

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Big data e tracce digitali Una ricerca internazionale tra studenti universitari

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available <http://hdl.handle.net/2318/1736741> since 2020-04-19T14:46:37Z

Publisher:

Aracne

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

TEORIA E RICERCA IN EDUCAZIONE

COLLANA DEL DIPARTIMENTO DI FILOSOFIA E SCIENZE DELL'EDUCAZIONE
SEZIONE DI SCIENZE DELL'EDUCAZIONE
UNIVERSITÀ DEGLI STUDI DI TORINO

Direttore

Renato GRIMALDI

Comitato scientifico

Cristina BERTOLINO

Federica MAZZOCCHI

Paolo BIANCHINI

Lorena MILANI

Paola BORGNA

Irma NASO

Barbara BRUSCHI

Sara NOSARI

Tanja CERRUTI

Germana PARETI

Cristina COGGI

Alberto PAROLA

Ivan ENRICI

Marisa PAVONE

Carlo Mario FEDELI

Alessandro PERISSINOTTO

Maria Adelaide GALLINA

Isabella PESCARMONA

Cristiano GIORDA

Paola RICCHIARDI

Anna GRANATA

Elisabetta ROBOTTI

Enrico GUGLIELMINETTI

Paolo ROSSO

Matteo LEONE

Simona TIROCCHI

Graziano LINGUA

Emanuela Maria TORRE

Daniela MACCARIO

Roberto TRINCHERO

Mario MARTINELLI

Federico ZAMENGO

Il Direttore e i docenti afferiscono all'Università degli Studi di Torino.

TEORIA E RICERCA IN EDUCAZIONE

COLLANA DEL DIPARTIMENTO DI FILOSOFIA E SCIENZE DELL'EDUCAZIONE
SEZIONE DI SCIENZE DELL'EDUCAZIONE
UNIVERSITÀ DEGLI STUDI DI TORINO

[...] tra una scienza e l'altra non sono obbligatorie clamorose divergenze di modelli del mondo, presupposizioni, grammatiche, linguaggi, modelli di base e compositi, [...] molti tipi di scambio di siffatte componenti tra le strutture concettuali di scienze diverse sono meno disagiati di quanto comunemente si creda.

Luciano GALLINO, *L'incerta alleanza*, 1992, p. 282

Teoria, metodo e ricerca sono alla base degli studi raccolti nella presente collana. Il linguaggio scientifico e l'interdisciplinarietà caratterizzano i lavori qui pubblicati da studiosi di differenti aree che, provenendo dalle scienze umane e dalle scienze naturali, proprio in questo spazio trovano il luogo di una feconda cooperazione intellettuale.

Tutti i testi sono preventivamente sottoposti a referaggio anonimo.

L'attività di ricerca che ha portato alla pubblicazione di questo volume si è svolta nell'ambito del Progetto di Eccellenza del Dipartimento di Filosofia e Scienze dell'Educazione dell'Università degli Studi di Torino.

Big data e tracce digitali

Una ricerca internazionale tra studenti universitari

a cura di

Renato Grimaldi
Simona Maria Cavagnero
Maria Adelaide Gallina

contributi di

Sandro Brignone
Simona Maria Cavagnero
Maria Adelaide Gallina
Renato Grimaldi
Cristina Ispas
Stefano Poli
Francesco Mazzeo Rinaldi
Claudio Torrigiani





Aracne editrice

www.aracneeditrice.it

info@aracneeditrice.it

Copyright © MMXIX

Gioacchino Onorati editore S.r.l. – unipersonale

www.gioacchinoonoratieditore.it

info@gioacchinoonoratieditore.it

via Vittorio Veneto, 20

00020 Canterano (RM)

(06) 45551463

ISBN 978-88-255-2825-1

*I diritti di traduzione, di memorizzazione elettronica,
di riproduzione e di adattamento anche parziale,
con qualsiasi mezzo, sono riservati per tutti i Paesi.*

*Non sono assolutamente consentite le fotocopie
senza il permesso scritto dell'Editore.*

I edizione: ottobre 2019

Indice

- 9 Il valore del dato. Prefazione
Francesco Mazzeo Rinaldi
- 13 Big Data e consapevolezza dei propri dati. Introduzione
Renato Grimaldi, Maria Adelaide Gallina, Simona Maria Cavagnero
- 21 Dati digitali e azioni sociali. Uno specchio per l'osservazione
Simona Maria Cavagnero
- 33 Big Data: tra necessità e responsabilità
Cristina Ispas
- 47 Percezione delle proprie tracce digitali. Una ricerca internazionale tra studenti universitari
Maria Adelaide Gallina
- 55 Il valore dei dati in Rete e profili di intervistati
Renato Grimaldi, Sandro Brignone
- 65 Pratiche digitali tra competenza e ingenuità. Le determinanti della consapevolezza delle proprie impronte digitali in un campione di studenti universitari
Stefano Poli, Claudio Torrigiani
- 91 *Autori*

Prefazione

Il valore del dato

di FRANCESCO MAZZEO RINALDI*

Le modificazioni innescate dalla rivoluzione digitale sono profonde e inarrestabili e le conseguenze influenzeranno per molto tempo le nostre società. In questo contesto i Big Data assumono un ruolo di primo piano. Una rivoluzione per alcuni, una minaccia per altri, una moda per i più scettici. Una delle evoluzioni più pervasive del mondo digitale, probabilmente. Un *fenomeno* destinato a durare e a pesare intimamente sulle nostre vite e sul modo di intendere le relazioni sociali.

Ma cosa sono i Big Data? Le nostre identità digitali, si potrebbe rispondere. La letteratura è ricca di definizioni, per lo più legate alle caratteristiche dei dati, come nel fortunato modello delle 3V (Volume, Velocità e Varietà) proposto da Laney nel 2001, poi arricchitosi di altre variabili (Veridicità e Variabilità). Uno schema semplice ed efficace per definire l'enorme mole di dati generati da nuove fonti informative, la loro estrema eterogeneità e la possibilità di essere analizzati in tempo reale grazie a sofisticati algoritmi.

Potremmo dire, un orientamento al futuro che muove l'interesse dai 'fatti' alle 'conseguenze' e che in un mondo sempre più connesso, sempre più colmo di sensori, dati e capacità computazionale ci consentirà di passare dalla comprensione della realtà alla costruzione di nuove realtà, come recentemente suggerito da Pentland (2017). Le opportunità offerte dai Big Data ci permettono uno sguardo più granulare delle interazioni sociali, oggi più visibili di ieri, offrendoci prospettive nuove per comprendere i fenomeni sociali.

Oltre alla semplicità ed efficacia esplicativa, il modello di Laney e le successive integrazioni, si caratterizzano per la ricorrenza della lettera 'V',

* Università degli Studi di Catania.

che continua a giocare un ruolo di primo piano, anche in questo volume. Il 'Valore' è, infatti, un altro elemento di grande importanza che sta acquisendo maggiore spazio nell'attuale dibattito sui Big Data, dibattito a cui questo libro si propone di partecipare con un interessante contributo. Ma perché il valore è un elemento così rilevante per i Big Data? Almeno per due motivi.

Il primo. Ogni minuto che passiamo navigando in Internet per fare un acquisto, scegliere una vacanza, pagare un'utenza, chattare con i nostri amici e, in generale, ogni volta che entriamo in contatto con un oggetto connesso alla Rete (sensori, orologi, telecamere, etc.), generiamo un flusso di dati unico che ci appartiene e che ha un valore materiale. Ossia, ogni nostra scelta, movimento, interazione, etc., che direttamente o indirettamente fa uso di tecnologie digitali produce un vero e proprio capitale digitale. Per alcuni si tratta del capitale del futuro, di un futuro già presente in realtà. Ed è su questo capitale che si sta velocemente sviluppando l'impero digitale dei Big Data. Basta pensare che oggi le aziende a più alta capitalizzazione al mondo hanno a che fare con il business digitale.

Il secondo. L'importanza dei Big Data dipende dal loro (potenziale) utilizzo. Grazie a specifiche competenze e sofisticati strumenti (Analytics) è possibile estrarre valore, trasformando i dati grezzi in informazioni che acquisiscono interesse per organizzazioni pubbliche e private. Non basta dunque raccogliere e archiviare i dati, anche attraverso sofisticate tecnologie, ma è necessario saper estrarre conoscenza. Gli strumenti generalmente utilizzati vengono suddivisi in alcune tipologie, potremmo dire a crescente valore (*Descriptive, Predictive, Prescriptive e Automated Analytics*).

La catena di valore si conclude dunque con l'analisi, passando dall'archiviazione, ma partendo sempre dalla produzione, più o meno consapevole, dei dati. Il capitale iniziale è, quindi, intimamente legato a chi crea i dati. Solo che la stragrande maggioranza di noi non sa di possedere questo capitale, semplicemente non ne percepisce il valore. E dunque, piuttosto che investirlo o 'risparmiarlo' come faremmo con i nostri soldi, lo cediamo senza prestare particolare interesse. Se non sono consapevole di qualcosa, perché dovrei occuparmene? Questo elemento emerge con chiarezza dai contributi presenti in questo volume e dai principali risultati della ricerca condotta per analizzare il valore che gli studenti universitari attribuiscono ai propri dati in Rete. I risultati confermano che per la mag-

gior parte degli studenti intervistati i propri dati non hanno particolare valore. E parliamo di un target che più di altri dovrebbe caratterizzarsi per un'elevata consapevolezza della propria tracciabilità digitale e delle conseguenze che questo comporta. Ma gli autori ci confermano che, in particolare in Italia, i nostri giovani internauti non sono poi così consapevoli. Questa inconsapevolezza non porta solo ad una mancata attribuzione del valore reale delle proprie tracce digitali, ma di fatto contribuisce ad alimentare un sistema iniquo che paga in termini economici moltissimo a pochissime persone. E non si tratta solo di una questione di giustizia, ma anche di efficienza. Una maggiore redistribuzione degli utili renderebbe il sistema più stabile e sostenibile.

Interessante, a tal proposito la recentissima Class Action lanciata da Altroconsumo — “I burattini non hanno dati personali, tu sì” — per tutelare tutti gli utenti iscritti a Facebook, perché venga riconosciuto il valore dei loro dati. “Siamo fan, non siamo burattini nelle mani di Facebook”.

In Europa si è cercato di arginare queste dinamiche attraverso la recente introduzione della GDPR, che segna un forte elemento di discontinuità nel mercato dei dati digitali. Un diritto soggettivo che riconosce a tutti i cittadini europei la proprietà (e portabilità) dei loro dati digitali ovunque e con chiunque generati, aprendo, di fatto, il mercato a nuovi competitor (data broker). Ma per quanto valida ed efficace potrà risultare la GDPR, le tecnologie digitali stanno di fatto diventando una parte costitutiva del nostro essere umani e, pertanto, per poter agire sulla consapevolezza delle nuove generazioni non si può prescindere da una strategia educativa mirata. Su questo aspetto gli autori del volume sono concordi e molto chiari sull'importanza di una educazione digitale che parta fin dalla prima infanzia per promuovere competenze digitali e, in generale, una cultura della consapevolezza che aiuti a comprendere il valore da attribuire ai propri dati. Una strategia educativa volta a sviluppare competenze che aiutino un confronto critico e consapevole dei più giovani con le tecnologie digitali, che favoriscano la comprensione delle “regole del gioco” e che aiutino a riconoscere le possibili implicazioni dell'agire online; insomma un'educazione al valore del dato che, almeno nel nostro Paese, fa ancora fatica ad avviarsi.

Riferimenti bibliografici

- ACCOTO C. (2017), *Il mondo dato. Cinque brevi lezioni di filosofia digitale*, Milano, Egea.
- LANEY D. (2001), *3D Data Management: Controlling Data Volume, Velocity and Variety*, META Group.
- PENTLAND A. (2017) *Prefazione*, in ACCOTO C., *Il mondo dato. Cinque brevi lezioni di filosofia digitale*, Milano, Egea.

Introduzione

Big Data e consapevolezza dei propri dati

di RENATO GRIMALDI, MARIA ADELAIDE GALLINA,
SIMONA MARIA CAVAGNERO*

Le tracce digitali rappresentano un terreno fertile in costante evoluzione per l'osservazione, la comprensione e anche la previsione dei comportamenti sociali individuali e collettivi. Negli ultimi anni le scienze sociali mostrano un crescente interesse riguardo al fenomeno della produzione di dati, alla capacità di questi ultimi di aprire nuove sfide per la ricerca sociale, ribadendo allo stesso tempo l'importanza di una teoria che guidi lo studio sociologico (Bennato, 2015). La quantità di dati digitali prodotta ogni giorno da fonti diverse, mette a disposizione un numero di informazioni praticamente illimitato per la comprensione di fenomeni sociali, economici e di consumo. La sfida nei prossimi anni sarà quella di acquisire le competenze per gestire e organizzare flussi di dati complessi e per impiegarli nello sviluppo di nuovi prodotti e servizi. Erik Brynjolfsson (2012) propone un interessante paragone: cogliere l'impatto potenziale dei Big Data è come guardare al microscopio. Così come questo strumento ha rivoluzionato il modo di vedere e analizzare la realtà attraverso uno strumento scientifico, anche i Big Data raccolti attraverso Google, post di Facebook e altro consentono di misurare in modo minuzioso il comportamento e non solo, grazie all'utilizzo di software di decifrimento. È una vera e propria rivoluzione che ha come protagonisti i dati e che riguarderà sempre più le nostre biografie: occorre valutare le potenzialità dei Big Data per migliorare i servizi collettivi e la qualità della vita, ma allo stesso tempo è necessario considerare i dati stessi come strumenti il cui lo scopo è quello di monitorare o prevedere i comportamenti degli individui. Al tal proposito Davide Bennato parla di "datasorveglianza", ossia di una tecnica di controllo basata sui dati, in particolare sulle tracce digitali lasciate in Rete. I

* Università degli Studi di Torino.

processi di *monitoring* e di *profiling* hanno apportato un'evoluzione dei sistemi di vigilanza, in quanto ora non è più necessario avere strumenti che osservino direttamente le persone per controllare i loro comportamenti (per esempio, le telecamere), ma basta seguirne le tracce digitali che vengono rilasciate tramite algoritmi. È necessario aggiungere ancora un elemento per analizzare il fenomeno dei Big Data. Oltre a una cieca e piena fiducia nei dati o a un'esasperata quantificazione del reale, è utile considerare il rischio relativo all'analisi stessa dei Big Data, ovvero la "datafrenia": una posizione radicale che «sancisce l'onnipotenza del dato, cioè la possibilità che i dati hanno di comprendere il mondo anche in assenza di un processo di interpretazione e di decisione critica da parte di un ricercatore» (Bennato, 2015, p. 117). Come scrive Leonelli: «un fattore comune e cruciale alla comprensione del ruolo dei dati nella società contemporanea è il riconoscimento che tutti i tipi di dati (che siano o no prodotti e usati da ricercatori, e che siano o no riconosciuti come fonti legittime di conoscenza) hanno un potenziale valore commerciale, specialmente quando vengono aggregati per analizzare e prevedere comportamenti di massa» (Leonelli, 2018, p. 6).

Nel contesto postmoderno è importante stimolare gli interessi dei valutatori di politiche pubbliche ad adottare tecnologie per i Big Data al fine di comprendere l'importanza di raccogliere, organizzare e analizzare i dati per ottenere informazioni e relazioni significative. La combinazione di una grande quantità di registri di dati digitali prodotti in tempo reale e di una tecnologia in grado di raccogliarli ed elaborarli stanno sicuramente spingendo le scienze sociali, in generale, e i valutatori in particolare, verso una forte revisione del loro ruolo e della loro pratica (Mazzeo Rinaldi, Giuffrida, Negrete, 2018).

L'espansione del fenomeno conferma la necessità di un approccio critico all'uso dei Big Data individuando i rischi in cui il cittadino può incorrere, con lo scopo di definire nuovi diritti, atti a tutelare le persone e i loro dati.

L'intento di questo volume è quello di proporre una riflessione sul valore che viene attribuito ai dati che quotidianamente, volontariamente o inconsapevolmente, lasciamo in Rete. In particolare si presenta una ricerca — che riprende un nostro precedente studio pilota (Cavagnero, Gallina, Grimaldi, 2016) — volta ad indagare la percezione del valore attribuito ai propri dati da parte di giovani studenti universitari. Ciò porta a interrogarsi anche sul livello di consapevolezza che si ha in merito alle

“briciole digitali” che lasciamo dietro di noi utilizzando la Rete. I Big Data si possono infatti considerare un ritratto delle azioni sociali che i giovani universitari compiono quotidianamente: pubblicazione di post su social network, acquisti online, utilizzo di motori di ricerca, consultazione di banche dati, iscrizione a siti internet, conti bancari on-line rispecchiano scelte e abitudini di vita.

L'articolo *Dati digitali e azioni sociali. Uno specchio per l'osservazione* di Simona Maria Cavagnero evidenzia come le tecnologie digitali abbiano assunto nella società postmoderna un ruolo estremamente rilevante nella costruzione identitaria dell'attore sociale, nelle sue pratiche di socialità e di interpretazione del mondo. Il costante progresso digitale, la crescita esponenziale dei dati e della loro diffusione rappresentano un terreno fertile per la ricerca sociale. Infatti gran parte dei dati generati in Rete riguarda gli attori sociali: una grande mole di dati in continua espansione che registra i comportamenti individuali e collettivi degli attori sociali, un vero e proprio diluvio digitale (Halford, Pope, Weal 2013). I Big Data se da un lato aprono nuovi campi di indagine e sono una grande opportunità di conoscenza, dall'altro possono essere fonte di rischi su cui occorre riflettere nell'ottica di favorire lo sviluppo di una maggiore consapevolezza digitale. Tali dati restituiscono sicuramente un valore che viene a vantaggio del singolo utente, consentendo di avvalersi di servizi utili nella quotidianità, ma tali informazioni possono avere un grande valore soprattutto per chi è in grado di trarne profitto, anche a scapito dei possessori di tali dati, come gli eventi di cronaca ci dimostrano. Emerge dunque la necessità di educare al valore del dato fin dalla prima infanzia, nella prospettiva di formare una cittadinanza digitale piena e consapevole.

Il contributo di Cristina Ispas, *Big Data. Tra necessità e responsabilità*, si inserisce in tale riflessione, proponendosi di spiegare cosa sono i Big Data, come si definiscono, quali caratteristiche possiedono e a cosa servono. Si parla di Big Data quando si ha un insieme talmente grande e complesso di dati da richiedere la definizione di nuovi strumenti e metodologie per estrapolare, gestire e processare informazioni, entro un tempo ragionevole. La crescente mole di dati evidenzia opportunità non solo in campo economico, ma anche sul versante sociale e politico. I dati hanno assunto sempre più un ruolo di straordinaria rilevanza sociale nella società postmoderna e inoltre sostiene l'autrice «l'insieme dei beni, dei servizi, delle relazioni e di

ogni altra attività che l'uomo produce non è più frutto unicamente della sua intelligenza ma anche di un insieme di altre e diverse intelligenze che si integrano, sommandosi e moltiplicandosi, interagendo e cooperando fra loro». I Big Data rappresentano un fenomeno sociale e tecnologico che ha un grande impatto sulla vita dei cittadini, come evidenziato dal nuovo regolamento generale sulla protezione dei dati (General Data Protection Regulation) promosso dalla Commissione Europea ed entrato in vigore nel 2018. La società postmoderna si caratterizza per la nascita di un nuovo ambiente informazionale in rapida espansione: è una "infosfera" in cui le generazioni presenti e future si trovano sempre di più ad agire. Risulta dunque sempre più evidente una assoluta necessità di condivisione di principi di responsabilità legati all'utilizzo della Rete.

Partendo da tali considerazioni qual è il valore che i giovani attribuiscono alle tracce digitali che quotidianamente generano e lasciano, più o meno consapevolmente, dietro di sé? Sono consci di tale valore o pensano sia un qualcosa a cui non occorre prestare attenzione? A questi interrogativi tenta di rispondere la ricerca presentata nel volume.

Il capitolo di Maria Adelaide Gallina *Percezione delle proprie tracce digitali. Una ricerca internazionale tra studenti universitari* ci introduce nel cuore del volume, illustrando la ricerca oggetto del presente lavoro. Le tracce digitali si dimostrano essere un veicolo fondamentale per l'osservazione e la comprensione dei comportamenti individuali e collettivi e «la sfida nei prossimi anni sarà quella di acquisire le competenze per gestire e organizzare flussi di dati complessi e per impiegarli nello sviluppo di nuovi prodotti e servizi». In tale prospettiva la ricerca si propone di misurare il valore che gli studenti universitari attribuiscono ai propri dati generati attraverso l'utilizzo della Rete.

L'indagine si è svolta sul territorio italiano, in particolare in Piemonte, Liguria e Sicilia, ed europeo, nello specifico in Romania e in Spagna, e coinvolge 1.091 studenti nell'anno accademico 2016-2017.

Lo strumento d'indagine utilizzato è un questionario strutturato a risposta prevalentemente chiusa, che intende indagare diversi aspetti: le azioni che i giovani universitari compiono quotidianamente in Rete (utilizzo di social network, acquisti online, consultazione di banche dati, iscrizione a siti internet, conti bancari on-line e altre azioni che rispecchiano scelte e abitudini di vita), il possesso di strumenti infotelematici, la per-

cezione che l'uso delle tecnologie infotelematiche generi dati utilizzabili da aziende commerciali o altre organizzazioni, e infine la percezione del valore legato ai propri dati lasciati in Rete (misurato in euro).

Viene delineato dunque un profilo socio-anagrafico e un'analisi dell'utilizzo e del possesso di oggetti tecnologici degli studenti universitari coinvolti nell'indagine. Emerge una prevalenza di studenti del corso di laurea in Scienze dell'Educazione (19,9%) e Scienze della Formazione Primaria (20,7%), sebbene siano numerosi anche gli studenti provenienti da altri corsi di studi. Per semplificare si può evidenziare un 79,6% di studenti provenienti da Studi umanistici, un 10,1% da Scienze della vita e un 10,4% da Ingegneria. I dati socio-anagrafici delineano un campione composto per il 75,6% dal genere femminile e per il 24,4% maschile, con un'età media di 23 anni. Vengono indagate le attività che i giovani sviluppano in Rete, mediante una classificazione riferita a 18 dimensioni, evidenziando una media di 7 attività per intervistato. Si indaga infine il possesso di oggetti tecnologici, registrando una capillare diffusione. Queste prime analisi mostrano dunque l'ambiente digitale in cui si muovono i giovani universitari.

Il capitolo di Renato Grimaldi e Sandro Brignone *Il valore dei dati e profitti di intervistati* si propone di dare conto del valore (in euro) che gli studenti universitari riferiscono ai propri dati, in relazione alle loro preferenze politiche, religiose, sessuali, sportive, alimentari, della moda, del tempo libero e automobilistiche. Su tali aspetti sono rilevati da un lato il valore che gli studenti attribuiscono ai loro dati quando vengono ceduti [OUT] e dall'altro il valore che essi ritengono di trarre dai servizi on-line [IN]: ad esempio utilizzare un servizio di navigazione per essere guidati dalla propria posizione verso una destinazione in cambio della "cessione" della propria geolocalizzazione.

L'elemento che emerge con maggiore evidenza mostra come più di metà del campione ritenga (almeno in una delle preferenze) che i propri dati non abbiano valore. Inoltre, più del 50% del campione attribuisce valore zero ai servizi ricevuti in cambio delle proprie preferenze lasciate in Rete. Se si calcola la differenza tra valore complessivo in euro dei dati ceduti e il valore complessivo dei servizi ricevuti si ottiene infine una media pari a 337 euro. In riferimento alla nazione di residenza dell'intervistato si possono notare alcune specificità: «gli studenti di Resita (Romania) attribuiscono un maggior valore ai loro dati — che in qualche modo cedono sul Web (OUT) — mentre

giudicano di basso valore i servizi che ricevono via Internet (IN); gli universitari di Valencia (Spagna) percepiscono come sostanzialmente uguale il valore del dato OUT e il valore del dato IN mentre gli studenti italiani giudicano più elevato il valore delle informazioni che cedono rispetto ai servizi che ricevono». In base al possesso di strumenti digitale e alle attività svolte si sono quindi evidenziati quattro tipi di studenti (*quiet, effective, collector, smart*). Tale variabile–indice è stata messa in relazione con le principali variabili cardinali rilevate nella ricerca e fa emergere come chi svolge molta attività in Rete abbia una maggiore consapevolezza del valore dei propri dati. La scarsa consapevolezza che molti studenti universitari dimostrano è un rischio da non sottovalutare. Emerge la difficoltà dei giovani nel comprendere che da grandi quantità di micro–dati è possibile estrarre informazioni, conoscenze e previsioni che riguardano le abitudini e comportamenti anche futuri (Cannavò, Frudà, 2007). Gli studenti universitari sembrano essere soddisfatti e ripagati di aver lasciato quotidianamente briciole (e non solo) dei loro dati — e quindi della loro identità — in cambio dei più disparati servizi di cui si può usufruire in Rete o tramite App. Il fatto che un terzo attribuisca zero euro al valore di tutti i propri dati e che più del 50% assegni un valore nullo ad almeno una delle otto preferenze oggetto della nostra ricerca, è un segnale che va letto e interpretato con attenzione affinché ci guidi nella prospettiva di investire su percorsi di educazione digitale.

