGIA, Professoressa ordinaria di diritto amministrativo nell'Università degli Studi di Perugia.

BENEDETTO PONTI, Professore associato di diritto amministrativo nell'Università degli Studi di Perugia.

ANDREA SIMONCINI, Professore ordinario di diritto costituzionale nell'Università degli Studi di Firenze.

CARLO TASSO, Professore ordinario di Sistemi di elaborazione dell'informazione Università di Udine.

STAMPATO IN ITALIA
nel mese di marzo 2021
da Rubbettino print
88049 Soveria Mannelli (Catanzaro)

A maggio 2019, pareva pionieristico parlare di *big data* e di intelligenza artificiale come strumenti di lavoro per le pubbliche amministrazioni; non così a poco più di un anno di distanza, dimostrando la lungimiranza nel sapere vedere una pubblica amministrazione capace di cogliere le sfide che l'innovazione tecnologica stava ponendo.

Questo libro raccoglie prospettive inedite, nate dal dialogo tra studiosi del diritto pubblico e dell'informatica che, a rileggerlo oggi, mostra come il connubio tra questi due mondi scientifici sia imprescindibile per la ricerca del XXI secolo.

Anche il luogo in cui si tenne questo convegno non è irrilevante: Torino. Città che a breve accoglierà l'Istituto Italiano per l'Intelligenza Artificiale, a dimostrazione che l'interesse per il tema ha assunto nel corso del tempo una notevole attualità, anche fuori dall'Accademia.

L'idea è che la potenza del diritto è capace di interagire con gli strumenti dell'innovazione tecnologica (come l'intelligenza artificiale), non per bloccarne il suo avanzare, ma per legittimarlo verso prospettive compatibili con i diritti e i principi di una millenaria tradizione giuridica.

Il "sapere nuovo" che emerge dalla lettura di questo libro è dunque un sapere capace di orientare il lettore tra le potenzialità e le sfide che la società dell'informazione e della comunicazione offre e pone alle Istituzioni pubbliche, alle quali si chiede di sapersi adattare con sapienza e capacità in vista del miglior perseguimento del fine dell'interesse pubblico.

Roberto Cavallo Perin è professore ordinario di diritto amministrativo presso l'Università degli studi di Torino e tra gli altri incarichi, Direttore per l'Italia del "Network Review on line" lus-publicum.com, nonché della rivista "Diritto amministrativo", Giuffrè Editore.

I suoi interessi di ricerca spaziano dalla teoria generale del diritto amministrativo, agli istituti che di questo sono tipici, come la concessione e gli appalti e non da ultimo i servizi pubblici. Di recente i suoi studi si sono concentrati sull'impatto che la rivoluzione tecnologica ha sull'organizzazione e l'attività della pubblica amministrazione.

ISBN 978-88-7590-180-6



9 788875 901806

€ 22,00