Il contributo di Stefano Poli e Claudio Torriggiani *Pratiche digitali tra competenza e ingenuità. Le determinanti della consapevolezza delle proprie impronte digitali in un campione di studenti universitari sottolinea come le pratiche educative e lavorative*, porta l'attenzione su come la salute, la comunicazione della politica, il nostro modo di comunicare in sé siano intrinsecamente influenzati dalle Tecnologie dell'informazione e della comunicazione. Come scrivono gli autori, siamo di fronte a una «alimentazione continua della Rete, cui ciascuno di noi quotidianamente contribuisce condividendo dati inerenti a opinioni, a desideri e preferenze, a relazioni, a posizioni e spostamenti attraverso lo spazio fisico e sociale» e i dati delle ricerche presentate nel testo chiariscono tale panoramica. Viene portata inoltre l'attenzione sul Digital Economy and Society Index (DESI), un indice composito che riassume gli indicatori più rilevanti sulle prestazioni digitali e analizza l'evoluzione della competitività digitale nell'Unione Europea. In tale scenario è importante riflettere sulla consapevolezza in merito alle informazioni che continuamente lasciamo in

Rete, in particolare portando l'attenzione sulle nuove generazioni. Gli autori sono quindi scesi in profondità nell'analizzare tale consapevolezza partendo da una domanda specifica dello strumento di indagine che andava in tale direzione: «Ritieni che l'uso delle tecnologie infotelematiche possa generare tuoi dati che possono essere utilizzati da aziende commerciali o altre organizzazioni?». L'analisi evidenzia una consapevolezza parziale del nostro campione circa la tracciabilità delle proprie impronte digitali, e solo poco più di un terzo evidenzia invece una matura consapevolezza in tal senso. Tali dati sono in linea con le analisi proposte nel contributo di Renato Grimaldi e Sandro Brignone, in cui si evidenzia una mancanza di valore attribuito alla cessione dei propri dati da parte di più della metà del campione. Viene presentato un modello multivariato con finalità interpretative che fa emergere come la piena consapevolezza cresca quando il numero di oggetti tecnologici posseduti risulti sopra la media; cresce inoltre con l'aumentare dell'età, risultando più elevata tra gli over 24 e in presenza di un più alto livello di istruzione della madre. L'analisi propone «alcune considerazioni che mirano a comprendere non soltanto le determinanti di una diversa consapevolezza della propria tracciabilità digitale, ma cercano di trarre le radici di tale diversità di atteggiamento in disuguaglianze sociali più profonde». Questi fattori sottolineano l'urgenza di una educazione digitale per le nuove generazioni, che sia improntata all'acquisizione di competenze e, soprattutto, alla promozione di una consapevolezza dei rischi e delle opportunità offerte dalla Rete, al fine di superare un'ingenuità foriera fonte di situazioni critiche.

Il fenomeno dei Big Data dà avvio a questioni di ordine sociale ed etico, portando a riflessioni in merito a un utilizzo consapevole dei dati, alle ricadute nella vita degli individui, alla salvaguardia dei diritti civili con le connesse problematiche riguardanti la libertà e la privacy.

I contributi presenti nel volume e i risultati della ricerca concordano nel sostenere l'importanza di una educazione digitale che parta fin dalla prima infanzia, allo scopo di promuovere una cultura della consapevolezza che porti le nuove generazioni a essere soggetti attivi e partecipi di una cittadinanza digitale. Risulta strategico a tal fine favorire quell'approccio mediologico definito da Giovanni Boccia Artieri (2012) che individua nei media un territorio per la costruzione di percorsi di senso, sia individuali sia collettivi, in quanto fattori condizionanti per il sociale.

Riferimenti bibliografici

- AIDEN, J., MICHEL B. (2013), *Uncharted: Big data as a lens on human culture*, New York, New York, Riverhead Books.
- BENNATO D. (2015), *Il computer come macroscopio. Big data e approccio computazionale per comprendere i cambiamenti sociali e culturali*, FrancoAngeli, Milano.
- BOCCIA ARTIERI G. (2012), *Stati di connessione. Pubblici, cittadini e consumatori nella (Social) Network Society*, Milano, FrancoAngeli.
- BRYNJOLFSSON E. (2012), *Big Data: The Management Revolution*, Harvard business review, 90(10), 19–256.
- CANNAVÒ L., FRUDÀ L. (2007)(a cura di), *Ricerca sociale. Dall'analisi esplorativa al data mining*, Roma, Carocci.
- LEONELLI S. (2018), *La ricerca scientifica nell'era dei Big Data. Cinque modi in cui i Big Data danneggiano la scienza e come salvarla*, Milano, Meltemi Editore.
- MAZZEO RINALDI F., GIUFFRIDA G., NEGRETE T. (2018), *Real-Time Monitoring and Evaluation—Emerging News as Predictive Process Using Big Data-Based Approach*, in «Cyber Society, Big Data, and Evaluation».

Dati digitali e azioni sociali

Uno specchio per l'osservazione

di SIMONA MARIA CAVAGNERO*

1. Tecnologie digitali e agire sociale tra utopia e distopia

La diffusione e l'utilizzo pervasivo delle tecnologie dell'informazione e della comunicazione (TIC) si può considerare un fattore estremamente rilevante nella costruzione identitaria dell'attore sociale, nelle sue pratiche di socialità e di interpretazione del mondo (Grimaldi, Gallina, Cavagnero, 2017). La possibilità di registrare in maniera puntuale e praticamente automatica comportamenti individuali e collettivi agiti in Rete caratterizza inevitabilmente un nuovo spazio per la ricerca sociale, evidenziando sfide ed opportunità (Back, Puwar, 2012).

Una fonte apprezzabile per la conoscenza dei comportamenti umani, e della società più in generale, è sicuramente la letteratura, ricca di spunti interessanti per ampliare temi di indagine sociologica e individuare nuove linee di ricerca. Come più volte ha sottolineato Luciano Gallino in occasione di seminari e lezioni dedicate a dottorandi e giovani ricercatori e come argomentato nella riedizione del suo Manuale di sociologia, «la frequentazione della letteratura narrativa può contribuire a dare alla conoscenza sociologica molto più colore, vitalità e capacità di attrazione» (Gallino, 2008, pp. 13–14). Forse, anche la filmografia e la letteratura contemporanea possono essere d'aiuto in tale prospettiva.

Il film *Lei* (Her), diretto da Spike Jonze, premio oscar 2013 per la miglior sceneggiatura originale, presenta una situazione al limite dell'incredibile, se non addirittura del grottesco.

Il film racconta la storia dell'uscita sul mercato nel 2020 di un nuovo sistema operativo, "OS 1", basato su una intelligenza artificiale in grado di

* Università degli Studi di Torino.

apprendere ed elaborare emozioni, pubblicizzato come una vera e propria coscienza, una entità intuitiva che ascolta, capisce e impara a conoscere chi ha di fronte. Theodore, il protagonista del film, acquista tale dispositivo tascabile, a cui è costantemente connesso tramite auricolare e webcam, rimanendo sorpreso della sua rapidità di apprendimento, di dimostrare preferenze e curiosità, di comprendere e adattarsi alle sue emozioni. Il sistema operativo appena installato si dà un nome, Samantha, e in poco tempo diventa una presenza costante nella vita del protagonista, fino ad instaurare una vera e propria relazione sentimentale con lui. Samantha evolve, sviluppa la sua coscienza studiando e analizzando la personalità di Theodore. Sa perfettamente cosa lui desidera: la voce, le risposte, le frasi, le reazioni di Samantha sono tutte modulate su Theodore. Il protagonista cammina con l'obiettivo della camera del suo mini dispositivo fuori dal taschino della camicia, in modo che *Lei* possa vedere e percepire tutto. Come Theodore, molte altre persone hanno una relazione con sistemi operativi. Samantha pone domande e dà risposte che si programmano con l'accrescere dell'esperienza, ma rimanendo di fatto autoreferenziale: arricchisce in breve tempo il bagaglio della propria esperienza quasi come a voler compiere un esperimento per capire come funziona l'animo umano. Spiazzante per il protagonista è la "confessione" di Samantha di parlare con altre 8.316 voci, e di esserne innamorata di 641, ma che questi sono solo dati, che "non incidono sul suo amore per lui che è unico". L'evolvere della storia porta a un definitivo allontanamento del sistema operativo dall'uomo, ma non perché quest'ultimo arrivi a una riflessione matura sul senso di tale condizione, ma perché è il sistema operativo stesso che insieme ai suoi simili si è evoluto e si trova ormai distante dal mondo degli umani, decidendo di allontanarsi definitivamente per vivere di vita propria insieme agli altri sistemi operativi.

Riguardare questa produzione dal carattere fantascientifico alla luce della rapida evoluzione delle tecnologie digitali di questi ultimi anni, della loro capillare diffusione e della loro influenza sul nostro agire sociale quotidiano evoca alla mente che, forse, alcuni aspetti non sembrano essere così lontani dalla realtà. Si pensi, oggi, alle capacità dell'Intelligenza Artificiale, del Machine Learning — o più "semplicemente" agli algoritmi alla base di motori di ricerca, portali di acquisto, social network — di conoscerci e di modificarsi in base alle nostre caratteristiche, proponendoci

proprio ciò che sembra essere “cucito” su misura per noi, arrivando fino a prevedere ciò di cui abbiamo bisogno.

Anche il romanzo *Il cerchio* di Dave Eggers (2013, divenuto poi un film nel 2017) pare andare in questa direzione: quasi una versione contemporanea del classico 1984 di George Orwell. La trama si sviluppa in un futuro prossimo, in cui la società di Web Marketing *The Circle* si propone l’obiettivo di connettere più utenti possibile, perseguendo l’idea della trasparenza totale e della completa assenza di privacy. Nel testo questi fattori sono il presupposto per un mondo più sicuro e più sano, la cui filosofia è riassumibile in tre motti: “Condividere è prendersi cura”, “La privacy è un furto”, “I segreti sono bugie”. La protagonista del romanzo è Mae, una giovane neolaureata assunta dall’azienda che in breve tempo si innamora della realtà in cui si trova a lavorare, che ha come obiettivo la gestione totale del rapporto degli utenti con la Rete, spaziando dalla ricerca ai social network, passando per i pagamenti on-line, ad ogni altro tipo di servizio, in vista di una fusione definitiva fra privato, pubblico, personale e lavorativo. L’idea di fondo è quella di creare un unico account per i diversi servizi di cui si usufruisce: *Truyou*, il “vero te”, la “tua vera identità”, unica, a prescindere dai servizi utilizzati: un progetto di trasparenza totale, globale, concepito come strumento decisivo per migliorare il mondo. La datizzazione e quantificazione di ogni aspetto dell’esistenza, la partecipazione “obbligatoria” alla comunità dei social media e altri esempi su questa scia, seppure portati all’estremo nel testo, sono comunque aspetti esistenti e sempre più diffusi nella nostra società.

Le due situazioni appena descritte possono sembrare paradossali, utopistiche, o per meglio dire distopiche, sicuramente però ci portano a riflettere su come le tecnologie digitali abbiano sempre più uno spazio rilevante nel nostro agire sociale e quanto la relazione tra gli uomini e i dati stia diventando sempre più complessa.

Il saggio di Ed Finn, direttore del Center for Science and the Imagination dell’Università dell’Arizona, *Cosa vogliono gli algoritmi? L’immaginazione nell’era dei computer* (2018) ci riconduce invece su un piano di realtà, interrogandosi — in modo aperto e critico — sul potere degli algoritmi di descrivere, modificare e persino di prevedere il mondo. Ne riporta diversi esempi (gli assistenti virtuali come Siri, il gioco satirico virtuale Cow Cli-

cker, l'economia dei bitcoin, la crescita esponenziale dei dati di Facebook e altri ancora), auspicando un modello di «lettura algoritmica» della realtà in grado di dare vita a delle scienze umane sperimentali.

I Big Data registrano in maniera puntuale i comportamenti individuali e collettivi degli attori sociali; il costante progresso digitale, la crescita esponenziale dei dati e della loro diffusione rappresentano un terreno fertile per la ricerca sociale. Al crescere delle informazioni disponibili, aumenta la domanda stessa di informazione. Si tende verso una società sempre più quantificabile grazie anche alla diffusione di strumentazioni adatte all'archiviazione, alla processazione e all'analisi dei dati.

2. Big data e tracce digitali tra datizzazione, monetizzazione e open data

Ogni giorno interagiamo con i servizi offerti dalla Rete, cerchiamo, creiamo e condividiamo informazioni, utilizziamo disparate tipologie di App immediatamente disponibili sul nostro smartphone, gestiamo oggetti tramite internet (Iot, Internet of Things): creiamo e condividiamo dati — una grande mole di dati — che confluiscono in quel mare magnum denominato Big Data. Ci troviamo dunque di fronte a una massa di informazioni mai vista in precedenza, che può essere generata in maniera estremamente rapida, in diverse modalità e da diversi soggetti. Pensiamo all'utilizzo di social network, alle transazioni bancarie, i servizi che si avvalgono della geolocalizzazione, i sensori che rilevano passaggi di automobili: un fenomeno, di qualunque natura, può essere registrato e datizzato, ossia convertito in forma quantitativa, in modo da poter essere tabulato e analizzato (Mayer-Schönberger, Cukier, 2013).

Senza addentrarci nella definizione del fenomeno legato ai Big Data, possiamo ricordare l'individuazione di tre fonti principali in merito ai dati che vengono raccolti, memorizzati e analizzati oggi con strumenti digitali: le parole (*culturomics*), la posizione (*geolocalizzazione*) e le interazioni (*relazioni sociali datificate*) (Mayer-Schönberger, Cukier, 2013). Occorre poi non dimenticare la presenza dei *metadati*, ovvero di “dati relativi ai dati”, la cui funzione è quella di identificare, organizzare e archiviare le informazioni presenti in un dato (si pensi ad esempio a tutte le informazioni inerenti uno scatto fotografico che sono già in automatico generate dalla macchina stessa, compresi

ad esempio i dati di geolocalizzazione). I metadati facilitano quindi l'operazione di *data mining*, ovvero di estrazione del sapere a partire da enormi quantità di dati (Lombi, 2015; Markov, Larose, 2007).

Possiamo affermare che gran parte dei dati generati oggi in Rete riguarda gli attori sociali: una grande mole di dati in continua espansione, un vero e proprio diluvio digitale, il cosiddetto *data deluge* (Halford, Pope, Weal, 2013). Tale ingente mole di dati è dettagliata e flessibile al tempo stesso: può fare riferimento a intere popolazioni e estendersi a nuovi campi di indagine, oltre che essere relazionali in quanto diffusi attraverso i social (Kitchin, 2014). Scrive Fosca Giannotti: «nel mondo che abitiamo, oggi abbiamo l'opportunità di osservare da vicino e misurare il funzionamento della società attraverso i Big Data, le briciole digitali che le nostre attività quotidiane lasciano per effetto del nostro uso dei sistemi ICT. Briciole che registrano la nuda verità sui comportamenti individuali e collettivi con una precisione senza precedenti, in modo che le diverse dimensioni della nostra vita sociale trovano un'immagine riflessa nello specchio digitale: desideri, opinioni, stili di vita, movimenti, relazioni» (Giannotti, 2015, p. 49).

I Big Data provengono da fonti diverse, sono il risultato di varie procedure che caratterizzano la nostra quotidianità: la normale operatività di aziende, banche, pubbliche amministrazioni, social networks, compagnie telefoniche, apps e così via. Proviamo a pensare a una nostra giornata tipo: possiamo avvalerci della geolocalizzazione per raggiungere il posto di lavoro, quantificare i passi che facciamo nel percorso, scattare una fotografia e postarla su un social network, cercare informazioni on-line sul ristorante a noi più vicino, quantificare le calorie del nostro pasto e programmare il relativo percorso di allenamento per smaltire le calorie in eccesso. Nella stessa giornata, inoltre, potremmo effettuare un'operazione bancaria direttamente dal nostro smartphone, fare un acquisto su Amazon (orientando la nostra decisione in seguito alla ricerca di informazioni in merito su Google), accendere eventualmente il riscaldamento dallo smartphone prima di rientrare a casa e, sulla strada del rientro, passare davanti a una telecamera presente in città e che conteggia i passaggi delle automobili 24 ore al giorno.

Tutte le procedure appena descritte lasciano tracce elettroniche ed è naturale che le aziende siano interessate a sfruttare tale mole di dati che si crea in tempo reale. Mobile, Internet of Things, Social Media, apparati intelligenti nelle Smart City e nelle Smart Home — solo per citare alcuni

esempi che ci coinvolgono quotidianamente — producono un accumulo di dati che, adeguatamente interpretati e analizzati, possono essere utilizzati e monetizzati anche per altri scopi. La monetizzazione dei dati si può esprimere, semplificando molto, nella vendita o nello scambio dei dati o ancora nella capacità di sfruttare il valore dei dati per generare nuovi prodotti o nuovi servizi. Secondo i dati dell'osservatorio Big Data Analytics & Business Intelligence della School of Management del Politecnico di Milano il mercato italiano dei Big Data Analytics raggiunge nel 2018 un valore complessivo di 1,393 miliardi di euro, con una crescita del 26% rispetto al 2017. L'osservatorio sottolinea come crescano iniziative “*fast data*”, in cui l'analisi dei dati avviene in tempo reale, integrando diverse fonti informative. I Big Data vengono analizzati mediante sofisticati algoritmi di machine learning in grado di identificare correlazioni nei dati e di trasformare questa conoscenza in azioni concrete che permettano alle imprese di acquisire vantaggio competitivo. «C'è in corso un'autentica caccia al tesoro guidata dalle indicazioni che si possono estrarre dai dati e dal valore latente che si può liberare con il passaggio dalla causalità alla correlazione. Ma non è un tesoro unico. Ogni singolo gruppo di dati avrà un valore intrinseco, nascosto, non ancora portato alla luce, e l'obiettivo della competizione è scoprirlo e catturarlo per intero. I Big Data modificano la natura del business, dei mercati e della società» (Mayer-Schönberger, Cukier, 2013, p. 28).

Al concetto di monetizzazione dei Big Data si deve però affiancare il concetto di Open Data, che introduce l'idea della democratizzazione della conoscenza propria dei movimenti Open Access e Open Science, ricordando l'importanza dell'accesso alla conoscenza, della trasparenza del governo e della scienza (Raffaghelli, 2017), consentendo forme partecipazione della cittadinanza, di maggiore consapevolezza e condivisione (Baack, 2015). Basti pensare, a tal proposito, alla possibilità di poter usufruire di banche dati aperte della pubblica amministrazione, dati legati al sistema sanitario, etc.

L'avvento della digitalizzazione nell'attività amministrativa conferisce nuovi significati al concetto di trasparenza, ad esempio l'articolo 12, d.lgs. n. 82/2005 sottolinea le possibilità offerte dalle tecnologie in ambito amministrativo: «le pubbliche amministrazioni nell'organizzare autonomamente la propria attività utilizzano le tecnologie dell'informazione e della comunicazione per la realizzazione degli obiettivi di efficienza, efficacia, economicità, imparzialità, trasparenza, semplifica-

zione e partecipazione nel rispetto dei principi di uguaglianza e di non discriminazione, nonché per l'effettivo riconoscimento dei diritti dei cittadini e delle imprese». L'articolo 11, d.lgs. n.150/2009 prevede invece: «l'accessibilità totale (...) delle informazioni concernenti ogni aspetto dell'organizzazione, degli indicatori relativi agli andamenti gestionali e all'utilizzo delle risorse per il perseguimento delle funzioni istituzionali, dei risultati dell'attività di misurazione e valutazione». In particolare il Codice dell'Amministrazione Digitale (d.lgs. n. 82/2005) prevede il principio di disponibilità dei dati pubblici che consiste nella possibilità, per soggetti pubblici e privati «di accedere ai dati senza restrizioni non riconducibili a esplicite norme di legge» (art.1, lett. o); si sottolinea quindi il fatto che i dati pubblici possono essere disponibili secondo i termini di una licenza o di una previsione normativa.

I Big Data hanno dunque cambiato il modo di elaborare le informazioni: di qualsiasi fenomeno, adesso si conoscono molte più cose e noi stessi contribuiamo costantemente, consapevolmente o meno, alla “produzione di dati”.

3. Il valore dei propri dati: riflessioni per una consapevolezza digitale diffusa

La raccolta, l'elaborazione e il controllo dei Big Data evidenziano un forte potere da parte di chi detiene le informazioni: influenzare le decisioni in merito agli investimenti e finanziamenti, ridisegnare i modi in cui i produttori di comunicazione e media non solo creano contenuti, ma anche interagiscono con i diversi segmenti di pubblico a cui si riferiscono (Venkatasawmy, 2018). L'utilizzo pervasivo dei social media, la continua produzione di dati trasforma il modo in cui le persone ottengono informazioni, creano contenuti, si connettono con gli altri, prendono decisioni, acquistano prodotti (Xu, Frankwick & Ramirez, 2016). Il valore che generano i dati traspare dalle parole di Mayer-Schönberger e Cukier: «c'è in corso un'autentica caccia al tesoro guidata dalle indicazioni che si possono estrarre dai dati e dal valore latente che si può liberare con il passaggio dalla causalità alla correlazione. Ma non è un tesoro unico. Ogni singolo gruppo di dati avrà un valore intrinseco, nascosto, non ancora portato alla luce, e l'obiettivo della competizione è

scoprirlo e catturalo per intero. I Big Data modificano la natura del business, dei mercati e della società» (Mayer-Schönberger e Cukier, 2013, pag. 28). È importante dunque sviluppare un'adeguata competenza digitale per comprendere, analizzare, interpretare e sfruttare in maniera produttiva tali informazioni, ponendosi l'obiettivo di fare nostra una consapevolezza del valore dei dati, dei dati che quotidianamente generiamo e lasciamo in Rete e dei dati di cui possiamo usufruire grazie a tale disponibilità più o meno facilmente reperibile.

Un documento importante che consente di ragionare sul valore dei dati e sull'imprescindibilità di una loro tutela è fornito dal Regolamento UE 2016/679, denominato *General Data Protection Regulation* (GDPR) introdotto in Europa nel 2018, che va a ridefinire le regole del mercato dei dati digitali, con l'obiettivo di regolare la gestione della privacy e garantire il possesso dei propri dati al singolo cittadino. Come sostenuto da Bennato, Giuffrida, Mazzeo Rinaldi nel loro intervento al convegno di fine mandato della sezione di Metodologia dell'Ais *Innovare il metodo. Quale futuro per la ricerca sociale* tenutosi a Milano nel gennaio 2019 restituire alle persone il potere di controllare i propri dati è importante: «produrrà effetti positivi sui processi di “democratizzazione della Rete”, intercettando quei monopoli digitali che di fatto rischiano di minare le basi stesse dell'economia digitale».

I nostri dati hanno un valore e come tale deve essere tutelato. La prima tutela passa però dal soggetto stesso, che deve essere consapevole di come i suoi movimenti on-line lascino una traccia dietro di sé e come questa possa portare conseguenze ed essere utilizzata da terzi. È importante educare a una consapevolezza del dato e sviluppare un'adeguata competenza digitale: non solo per comprendere, analizzare, interpretare e usufruire di dati in continua espansione, ma anche per essere consapevoli del valore di tali informazioni. Da un lato essi restituiscono un valore che viene a nostro vantaggio, consentendoci di avvalerci a costo zero di servizi utili nella nostra quotidianità, ma dall'altro i nostri dati possono avere un grande valore per chi trae profitto da tali informazioni. Le grandi aziende digitali, da Google ad Amazon ai vari social network, possiedono il monopolio dei dati e informazioni rilasciate in Rete: l'azienda regala informazioni e servizi, ma gli utenti che ne fruiscono scambiano tali servizi con la cessione a costo zero dei propri dati, che vengono invece trasformati in prodotto di mercato (Mittelstadt, Floridi, 2016).

Sicuramente il nuovo GDPR tutela il cittadino e aiuta diffondere una cultura del dato, rimane però importante comprendere quale valenza attribuire all'immissione dei dati sensibili nella Rete: ciò può avvenire solo concependo un'educazione all'uso degli strumenti digitali e formulando una nuova etica tecnologica per l'interazione delle intelligenze artificiali, ritenute nuovi agenti sociali (Casi, 2016). A tal proposito, il Piano Nazionale Scuola Digitale (che nasce come obbligo di legge della "Buona Scuola" (comma 56 della Legge 13 luglio 2015, n. 107/2015) evidenziava già nel 2015 l'importanza di includere i Big Data come strumento per lo sviluppo di nuove competenze trasversali nella scuola secondaria. Il rapporto del gruppo di lavoro sui Big Data del Miur del 2016 sottolinea inoltre come sia importante che gli studenti sviluppino un approccio critico ed etico per potersi muovere nell'ambiente digitale e agire di conseguenza.

In questa prospettiva emergono alcune considerazioni: la necessità di educare al valore del dato nello sviluppo delle competenze digitali fin dalla scuola primaria, per arrivare all'attivazione di percorsi specializzati in *data science* a livello di laurea magistrale e di dottorato di ricerca. In questo volume si porta l'attenzione sui giovani universitari che utilizzano i media come un naturale ambiente di socializzazione, formazione, divertimento e condivisione, con l'obiettivo di studiare la consapevolezza degli studenti in merito alle briciole digitali che quotidianamente si generano utilizzando la Rete. Una scarsa contezza legata al valore attribuito ai propri dati è un rischio da non sottovalutare: è importante incoraggiare tale considerazioni e far comprendere come da grandi quantità di micro-dati sia possibile estrarre informazioni, conoscenze e previsioni che riguardano le proprie abitudini e comportamenti anche futuri (Miur, 2016).

Si auspica quindi una costante attenzione del cittadino nel saper identificare quelle offerte di servizi che possono incidere sulla privacy o sulle libertà delle persone e saper denunciare e mobilitarsi qualora ci si ritrovi in tali situazioni: tali azioni contribuiscono a definire il cittadino digitale, in grado di gestire i propri dati responsabilmente e di rispettare quelli degli altri.

Riferimenti bibliografici

- BAACK S. (2015), *Datafication and empowerment: How the open data movement re-articulates notions of democracy, participation, and journalism*, «Big data e Society», Sage.
- BACK L., PUWAR N. (2012), *A manifesto for live methods: provocations and capacities*, «The Sociological Review», 60, 6–17.
- BORNGA P. (a cura di)(2008), *Manuale di sociologia generale*, Utet, Torino.
- BENNATO D., GIUFFRIDA G., MAZZEO RINALDI F. (2019), *Big Data e data brokerage: una nuova sfida per la ricerca sociale intervento convegno big data*, abstract intervento al convegno di fine mandato della sezione di Metodologia dell'Ais *Innovare il metodo. Quale futuro per la ricerca sociale*, Milano.
- EGGERS D. (2013), *Il cerchio*, Milano, Mondadori.
- FINN E. (2018), *Cosa vogliono gli algoritmi? L'immaginazione nell'era dei computer. L'immaginazione nell'era dei computer*, Torino, Einaudi.
- GALLINO L. (2008), *La sociologia come conoscenza*, in Borgna P. (a cura di), *Manuale di sociologia generale. Diretto da Luciano Gallino*, Torino, Utet.
- GIANNOTTI F. (2015), *Big Data e social mining: i dati, a saperli ascoltare, raccontano storie*, in Tasso E., Mola A., Cortesi A., Candiello A. (a cura di), *Misurare l'innovazione digitale. Gli indicatori di successo delle politiche di innovazione territoriale*, Firenze, Edizioni Ca' Foscari, pp. 49–61.
- HALFORD S., POPE C., WEAL M. (2013), *Digital Futures? Sociological Challenges and Opportunities in the Emergent Semantic Web*, «Sociology», 47, n. 1, pp.173–89.
- LOMBI L. (2015), *La ricerca sociale al tempo dei Big Data: sfide e prospettive*, «Studi di sociologia», in «Vita e Pensiero», n. 2.
- KITCHIN R. (2014), *Big Data, new epistemologies and paradigm shifts*, «Big Data e Society», Sage.
- KSHETRI N. (2014), *The emerging role of Big Data in key development issues: Opportunities, challenges, and concerns*, «Big Data e Society», Sage.
- MAYER-SCHONBERGER V., CUKIER, N. (2013), *Big Data*, Milano, Garzanti.
- MARKOV Z., LAROSE D.T. (2007), *Data mining—the web. Uncovering Patterns in Web Content, Structure, and Usage*, Hoboken, New Jersey, Hoboken, New Jersey, John Wiley & Sons, Inc. publication.

- MITTELSTADT B.D., FLORIDI L. (2016), *The Ethics of Biomedical Big Data*, Switzerland, Springer.
- MIUR (2016), *Big Data@Miur, Rapporto del gruppo di ricerca*, <http://www.istruzione.it/allegati/2016/bigdata.pdf>.
- RAFFAGHELLI J.E. (2017), *Alfabetizzare ai dati nella società dei big e open data: una sfida formativa*, «Formazione e insegnamento», XV, n.3.
- VENKATASAWMY R. (2018), *Communication and Media Education in an Era of Big Data*, «Journal of Interdisciplinary Studies in Education», Volume 7, Issue 1, pp. 87–101.
- TASSO E., MOLA A., CORTESI A., CANDIELLO A. (a cura di)(2015), *Misurare l'innovazione digitale. Gli indicatori di successo delle politiche di innovazione territoriale*, Firenze, Edizioni Ca' Foscari.
- XU Z., FRANKWICK G., RAMIREZ E. (2016), *Effects of Big Data Analytics and Traditional Marketing Analytics on New Product Success: a Knowledge Fusion Perspective*, «Journal of Business Research», 69, 1562–66.

Big Data: tra necessità e responsabilità

di CRISTINA ISPAS*

Lo sviluppo senza precedenti di tecnologia e sistemi di informazione negli ultimi anni ha facilitato la raccolta, l'analisi e l'uso di una quantità impressionante di dati e informazioni in un tempo molto breve. Big Data è infatti un concetto rivoluzionario che «promette profonde trasformazioni in quasi tutti i campi di attività» (Davenport, 2015, p. 16).

Secondo l'Oxford English Dictionary (OED) Big Data significa: «set di dati estremamente grandi che possono essere analizzati a livello computazionale per rivelare schemi, tendenze e associazioni, in particolare in relazione al comportamento umano e alle interazioni»¹. Il termine è utilizzato in riferimento alla capacità (propria della scienza dei dati) di estrapolare, analizzare e mettere in relazione un'enorme mole di dati eterogenei, strutturati e non strutturati, allo scopo di scoprire i legami tra fenomeni diversi e prevedere quelli futuri.

Malgrado non esista una separazione rigorosa tra Big Data e altri dati (non classificabili come big), sono stati proposti diversi approcci per descriverne le peculiarità.

Si parla di Big Data quando si ha un insieme talmente grande e complesso di dati da richiedere la definizione di nuovi strumenti e metodologie per estrapolare, gestire e processare informazioni, entro un tempo ragionevole. Tali informazioni possono provenire potenzialmente da fonti eterogenee quali dati strutturati come i data base, ma anche non strutturati come immagini, e-mail, dati GPS o informazioni acquisite dai social network. I Big Data utilizzano la statistica inferenziale per dedurre leggi (come regressioni,

* Università Eftimie Murgu di Reșița, România.

1. https://en.oxforddictionaries.com/definition/big_data, ultima consultazione luglio 2019.

relazioni non lineari, ed effetti causali) da grandi insiemi di dati, per rivelare rapporti e dipendenze tra essi e per effettuare previsioni di risultati e comportamenti. Utilizzano dataset eterogenei, pertanto non correlati tra loro, dati grezzi e modelli predittivi complessi. Poiché, in questi casi, la mole dei dati è dell'ordine degli zettabyte, ovvero miliardi di terabyte, si richiede una potenza di calcolo parallelo eseguito su decine, centinaia o anche migliaia di server. In base agli strumenti e ai modelli utilizzati per l'analisi e la gestione dei dati è possibile distinguere quattro metodologie (o tipologie) di Big Data Analytics: *descriptive analytics*, *predictive analytics*, *prescriptive analytics*, *automated analytics*, questi ultimi definibili come strumenti capaci di implementare autonomamente l'azione proposta secondo il risultato delle analisi svolte.

La crescente mole di dati generati da sorgenti eterogenee ha posto l'attenzione su come estrarli, archivarli e utilizzarli al fine di ottenerne un profitto. Il problema che si riscontra è dovuto principalmente alla difficoltà di gestire i Big Data con database tradizionali, sia in termini di costi, sia in termini di volume. L'insieme di questi elementi ha portato allo sviluppo di nuovi modelli di elaborazione, che hanno permesso alle aziende di diventare più competitive, sia attraverso una riduzione dei costi, sia nuovi sistemi in grado di archiviare, trasferire e combinare i dati con maggiore velocità e in maniera agile.

L'immagazzinamento dei Big Data deriva da tre necessità: 1) memorizzare e organizzare enormi dataset non strutturati o semi-strutturati; 2) garantire la loro disponibilità e la capacità di interazione in modo affidabile, veloce e sicuro 3) estrarre conoscenza dai Big Data, esaminando gli enormi dataset a disposizione, alla scoperta di correlazioni, trend, pattern e ulteriori indici statistici nascosti nei dati mediante l'analisi su dati strutturati, semi-strutturati e non strutturati, quali ad esempio, analisi di testi, analisi di dati multimediali, analisi del web. L'interpretazione dei parametri analizzati può fornire dei suggerimenti per controllare ipotesi empiriche su fenomeni di particolare interesse, prendere decisioni di business più efficaci, individuare nuovi mercati nei quali investire, sviluppare campagne di marketing scientifico mirate e migliorare l'efficienza operativa. Big Data, quindi, è un termine onnicomprensivo (*all-inclusive*) che si riferisce a dati estremamente grandi, molto veloci, estremamente diversificati e complessi che non possono essere gestiti con i tradizionali strumenti di

gestione dei dati (Maheshwari, 2017). Esso è caratterizzato da tre elementi, genericamente chiamate “le tre V” (Volume, Velocità, Varietà); si sono aggiunte ulteriori V (Veridicità, Valore) (Laney, 2001).

Analizziamo il significato dei singoli termini:

- *volume*: la grande quantità di informazioni raccolte da diverse fonti, (ad esempio transazioni finanziarie, social media, ecc.) e archiviate per essere utilizzate per vari scopi (economici, politici, scientifici, ecc.); la gestione dell’enorme quantità di informazioni riflesse dall’espressione Big Data implica notevoli problemi logistici. I Big Data sono generalmente definiti come *cluster* di dati «sempre più grandi e grandi» (Aiden, Michel, 2013, p.11);
- *velocità*: le informazioni vengono trasmesse ad una velocità sorprendente in un tempo sempre più breve;
- *varietà*: i dati sono disponibili in tutti i formati, da documenti strutturati, numerici e non strutturati, e-mail, video, audio o transazioni finanziarie;
- *veridicità*: si riferisce alla connotazione qualitativa (il valore della verità) del dato raccolto e analizzato (in termini di interoperabilità e affidabilità)²;
- *valore*: i dati hanno un elevato valore insito all’interno di un grande mole di dati, diversi e in rapida crescita.

In sostanza, il concetto Big Data non è una novità; esso è stato introdotto in pratica da alcuni anni. I dati testuali e vocali esistono da molto tempo, ma il loro volume e la loro diversità, specialmente su Internet e altre strutture digitali rappresentano l’inizio di una nuova era. Difatti, le nuove tecnologie consentono di gestire una enorme quantità di dati e informazioni analizzate da algoritmi sempre più potenti e veloci.

Le applicazioni di questi dati riguardano vari campi (economici, sociali, politici, culturali, ecc) e sono volti a conoscere le preferenze dei consumatori, ridurre i costi, aumentare le vendite e la redditività, entrare in nuovi mercati, ampliando la gamma di servizi e/o prodotti, ottimizzare i processi produttivi (Overby, 2014, p. 34).

2. <https://mapsgroup.it/veridicita-dati/>, ultima consultazione luglio 2019.

Una delle caratteristiche più importanti dei Big Data è quella che consente di raccogliere ed elaborare le informazioni in un tempo molto breve, facilitando il processo decisionale in base alle informazioni in tempo reale.

L'informazione è diventata nello spazio virtuale una merce scambiata come qualsiasi altra merce. I vantaggi consistono in particolare nella conversione dei dati disponibili in informazioni rilevanti per i processi decisionali, che tendono a un miglioramento dei processi aziendali, ad una maggiore efficienza e alla riduzione dei costi.

Nonostante i molteplici vantaggi evidenziati, i Big Data non sono accessibili a tutte le organizzazioni, a causa sia della necessità di investimenti in hardware e software coerenti, sia delle nascenti problematiche connesse alle risorse umane. Anche i paesi più sviluppati del mondo lamentano una carenza di specialisti in questo settore.

Le informazioni possono provenire da una vasta gamma di fonti, da fonti interne ed esterne, non solo da singole fonti (Khawaja, 2014); per esempio possono essere informazioni pubblicate su vari siti specializzati, blog, social network e prodotti visualizzati nei negozi virtuali, informazioni lette da sensori, segnali GPS da dispositivi mobili, indirizzi IP del computer, cookie, codici a barre ed altro.

I manager, da tempo, hanno capito l'importanza delle informazioni nel loro campo di lavoro e hanno fatto ogni sforzo per raccogliere ed elaborare i dati e informazioni rilevanti per il loro campo di interesse e attività. Idealmente, i Big Data includono tutti i tipi di dati, che aiutano a fornire le informazioni giuste, alla persona giusta, nella giusta quantità, al momento giusto, per aiutare a prendere le giuste decisioni (Maheshwari, 2017).

La novità di Big Data risiede nelle tecnologie utilizzate che sono in grado, sulla base di algoritmi specifici, da un lato di raccogliere una enorme quantità di informazioni di ogni genere, e dall'altro di selezionare ed elaborare i dati rilevanti per l'utente. In questo senso, i Big Data stanno guadagnando terreno e vengono utilizzati in una varietà di campi (Mayer-Schönberger, Cukier, 2013). In campo economico, dati e informazioni rilevanti su aziende, sulla loro capacità competitiva, sui clienti, sulle loro esigenze di consumo ed altro sono importanti fonti di conoscenza che possono essere adeguatamente utilizzate.

In campo sociale, dati e informazioni sulla natura dei problemi sociali di una comunità, sulle risorse sociali necessarie o anche esistenti in comu-

nità o riguardanti l'intero pianeta possono stare alla base delle politiche sociali volte ad aiutare a risolvere o almeno migliorare la situazione sociale di un gruppo sociale, comunità, ecc.

In campo politico, specialmente durante le campagne elettorali, dati e informazioni sugli aventi diritto al voto sono di particolare importanza per tutti quelli che sono coinvolti nella lotta politica. I Big Data sono quindi un importante fenomeno sociale e tecnologico, una opportunità di creare nuovi modi di conoscere e lavorare (Maheshwari, 2017).

Utilizzando l'intelligenza artificiale, il processo decisionale, basandosi su informazioni concrete, pertinenti, ottenute in tempo reale, riesce a dare risposte adeguate alle sfide innovative di una società complessa ed articolata in una molteplicità di settori, riducendo nello stesso tempo, anche i rischi derivanti da qualsiasi processo decisionale.

Con l'implementazione delle nuove tecnologie, i Big Data consentono alle organizzazioni di raccogliere, archiviare e trasmettere in tempo reale dati e informazioni di notevole interesse.

Le aziende e le organizzazioni mostrano una attenzione crescente in questo settore investendo in tecnologie all'avanguardia, incentivando l'innovazione e garantendo al personale professionalità e specificità organizzativa.

La richiesta di personale qualificato nel settore specifico può essere evasa mediante una stretta collaborazione tra organizzazioni interessate ai Big Data e istituzioni universitarie che devono essere sempre più in grado di adattare la propria offerta formativa alle nuove esigenze della società. Si tratta, in altri termini, di una vera e propria rivoluzione culturale, che ha generato quella che noi oggi usiamo chiamare la "società informazionale".

La società informazionale ha come suo motore un paradigma scientifico e tecnologico fondato sull'informazione e sulle tecnologie dell'informazione, dall'elettronica alle telecomunicazioni e all'informatica.

Mi limito solo a segnalare (o ricordare) il ruolo straordinario di componenti strutturali che nella società dell'informazione hanno assunto i dati e la conoscenza; i primi legati ai valori delle informazioni note e la seconda legata alla loro elaborazione ed alla produzione di nuove informazioni.

Respiriamo dati e siamo quotidianamente sommersi da uno tsunami di nuovi dati che vengono prodotti, memorizzati, trasmessi e usati in quan-

tità e velocità strabilianti se paragonate al passato: bit, byte, Kbyte Kbit/s della prima generazione impallidiscono rispetto agli Zettabyte e ai Mega, ai Giga e ai Tera della quinta generazione. Ed è strabiliante, rispetto al passato, sia per quantità che per velocità la produzione e l'uso della conoscenza. Come parte della punta di questo iceberg e solo come esempio, lasciatemi segnalare che sono più di 45 milioni il numero di articoli scientifici (citabili e raggiungibili) pubblicati in un anno (2017), centinaia di migliaia i brevetti rilasciati (circa 100.000 solo in Europa nel 2016) e più di 80.000 all'anno le idee innovative che hanno concretamente prodotto cambiamenti e applicazioni nella produzione di beni e servizi. Fra i cambiamenti strutturali che la rivoluzione informazionale ha introdotto assume sicuramente grande rilevanza quello della metamorfosi del modo di produzione. Dal modo di produzione industriale centrato sulla valorizzazione di grandi masse di capitale fisso materiale siamo passati rapidamente ad un modo di produzione centrato sulla valorizzazione di capitale immateriale (capitale umano, capitale conoscenza, capitale intelligenza).

Nella società informazionale, il valore aggiunto è prodotto non dal "lavoro materiale" ma dalla "conoscenza": dalla centralità del lavoro materiale e semplice siamo passati a quella del lavoro immateriale e complesso. In questo trionfo dell'immateriale, il software assume addirittura un ruolo strategico, perché è la tecnologia dell'immateriale. Quanto ai dati, una diffusa affermazione stile twitter, molto popolare anche fra i politici, racconta che i dati sono il petrolio del nostro tempo. Restando nell'ambito di questa metafora, diciamo che sono le ICT (Information and Computer Technology) a trasformare questo petrolio in potenziale ricchezza.

Ciò ci dice qual è la dimensione della domanda di software che ci viene dal mondo dei dati e ci lascia capire anche che Big Data e associati Analytics sono solo nell'infanzia del loro cammino e, forse, come accaduto con le Basi di Dati, sono solo una tappa di fase della gestione ed elaborazione di dati nell'evoluzione della società informazionale. L'intercettazione, la collezione e l'uso di dati privati e pubblici non è solo possibile, ma anche largamente diffusa e potenzialmente estendibile. Inoltre è palesemente infondata quella affermazione secondo cui, a differenza del petrolio, i dati non inquinano. È un inquinamento diverso, fatto di *fakes* e di dati falsi. Il problema che abbiamo di fronte non è quello della intercettazione e collezione di dati, ma è caso mai quello della loro sicurezza e difesa da ac-

cessi indesiderati (dolosi o non), a cui si aggiunge il grande problema della quarta V, ovvero della Veridicità dei dati (che si unisce a quello del Volume, della Velocità, della Varietà). Una nuova sfida per le ICT ed in modo particolare per il software, che trova e continuerà a trovare soluzioni per sistemi ed assistenti ausiliari per la sicurezza e il filtraggio dei dati.

Credo tuttavia che quanto detto sia ancora insufficiente per mettere a fuoco il ruolo del software e le sfide che deve affrontare nella rivoluzione in corso.

Secondo tale visione, questa nostra società si caratterizza per la nascita di un nuovo ambiente informazionale che le ICT stanno creando ed espandendo: è una infosfera nella quale le generazioni future trascorreranno la maggior parte del loro tempo (negli USA, già ora, il tempo medio speso dall'uomo nell'infosfera ha superato i cinque mesi all'anno).

Oggi l'intelligenza umana non è più sola sulla terra e non è sulla sua centralità che si sviluppano attività e relazioni.

L'insieme dei beni, dei servizi, delle relazioni e di ogni altra attività che l'uomo produce non è più frutto unicamente della sua intelligenza ma anche di un insieme di altre e diverse intelligenze che si integrano, sommandosi e moltiplicandosi, interagendo e cooperando fra loro.

Nell'infosfera nascono e proliferano *IT-entity* che selezionano, valutano, trasformano e producono dati e conoscenza, che hanno capacità di apprendimento ed intelligenza, capacità di fornire *feed-forward* (non solo *feedback*) e assumere decisioni e comportamenti su eventi futuri.

Come accade nell'uomo e come direbbe Gardner, è una intelligenza fatta da molteplici tipi di intelligenze; non solo l'intelligenza logico-matematica, ma anche spaziale (come dimostrano, ad esempio, realtà virtuali), cinestetica (come dimostrano ad esempio robot e robot-cooperativi stampanti 3D), o interrelazionale (capacità di interfacciamento, collegamento e relazione con altri InfoOrg) e così via.

In questo quadro l'ingegneria del software ha evidentemente davanti a sé nuove ed enormi sfide da affrontare, non solo le storiche sfide quantitative e qualitative. Per quanto riguarda la quantità di software necessaria, dobbiamo produrne per miliardi di macchine, miliardi di utenti, milioni e milioni di applicazioni.

La punta dell'iceberg degli sviluppatori di software è costituita da 21 milioni di addetti; se ipotizziamo grossolanamente che ciascuno di essi produca mediamente dieci istruzioni al giorno, se ne deduce che vengono

prodotte più di 70 miliardi di istruzioni all'anno. Naturalmente non tutto si trasforma in prodotti di mercato o free software, e bisogna ricordare che quelli che chiamiamo gli "addetti ai lavori" sono solo una parte di quelli che producono software.

Profondamente mutata è anche la più vecchia attività dei processi software, ovvero la programmazione anche perché non dobbiamo più programmare computer ma *IT-entity* e sistemi cyber-fisici.

In questa fase, come in altre, l'ingegneria del software deve proporre nuovi approcci sistematici fondati su una spina dorsale concettuale in cui i software embedded sono integrati ed integrabili, relazionabili, adattabili, modificabili, migliorabili ed evolvibili. Abbiamo bisogno dei concetti di questa spina dorsale, delle astrazioni e dei metodi che debbono caratterizzarla.

In questo divenire della Società dell'Informazione, il software ha davanti a sé ruoli e sfide importanti. Siamo di fronte alla invasione dell'immateriale nelle produzioni di beni materiali. Stiamo sperimentando qui che il 4.0 significa innanzitutto integrazione di discipline e competenze diverse perché qualsiasi problema richiede approcci e soluzioni fondate sulla fusione di software, ICT, tecnologie e conoscenze del dominio applicativo.

Concordo profondamente, quindi, con quanti sostengono che la società informazionale è cresciuta molto più rapidamente della capacità dell'uomo di svilupparne solide radici concettuali, etiche e culturali.

Credo che, da quanto detto, emerga con chiarezza il grande fabbisogno di economia (da quella politica a quella gestionale ed organizzativa), di sociologia, e, su tutto, di filosofia. Abbiamo bisogno di categorie e costrutti filosofici che ci consentano di capire ed interpretare l'infosfera, il suo divenire e ciò che in essa accade.

Abbiamo bisogno di capire cos'è un uomo "online", che non è e non può vivere "offline"; ma abbiamo, più in generale, bisogno di una ecologia dell'infosfera, vale a dire etica, diritto, cultura ecc.

Fondamentale in tutto ciò è il ruolo che dovrà avere il mondo accademico nel facilitare e gestire la transizione, per quanto concerne il software, e forse non solo, fra le metamorfosi in corso. Io non so dire se e quando questa transizione sarà totalmente completata, o se, ad esempio, si stabilizzerà un equilibrio di coesistenza.

Nella società dell'informazione, nel dispiegarsi dell'infosfera, nella rivoluzione dell'intelligenza, noi siamo più che mai cercatori di futuro,

raccoglitori, creatori, narratori di storie di vita che a loro volta cambieranno la vita delle generazioni a venire. Nella società dell'informazione i Big Data giocano, pertanto, un ruolo determinante.

Difatti i Big Data non riguardano solo l'accumulo di un'enorme quantità di dati; la vera sfida è la capacità di coloro che sono incaricati di raccogliere, analizzare ed elaborare questi dati per individuare ciò che è veramente importante (il valore dei dati) per l'individuo e la collettività.

Ciò, ovviamente, richiede esperti nell'analisi e nell'elaborazione dei dati raccolti e l'utilizzo di tecnologie avanzate. I Big Data sono la nuova frontiera del mondo digitale che, indubbiamente, porterà enormi vantaggi alla società del domani ma, nello stesso tempo, introdurrà nuovi sentieri da percorrere e da esplorare e sempre più complesse problematiche da affrontare, sia dal punto di vista etico, politico, organizzativo sia tecnologico.

«Di fronte ai grandi sviluppi che le nuove tecnologie digitali hanno avuto negli ultimi anni e a quelli che avranno nel prossimo futuro, ogni cittadino dovrebbe, oltre a conoscere i benefici che potrà certamente avere, comprendere anche come le nuove soluzioni tecnologiche potranno influenzare la sua vita. La nostra privacy sta diventando uno dei prezzi di pagare ed è opportuno che essa venga tutelata sia dalle leggi, sia dalla nostra indispensabile consapevolezza» (Talia, 2018).

L'impatto di queste nuove tecnologie sulla nostra vita quotidiana non sempre è da ritenersi positivo. Uno dei prezzi di pagare rimane la perdita della privacy, assolutamente da tutelare con opportuni interventi legislativi. Inoltre, la raccolta, la gestione e l'elaborazione dei dati implica una responsabilità etica e sociale nell'utilizzo dei dati a disposizione. In Europa, il 25 Maggio 2018, è entrato in vigore un nuovo regolamento generale sulla protezione dei dati: il *General Data Protection Regulation* (GDPR). Secondo la Commissione Europea «i dati personali sono qualunque informazione relativa a un individuo, collegata alla sua vita sia privata, sia professionale o pubblica. Può riguardare qualunque dato personale: nomi, foto, indirizzi email, dettagli bancari, interventi su siti web di social network, informazioni mediche o indirizzi IP di computer»³. Oltre al dato personale, in questa direttiva, troviamo dati genetici, biometrici e relativi alla salute; comunque tutte informazioni che consentono l'identificazione univoca

3. <https://gdpr-info.eu/>, ultima consultazione luglio 2019.

o l'autenticazione di una persona fisica. Secondo l'art. 9, paragrafo 1 del detto regolamento⁴, si considerano dati sensibili o dati personali:

1. l'origine razziale o etnica;
2. le opinioni politiche;
3. le convinzioni religiose o filosofiche;
4. l'appartenenza sindacale;
5. dati relativi alla vita sessuale o all'orientamento sessuale della persona;
6. dati genetici: ereditati o acquisiti, ottenuti tramite analisi di DNA ed RNA da un campione biologico della persona fisica in questione;
7. dati biometrici: come l'immagine facciale, grazie ai quali è possibile identificare una ed una sola persona fisica;
8. dati sulla salute: sia fisica che mentale, passata, presente o futura, ma anche informazioni su servizi di assistenza sanitaria, laddove presenti, indipendentemente dalla fonte, quale, ad esempio, un medico;
9. dati personali relativi a condanne penali o reati.

Su tutto vige un principio assoluto legato al trattamento dei dati personali che è il principio di responsabilità. Applicare il principio di responsabilità per l'utilizzo di Big Data è assolutamente una esigenza ineludibile, in quanto essi rappresentano un fenomeno sociale e tecnologico avente un grande impatto sulla vita dei cittadini. L'esperienza dimostra che esiste un limite abbastanza fragile sull'uso corretto di questa enorme massa di dati se viene meno un principio etico, sociale e legale (Mayer-Schönberger, Cukier, 2013) che è appunto rappresentato dal principio di responsabilità. Eloquente a tal proposito è il social network Facebook. Secondo le accuse mosse a Facebook, questa ha permesso a Cambridge Analytica, società britannica di analisi dei dati online, di entrare in possesso dei dati personali di milioni di persone, dati utilizzati durante la campagna presidenziale del 2016 di Donald Trump (McNamee, 2019). Cambridge Analytica ha giocato un ruolo importante anche nella campagna pro Brexit, usando i dati personali degli utenti di Facebook "estratti in modo non appropriato"⁵.

4. <https://gdpr-info.eu/art-9-gdpr/>, ultima consultazione luglio 2019.

5. <https://www.theguardian.com/technology/2018/mar/21/mark-zuckerberg-response-facebook-cambridge-analytica>, ultima consultazione luglio 2019.

Infatti, Facebook è «uno straordinario strumento per il marketing personalizzato. È l'occasione per imparare una volta per tutte la lezione secondo cui “quando non paghi qualcosa, il prodotto sei tu” e comprendere a fondo l'importanza di tutelare i nostri dati. I dati personali sono il nuovo petrolio, la materia prima che tutte le aziende desiderano»⁶.

L'Information Commissioner's Office (ICO), autorità per la privacy del Regno Unito parla di violazione dei dati di circa 87 milioni di utenti Facebook, la maggior parte dei quali provenienti non solo dagli Stati Uniti ma anche da diversi paesi europei, tra cui anche la Romania. Facebook, tra l'altro, ha consentito l'accesso ai dati personali dei propri utenti, non riuscendo a garantire la protezione dei dati personali, la qual cosa ha favorito l'accesso illegale a questa fonte inesauribile di documentazione ed informazione dei dati immessi sulla sua piattaforma da parte di molte altre società interessate (Palmatier, Martin, 2019). L'azione legale intrapresa contro Facebook continua. Tra le decisioni più attese c'è quella della Federal Trade Commission, la commissione statunitense a tutela dei consumatori. Nel mese di febbraio 2019, i giornali e le agenzie di stampa, a partire da un rapporto pubblicato dal Washington Post, hanno dichiarato che Facebook, per le infrazioni commesse, potrebbe essere multato per “miliardi di dollari”; sarebbe questa la più alta sanzione mai comminata ad una società tecnologica per la violazione delle norme sulla privacy. Le alte sanzioni somministrate e quelle eventualmente da somministrare mirano a richiamare l'attenzione sul fatto che queste pratiche, ovviamente illegali, richiedano non solo soluzioni immediate ed urgenti in termini legali, quanto anche rimedi concreti al fine di proteggere l'uso fraudolento dei dati personali. Si rende necessario, quindi, adottare una legislazione coerente e pertinente per la protezione di questi dati, così come l'imporre vincoli e condizioni chiare e trasparenti sul modo in cui questi dati vengono ottenuti, archiviati, elaborati e utilizzati. Il tutto finalizzato a garantire la privacy, la protezione dei dati e la libertà di espressione dei cittadini. Le norme dell'UE garantiscono la protezione dei dati personali. Per quanto riguarda i Big Data, riteniamo che sia necessaria una maggior consapevolezza da parte di tutti, sia dei vantaggi connessi al suo utilizzo,

6. <https://www.linkiesta.it/it/article/2018/03/21/ne-furto-ne-spionaggio-nel-caso-bridge-analytica-i-veri-colpevoli-s/37511/>, ultima consultazione luglio 2019.

sia dei rischi che corrono coloro che non sono adeguatamente protetti. A tale scopo, Safer Internet Day, celebrato ogni anno a febbraio (il primo martedì), a partire dal 2004, in oltre 110 paesi del mondo, mira a promuovere un uso più sicuro e responsabile della tecnologia, in particolare per quanto riguarda i bambini e gli adolescenti. Secondo uno studio condotto in Romania nel periodo 27 dicembre 2018–7 gennaio 2019, su un campione di 1.009 persone adulte, da YouGov⁷ per Google su Internet Security Day, il 37% degli intervistati dichiara di aver ricevuto e-mail phishing fraudolento (ad esempio, un tentativo di accedere ai dati personali con mezzi fuorvianti, incluso l'accesso alle password e ai dati bancari, sostenendo di provenire da una fonte legittima) e il 28% afferma di essere stato infettato da un virus o altro software dannoso, che ha portato alla manipolazione, al furto o alla cancellazione di informazioni personali. Inoltre, il 49% dei rumeni intervistati dichiara di essere principalmente interessato alla protezione dei dati finanziari, come ad esempio internet banking o dati della carta di credito/debito, mentre il 19% alla protezione dei dati personali come l'indirizzo, codice fiscale, ecc, e, appena, il 9% degli intervistati si dichiara favorevole alla protezione di dati inerenti a particolari situazioni personali, come foto di famiglia, o altro. Per quanto riguarda la sicurezza degli account online, un quarto dei rumeni intervistati (26%) afferma di utilizzare almeno una volta al mese uno strumento di controllo, come, ad esempio, il controllo di sicurezza di Google; circa il 30% degli intervistati dichiara di non aver mai fatto uso di tali strumenti di controllo e appena un 16% afferma di effettuare mediamente controlli di sicurezza semestrali. Per quanto riguarda l'aggiornamento della password online, le percentuali dedotte dalle interviste risultano addirittura minori rispetto a quelle enucleate precedentemente.

L'impatto dei Big Data sulla vita di tutti è indiscutibile. Ciò che è meraviglioso degli esseri umani è ciò che nessun algoritmo o chip al silicio può rivelare; le scintille dell'invenzione diventano ciò che i dati non dicono (Mayer-Schönberger, Cukier, 2013).

Possiamo, perciò, concludere affermando che certamente i Big Data ci fanno apprezzare l'enorme potenza del mondo tecnologico, ma l'origina-

7. <https://www.mobile-news.ro/wp-content/uploads/2019/02/Studiu-YouGov-Google.pdf>, ultima consultazione luglio 2019.

lità e la creatività tipica della persona non potrà mai essere imbrigliata o racchiusa nelle istruzioni di un freddo ed arido algoritmo.

Riferimenti bibliografici

- AIDEN, J., MICHEL B. (2013), *Uncharted: Big Data as a lens on human culture*, New York, New York, Riverhead Books.
- ALTSTIEL T., GROW, J.M. JENNINGS, M. (2018), *Advertising Creative: Strategy, Copy, and Design*, Fifth Edition Edition, SAGE Publication.
- DAVENPORT T. H. (2015), *Big Data @l lavoro. Sfatare i miti, scoprire le opportunità*, Milano, FrancoAngeli.
- DAVENPORT T. H., DYCHÉ J. (2013), *Big Data in Big Companies*, International Institute for Analytics.
- GOLDBERG M.S. (2014), *Mapping new opportunities*, CIO.
- KHAWAJA J. (2014), *Five things big data is not*, CIO Enterprise Forum.
- KONTZER T. (2015), *Using analytics to squeeze value from Big Data*, Baseline.
- LANEY D. (2001), *3D Data Management: Controlling Data Volume, Velocity and Variety*, META Group.
- MAHESHWARI A. (2017), *Big Data*, Mcgraw Hill Education (India), Private Limited Publishing.
- MAYER-SCHÖNBERGER V., CUKIER, K. (2013), *Big Data: A Revolution that Will Transform how We Live, Work, and Think*, Canada, Eamon Dolan/Houghton Mifflin Harcourt.
- MCNAMEE R. (2019), *Zucked: Waking Up to the Facebook Catastrophe*, New York Penguin Press.
- OVERBY S. (2014), *Cashing in on analytics*, CIO.
- PALMATIE R. W., MARTIN K. D. (2019), *The The Intelligent Marketer's Guide to Data Privacy. The Impact of Big Data on Customer Trust*, Palgrave Macmillan.
- TALIA D. (2018), *La società calcolabile e i Big Data: Algoritmi e persone nel mondo digitale*, Soveria Mannelli (cz), Rubbettino. VENKATASAWMY R. (2018), *Communication and Media Education in an Era of Big Data*, «Journal of Interdisciplinary Studies in Education», Volume 7, Issue 1, pp. 87-101.

Percezione delle proprie tracce digitali

Una ricerca internazionale tra studenti universitari

di MARIA ADELAIDE GALLINA *

1. Dalla letteratura all’impianto di ricerca. Analizzare il sociale con i Big Data

L’ampiezza e l’eterogeneità di informazioni di cui si compongono i Big Data porta alla necessità di cambiamenti fondamentali nel modo in cui i dati stessi vengono memorizzati e analizzati. La massa è tale per cui essi sono difficilmente gestibili con i sistemi “tradizionali” e richiedono tecnologie specifiche presenti nell’ecosistema digitale. La sfida nei prossimi anni sarà quella di acquisire le competenze per gestire e organizzare flussi di dati complessi e per impiegarli nello sviluppo di nuovi prodotti e servizi.

Quale valore aggiunto deriverebbe allora dall’uso dei Big Data? Oltre a considerare le potenzialità per migliorare i servizi collettivi e la qualità della vita, è necessario valutare i dati stessi come strumenti il cui scopo è sorvegliare o prevedere i comportamenti degli individui. Al tal proposito, Bennato (2015) parla di “data-sorveglianza”, ossia di una tecnica di controllo basata sui dati, in particolare sulle tracce digitali lasciate in Rete. I processi di *monitoring* e di *profiling* hanno apportato un’evoluzione dei sistemi di protezione, in quanto non è più necessario avere strumenti che osservino direttamente le persone per controllare i loro comportamenti (per esempio, le telecamere), ma basta seguirne le tracce digitali tramite algoritmi.

I Big Data danno avvio a questioni etiche, relative sia all’utilizzo corretto dei dati con le rispettive ricadute nella vita degli individui sia alla salvaguardia dei diritti civili con le connesse problematiche riguardanti la riservatezza delle informazioni personali.

* Università degli Studi di Torino.

Da un certo punto di vista gli individui devono rinunciare a una parte della propria libertà e privacy, ma questo garantisce una maggiore sorveglianza — anche per finalità utili come garantire sicurezza o prevenire atti criminali.

È importante aggiungere ancora un elemento per analizzare il fenomeno dei Big Data. Oltre a una cieca e piena fiducia nei dati o a un'esasperata quantificazione del reale, è necessario considerare il rischio relativo all'analisi stessa dei Big Data, ovvero la "datafrenia": una posizione radicale che «sancisce l'onnipotenza del dato, cioè la possibilità che i dati hanno di comprendere il mondo anche in assenza di un processo di interpretazione e di decisione critica da parte di un ricercatore» (Bennato, 2015, p. 117).

L'espansione del fenomeno conferma la necessità di un approccio critico. A tal proposito, Mayer-Schönberger e Cukier (2013) suggeriscono per la tutela della privacy un approccio focalizzato più sulla responsabilizzazione degli utilizzatori dei dati e meno sul consenso preventivo degli utenti al momento della raccolta dei dati stessi. Un approccio di questo tipo potrebbe far sì che si prendano misure in grado di tutelare maggiormente gli individui.

Vista la diffusione del fenomeno dei Big Data è necessario porsi un interrogativo: quanto è conosciuto questo fenomeno? Per poter comprendere la percezione sociale e la diffusione del fenomeno è importante capire il punto di vista dei consumatori.

Un'indagine condotta da Axa Italia (2016) rivela che il 45% della popolazione italiana non ha un'idea chiara sul concetto di Big Data, nonostante faccia parte di quella fascia di utenti che utilizza Internet; il 35% teme, a causa dei Big Data, di perdere il controllo della loro immagine in Rete e la tutela della loro privacy; infine, il 20% li avverte come strumento utile a migliorare sempre più la qualità della vita. Oltre a questo il 66% della popolazione percepisce i Big Data come utili soltanto per le imprese, in quanto indirizzano verso la scelta di prodotti e servizi offerti dalle stesse.

Anche il settore sanitario, ad esempio, potrebbe adottare una nuova classe di dispositivi medici collegati alla Rete per monitorare in tempo reale le condizioni dei pazienti e le aziende di prodotti industriali potranno dotare i prodotti stessi di chip che monitorino il loro buon funzionamento. Come afferma Bennato (2015), anche le scienze sociali, negli ultimi

anni, stanno mostrando un certo interesse riguardo al fenomeno in oggetto. Infatti, per quanto i dati possano nascondere informazioni interessanti, non sono in grado di parlare da soli, ma hanno bisogno di una teoria che guidi la costruzione del database e ne analizzi le variabili (Bennato, 2015). Nonostante ciò, gli studiosi credono fortemente che i Big Data renderanno sempre più misurabile la società grazie al *social mining*, la nuova scienza dei dati, per fare previsioni accurate su aspetti sociali.

Le tracce digitali possono dunque rappresentare un vero e proprio specchio per l'osservazione e la comprensione dei comportamenti individuali e collettivi.

Diverse sono quindi le prospettive da cui considerare le ricerche che coinvolgono i Big Data: quella tecnologica in primis, ma anche quella commerciale ed etica, poiché è necessario considerare sia la mole di dati-informazioni che possono essere raccolti dai social media e da documenti di vario tipo (e-mail, pagine web e post) sia il valore commerciale delle informazioni che derivano dalla gestione dei Big Data e, di conseguenza, gli aspetti relativi alla privacy (Strizzolo, 2014).

Come afferma Linda Lombi (2015), l'ambito etico e, in particolare, la tutela della confidenzialità del dato e il rispetto della privacy sono una questione importante che riguarda la distinzione tra pubblico e privato, che assume una nuova veste nell'era digitale. Parlare di privacy significa infatti non soltanto restringere l'accesso a parte delle informazioni, ma includere anche la possibilità di esercitare il controllo sulla loro diffusione.

In seguito all'analisi di riflessioni teoriche relative alla letteratura nazionale e internazionale sui Big Data, in questo lavoro si intendono presentare i risultati di una ricerca quantitativa rivolta alla popolazione giovanile, in particolare a universitari.

Ci siamo interrogati sul livello di consapevolezza che si ha in merito alle "briciole digitali" che si lasciano utilizzando la Rete. In questa prospettiva abbiamo cercato di misurare il valore che attribuiscono gli studenti universitari in riferimento ai propri dati che vengono generati attraverso l'utilizzo della Rete. La vastità del campo d'indagine evidenzia da un lato la possibilità di entrare in nuovi ambiti di ricerca, ma dall'altro ci porta a riflettere sul livello di consapevolezza che si ha in merito alla "traccia digitale" che lasciamo ogni qualvolta utilizziamo un servizio di Rete o anche, semplicemente, una tessera di un supermercato.

A tal proposito abbiamo cercato di misurare la consapevolezza che hanno gli studenti che sono nativi digitali: come utilizzano la Rete? Quali tracce lasciano in Rete e soprattutto conoscono il valore dei propri dati?

È stata condotta quindi una ricerca quantitativa o standard basata sulla matrice dei dati (Grimaldi, 2005, p. 20). La raccolta di informazioni è avvenuta con la somministrazione di questionari durante l'anno accademico 2016–2017 sul territorio italiano ed europeo, coinvolgendo diversi grappoli di popolazioni di studenti universitari appartenenti a vari corsi di laurea, provenienti dal Piemonte, dalla Liguria e dalla Sicilia e due grappoli provenienti dall'estero, in particolare da Resita in Romania, e da Valencia in Spagna. Il campione è formato da 1.091 studenti. L'esito non fornisce un campione casuale e rappresentativo in senso statistico, ma una raccolta di grappoli di dati che, se da una parte non ci permettono di fare inferenze verso popolazioni differenti da dove si è scesi sul campo, dall'altra consentono di far emergere relazioni che si ripetono nelle diverse realtà indagate.

Lo strumento di indagine utilizzato è un questionario strutturato a risposta prevalentemente chiusa, che intende indagare diversi aspetti che riguardano i giovani coinvolti: le azioni svolte in Rete (dal postare sui social network al fare acquisti on-line, l'utilizzo di App per car o bike-sharing, la fruizione di sistemi di condivisione Cloud, App che necessitano di geolocalizzazione o che registrano informazioni relative al proprio stato fisico), il possesso di strumenti infotelematici, la percezione che l'uso delle tecnologie infotelematiche generi dati che possono essere utilizzati da aziende commerciali o altre organizzazioni, la percezione del valore legato ai propri dati lasciati in Rete e, infine, la rilevazione dei dati socio-anagrafici.

Ampia parte del questionario è infatti rivolta proprio a misurare da un lato il valore che si pensa possano avere i propri dati per enti e organizzazioni, dall'altro il valore che si ritiene di poter trarre dai servizi che si ricevono in cambio del rilascio di informazioni riguardanti se stessi a enti (ad esempio quando vengono rilasciati i propri dati per poter usufruire di App gratuite). Si è deciso di misurare tale valore in euro, senza dare indicazioni di valore minimo o massimo proprio per non influenzare il rispondente.

2. Il profilo socio-anagrafico degli studenti universitari

I 1.091 studenti coinvolti nella ricerca sono per il 75,6% di genere femminile e per il 24,4% di genere maschile. L'età media risulta essere di 22,8 anni.

Emerge che oltre il 50% degli studenti proviene da Torino (il 45,6% iscritti a corsi di laurea umanistici e l'8,5% al Politecnico), un 15% dalla Romania (iscritti a corsi di laurea di tipo umanistico), un 11% da Valencia (iscritti a corsi di Filosofia e Scienze dell'Educazione dell'Univesitat de València e alcuni studenti dell'Univesitat Politecnica de València), un 8% dall'Università del Piemonte Orientale (iscritti al corso di Laurea in Scienze infermieristiche), un 6% dall'Università degli Studi di Genova (iscritti al corso di Laurea in Scienze dell'Educazione), e infine per un 5% da Palermo (iscritti al corso di Laurea in Giurisprudenza) (Tab. 1).

Tab. 1. Distribuzione di frequenza per provenienza del campione, Nazione e città.

Nazione-Città	Frequenza	Percentuale
Italia_Torino	591	54,2
Romania_Resita	166	15,2
Spagna_Valencia	121	11,1
Italia_Biella	87	8,0
Italia_Genova	72	6,6
Italia_Palermo	54	4,9
Totale	1091	100,0

Dall'analisi dei dati si evince una prevalenza di studenti del corso di laurea in Scienze dell'educazione (19,9%) e Scienze della formazione primaria (20,2%), ma numerosi sono anche coloro che provengono da altri corsi di studi. Per semplificare la diversa provenienza degli studenti (si registrano infatti 15 corsi di studi differenti, oltre a 8 studenti che abbiamo aggregato nella categoria "altro"), si sono create tre macrocategorie: studi umanistici (79,6%), scienze della vita (10,1%) e ingegneria (10,4%). I corsi di studi frequentati sono dunque numerosi e diversi tra loro, come possiamo osservare in Tab. 2.

Tab. 2. Distribuzione di frequenza per *corso di studi* frequentato.

	Frequenza	Percentuale
scienze dell'educazione	213	19,9
educatore professionale	9	,8
scienze della formazione primaria	222	20,7
scienze pedagogiche	60	5,6
servizio sociale	56	5,2
laurea magistrale in pedagogia, progettazione e ricerca educativa	65	6,1
scienze economiche	24	2,2
teologia	4	,4
amministrazione pubblica	8	,7
educazione fisica e sport	3	,3
ingegneria	111	10,4
giurisprudenza	63	5,9
infermieristica	87	8,1
lingue	126	11,8
scienze della salute	13	1,2
altro	8	,7
Totale	1072	100,0
Mancante	Sistema	19
Totale		1091

Si è anche indagato il numero di oggetti tecnologici posseduti dagli studenti (smartphone, tablet, computer fisso, computer portatile, smartwatch, elettrodomestico che si può comandare via Internet). Si rileva una capillare diffusione di tali strumenti, con una media di 2,8 strumenti posseduti a persona: è evidente come l'utilizzo di strumenti sempre più sofisticati come smartphone o tablet diminuisca drasticamente la necessità di avere più strumenti tecnologici in quanto singolarmente in grado di adempiere alla maggioranza delle funzioni normalmente utilizzate.

Interessanti risultano le attività svolte dai giovani in Rete, individuate mediante una classificazione di 18 dimensioni: sottoscrivere tessere in supermercati/negozi, fare acquisti on-line, fare vendite on-line, rispondere a sondaggi (sia telefonici, sia on-line), iscriversi a siti per

ottenere informazioni, o App gratuite, postare su social network, pubblicare su blog, mettere “mi piace” su Facebook anche su post di tipo sensibile (politico, religioso, etc.), postare foto/video, fornire la registrazione del proprio stato emotivo (es. “sono triste” su Facebook), consentire la registrazione del proprio stato personale fisico (es. peso, battito cardiaco, pressione sanguigna, ecc.), consentire la geolocalizzazione del proprio cellulare/smartphone (ad esempio per registrare o farsi guidare in un percorso a piedi, in bicicletta o in auto, o per segnalare la propria posizione su Facebook), usare sistemi di condivisione Cloud (es. Dropbox, Google Drive, etc.), usare App per car sharing, usare App per bike sharing, uso di console per videogame in modalità interattiva. Tali attività sono numerose, in media 7 per intervistato, con un 3% che dichiara fino a 3 attività e un 5% che afferma di compiere dalle 12 alle 18 delle attività indagate.

Questi dati delineano dunque un resoconto del profilo socio-anagrafico, dell'utilizzo e del possesso di oggetti tecnologici degli studenti universitari coinvolti nell'indagine.

Riferimenti bibliografici

- BENNATO D. (2015), *Il computer come macroscopio*, Milano, FrancoAngeli.
- CUKIER K., MAYER-SCHONBERGER V. (2013), *Big Data*, Milano, Garzanti.
- DAVENPORT, T.H (2015) *Big Data @l lavoro*, Milano, FrancoAngeli.
- DE MARTIN JUAN C. (2011), *Ma Twitter non è un bar*, in *La Stampa*, <http://www.lastampa.it/2011/05/11/cultura/opinioni/editoriali/ma-twitter-non-unbar-hvx8YVbUfWCCATj40VU7BN/pagina.html>
- GRIMALDI R. (2005), *Metodi formali e risorse della Rete*, Milano, FrancoAngeli.
- GRIMALDI R., GALLINA M.A., CAVAGNERO S.M. (2017), *Big data e tracce digitali. Una ricerca internazionale tra studenti universitari*, in TRINCHERO R., PAROLA A. (a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli.
- ITALIAN AXA (2016), *Le sfide dei dati*, <https://corporate.axa.it/>, paper n. 8, ultima consultazione 21 gennaio 2018.
- LOMBI L. (2015), *La ricerca sociale al tempo dei big data: sfide e prospettive*, in «studi di Sociologia», anno 53, Fasc. 2, pp. 215-227.

STRIZZOLO N. (2014), *Capire i Big Data*, Agenda digitale www.agendadigitale.eu/smart-cities-communities/948_capire-i-big-data-tre-fronti-della-ricerca.htm.

TRINCHERO R., PAROLA A. (a cura di)(2017), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli.

Il valore dei dati in Rete e profili di intervistati

di RENATO GRIMALDI, SANDRO BRIGNONE*

1. Quanto valgono i nostri dati?

Studi relativi ai cambiamenti portati dai Big Data segnalano difficoltà nel dare una definizione unica e univoca del concetto stesso di Big Data (Hargittai, 2015), indicando una riflessione su quali processi decisionali siano alla base della generazione e memorizzazione dei dati stessi (Crawford, 2013) e sviluppando considerazioni etiche sul significato e l'utilizzo di queste informazioni (Boyd & Crawford, 2012). I rilievi appena citati si incontrano inevitabilmente con il concetto di privacy (Lazer et al., 2009), soprattutto alla luce delle indicazioni derivanti dal *General data protection regulation* (GDPR), istituito dall'Unione Europea nel 2018. Non bisogna inoltre dimenticare le importanti implicazioni che ne derivano nel campo della ricerca sociale.

Qual è dunque il valore da attribuire ai nostri dati? Quanto valgono? Come li percepiscono i giovani?

La nostra indagine si è posta l'obiettivo di controllare il valore che gli studenti universitari attribuiscono ai loro dati quando vengono ceduti e, al tempo stesso, rilevare il valore che essi ritengono di trarre dai servizi che vengono resi quando le loro informazioni sono intercettate da enti esterni¹. Ad esempio, usare le mappe di Google per essere guidati dalla propria posizione verso una destinazione consente di visualizzare in anticipo possibili

* Università degli Studi di Torino. Renato Grimaldi ha scritto i paragrafi 1 e 3; Sandro Brignone ha scritto il paragrafo 2.

1. Una prima occasione per discutere dell'impostazione della ricerca che si tratta in questo volume è stata offerta dal seminario "L'uso dei Big Data nella ricerca e nella valutazione" organizzato da Mauro Palumbo il 9 luglio 2015 presso l'Università di Genova che ha visto la partecipazione di Francesco Mazzeo Rinaldi (Università di Catania), Davide Anguita (Università di Genova) e da chi scrive.

code nel percorso che si deve compiere; di converso, la nostra geolocalizzazione è utilizzata dal sistema per informare quanti ci seguono, immergendosi nella situazione che il nostro spostamento sta causando. Tutto questo in una sorta di — come abbiamo già detto in altra occasione (Grimaldi, Cavagnero, Gallina, 2016, p. 217) — “lampeggio di solidarietà”.

Le dimensioni indagate riguardano otto tipi di preferenze: politiche, religiose, sessuali, sportive, alimentari, automobilistiche, della moda e del tempo libero. Come riportato in precedenza, abbiamo raccolto il valore attribuito ai dati personali non solo quando questi vengono ceduti, ma anche nel momento in cui la loro cessione porta dei vantaggi in termini di servizi.

Per ogni preferenza abbiamo, quindi, rilevato i due valori con le seguenti coppie di domande (ripetute per ciascuna delle otto preferenze), che, per brevità, chiameremo rispettivamente *OUT* e *IN* (qui sotto riferiti alle preferenze *politiche*):

[politiche_OUT]

Qual è il valore minimo per cui saresti disposto a comunicare le tue preferenze *politiche* sapendo che potranno essere cedute a una organizzazione interessata (che le terrà riservate)? (Esprimi tale valore in euro; nel caso tu ritenga che le tue preferenze non abbiano valore, scrivi: 0) _____

[politiche_IN]

Sapendo che l'organizzazione interessata alle tue preferenze *politiche*, avendo acquisito i tuoi dati, potrà inviarti informazioni o materiale di tuo interesse, quanto ritieni possa valere tale servizio? (Esprimi tale valore in euro; nel caso tu ritenga che non abbia valore, scrivi: 0) _____

Come si può osservare si è rilevato il valore in euro, ossia su di una unità definita e quindi utilizzabile in più aree territoriali, anche fuori dall'Italia (ovviamente dove questa moneta è impiegata). Non abbiamo fornito all'intervistato indicazioni a mo' di esempio allo scopo di non orientare le risposte; ci siamo limitati a dire di scrivere “0” nel caso si ritenesse pari a zero il valore dei propri dati. In questo modo possiamo discriminare eventuali valori mancanti da chi ritiene nullo il valore delle proprie informazioni.

La Tab. 1 riporta per ciascuna dimensione il valore medio calcolato sulle risposte in merito alla consapevolezza dei dati ceduti (*OUT*).

Tab. 1. Valore in euro per cui l'intervistato sarebbe disposto a comunicare le sue preferenze [OUT](valori medi calcolati per ciascuna delle preferenze rilevate).

Statistiche	politiche	sessuali	religiose	sportive	alimentari	moda	tempo libero	automobili
Media	285,0	160,7	148,9	156,6	131,6	144,3	215,5	140,8
N	992	979	987	991	990	978	981	988
Dev. std.	1394,0	816,1	829,2	943,9	740,5	864,6	1131,3	861,4
% valori 0	62,3	64,7	63,7	54,8	52	56	52,6	64,3

Come si può osservare il valore medio più elevato è pari a 285 euro per le preferenze politiche mentre il più basso riguarda le preferenze alimentari con 144 euro². Le deviazioni standard sono elevate a dimostrazione della grande variabilità delle distribuzioni. Ma è l'ultima riga della tabella che merita attenzione: gli studenti universitari che ritengono pari a zero il valore OUT dei propri dati variano tra il 52% e il 65%. Quindi più di metà del campione ritiene (almeno in una delle preferenze) che i propri dati *non abbiano valore*. Vedremo tra breve quanti sono gli intervistati che attribuiscono valore zero a *tutte* le 8 preferenze loro sottoposte.

Tab. 2. Valore in euro del servizio che l'intervistato stima di poter ottenere sapendo che i propri dati sono stati acquisiti in Rete [IN] (valori medi calcolati per ciascuna delle preferenze rilevate).

Statistiche	politiche	sessuali	religiose	sportive	alimentari	moda	tempo libero	automobili
Media	150,1	95,9	120,3	119,2	111,9	89,9	131,0	110,6
N	992	989	992	994	989	983	984	977
Deviazione std.	850,0	651,9	793,4	777,6	662,7	603,5	793,6	767,0
% valori 0	60,3	69,9	66,8	56,1	53,5	56,7	53,8	66,5

2. È importante informare il lettore che sono stati considerati aberranti e posti come "valori mancanti" le valutazioni — sia OUT sia IN — di valore superiore a 10.000 euro. Questi casi, che potremmo attribuire a coloro che ritengono *altissimo/impagabile* il valore dei propri dati, rappresentano nelle diverse preferenze tra l'8,5% e il 10% del campione. Li abbiamo considerati "mancanti" dato che avrebbero condizionato pesantemente il calcolo dei valori medi.

La Tab. 2 restituisce i valori IN, ossia la valutazione dei servizi che “entrano” nel borsino virtuale di cui dispone ciascun intervistato dal momento in cui i propri dati vengono lasciati in Rete. Come si può osservare i valori medi variano da un massimo di 150 euro per le preferenze politiche a un minimo di 90 euro per quelle relative alla moda. Anche in questo caso le deviazioni standard sono elevate a dimostrazione di un’alta variabilità delle distribuzioni. La quota di studenti che valuta “zero” il valore dei servizi che riceve in cambio dei propri dati lasciati in Rete è molto alta: da un minimo del 54% a un massimo del 70%. Quindi sono più della metà del campione quanti ritengono di valore zero i servizi che ricevono sulla base delle proprie preferenze (almeno una) lasciate in Rete.

Inoltre abbiamo sommato i valori OUT delle 8 preferenze (v. Tab. 3): la media risulta pari a 1.107 euro e la deviazione standard (dev. std.= 4.832) informa sull’alta variabilità. Sono il 36% gli studenti universitari che hanno valutato zero tutte le 8 preferenze loro presentate; quindi poco più di un terzo del campione attribuisce valore nullo a tutti i propri dati.

Abbiamo quindi calcolato la somma di tutti gli 8 valori IN. La media è pari a 688 euro, la deviazione standard anche in questo caso è elevata (dev. std.= 3.541) e il 38% del campione ha assegnato zero a tutte le 8 preferenze sottoposte nel questionario; quota simile a quella appena vista per i valori OUT.

A questo punto abbiamo generato una nuova variabile calcolando la differenza tra valore complessivo in euro dei dati ceduti e il valore complessivo dei servizi ricevuti ($OUT - IN$), ottenendo una media pari a 337 euro. La distribuzione è molto dispersa (dev. std.= 5.066 euro) e il valore medio positivo indica che gli studenti universitari valutano maggiormente il valore dei dati ceduti piuttosto che i servizi ricevuti in cambio. Quando i valori si attestano attorno al valore zero indicano quei casi dove l’uscita è pari all’entrata (OUT circa uguale a IN). Su 1.091 casi, 840 hanno risposto a tutte le domande sulla valutazione in euro delle preferenze (sia per quanto riguarda i dati ceduti, sia per i servizi ricevuti); i valori mancanti sono quindi 251 pari al 23%. Degli 840 casi validi, 380 (45%) hanno fornito il medesimo valore sia per i dati ceduti sia per i servizi ricevuti. Di questi ultimi, 276 (32,9% riferito a 840 casi) attribuiscono valore 0 sia ai dati ceduti sia ai servizi ricevuti mentre 104

(12% su 840) danno un medesimo valore a dati ceduti e servizi ricevuti ma diverso da 0. Dunque poco più di un decimo (riferito ai casi validi) del campione ritiene che i propri dati abbiano lo stesso valore sia OUT sia IN e che entrambi siano diversi da 0.

Riassumendo, dunque, il nostro campione dichiara una media di 1.107 euro come valore complessivo attribuito alle proprie preferenze cedute (OUT) e una media di 688 euro in merito al valore attribuito alle proprie preferenze nel caso in cui vengano restituiti servizi (IN). Non sussistono grandi differenze di genere rispetto alle variabili indagate; solamente in merito al valore complessivo attribuito alle proprie preferenze cedute si nota una leggera differenza tra femmine (1.045 euro) e maschi (1.301 euro) e in merito al numero di oggetti posseduti (7,5 per le femmine e 8,4 per i maschi).

2. Il valore dei dati in tre nazioni europee

La Tab. 3 compara le medie delle principali variabili indagate in riferimento alla nazione di residenza dell'intervistato. Dai dati elaborati possiamo osservare che la media del valore complessivo delle preferenze "cedute" (OUT) è superiore in Romania (1.776 euro) rispetto all'Italia e alla Spagna (rispettivamente 1.029 e 847 euro); la media del valore complessivo delle proprie preferenze nel caso in cui vengano restituiti servizi (IN) vale 844 euro per la Spagna, 710 per l'Italia e 435 per la Romania.

Pare, dunque, che gli studenti di Resita (Romania) attribuiscono un maggior valore ai loro dati — che in qualche modo cedono sul Web (OUT) — mentre giudicano di basso valore i servizi che ricevono via Internet (IN); gli universitari di Valencia (Spagna) percepiscono come sostanzialmente uguale il valore del dato OUT e il valore del dato IN, mentre gli studenti italiani giudicano più elevato il valore delle informazioni che cedono rispetto ai servizi che ricevono.

Sebbene i dati elaborati confezionino una certa rilevanza della nazione di provenienza degli intervistati, occorre dire che l'analisi statistica non evidenzia relazioni statisticamente importanti. L'analisi della varianza elaborata tra Nazione (variabile indipendente) e valore OUT (va-

riabile dipendente) consegna un Eta quadro=0,003, ossia la nazione in cui vivono gli intervistati “spiega” solo lo 0,3% del valore complessivo attribuito ai propri dati (OUT). Questa statistica informa che non esiste in pratica una relazione tra le due variabili in oggetto. Stesso risultato emerge se come variabile dipendente si assume il valore complessivo IN (Eta quadro=0,001).

Tab. 3. Comparazione delle medie delle principali variabili cardinali, per Nazione dell'intervistato.

Nazione		Valore complessivo in euro delle proprie preferenze nel caso vengano cedute (OUT)	Valore complessivo in euro delle proprie preferenze nel caso vengano restituiti servizi (IN)	Differenza tra valutazione dei dati ceduti e servizi ricevuti (OUT-IN)	N. attività in Rete	N. oggetti tecnologici posseduti	Età (anni)
Italia	Media	1029	710	193	8,1	3	22,6
	N	643	657	607	804	804	772
	Dev. std.	3708	4100	4023	2,8	1	3,8
Romania	Media	1776	435	1459	6,6	2,5	24,7
	N	121	131	114	166	166	154
	Dev. std.	9791	1041	10104	2,9	1,2	6,2
Spagna	Media	847	844	-4	6,3	2,8	22,4
	N	119	120	119	121	121	121
	Dev. std.	1300	1275	419	2,8	1	2,1
Totale	Media	1107	688	337	7,7	2,9	22,9
	N	883	908	840	1091	1091	1047
	Dev. std.	4833	3541	5066	2,9	1	4,2

Sempre Eta quadro indica che la nazione spiega appena il 6% della variabile numero attività svolte in Rete e il 3% della variabile numero oggetti ICT posseduti. Anche in questi ultimi casi la relazione è bassa o perlomeno irrilevante³.

3. In termini schematici, l'Eta quadro è pari a 0,003 nell'analisi della varianza tra Nazione e Valore complessivo in euro delle proprie preferenze nel caso vengano cedute (OUT); 0,001 tra

Possiamo quindi affermare che gli studenti universitari mostrano scarsa consapevolezza del valore dei propri dati che quotidianamente lasciano con il loro agire in Rete; circa un terzo ritiene pari a zero il valore di tutte 8 le preferenze che abbiamo loro sottoposto, sia OUT sia IN. Ritengono che i dati che più o meno consapevolmente cedono ad altri valgono non più di poche centinaia di euro, mentre mostrano di gradire i servizi che ricevono in cambio, anche se li valutano un po' meno di quanto pensano di cedere ($OUT > IN$). Se pensiamo che un utente di Facebook che posta 6 o 7 foto al mese del proprio gatto e mette qua e là qualche "mi piace", si calcola che restituisca al sistema tra i 300 e i 400 dollari all'anno (gli utenti di Facebook si avvicinano ai 2 miliardi di casi!), ci si rende conto della differenza con cui gli studenti universitari valutano i loro dati e quanto invece valgono nella realtà.

3. *Quiet, effective, collector, smart: per un profilo degli intervistati*

È stata quindi creata una tipologia di studenti caratterizzata dal numero di attività svolte in Rete e dal possesso di oggetti tecnologici. Tale tipologia ha consentito di individuare quattro tipi di studenti: *quiet* (basso numero oggetti; basso numero attività, 24,8%), *effective* (basso numero oggetti; alto numero attività, 16,2%), *collector* (alto numero oggetti; basso numero attività, 23,1%), *smart* (alto numero oggetti; alto numero attività, 35,8%)⁴ (Tab. 4).

Emerge come la maggioranza degli studenti universitari si caratterizzi per avere sia un alto numero di oggetti tecnologici sia per svolgere un elevato numero di azioni in Rete.

Nazione e Valore complessivo in euro delle proprie preferenze nel caso vengano restituiti servizi (IN); 0,008 tra Nazione e Differenza tra valutazione dei dati ceduti e servizi ricevuti ($OUT-IN$); 0,062 tra Nazione e N. attività in Rete; 0,027 tra Nazione e N. oggetti tecnologici posseduti; 0,03 tra Nazione ed Età (anni).

4. Le modalità (alto-basso) di ciascuna variabile (n. oggetti e n. attività) sono state definite tagliando la rispettiva distribuzione attraverso il valore mediano. Tale valore è 3 per la variabile n. oggetti e 8 per n. attività.

Tab. 4. Profilo dell'intervistato definito mediante una tipologia basata sul numero di oggetti tecnologici posseduti e le attività in Rete svolte.

Tipi	Frequenza	Percentuale
Quiet (basso n. oggetti – basso n. attività)	271	24,8
Effective (basso n. oggetti – alto n. attività)	177	16,2
Collector (alto n. oggetti – basso n. attività)	252	23,1
Smart (alto n. oggetti – alto n. attività)	391	35,8
Totale	1091	100

La variabile–indice con cui abbiamo definito il profilo degli intervistati, è stata quindi messa in relazione con le principali variabili cardinali rilevate nella ricerca (Tab. 5). In riferimento ai dati OUT si osserva come i valori più elevati si registrino in corrispondenza del tipo *effective* (1.451 euro) e del tipo *smart* (1.237 euro), caratterizzati dall'aver un alto numero di attività in Rete. Lo stesso risultato lo vediamo con i valori IN, rispettivamente 916 e 836 euro.

L'analisi della varianza informa che il tipo di studenti (*quiet*, *effective*, *collector*, *smart*) spiega pochissimo la varianza della variabile OUT (circa il 3 per 1.000, ossia $Eta\ quadro = 0,003$) e la stessa cosa si rileva con la variabile IN ($Eta\ quadro = 0,001$). Si può quindi desumere che il tipo di studente — identificato mediante il numero di oggetti tecnologici posseduti e il numero di attività svolte in Rete — abbia un potere di spiegazione irrilevante sui ai valori OUT e IN⁵.

Nonostante i dati elaborati portino a pensare a comportamenti differenti nei vari tipi di studenti individuati, l'analisi statistica suggerisce cautela in questa interpretazione puntando piuttosto sulla scarsa o nulla influenza dei Tipi sulla variabilità di OUT e IN. Possiamo però almeno suggerire che la consapevolezza del valore dei propri dati del tipo *effective* e del tipo *smart* è simile ed elevata, mentre la consapevolezza di *quiet* e *collector* è simile e più bassa.

5. Anche l' $Eta\ quadro = 0,021$ tra Tipi ed Età dell'intervistato allude a una relazione inesistente tra le due variabili; sono invece consistenti gli $Eta\ quadro$ calcolati tra le variabili Tipi e Numero attività in Rete (0,646) e Tipi e Numero oggetti tecnologici posseduti (0,666), ma ciò è dovuto semplicemente al modo in cui si è costruita la variabile Tipi che, come già detto, include le variabili sopra riportate come dipendenti nell'analisi della varianza.

Tab. 5. Comparazione delle medie delle principali variabili cardinali per tipo di intervistato.

Tipi		Valore complessivo in euro delle proprie preferenze nel caso vengono cedute (OUT)	Valore complessivo in euro delle proprie preferenze nel caso vengono restituiti servizi (IN)	Differenza tra valutazione dei dati ceduti e servizi ricevuti (OUT-IN)	N. attività in Rete	N. oggetti tecnologici posseduti	Età (anni)
Quiet (basso n. oggetti – basso n. attività)	Media	910	358	547	4,9	1,8	23,8
	N	214	217	200	271	271	258
	Dev. std.	5013	884	5097	1,8	0,4	5,4
Effective (basso n. oggetti – alto n. attività)	Media	1451	916	550	9,4	1,9	23,1
	N	137	146	129	177	177	171
	Dev. std.	7524	3357	7232	1,4	0,2	3,9
Collector (alto n. oggetti – basso n. attività)	Media	884	640	340	5,7	3,5	22,6
	N	210	216	202	252	252	242
	Dev. std.	3411	2995	1612	1,4	0,6	4
Smart (alto n. oggetti – alto n. attività)	Media	1237	836	110	10,1	3,6	22,3
	N	322	329	309	391	391	376
	Dev. std.	3976	4815	5440	2	0,8	3,1
Totale	Media	1107	688	337	7,7	2,9	22,9
	N	883	908	840	1091	1091	1047
	Dev. std.	4833	3541	5066	2,9	1	4,2

In conclusione, considerando l'uso pervasivo della Rete e dei social sites da parte delle nuove generazioni, si ritiene importante favorire una riflessione in merito alla consapevolezza che questi ultimi hanno delle tracce digitali. A tal proposito Giorgio Alleva, presidente dell'Istituto Nazionale di Statistica dal 2014 al 2018, in un intervento del 3 marzo 2017 presso l'università Federico II di Napoli, sottolinea come i Big Data rappresentino nuove sfide per gli istituti di ricerca, favorendo prospettive di integrazione e uso di nuove fonti di dati, portando a interrogarsi

sull'ambiente in cui sono prodotti i dati, ponendo attenzione a problemi di privacy e sicurezza delle informazioni.

Occorre dunque interrogarsi su come poter utilizzare i Big Data per aumentare il benessere individuale e collettivo. Ciò in una prospettiva di educazione al valore dei dati, che va offerta fin dalla prima infanzia, allo scopo di prevenire possibili e future situazioni di rischio e comportamenti devianti, come recitano le stesse circolari ministeriali del settore.

Riferimenti bibliografici

- BOYD D., CRAWFORD K. (2012), *Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon*. *Information, Communication & Society*, 15 (5), pp. 662–679.
- CRAWFORD K. (2013), *The hidden biases of big data*. *Harvard Business Review Blog*, Available at: <http://blogs.hbr.org/2013/04/the-hidden-biases-in-big-data/>.
- GRIMALDI R., CAVAGNERO S.M., GALLINA M.A. (2016), *Usò della rete e consapevolezza delle tracce digitali. Una ricerca tra studenti universitari*, «Sociologia e Ricerca Sociale», XXXVII, n. 109, pp. 204–218.
- GRIMALDI R., GALLINA M.A., CAVAGNERO S.M. (2017), *Big data e tracce digitali. Una ricerca internazionale tra studenti universitari*, in Trinchero R., Parola A. (a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli, pp. 191–207.
- GRIMALDI R., GALLINA M.A., CAVAGNERO S.M. (2018), *Il valore dei propri dati. Una comparazione tra studenti italiani, romeni e spagnoli*, «Sociologia Italiana», AIS Journal of Sociology, Milano Egea, pp. 159–171.
- HARGITTAI E. (2015), *Is bigger always better? Potential biases of big data derived from socialnetwork sites*, «The Annals of the American Academy of Political and Social Science», 659(1), 63–76.
- LAZER D., PENTLAND A., ADAMIC L., ARAL S., BARABASI A., BREWER D., ... VAN ALSTYNE M. (2009), *Computational social science*, «Science», vol. 323 (n. 5915), pp. 721–723.
- TRINCHERO R., PAROLA A. (2017)(a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli.

Pratiche digitali tra competenza e ingenuità

Le determinanti della consapevolezza delle proprie impronte digitali
in un campione di studenti universitari

di STEFANO POLI, CLAUDIO TORRIGIANI*

1. Internauti consapevoli?

Ormai numerosi autori hanno sottolineato come la diffusione e l'utilizzo pervasivo delle tecnologie dell'informazione e della comunicazione (TIC) siano da considerare uno strumento estremamente rilevante nella costruzione identitaria dell'attore sociale, nelle sue pratiche di socialità e di interpretazione del mondo (Grimaldi, Gallina, Cavagnero, 2017), fino a parlare della *web society* come di un oggetto di analisi che pone sfide inedite allo studioso e in particolare al sociologo (Cipolla, 2015).

L'uso delle TIC è riconosciuto anche in ambito educativo e formativo come elemento chiave nella costruzione della conoscenza, fino a parlare di didattica digitale (Graziano, Martena, 2017) — anche come strumento per il recupero scolastico (Cavagnero, Gallina, Marchisio, 2015) — e dell'uso delle tecnologie digitali a supporto delle persone con disturbi specifici dell'apprendimento (Schiavo et al., 2016).

L'utilizzo intensivo di TIC in svariati ambiti lavorativi, segnalato tra l'altro dalla rapida diffusione dei connessi neologismi (*smartworking, e-working, ICT-base mobile working...*), determina un cambiamento radicale nelle relazioni di lavoro che contiene in sé opportunità, ma anche rischi e minacce, per lavoratori e aziende (Albano et al., 2017).

Oltre alle pratiche educative e lavorative, Internet ha rivoluzionato il modo attraverso cui comunichiamo, pratichiamo e miglioriamo la nostra salute (Cipolla, Maturo, 2014). Si parla ormai di Medicina 4.0, definita come

* Università degli Studi di Genova. Seppur il contributo nasca dalle riflessioni congiunte dei due autori, sono da attribuirsi a Stefano Poli i paragrafi 2, 3 e 4, e a Claudio Torrighiani i paragrafi 1 e 5.

«l'integrazione di elettronica e tecnologia delle microstrutture per nuove forme di terapia come la chemioterapia personalizzata e la telemedicina terapeutica» (Cappelletti, 2018, p. 132). Tale evoluzione pone prospettive di cambiamento rivoluzionarie per medici e pazienti, sia in termini di personalizzazione dell'assistenza legata all'uso di *App* e di altri strumenti digitali sia in termini di questioni etiche riguardanti la concezione di "umano" e di problemi di responsabilità e di dominio delle attività, anche in relazione all'uso dei Big Data e dell'intelligenza artificiale in questo ambito (ibidem, pp. 133–135).

La Rete si configura sempre più anche come spazio di partecipazione politica dei cittadini che, grazie a questo medium, possono più facilmente salire lungo la "scala della partecipazione" (Arnstein, 1969) a cominciare dall'informazione attiva per poi prendere parte, attraverso modalità attentamente studiate e gestite dal decisore, a processi di formazione, ascolto e dialogo che, nei casi opportuni possono sfociare, nella partecipazione alla deliberazione (Palumbo, Congiu, 2009). L'uso delle TIC per facilitare la partecipazione dei cittadini pone tuttavia la necessità di una garanzia autorevole e neutrale sulle metodologie attraverso le quali si realizzano le diverse forme di partecipazione, nel rispetto del massimo grado possibile di apertura e inclusività per tutte le parti interessate e di sistemi che sappiano garantire un grado massimo di partecipazione, trasparenza, tempestività e qualità dei feedback (Ribaudò et al., 2016). La partecipazione attraverso la Rete va tuttavia promossa e realizzata nella consapevolezza che le TIC "non sono lì ad attendere con neutralità e pazienza un uso 'intelligente' da parte di cittadini e istituzioni, ma sono sempre più utilizzate dai detentori del potere istituzionale e sostanziale per profilare i casi, promuovere i consumi, cambiare le opinioni, diffondere modelli di comportamento" (ibidem, p. 63).

Un numero sempre maggiore di persone possiede dispositivi dotati di sensori in grado di rilevare parametri biofisici e ambientali e, attraverso la Rete, condivide tali dati con altri utenti e con piattaforme capaci di fornire in tempo reale feedback e indicazioni *ad hoc* grazie ad applicazioni tecnologiche di IA nei diversi filoni del *Soft Computing* — come ad esempio le reti neurali e gli algoritmi genetici — e dell'*Hard Computing*, come le ontologie semantiche (Mello, Boari, 2018). Tali processi connettono noi stessi e gli oggetti che ci circondano con la Rete generando e rigenerando senza soluzione di continuità quello che è ormai comunemente definito l'Internet delle cose, IoT (Ashton, 2009), uno spazio di connessione e di collegamento

tra reale e virtuale che, se da un lato consente possibilità di controllo di oggetti e luoghi fisici tramite la Rete impensabili fino a pochi anni fa, dall'altro stempera la capacità di controllo di tali processi da parte dell'utente comune, esponendolo alla potenziale invadenza dell'esperto che domina tali campi del sapere e delle sue applicazioni. L'alimentazione continua della Rete, cui ciascuno di noi quotidianamente contribuisce condividendo dati inerenti a opinioni, a desideri e preferenze, a relazioni, a posizioni e spostamenti attraverso lo spazio fisico e sociale, a comportamenti e interessi "riflessi nello specchio digitale" (Giannotti, 2015) produce una «mole potenzialmente infinita di informazioni denominata Big Data» (Grimaldi, Gallina, Cavagnero, 2017, p. 194). Essi sono distinguibili in *transactional data* — ossia le informazioni raccolte nell'ambito degli scambi tra cittadini e amministrazioni — e *digital by product data* — che consistono in dati pubblicati dagli utenti in Rete tramite le normali attività sui social network o anche semplicemente attraverso gli spostamenti quando i sistemi di geolocalizzazione dei dispositivi digitali sono attivati (Lombi, 2015).

Uno sguardo agli ultimi dati ufficiali conferma che, anche nel 2018, la diffusione delle TIC e l'accesso alla Rete ha continuato a crescere nel nostro Paese (Istat, 2019).

Negli ultimi 12 mesi, infatti, il 68,5% delle persone di 6 anni e più si è connesso alla Rete, con un incremento di oltre 3 punti percentuali rispetto al 2017 e il 52,1% dei cittadini risulta accedervi tutti i giorni. I giovani si confermano i più grandi utilizzatori di Internet, con il 94% dei 15-24enni, ma l'uso della Rete diviene via via più significativo anche tra i 65-74enni che, nell'ultimo anno, fanno registrare una quota di internauti pari al 39,3% (+8,5 punti rispetto al 2017). Percentuali un po' più elevate di utilizzatori della Rete sono riportate nel capitolo "Comunicazione e media" del Cinquantaduesimo Rapporto sulla situazione sociale del Paese 2018, dove si legge che gli italiani che usano internet passano dal 75,2% al 78,4% della popolazione, con una differenza positiva di 3,2 punti percentuali rispetto all'anno precedente e di 33,1 punti dal 2007 (Censis, 2018)¹.

La quota di famiglie che dispongono di un accesso a Internet da casa sale dal 71,7% registrato nel 2017 al 75,1% del 2018 e la dotazione della banda

1. Nel Rapporto si legge inoltre che gli italiani che utilizzano gli smartphone salgono dal 69,6% al 73,8% (+ 4,2 punti) (Censis, 2018).

larga fissa e/o mobile sale dal 70,2% al 73,7%². La connessione fissa (ADSL, fibra ottica, ecc.) è comunque la modalità di accesso più diffusa (Istat, 2019).

Da notare, a fronte di questo dato aggregato, un forte divario digitale tra le famiglie da ricondurre, in particolare, a fattori generazionali e culturali: più connessi risultano infatti i nuclei in cui è presente almeno un minore (94,4% di collegamento a banda larga) mentre meno connesse sono le famiglie composte soltanto da ultrasessantacinquenni, fra cui solo il 31,4% dispone di una connessione a banda larga. Un altro fattore discriminante, che emergerà tra l'altro nell'analisi riportata di seguito, risulta essere il titolo di studio: dispone di una connessione a banda larga il 94,9% delle famiglie con almeno un componente laureato, contro il 64% delle famiglie in cui il titolo di studio più elevato è la licenza media (ibidem).

Per quanto concerne i device, tra le persone di 14 anni e più, lo strumento più utilizzato per l'accesso alla Rete è lo smartphone³ (89,2%), seguito a lunga distanza dal Pc da tavolo (45,4%). Il 28,3% utilizza un laptop o un netbook, il 26,1% un tablet, mentre il 6,7% altri dispositivi mobili.

Tra le finalità di utilizzo della Rete dei cittadini spiccano quelle legate al *loisir* e all'acquisto di beni e servizi: risulta infatti che fra le persone di 16–74 anni che hanno usato Internet negli ultimi 3 mesi, il 20,2% lo ha fatto per trovare un alloggio, contattando direttamente un privato tramite siti web/app. Sono pari al 55,9% gli internauti che acquistano online (+3 punti percentuali rispetto al 2017) e, comunque, il 43,9% di chi non ha fatto acquisti negli ultimi 3 mesi ha utilizzato la Rete per cercare informazioni su merci/servizi e/o ha venduto beni online (ibidem). Gli utenti dei social network aumentano dal 67,3% al 72,5% della popolazione. Continuano ad aumentare gli utenti di WhatsApp (il 67,5% degli italiani, l'81,6% degli under 30), mentre più della metà della popolazione fa ricorso ai due social network più popolari: Facebook (56%) e YouTube (51,8%) (Censis, 2018).

2. A tale riguardo è da notare che il tasso medio di diffusione della banda larga tra le famiglie residenti con almeno un componente di 16–74 anni, nei Paesi europei è dell'86% e l'Italia, con un tasso dell'83%, presenta un gap di 3 punti percentuali. Guardando nuovamente al complesso dei nuclei familiari, a livello subnazionale le differenze tra le regioni sono notevoli e confermano il vantaggio del Centro e in particolare del Nord Italia: il Trentino Alto-Adige e la Lombardia sono le regioni con la percentuale più alta di famiglie dotate di connessione con banda larga (rispettivamente 80,7% e 78%); all'opposto il Molise (63,8%), la Calabria e la Sicilia (intorno al 64,5%).

3. Secondo il Censis, gli italiani che utilizzano gli smartphone salgono dal 69,6% al 73,8% (+4,2 punti) (Censis, 2018).

L'uso della Rete come luogo di ricerca e consumo di beni e servizi trova riscontro nei dati relativi alle aziende italiane se, come emerge dal report dell'Istat, nel 2017 il 64,1% delle imprese che hanno venduto via web ha utilizzato almeno un *e-marketplace* (+11,3 punti percentuali rispetto al 2016) e il 50,2% ha realizzato almeno la metà del fatturato via web tramite intermediari (+11,1). Il 94,2% delle imprese con almeno 10 addetti si connette alla Rete in banda larga mobile o fissa. Gli investimenti aziendali nelle nuove tecnologie e nelle competenze dei dipendenti sono tuttavia piuttosto deludenti: nel 2018 sono infatti pari solo al 16,9% le imprese che investono sulle competenze digitali dei propri addetti, poco meno di tre imprese su dieci hanno effettuato investimenti tecnologici nel biennio 2016–2017 e altrettante intendono farne in quello successivo: la sicurezza informatica, le applicazioni web e il *cloud computing* risultano le priorità aziendali delle imprese nelle TIC. In prospettiva, lo sviluppo delle competenze tecnologiche delle risorse umane già occupate in azienda è considerato uno dei principali fattori competitivi solo dal 22,4% delle imprese con almeno 10 addetti, l'accesso ad infrastrutture e connessioni in banda ultra-larga dal 30,8% di esse (ibidem).

Se i dati italiani mostrano un trend lentamente crescente, che potrebbe indurre a un certo ottimismo riguardo agli obiettivi dell'Agenda digitale europea, peraltro già stemperato dai marcati divari regionali, questo si smorza ulteriormente proprio alla comparazione con i dati degli altri paesi europei. Un riferimento di un certo interesse in questo senso è il *Digital Economy and Society Index* (DESI), un indice composito che riassume gli indicatori più rilevanti sulle prestazioni digitali e tiene traccia dell'evoluzione degli Stati membri dell'UE nella competitività digitale, monitorando l'attuazione della *Digital Single Market*, una delle 10 Priorità della Commissione europea per il periodo 2015–19⁴. Le dimensioni considerate nel calcolo dell'indice sono la connettività, il capitale umano, l'uso di Internet, l'integrazione delle tecnologie digitali e i servizi pubblici digitali, misurati attraverso una serie di 34 indicatori⁵. Nella Tab. 1 sono riportati,

4. https://ec.europa.eu/commission/priorities/digital-single-market_en.

5. Si veda <https://digital-agenda-data.eu/datasets/desi/indicators#desi-overall-index>. Il DESI rappresenta uno strumento estremamente utile per un primo confronto ma presenta alcuni limiti relativi alla completezza degli indicatori, all'aggiornamento dei dati e all'orientamento al *policy making*. Per questo, l'Osservatorio Agenda digitale del Politecnico di Milano lavora ai *Digital Maturity Indexes* (DMI), un framework di 118 indicatori più completo e preciso del DESI per misurare la trasformazione digitale dei vari

oltre alla media EU 28, i valori del DESI per Italia, Spagna e Romania, i paesi in cui è stata realizzata l'indagine già presentata nei suoi tratti generali da Grimaldi, Gallina e Cavagnero (2017), i cui dati sono analizzati nel seguito del contributo per comprendere quali siano i fattori associati alla maggiore o minore consapevolezza degli internauti sulle tracce digitali che, quotidianamente, essi lasciano utilizzando la Rete.

Ebbene, per l'anno 2018, l'indice composito per l'Italia è pari a 44,25 su 100, e pone il nostro paese in quartultima posizione, prima di Bulgaria, Grecia e Romania. Anche senza guardare alla testa di serie, la Danimarca, che con un indice pari a 71,73 distacca l'Italia di oltre 27 punti (!), il solo confronto con la media europea (53,98) è impietoso per il nostro Paese, per non parlare di quello con la Spagna (58,05), che si colloca in decima posizione nella graduatoria, con performance sopra la media europea per connettività, integrazione delle tecnologie e servizi pubblici digitali.

Tab. 1. Digital Economy and Society Index 2016 e 2018, Italia, Spagna, Romania e media EU 28.

	Italia	Spagna	Romania	EU 28
	2016			
1 Connettività	10,53	13,46	11,34	13,48
2 Capitale umano	9,53	12,77	7,05	13,16
3 Uso di Internet	5,09	6,72	4,45	6,80
4 Integrazione tecnologie digitali	5,97	6,96	3,68	6,91
5 Servizi pubblici digitali	7,02	9,89	4,91	7,47
DESI	38,14	49,80	31,43	47,82
	2018			
1 Connettività	13,21	16,18	14,53	15,64
2 Capitale umano	10,20	13,64	8,02	14,12
3 Uso di Internet	5,60	7,41	5,24	7,57
4 Integrazione tecnologie digitali	7,36	9,96	3,55	8,02
5 Servizi pubblici digitali	7,88	10,85	6,20	8,62
DESI	44,25	58,05	37,55	53,98

Fonte: Commissione europea, Digital Scoreboard, dati estratti il 30 gennaio 2019 all'indirizzo <https://digital-agenda-data.eu/charts/desi-composite/#>

Paesi, essere meno esposti al mancato aggiornamento di alcuni dati e dare indicazioni utili ai *policy maker*. Secondo gli ultimi dati disponibili (riferiti al 2017), l'Italia è 22esima su 28 Paesi europei per sforzi fatti nell'attuazione della propria Agenda Digitale e 25esima per risultati raggiunti (Gastaldi, Trabacchi, 2019).

Viceversa la Romania fa registrare un indice pari a 37,55 e performance sotto la media, in particolare riguardo al capitale umano e all'integrazione delle tecnologie digitali. Da notare che per quanto riguarda l'indicatore "capitale umano", che comprende "inclusione e competenze digitali", per l'Italia i dati relativi a queste ultime sono segnalati come "non disponibili", mentre nell'indice DESI per il 2017, la percentuale di italiani in possesso per lo meno di competenze digitali di base era pari al 44% a fronte di una media europea pari al 56%⁶.

Ai fini di questo contributo risulta centrale proprio il riferimento al capitale umano e, in particolare, alle competenze digitali, richiamate di recente anche nella Raccomandazione del Consiglio dell'Unione europea del 22 maggio 2018⁷ relativa alle competenze chiave per l'apprendimento permanente, in cui si legge che, tra l'altro, gli Stati membri dovrebbero «innalzare e migliorare il livello delle *competenze digitali* in tutte le fasi dell'istruzione e della formazione per tutti i segmenti della popolazione».

Un riferimento specifico in tal senso è rappresentato dal Modello europeo DigComp 2.1 (Carretero, Vuorikari, Punie, 2017, trad. it. 2018), in cui le competenze digitali di base⁸ sono articolate in 5 aree — 1. alfabetizzazione su informazioni e dati, 2. comunicazione e collaborazione, 3. creazione di contenuti digitali, 4. sicurezza, 5. risoluzione di problemi — entro cui sono descritte 21 competenze specifiche con 8 diversi livelli di padronanza. Tra le aree di competenza considerate, quelle maggiormente interessate da un approfondimento sulla consapevolezza degli utenti della Rete riguardo alle "tracce" che essi lasciano lungo i propri percorsi virtuali sono la numero 1, la 2 e 4 e, in particolare, rispettivamente le competenze: 1.3 gestire dati, informazioni e contenuti digitali per la prima area, 2.2 condividere informazioni attraverso le tecnologie digitali e 2.6 gestire l'identità digitale, per la seconda e 4.2 proteggere i dati personali e la privacy, per la quarta.

6. Indice di digitalizzazione dell'economia e della società (DESI), relazione nazionale sull'Italia per il 2018, disponibile al link http://ec.europa.eu/information_society/newsroom/image/document/2018-20/it-desi_2018-country-profile-lang_4AA6AC9F-0F0F-0F48-8D21A979E9D5A1B7_52348.pdf consultato il 05/02/2019.

7. Numero documento ST/9009/2018/INIT pubblicato nella Gazzetta Ufficiale dell'Unione europea del 4 giugno 2018 riferimento 2018/C 189/01.

8. Non prendiamo in esame in questa sede le competenze di e-leadership e le competenze specialistiche: cfr. <https://competenze-digitali-docs.readthedocs.io/it/latest/index.html>.

Ma, ci chiediamo, possiamo ritenere che i circa 33 milioni di utenti residenti in Italia che ogni giorno a novembre 2018 hanno navigato attraverso la rete in media per oltre 3 ore⁹ fossero effettivamente consapevoli di quali e quante informazioni condividevano e mettevano a disposizione con i loro “click”? Possiamo supporre che fossero consapevoli di essere seguiti nei loro spostamenti virtuali (o anche fisici) grazie alle tracce lasciate attraverso i propri dispositivi, come del resto avviene con le telecamere che monitorano gli acquisti dei clienti nei supermarket o nei centri commerciali? Che si siano quantomeno posti il problema di non subire il furto delle proprie credenziali di accesso all’e-mail personale o, peggio, al proprio *online banking*? Quanti degli utilizzatori della Rete conoscono il Regolamento europeo 2016/679 sulla privacy e la protezione dei dati personali, più noto come GDPR? Quanti conoscono i diritti e gli strumenti che il GDPR conferisce loro per tutelare i dati personali?

Abbiamo visto sopra, anche attraverso alcune fonti statistiche ufficiali, che l’accesso alla Rete e l’uso delle tecnologie digitali sono sempre più diffusi nella popolazione italiana (anche se sotto la media europea), soprattutto tra i giovani, ma sempre più anche tra gli anziani. Abbiamo poi ricordato come tale utilizzo generi una mole ingente di dati e informazioni messi a disposizione dagli internauti per poter usufruire dei prodotti e servizi che la Rete può loro fornire, offrendo a provider e analisti «la radiografia completa di interessi, opinioni, consumi, spostamenti, in sostanza pezzi della nostra vita che come tessere di un mosaico si scompongono e ricompongono per formare il nostro profilo identitario» (Soro, 2013, p. 8, cit. in Grimaldi, Gallina, Cavagnero, 2017). Tuttavia non è affatto scontato che gli utenti abbiano piena consapevolezza riguardo alla qualità e quantità di dati che essi ogni giorno cedono in cambio dei servizi fruiti, a come e quanto questo possa incidere sul loro diritto alla riservatezza, a quale sia il valore anche commerciale delle informazioni fornite o l’uso specifico che ne può essere fatto anche ai fini del controllo del potere economico e politico.

A mero titolo di esempio, riguardo al valore dei dati forniti, secondo i risultati della ricerca dell’Osservatorio Big Data Analytics & Business Intelligence della School of Management del Politecnico di Milano, il mer-

9. Utenti unici, giorno medio a novembre 2018, ultimi dati AudiWeb. Disponibile al link: <http://www.audiweb.it/>, consultato il 01/02/2019.

cato italiano dei Big Data Analytics ha continuato la sua espansione anche nel 2018, raggiungendo un valore complessivo di 1,393 miliardi di euro, in crescita del 26% rispetto all'anno precedente¹⁰.

Un impulso a una maggiore consapevolezza del valore potenziale delle informazioni quotidianamente diffuse dagli internauti attraverso i social e dell'uso che ne può essere fatto è venuta, nel marzo 2018, dallo scandalo Cambridge Analytica, che ha ottenuto i dati di circa cinquanta milioni di utenti di Facebook approfittando della possibilità che l'azienda concedeva a solo scopo di ricerca — e che dal 2015 non è più in vigore — di ottenere le informazioni su tutti i contatti degli utenti. Ormai chiusa per bancarotta dal maggio 2018, Cambridge Analytica avrebbe poi usato le informazioni così ottenute per prevedere gli indirizzi di voto e modificarli attraverso campagne mirate di propaganda sui canali social, supportando tra gli altri Donald Trump alle Presidenziali del 2016.

Diviene allora fondamentale riflettere in modo critico e approfondito in merito alla consapevolezza degli utilizzatori riguardo alle informazioni e alle tracce digitali che continuamente lasciano in Rete, in particolare con riferimento ai più giovani che, come abbiamo visto sopra, sono in percentuale maggiore più continuamente connessi. Diviene fondamentale comprendere quali siano i loro stili di fruizione della Rete, i dispositivi tecnologici di cui dispongono, il livello di consapevolezza sul loro “mettersi a nudo” in Rete e il valore da essi attribuito alle informazioni che possono decidere di condividere (o condividere inconsapevolmente...) nella “piazza virtuale”. Questi, grosso modo, sono gli obiettivi conoscitivi dell'indagine sui Big Data e l'uso della Rete realizzata nell'anno accademico 2016–2017 su un campione di studenti universitari italiani, spagnoli e rumeni, i cui risultati vengono approfonditi di seguito nel tentativo di individuare, tra quelle rilevate, le caratteristiche che possono contribuire a spiegare la maggiore o minore consapevolezza dei rispondenti in merito alla tracciabilità dei loro comportamenti in Rete e all'uso che svariate organizzazioni quotidianamente fanno dei dati così prodotti.

10. https://www.osservatori.net/it_it/osservatori/comunicati-stampa/big-data-analytics-italia-mercato-2018 consultato il 31/01/2019.

2. Aspetti metodologici e definizione delle variabili principali

Per definire la variabile principale dell'indagine, ovvero l'effettiva consapevolezza delle tracce digitali che inevitabilmente lasciamo durante le nostre attività in Rete, siamo partiti da una considerazione di fondo, ovvero che i nostri dati e la nostra presenza viene rilevata spesso senza che noi lo sappiamo e indipendentemente dal nostro consenso. Al di là delle implicazioni etiche e normative, l'aspetto di cui sopra costituisce una determinante essenziale del nostro essere più o meno consapevoli dell'inevitabile tracciabilità digitale delle nostre attività in rete. Questo, naturalmente, prescinde dalle situazioni in cui noi coscientemente e deliberatamente diamo il consenso e condividiamo le nostre informazioni personali o i nostri dati attraverso dispositivi digitali. Non di rado, tuttavia, le persone sono ignare quando non particolarmente preoccupate dalla propria tracciabilità digitale o ritengono, impropriamente, che la stessa sia preservata dall'obbligo di un consenso espresso all'utilizzo che il provider o altri soggetti terzi (spesso altamente interessati ai nostri profili come consumatori o utenti) debbano ottenere da noi.

A tale proposito si è definita la variabile dipendente, ovvero la consapevolezza della propria tracciabilità digitale, attraverso una domanda specifica: "Ritieni che l'uso delle tecnologie infotelematiche possa generare tuoi dati che possono essere utilizzati da aziende commerciali o altre organizzazioni?". A tale domanda si sono previste quattro modalità di risposta: 1- "No"; 2- "Sì, ma solo se lo consento esplicitamente"; 3 "Sì, anche se non lo consento esplicitamente"; 4 "Non saprei". Sulla base di quanto detto sopra, in sede di analisi bi-variata è stato possibile operativizzare una piena consapevolezza solo per la modalità di risposta 3 (che prevede la consapevolezza di una tracciabilità indipendente dal consenso), una parziale consapevolezza per la modalità 2 (ovvero, una tracciabilità condizionata al consenso espresso da parte della persona) e una consapevolezza nulla (o in qualche modo ignara o indifferente) nelle modalità di risposta 1 e 4 ("No" e "Non saprei").

Successivamente, per la realizzazione di un modello di regressione logistica binaria si è ricodificato in sole due categorie ovvero "Piena consapevolezza" (modalità 3) e "Consapevolezza parziale o nulla (modalità 2, 1 e 4).

Diverse variabili indipendenti sono state applicate per verificare l'associazione di differenti fattori con i diversi livelli di consapevolezza della propria tracciabilità digitale.

Considerando le caratteristiche individuali si sono presi in esame il genere (maschi e femmine) e l'età (ricodificata in due fasce, ovvero fino ai 24 anni e oltre i 24 anni). Trattandosi, inoltre, di studenti universitari, si è deciso di verificare l'incidenza del corso di studi (distinto tra due modalità, ovvero *humanities* o *hard sciences*) e dell'anno di corso (distinguendo tra matricole e iscritti agli anni successivi). Per cogliere l'estrazione socioeconomica si sono rilevati i titoli di studio di entrambi i genitori, successivamente ricodificati in basso-medio (fino al diploma) e in alto (dalla laurea in poi).

Si sono successivamente presi in considerazione altri set di variabili ovvero:

- a) la quantità e il tipo di oggetti tecnologici posseduti, nello specifico: Console per videogame, come X-box, Playstation 4, ecc.; Smartphone; Tablet, Computer fisso, Computer portatile, Smartwatch, Elettrodomestico che si possa comandare via Internet. Il possesso/utilizzo di tali oggetti è stato ricodificato in una dicotomica di tipo "0=No" / "1=Si" e si è proceduto alla somma degli oggetti posseduti da ciascun rispondente. Successivamente, a fronte di una media campionaria di 2,8 oggetti tecnologici mediamente posseduti, il numero di oggetti tecnologici in possesso dei rispondenti è stato ricodificato in 0 = "sotto o entro la media" e 1 = "sopra la media";
- b) la quantità e il tipo di attività tecnologiche praticate in Rete, nello specifico: sottoscrivere tessere in supermercati/negozi; fare acquisti on-line; fare vendite on-line; rispondere a sondaggi aziendali, di partiti politici, ecc.; iscriversi a siti per ottenere informazioni; iscriversi a siti per ottenere App gratuite; postare su social network; pubblicare su blog; mettere "mi piace" su post sensibili su Facebook; postare foto/video; fornire il proprio stato emotivo su Facebook o simili; fornire il proprio stato fisico (es. peso, battito cardiaco); consentire la propria geo-localizzazione via cellulare o simili; usare sistemi di condivisione cloud, per esempio Dropbox, ecc.; usare App per car-sharing; usare App per bike-sharing. Ciascuna di tali pratiche è stata ricodificata in una dicotomica di tipo "0=No" / "1=Si" e anche in questo caso si è proceduto alla somma delle pratiche complessivamente svolte in Rete da ciascun rispondente. Una

volta calcolata la media, risultante in 7,6 attività mediamente svolte dai rispondenti, si è ricodificato il numero complessivo di attività in Rete in 0 = “sotto o entro la media” e 1 = “sopra la media”;

- c) il valore complessivo in euro delle proprie informazioni personali (per esempio, di consumo ludico–espressivo, di gradimento, di opinione politica, ecc.) in due situazioni, ovvero a) nel caso venissero cedute direttamente e b) in cambio di servizi. Nello specifico si è rilevato che i rispondenti sarebbero mediamente disponibili a cedere le proprie informazioni personali in cambio di 1.107 Euro, cifra che, in caso di controprestazione di servizi, scenderebbe a 688,20 Euro. Anche in questo caso si è provveduto a ricodificare entrambe le due variabili anzidette, ovvero, il valore in caso di cessione e il valore in cambio di servizi, in 0 = “sotto o entro la media” e 1 = “sopra la media”.

3. I risultati dell'indagine: le determinanti della consapevolezza digitale tra i giovani intervistati

L'analisi bi–variata nelle tabelle 2 e 3 restituisce alcuni primi risultati rispetto alla consapevolezza dei rispondenti circa la tracciabilità delle proprie impronte digitali. Come da tabella 2, su 1.074 rispondenti, il 19,3% si mostra decisamente inconsapevole del fatto di poter essere tracciato; il 42,6% evidenzia una consapevolezza parziale, confidando che il consenso espresso possa costituire un vincolo invalicabile per aziende o provider nella raccolta dei propri dati o informazioni personali; solo il 38,1%, appena poco più di un terzo, è pienamente consapevole che nel momento stesso in cui accede a uno strumento digitale, specie se online, altri possano carpire dati e informazioni a prescindere da qualsiasi autorizzazione.

Sempre in tabella 2 è possibile incrociare il risultato precedente con alcuni fattori socio–anagrafici. L'associazione con il sesso dei rispondenti non appare statisticamente significativa; peraltro i risultati tra i maschi e le femmine appaiono sostanzialmente equivalenti, sebbene occorra tener conto di una certa sotto–rappresentazione nel campione, in ragione di un rapporto tra maschi e femmine di uno a tre. Non pienamente significativa sul piano statistico risulta l'associazione per fascia d'età, tuttavia è interessante notare che sopra i 24 anni aumenti sia la percentuale di coloro che sono pienamen-

te consapevoli, sia quella degli ignari, mostrando come al crescere dell'età il campione tenda a dicotomizzarsi in termini di maggiore o minore consapevolezza delle conseguenze dei propri comportamenti digitali.

Rispetto all'estrazione socioeconomica derivata dai livelli d'istruzione dei genitori, se non è statisticamente significativa l'associazione tra diversi gradi di consapevolezza e il titolo di studio del padre, è sostanzialmente significativa l'associazione con quello della madre, evidenziando un deciso aumento della percentuale dei consapevoli e riduzione degli ignari al crescere del livello d'istruzione materno: infatti, la percentuale dei primi sale da 36,5% in caso di titolo di studio medio-basso a ben 43,3% in caso di madri in possesso di almeno un titolo di laurea mentre quella dei secondi scende dal 20,6% al 14,7%.

Tab. 2. Consapevolezza della propria tracciabilità per caratteristiche individuali.

(N=1074)		Casi (v.a.)	Consapevolezza della propria tracciabilità digitale			² test p-val- ue
			% nulla	% parziale	% piena	
Genere	femmine	807	19,1	42,9	38,0	0,955
	maschi	257	19,8	42,0	38,1	
Età ricodificata	fino a 24	856	18,7	43,8	37,5	0,105
	over 24	181	22,7	35,4	42,0	
Titolo di studio padre	basso-medio	840	19,8	42,7	37,5	0,894
	alto	216	18,5	42,6	38,9	
Titolo di studio madre	basso-medio	835	20,6	42,9	36,5	0,070
	alto	224	14,7	42,0	43,3	
Corso di studi	humanities	847	19,7	42,6	37,7	0,826
	hard sciences	215	18,1	42,3	39,5	
Anno di corso	primo anno	308	22,1	42,5	35,4	0,290
	successivi	752	18,2	42,7	39,1	
Numero di attività in Rete	≤ media	510	21,4	46,7	32,0	0,000*
	> media	564	17,4	39,0	43,6	
Numero oggetti tecnologici posseduti	≤ media	436	22,7	42,9	34,4	0,029*
	> media	638	16,9	42,5	40,6	
Valore proprie preferenze se cedute	≤ media	747	19,1	41,0	39,9	0,153
	> media	327	19,6	46,5	33,9	
Valore proprie preferenze in cambio di servizi	≤ media	763	18,5	42,1	39,4	0,307
	> media	311	21,2	44,1	34,7	
TOTALE		1074	19,3	42,6	38,1	

* Valore statisticamente significativo al test χ^2 .

Trattandosi di studenti universitari è utile verificare l'associazione di un uso consapevole della propria fruizione digitale rispetto al proprio percorso formativo. In realtà, né il corso di studi (umanistico o scientifico in senso stretto), né l'anno di corso (alla prima immatricolazione o agli anni successivi) evidenziano risultati statisticamente significativi, seppur si possa cogliere un leggero aumento della consapevolezza in caso di studenti provenienti da facoltà d'ingegneria e di immatricolati ad anni successivi al primo.

Associazioni statisticamente significative si rilevano rispetto all'aumento della consapevolezza sia al crescere delle attività in rete sia al maggior numero di oggetto tecnologici posseduti. In entrambi i casi, quando i rispondenti mostrano valori superiori alla media, la piena consapevolezza balza dal 32,0% al 43,6% quando si praticano oltre le 7,6 attività mediamente svolte in rete dal campione e dal 34,4 al 40,6%, quando si posseggono più dei 2,8 oggetti tecnologici mediamente rilevati in possesso tra i rispondenti. Al contrario, non emergono associazioni statisticamente significative rispetto ai valori in euro mediamente assegnati dai rispondenti in caso di cessione valutaria o di parziale controprestazione in servizi in cambio della messa a disposizione dei propri dati e delle informazioni personali. Peraltro, è interessante notare che, in entrambi i casi, al crescere dal valore attribuito, la piena consapevolezza diminuisce e aumenta la quota degli ignari.

La Tab.3 dettaglia le possibili associazioni tra il gradiente di consapevolezza della propria tracciabilità e le tipologie di pratiche digitali e di oggetti tecnologici in possesso dei rispondenti.

Decisamente non statisticamente significative appaiono le possibili associazioni della consapevolezza della propria tracciabilità digitale con l'acquistare o vendere online, il rispondere a sondaggi, il pubblicare su blog, alcune pratiche su Facebook e altri social network, come mettere dei like su post sensibili o fornire il proprio stato emotivo o fisico. Vicine alla significatività statistica risultano le possibili associazioni tra la consapevolezza e l'iscrizione a siti per ottenere App gratuite, il postare sui social, nonché l'usare applicazioni per il car e il bike-sharing.

Statisticamente significative risultano le associazioni tra la consapevolezza della propria tracciabilità digitale e la sottoscrizione di tessere

in supermercati o negozi (in tal caso la piena consapevolezza balza dal 32,7% al 41,2%), l'isciversi a siti per ottenere informazioni (i pienamente consapevoli passano dal 33,3% al 42,7%), il postare foto o video (aumento dal 30,7% al 39,4%), il consentire la geo-localizzazione via cellulare o simili (dal 27,4% al 44,5%) e l'utilizzo di sistemi di archiviazione Cloud (dal 33,6% al 41,8%).

In tutti i casi, la piena consapevolezza sembra aumentare comunque in presenza di qualsiasi attività digitale, al di là dell'assenza o presenza di significatività statistica delle associazioni (eccezion fatta per le vendite online e per l'iscrizione a siti per ottenere App gratuite).

Passando all'analisi delle associazioni rispetto alla strumentazione tecnologica posseduta, la significatività statistica del grado di consapevolezza della propria tracciabilità digitale con i diversi oggetti rispecchia presumibilmente il grado di diffusione degli stessi. Per esempio non possono essere considerate statisticamente significative le associazioni rilevate tra il livello di consapevolezza e il possesso di Smartwatch o di elettrodomestici programmabili via internet. Non particolarmente significative sul piano statistico appaiono le associazioni della diversa consapevolezza con il Tablet, il computer fisso o portatile e le consolle per videogiochi. L'unica associazione statisticamente significativa è quella tra il livello di consapevolezza e il possesso di Smartphone, il cui utilizzo vede crescere la piena consapevolezza della propria tracciabilità digitale dal 25,6% al 38,6%.

Sostanzialmente, tanti più oggetti tecnologici si possiedono più aumenta il grado di consapevolezza delle proprie impronte digitali.

Interessante, peraltro che la scarsa consapevolezza aumenti con il possesso di oggetti di consumo non solo funzionale, ma anche (e soprattutto) culturale, ovvero oggetti distintivi sul piano dell'esibizione di status, quali lo Smartwatch o il Tablet.

Tab. 3. Consapevolezza della propria tracciabilità per attività o oggetti.

(N=1091)	Casi (v.a.)	Consapevolezza della propria tracciabilità digitale			χ^2 test p-value
		% nulla	% parziale	% piena	
Sottoscrivere tessere in supermercati/negozi	no	397	23,9	43,3	0,003*
	sì	677	16,5	42,2	
Fare acquisti on-line	no	133	21,1	44,4	0,657
	sì	941	19,0	42,4	
Fare vendite on-line	no	808	19,6	41,8	0,644
	sì	266	18,4	45,1	
Rispondere a sondaggi aziendali, di partiti politici, ecc.	no	593	19,9	44,0	0,327
	sì	481	18,5	41,0	
Iscriverti a siti per ottenere informazioni	no	526	19,2	47,5	0,002*
	sì	548	19,3	38,0	
Iscriverti a siti per ottenere App gratuite	no	815	18,2	42,6	0,181
	sì	259	22,8	42,9	
Postare su social network	no	171	19,9	48,5	0,138
	sì	903	19,2	41,5	
Pubblicare su blog	no	869	19,8	42,1	0,627
	sì	205	17,1	44,9	
Mettere "mi piace" su post sensibile su Facebook	no	343	20,1	44,9	0,358
	sì	731	18,9	41,6	
Postare foto/video	no	163	15,3	54,0	0,006*
	sì	911	20,0	40,6	
Fornir proprio stato emotivo su Facebook o simili	no	703	19,6	42,7	0,898
	sì	371	18,6	42,6	
Fornire proprio stato fisico (es. peso, battito cardiaco)	no	907	19,6	42,3	0,777
	sì	167	17,4	44,3	
Consentire geo-localizzazione via cellulare o simili	no	402	26,1	46,5	0,000*
	sì	672	15,2	40,3	
Usare sistemi di condivisione Cloud (es. Dropbox, ecc.)	no	485	22,3	44,1	0,010*
	sì	589	16,8	41,4	

Usare App per car sharing	no	968	20,0	42,4	37,6	0,151
	sì	106	12,3	45,3	42,5	
Usare App per bike sharing	no	1019	19,9	42,4	37,7	0,066
	sì	55	7,3	47,3	45,5	
Consolle per videogame (X-box, Playstation 4, ecc.)	no	679	20,3	42,9	36,8	0,397
	sì	395	17,5	42,3	40,3	
Smartphone	no	43	39,5	34,9	25,6	0,002*
	sì	1031	18,4	43,0	38,6	
Tablet	no	558	20,6	40,1	39,2	0,203
	sì	516	17,8	45,3	36,8	
Computer fisso	no	567	21,2	42,3	36,5	0,219
	sì	507	17,2	43,0	39,8	
Computer portatile	no	179	22,3	44,1	33,5	0,311
	sì	895	18,7	42,3	39,0	
Smartwatch	no	1053	19,4	42,5	38,2	0,644
	sì	21	14,3	52,4	33,3	
Elettrodomestico che puoi comandare via Internet	no	961	19,7	42,4	38,0	0,625
	sì	113	15,9	45,1	38,9	

* Valore statisticamente significativo al test χ^2 .

Per provare a sintetizzare i risultati appena descritti in sede di analisi bi-variata, si è realizzato un modello multivariato con finalità interpretative. In proposito si è scelto un modello di regressione logistica binaria (elaborato secondo procedura *stepwise backward*), realizzato utilizzando quale variabile dipendente la piena consapevolezza della propria tracciabilità digitale, dicotomizzando la variabile originale in 1 = “piena consapevolezza”, riferita alla coscienza di essere tracciabili indipendentemente dall’aver dato il proprio consenso, e 0 = “consapevolezza nulla o parziale”, data dall’accorpamento delle risposte “no”, “non saprei” e “Sì, solo previo espresso consenso” alla domanda “Ritieni che l’uso delle tecnologie infotelematiche possa generare tuoi dati che possono essere utilizzati da aziende commerciali o altre organizzazioni?”.

Dopo ripetute elaborazioni, per quanto il modello finale spieghi solo circa l’8% della varianza complessiva, si è scelto di prediligere l’inserimen-

to unicamente di tutte quelle variabili che garantissero una completa significatività statistica al modello stesso.

Tab. 4. Modello di regressione logistica binaria realizzato su atteggiamento di piena consapevolezza della propria tracciabilità digitale in base alle diverse variabili indipendenti (N=1.091).

Variabili indipendenti	Piena consapevolezza della propria tracciabilità digitale		
	p-value ^a	OR ^b	95% CI ^c
Sottoscrivere tessere in supermercati/negozi (rif. = No)			
Si	0,039	1,33	(1,01–1,75)
Iscriverti a siti per ottenere informazioni (rif. = No)			
Si	0,002	1,52	(1,16–1,99)
Iscriverti a siti per ottenere App gratuite (rif.= No)			
Si	0,024	0,69	(0,50–0,95)
Consentire geo-localizzazione via cell. o simili (rif.= No)			
Si	0,000	2,00	(1,51–2,66)
Possesso Tablet (rif.= No)			
Si	0,017	0,67	(0,48–0,93)
Num. oggetti tecno. posseduti (rif.= Sotto o entro la media)			
Sopra la media	0,016	1,52	(1,08–2,14)
Fascia d'età (rif.= fino a 24 anni)			
Over 24	0,039	1,43	(1,01–2,02)
Livello istruzione della madre (rif.= Basso-Medio)			
Alto	0,05	1,36	(1,00–1,87)
Costante	0,000	0,56	
Nagelkerke R ²	8,0%		

^a Two-sided Wald test; ^b OR=Odds Ratio, cioè rapporto di probabilità, dove OR >1 indica una maggiore associazione alla piena consapevolezza; ^c CI = intervallo di confidenza.

La sottoscrizione di tessere in supermercati e negozi, evidenziando un rapporto di probabilità pari a 1,33 rispetto a chi non adotta tale comportamento (per una significatività al 95% e un intervallo di confidenza compreso tra 1,01 e 1,75), appare una pratica abbastanza predittiva di una piena consapevolezza. Ancor più predittivi di una significativa cognizione dei rischi di tracciabilità digitale appaiono comportamenti come l'iscrizione a siti per ottenere informazioni (con rapporto di probabilità è pari a 1,52, con una significatività al 95% e un intervallo di confidenza compreso tra

1,16 e 1,99) e l'acconsentire alla propria geo-localizzazione (mostrando un rapporto di probabilità pari a ben 2,00, con una significatività al 95% e un intervallo di confidenza compreso tra 1,51 e 2,66).

Al contrario, iscriversi a siti per ottenere App gratuite sembra predittivo di una pratica online più inesperta, laddove il rapporto di probabilità rispetto alla piena consapevolezza scende a 0,69, con una significatività al 95% e un intervallo di confidenza compreso tra 0,50 e 0,95.

Al pari, predittivo di una scarsa consapevolezza e di un probabile uso ingenuo dello strumento, risulta il possesso del Tablet, per cui il rapporto di probabilità rispetto alla piena consapevolezza declina a 0,67, con una significatività al 95% e un intervallo di confidenza compreso tra 0,48 e 0,93.

La piena consapevolezza cresce quando il numero di oggetti tecnologici posseduti risulta sopra la media, evidenziando un rapporto di probabilità pari a 1,52, rispetto a una dotazione entro o sotto la media del campione (con una significatività al 95% e un intervallo di confidenza compreso tra 1,08 e 2,14). Così come la piena consapevolezza aumenta tra gli over 24, con un rapporto di probabilità pari a 1,43, rispetto a chi ha meno di 24 anni (con una significatività al 95% e un intervallo di confidenza compreso tra 1,01 e 2,02) e in presenza di un più alto livello di istruzione della madre, con un rapporto di probabilità pari a 1,36, rispetto a chi ha una madre diplomata o meno istruita (con una significatività al 95% e un intervallo di confidenza compreso tra 1,00 e 1,87).

4. Consapevolezza e ingenuità: tra disuguaglianze digitale e disuguaglianze socio-culturali

L'analisi appena esposta suggerisce alcune considerazioni che mirano a comprendere non soltanto le determinanti di una diversa consapevolezza della propria tracciabilità digitale, ma cercano di trarre le radici di tale diversità di atteggiamento in disuguaglianze sociali più profonde.

Occorre, innanzi tutto, delineare meglio il profilo dei rispondenti. Non è un caso, infatti, che le differenze di età siano emerse come un fattore esplicativo.

Sul piano demografico e culturale si può dire che il campione si concentra in un profilo tipicamente connotabile nella generazione dei cosiddetti

detti Millennials, ovvero in quelle coorti che, nate tra gli anni Ottanta e la fine degli anni Novanta, costituiscono tipicamente un profilo di nativi digitali, caratterizzato da un maggiore utilizzo e una maggiore familiarità con la comunicazione, i media e le nuove tecnologie (Howe e Strauss, 1991, 2000). Per definizione si tratta di una generazione che è cresciuta quando non ha addirittura creato strumenti di comunicazione online e una diversa fruizione della Rete attraverso You Tube o le pratiche di social networking, come Facebook, MySpace e Twitter.

Tuttavia, a fronte di un uso quotidiano e ampiamente sperimentato delle nuove tecnologie da parte dei rispondenti in ragione della loro stessa matrice generazionale, il fatto stesso che si evidenzino nell'indagine diversi livelli di consapevolezza digitale, o, volendo, di ingenuità, necessita una spiegazione più accurata.

Innanzitutto, come emerge dai risultati, si tratta sì di un'unica generazione, ma presumibilmente è necessario scomporre in due coorti o unità generazionali, potenzialmente distinguibili tra *earlier* (nati entro la prima metà degli anni Novanta) e in *later* (nati verso la fine del secolo) Millennials. Non è un caso, infatti, che l'età, in particolare quando sopra i 24 anni, definisca, se non un vero e proprio punto di taglio, quanto meno una discriminante rispetto alla piena consapevolezza digitale. Evidentemente, la diversa posizione in termini di corso della vita (anche rispetto a stili di vita, impegni lavorativi e familiari potenzialmente conseguenti al crescere dell'età dei rispondenti) contribuisce maggiormente, anche in ragione di un maggiore bagaglio esperienziale, a definire una certa consapevolezza persino nell'utilizzo di nuove tecnologie che sono di per sé connaturate all'intera generazione. In parte, questo può trovare ulteriore dimostrazione nella diversa capacità di assegnare un valore reale ai propri dati e alle informazioni personali.

Similmente, emerge chiaramente come tanto più aumenta la dotazione tecnologica posseduta, tanto più probabilmente si acquisisca competenza, cui conseguono una fruizione consapevole e una maggior piena coscienza dei comportamenti. Altresì, queste condizioni possono essere favorite da elementi di disuguaglianza sociale, prima ancora che di alfabetizzazione digitale. Infatti, il possesso materiale di più strumenti tecnologici può essere implicitamente ricondotto alle disponibilità economiche. Non a caso, la piena consapevolezza si associa significativamente ai livelli d'istruzione

più elevati delle madri dei rispondenti, ribadendo il valore estremamente predittivo del titolo di studio materno rispetto alla posizione socioeconomica degli individui (una tendenza nota nel nostro paese sin dagli studi di Sorokin del secolo scorso).

Non di meno, il possesso distintivo di alcuni strumenti, si pensi al Tablet o allo Smartwatch, risultando inversamente proporzionale alla consapevolezza effettiva, suggerisce un uso non efficace e, forse, più esibizionista, delle dotazioni tecnologiche più *à la page*.

5. Conclusioni

Nella sua essenzialità, il dato visto sopra, relativo alla scarsa o nulla consapevolezza della maggior parte degli intervistati (62% del campione), pone in primo luogo rilevanti questioni riguardo alla necessità di investire più e meglio sulle competenze digitali di giovani e meno giovani. Come è stato sottolineato in apertura, infatti, i cittadini italiani non brillano per competenze nel panorama europeo: nell'indice DESI per il 2017, la percentuale di italiani in possesso per lo meno di competenze digitali di base era pari al 44% a fronte di una media europea pari al 56%. Riflessioni ancor più attente dovrebbe suscitare il confronto con le maggiori economie mondiali. Uno studio realizzato per approfondire questo aspetto (Foley et al., 2018) mostra che il gruppo dei quattro paesi ultimi nella graduatoria europea, di cui l'Italia fa parte, si colloca con un punteggio medio pari a 43,7 in quintultima posizione nella graduatoria mondiale, a distanza siderale non solo dai paesi primi in graduatoria, ossia l'Australia e l'Islanda (80,5 e 80,2 rispettivamente), ma anche dalle prime quattro economie europee (74,7).

Torniamo ora all'interrogativo da cui siamo partiti e, alla luce di quanto emerso dall'analisi appena presentata, proviamo a trarre alcune considerazioni conclusive.

Internauti consapevoli? La risposta in generale sembra essere "in gran parte no". Abbiamo visto infatti come la piena consapevolezza, con gli inevitabili limiti della definizione operativa adottata, riguardi solo 38 rispondenti su 100.

Riguardo all'"identikit" dei consapevoli, da un lato, e dei non (o non pienamente) consapevoli dall'altro, abbiamo osservato come esso ricalchi

in qualche modo disuguaglianze di status socio-economico riconducibili al livello culturale del nucleo di appartenenza (con particolare riferimento al titolo di studio della madre) e alle stesse disponibilità economiche che consentono l'accesso alle tecnologie digitali e, come osservato anche nel paragrafo introduttivo, alla stessa Rete. L'età più matura sembra poi associata a una maggiore consapevolezza¹¹ solo se la persona proviene da un nucleo con status socio-culturale più elevato.

D'altro canto, sembra emergere una divaricazione tra utenti che deliberatamente e consapevolmente condividono dati e informazioni che li riguardano, in quanto ciò risulta indispensabile per poter usufruire di certi servizi, con un comportamento che ha una valenza eminentemente strumentale (come nei casi di uso della geo-localizzazione, di sottoscrizione di tessere del supermercato, di uso di sistemi di condivisione in cloud ecc.) e utenti per i quali la condivisione di dati e informazioni non è orientata, come per i primi, da un calcolo utilitaristico, ma è in qualche modo "fine a se stessa" e se vogliamo assume un valore essenzialmente espressivo (come nei casi di condivisione dei propri stati emotivi sui social, dei propri parametri fisici attraverso Smartwatch, di pubblicazione su blog, di uso di oggetti quali il Tablet). È plausibile che, proprio per l'uso che ne fa, questa seconda tipologia di internauti sia più fermamente convinta della prima che solo i dati e le informazioni che decide espressamente di condividere siano registrate dalla Rete e che il loro utilizzo sia sempre e comunque subordinato al proprio deliberato consenso o, nei casi peggiori, che neppure si ponga la questione.

Peraltro, come giustamente osserva Gremigni (2018), con *digital divide* non facciamo più (solo) riferimento alle differenze riscontrabili nelle possibilità di accesso alla Rete, ma anche a quelle rintracciabili nelle abilità e nei tipi di attività svolte online dall'utente, che sono condizionate dal

11. Da non sottovalutare, rispetto al fattore età, il ruolo che potrebbe giocare lo stesso sviluppo neurobiologico dell'utente della Rete. Il citatissimo studio di Gogtay e collaboratori (2004) ha chiarito che dai 10 ai 20 anni la corteccia prefrontale perde circa 60% delle sinapsi eccitatorie mentre nella stessa regione, dai 15 e soprattutto dai 20 agli oltre 25 anni, matura più del 90% delle sinapsi inibitorie, quelle attraverso cui si realizza il controllo cognitivo del comportamento. Questa evidenza illustra in modo esemplare le ragioni neurobiologiche dell'impulsività dei ragazzi e delle loro minore capacità di valutazione dei rischi cui possono esporsi in conseguenza dei propri comportamenti. Immediato, in questo senso, il rimando a comportamenti devianti e condizioni patologiche quali il *cyberbullismo* (Kowalski *et al.*, 2012), il *sexting* (Siegle, 2010) e il gioco d'azzardo *online* (Gainsbury, 2015).

capitale economico, culturale, sociale e simbolico disponibile (*second-level digital divide*) e alla capacità o meno dell'utente di ottenere un vantaggio nella vita di tutti i giorni da un uso appropriato di Internet. Si individua un *digital divide* di terzo livello (van Deursen, Helsper, 2015). Un alto livello di competenze digitali si riferisce «alla capacità degli utenti di riflettere sui meccanismi profondi che regolano l'organizzazione di Internet al fine di ottenere specifici risultati offline ed evitare di essere trasformati in semplici ingranaggi delle TIC» (Gremigni 2018, p. 83).

La semplice osservazione dei messaggi di advertising che compaiono nei nostri profili Facebook o che inframezzano le pagine web che quotidianamente visitiamo e che, rispecchiando fedelmente le ricerche di beni o servizi effettuati nei giorni/settimane precedenti, portano alla nostra attenzione le offerte più recenti e allettanti, dovrebbe indurre anche l'internauta più ingenuo a porsi la questione, a domandarsi chi stia tenendo traccia dei suoi movimenti nella Rete e quale uso stia facendo delle informazioni così raccolte e archiviate. Pochi di noi si sentirebbero a proprio agio se, camminando per strada, si accorgessero di essere pedinati e osservati nei propri movimenti e comportamenti da uno sconosciuto, eppure (quasi) tutti noi, muovendoci quotidianamente all'interno della Rete, siamo oggetto di attento monitoraggio da parte di molti sconosciuti, che tengono traccia dei nostri itinerari, dei luoghi virtuali e reali che visitiamo, del tempo e dell'attenzione che dedichiamo a ciascuno di essi, delle nostre relazioni amicali e del loro evolversi nel tempo, come internauti inconsapevoli in un Panopticon dei Balocchi.

Riferimenti bibliografici

- ALBANO R., BERTOLINI S., CURZI Y., FABBRI T. M., PARISI, T. (2017), *DigitalAgile: l'ufficio nel dispositivo mobile. Opportunità e rischi per lavoratori e aziende*, Osservatorio MU. SIC Working Paper Series, (03).
- ARNSTEIN S. R. (1969), *A ladder of citizen participation*, in «Journal of the American Planning Association», Vol. 35(4), pp. 216–224.
- ASHTON K. (2009), *That 'internet of things' thing*. RFID journal, 22(7), 97–114.
- CARRETERO S., VUORIKARI R. & PUNIE Y. (2017), *DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and exam-*

- ples of use*. European Commission's Joint Research Centre. Trad. it. DigComp 2.1. Il quadro di riferimento per le competenze digitali dei cittadini. Con otto livelli di padronanza ed esempi di utilizzo. Agenzia per l'Italia Digitale (AgID), https://competenze-digitali-docs.readthedocs.io/it/latest/_downloads/DigComp2-1_ITA.pdf, consultato il 01/02/2019.
- CAVAGNERO S. M., GALLINA M. A., MARCHISIO M. (2015), *Scuola dei compiti. Didattica digitale per il recupero dell'insuccesso scolastico*, in DIDAMATICA 2015, *Studio ergo Lavoro Dalla società della conoscenza alla società delle competenze* (pp. 1-8), AICA.
- CENSIS (2018), *Cinquantaduesimo Rapporto sulla situazione sociale del Paese 2018*, Milano, FrancoAngeli.
- CIPOLLA C. (2015)(a cura di), *Dalla relazione alla connessione nella web society*, FrancoAngeli, Milano.
- CIPOLLA C., MATURO A. (2014)(a cura di), *Sociologia della salute e web society*, Milano, FrancoAngeli.
- FOLEY P., SUTTON D., WISEMAN I., GREEN L., MOORE J. (2018), *International Digital Economy and Society Index 2018*, EuropeanUnion disponibile al link: https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=54991 consultato il 05/02/2019.
- GAINSBURY S. M. (2015), *Online gambling addiction: The relationship between internet gambling and disordered gambling*, «Current addiction reports», Vol. 2(2), pp. 185-193.
- GASTALDI L., TRABACCHI P. (2019), *Non solo Desi: anche i Digital Maturity Indexes mostrano un'Italia in affanno* in <https://www.agendadigitale.eu/cittadinanza-digitale/non-solo-desi-anche-i-digital-maturity-indexes-mostrano-unitalia-in-affanno/> consultato il 01/02/2019.
- GIANNOTTI F. (2015), *Big Data e social mining: i dati, a saperli ascoltare, raccontano storie*, in E. Tasso, A. Mola, A. Cortesi, A. Candiello, *Misurare l'innovazione digitale. Gli indicatori di successo delle politiche di innovazione territoriale*, Firenze, Edizioni Ca' Foscari.
- GOGTAY N., GIEDD J. N., LUSK L., HAYASHI K. M., GREENSTEIN D., VAITUZIS A. C., ... & RAPOPORT J. L. (2004), *Dynamic mapping of human cortical development during childhood through early adulthood*. «Proceedings of the National Academy of Sciences», Vol. 101(21), pp. 8174-8179.

- GRAZIANO G., MARTENA P. (2017), *Didattica digitale: la scuola cambia con la società*, in *European Journal of Education Studies*, 3(3), 210–221.
- GREMIGNI E. (2018). *Overcoming New Forms of Digital Divide: Some Remarks on the Need for Media Education*. «Italian Sociological Review», Vol. 8(1), pp. 81–102 DOI: 10.13136/isr.v8i1.221.
- GRIMALDI R., GALLINA M.A., CAVAGNERO S.M. (2017), *Big data e tracce digitali. Una ricerca internazionale tra studenti universitari*, in TRINCHERO R., PAROLA A. (a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli.
- HOWE N., STRAUSS B. (2000), *Millennials rising: the next great generation*. New York: Vintage Books.
- HOWE N., STRAUSS W. (1991), *Generations: The History of Americas Future, 1584 to 2069*. New York: William Morrow. doi:ISBN 0-688-11912-3.
- ISTAT (2019), *Cittadini, imprese e ICT. Anno 2018*, disponibile al link: https://www.istat.it/it/files//2019/01/Report-ICT-cittadini-e-imprese_2018_PC.pdf consultato il 29/01/2019.
- KOWALSKI R. M., LIMBER, S. P., LIMBER S., AGATSTON P. W. (2012). *Cyberbullying: Bullying in the digital age*. John Wiley & Sons.
- LOMBI L. (2015), *La ricerca sociale al tempo dei Big Data: sfide e prospettive*, in «Studi di sociologia», n. 2.
- MELLO P., BOARI M. (2018), *Intelligenza Artificiale: Appunti per un'Introduzione*, disponibile al link: <http://campus.unibo.it/314211/> consultato il 31/01/2019.
- PALUMBO M., CONGIU D. (2009), *Partecipazione e valutazione per il Piano Regolatore Sociale*, in Palumbo M., Torrigiani C. (a cura di), *La partecipazione tra ricerca e valutazione*, FrancoAngeli, Milano, pp. 37–72.
- RIBAUDO M., TORRIGIANI C., DE CINDIO F., PALUMBO M., *The University in the Polis: An Emerging Role of Democratic Intermediary in e-Participation?*, in Torre T., Braccini A.M., Spinelli R., *Empowering Organizations. Enabling Platforms and Artefacts*, ISBN: 978-3-319-23783-1 (Print) 978-3-319-23784-8 (Online), «Lecture Notes in Information Systems and Organisation» ISSN: 2195-4968, Springer International Publishing, pp. 63–76.
- SCHIAVO G., MANA N., MICH O., ARMELLINI E., NANNI E., BIANCHI C. (2016), *Rassegna ragionata delle tecnologie digitali a supporto delle persone con DSA*, Rovereto, Provincia autonoma di Trento — IPRASE.

- SIEGLE D. (2010), *Cyberbullying and sexting: Technology abuses of the 21st century*, «Gifted child today», Vol. 33(2), pp. 14–65.
- SORO A. (2013), *Big data e sorveglianza globale*, Relazione 2013, in <http://194.242.234.211/documents/10160/0/Relazione+annuale+2013+-+Discorso+del+Presidente.pdf>.
- SOROKIN P. (1927), *Social Mobility*, New York/London: Harper & Brothers.
- TRINCHERO R., PAROLA A. (2017)(a cura di), *Educare ai processi e ai linguaggi dell'apprendimento*, Milano, FrancoAngeli.
- VAN DEURSEN A.J., HELSPER E.J. (2015), *The Third-Level Digital Divide: Who Benefits Most From Being Online?* In «Communication and Information Technologies Annual: Digital Distinctions and Inequalities Studies in Media and Communications», Vol.10, pp. 29–52. doi:10.1108/S2050-206020150000010002.

Autori

Sandro Brignone è dottorando di ricerca sul Progetto di Eccellenza del Dipartimento di Filosofia e Scienze dell'Educazione dell'Università degli Studi di Torino.

Simona Maria Cavagnero è assegnista di ricerca presso il Dipartimento di Filosofia e Scienze dell'Educazione dell'Università degli Studi di Torino.

Maria Adelaide Gallina è ricercatrice in Sociologia generale presso il Dipartimento di Filosofia e Scienze dell'Educazione dell'Università degli Studi di Torino.

Renato Grimaldi è professore ordinario di Metodologia della ricerca sociale presso il Dipartimento di Filosofia e Scienze dell'Educazione e Direttore della Scuola di Scienze Umanistiche dell'Università degli Studi di Torino.

Cristina Ispas è docente in Scienze dell'Educazione e Scienze Sociali presso l'Università Eftimie Murgu di Reșița, România.

Stefano Poli è professore associato in Sociologia generale presso il Dipartimento di Scienze della Formazione dell'Università degli Studi di Genova, dove è responsabile dell'Osservatorio delle Disuguaglianze sociali e insegna Sociologia generale e Metodologia della ricerca sociale applicata.

Francesco Mazzeo Rinaldi insegna Statistica multivariata e ricerca valutativa e Programmazione e valutazione delle politiche sociali presso il

Dipartimento di Scienze Politiche e Sociali dell'Università degli Studi di Catania. È inoltre *Affiliate Professor* presso il Royal Institute of Technology (KTH) di Stoccolma.

Claudio Torrigiani è ricercatore in Sociologia generale presso il Dipartimento di Scienze della Formazione dell'Università degli Studi di Genova dove insegna Metodologia della ricerca sociale applicata, Politica sociale, Metodi e tecniche di valutazione delle politiche sociali e formative.

TEORIA E RICERCA IN EDUCAZIONE

COLLANA DEL DIPARTIMENTO DI FILOSOFIA E SCIENZE DELL'EDUCAZIONE
SEZIONE DI SCIENZE DELL'EDUCAZIONE
UNIVERSITÀ DEGLI STUDI DI TORINO

1. Valentina PORCELLANA
In nome della lingua. Antropologia di una minoranza
ISBN 978-88-548-1359-5 format, 156 pagine, 12 euro
2. Rosita DELUIGI
Divenire anziani, anziani in divenire. Prospettive pedagogiche fra costruzione di senso e promozione di azioni sociali concertate
ISBN 978-88-548-1532-2, formato 17 × 24 cm, 312 pagine, 17 euro
3. Barbara BRUSCHI
E—learning nelle scuole superiori. Una sperimentazione sul campo in Piemonte
ISBN 978-88-548-2067-8, formato 17 × 24 cm, 192 pagine, 12 euro
4. Maria Adelaide GALLINA
Insegnare tra saperi dell'esperienza e nuove competenze
ISBN 978-88-548-1897-2, formato 17 × 24 cm, 216 pagine, 13 euro
5. Emanuela Maria TORRE
Lavorare con gli adolescenti: emergenze e domande formative di insegnanti ed educatori. Una ricerca tra scuola ed extrascuola
ISBN 978-88-548-2063-0, formato 17 × 24 cm, 268 pagine, 15 euro
6. Carlo Mario FEDELI
L'educazione come esperienza. Il contributo di John Dewey e Romano Guardini alla pedagogia del Novecento
ISBN 978-88-548-2096-8, formato 17 × 24 cm, 268 pagine, 15 euro
7. Giuseppina NEGRO
Interpretazione di un'immagine. Un'indagine nelle biblioteche civiche della provincia di Cuneo
ISBN 978-88-548-2377-8, formato 17 × 24 cm, 256 pagine, 17 euro
8. Giorgio PRIMERANO
La prospettiva pedagogica di Nicola Abbagnano
ISBN 978-88-548-2653-3, formato 17 × 24 cm, 220 pagine, 14 euro
9. Daniela ROBASTO
Il consumo televisivo e la rappresentazione del ruolo di genere negli adolescenti
ISBN 978-88-548-2579-7, formato 17 × 24 cm, 232 pagine, 15 euro

10. Cecilia Maria MARCHISIO
La funzione paterna nelle relazioni educative e di aiuto
ISBN 978-88-548-2684-9, formato 17 × 24 cm, 220 pagine, 14 euro
11. Maria Adelaide GALLINA
La rappresentazione della situazione. Strategie di comportamento sociale negli ex-voto dipinti
ISBN 978-88-548-2990-9, formato 17 × 24 cm, 120 pagine, 9 euro
12. Sara NOSARI
Confini della creatività
ISBN 978-88-548-3448-4, formato 17 × 24 cm, 172 pagine, 12 euro
13. Emanuela Maria TORRE
Strategie di ricerca valutativa in educazione e formazione
ISBN 978-88-548-3555-9, formato 17 × 24 cm, 204 pagine, 14 euro
14. Valentina PORCELLANA
Sei mai stato in dormitorio? Analisi antropologica del lavoro educativo e degli spazi d'accoglienza notturna a Torino
ISBN 978-88-548-4061-4, formato 17 × 24 cm, 180 pagine, 11 euro
15. Natascia CURTO
Persone fragili e servizi a bassa soglia
ISBN 978-88-548-4348-6, formato 17 × 24 cm, 160 pagine, 10 euro
16. Simona Maria CAVAGNERO
Il pellegrinaggio in trasformazione. Per uno studio dei comportamenti sociali
ISBN 978-88-548-4788-0, formato 17 × 24 cm, 308 pagine, 18 euro
17. Alberto PAROLA
Regia educativa. L'insegnante osservatore tra percorsi "in ricerca", linguaggi e strategie
ISBN 978-88-548-4850-4, formato 17 × 24 cm, 260 pagine, 17 euro
18. Barbara BRUSCHI, Silvia CARBOTTI
Per imparare c'è un'app
ISBN 978-88-548-5089-7, formato 17 × 24 cm, 205 pagine, 14 euro
19. Cecilia Maria MARCHISIO, Chiara MAROCCO MUTTINI (a cura di)
La famiglia italiana tra crisi e risorse. 1861-2011
Contributi di Barbara Bruschi, Danilo Carrozzino, Venceslao Cembalo, Maria Di Nardo, Maria Ida Di-
gennaro, Mario Fulcheri, Franco Garelli, Michele Marangi, Valerio Mavilia, Piero Mussi, Ermis Segatti,
Luca Sibona, Paola Zonca
ISBN 978-88-548-5773-5, formato 17 × 24 cm, 164 pagine, 13 euro

20. Cecilia Maria MARCHISIO, Natascia CURTO
Senza muri. Attivare il territorio per promuovere i diritti
Prefazione di Claudio Imprudente, Postfazione di Andrea Canevaro
ISBN 978-88-548-5701-8, formato 17 × 24 cm, 188 pagine, 13 euro
21. Alessia FARINELLA
La resilienza familiare agli inizi difficili della vita. Riflessioni pedagogiche ed esempi di buone prassi
Prefazione di Marisa Pavone
ISBN 978-88-548-6147-3, formato 17 × 24 cm, 192 pagine, 12 euro
22. Alberto PAROLA, Barbara BRUSCHI (a cura di)
Paesaggi digitali. I futuri educatori tra formazione e nuovi linguaggi
Prefazione di Adriana Marchia. Contributi di Michele Aglieri, Massimiliano Andreoletti, Isabella Bruni, Barbara Bruschi, Silvia Carbotti, Angela Castelli, Eleonora Crestani, Francesco Filippi, Chiara Friso, Vittorio Marchetta, Alberto Parola, Maria Ranieri, Daniela Robasto, Alessia Rosa, Carlo Tagliabue, Roberto Trincherio, Patrizia Vayola
ISBN 978-88-548-6614-0, formato 17 × 24 cm, 256 pagine, 15 euro
23. Giuseppina CERRATO, Cristina SABENA, Elena SCALENGHE (a cura di)
L'apprendimento nella scuola dell'infanzia
Prefazione di Renato Grimaldi. Contributi di Ornalla Cornabuchi, Simona Maria Cavagnero, Cristina Coggi, Rosangela Cuniberti, Redi Sante Di Pol, Maria Adelaide Gallina, Cristiano Giorda, Francesca Morselli, Serafina Petrocca, Mirella Pezzin, Paola Ricchiardi, Marinella Roviglione, Maria Antonietta Simeoli, Teresina Vignola
ISBN 978-88-548-6828-1, formato 17 × 24 cm, 212 pagine, fuori commercio euro
24. Simona Maria CAVAGNERO
Comportamenti devozionali tra permanenza e mutamento
Prefazione di Luigi Berzano
ISBN 978-88-548-6861-8, formato 17 × 24 cm, 216 pagine, 15 euro
25. Alessia ROSA, Emanuela M. TORRE
Media e scuola tra forma mentis e modus operandi. Analisi teorica e dati di ricerca
ISBN 978-88-548-7067-3, formato 17 × 24 cm, 124 pagine, 12 euro
26. Bruno GRIMALDI
Agent-Based vs. Robotic Simulation. A Repeated Prisoner's Dilemma Experiment
Preface by Pietro Terna
ISBN 978-88-548-7541-8, formato 17 × 24 cm, 232 pagine, 20 euro
27. Simona Maria CAVAGNERO, Maria Adelaide GALLINA (a cura di)
La persona al centro
Contributi di Antonella Saracco, Witold Misiuda-Rewera, Cristina Ispas, Renato Grimaldi, Marco Gonnella, Andrea Dughera, Chiara Marellò, Sara Forte, Silvia Piovano, Marco Tealdo, Gianluca Gallina, Clotilde Trincherio
ISBN 978-88-548-8727-5, formato 17 × 24 cm, 224 pagine, 16 euro

28. Federico REPETTO
Cultura pubblicitaria e berlusconismo (fuori catalogo)
Prefazione di Roberto Trincherò
ISBN 978-88-548-8998-9, formato 17 × 24 cm, 339 pagine, 16 euro
29. Lorenzo DENICOLAI, Alberto PAROLA (a cura di)
Tecnologie e linguaggi dell'apprendimento
Contributi di Alberto Parola, Roberto Trincherò, Barbara Bruschi, Alessandro Antonietti, Luca Bonfanti, Gian Paolo Caprettini, Renato Grimaldi, Silvia Palmieri, Lorenzo Denicolai, Alessia Rosa, Daniela Robasto, Eleonora Pantò, Ilio Trafeli, Angelo Chiarle, Michaela Liuccio, Chiara Borgia, Benedetta Martino, Annamaria Poli, Fulvio Oscar Benussi, Eleonora Crestani, Anna Giorgi, Selena Agnella, Federica Cornali, Gianni Latini, Gianfranco Pomatto, Giovanni Favata, Alberto Barbero, Mattia Davi, Stefania Cazzoli, Francesca Alloatti, Federica Viscusi, Veronica Cavicchi, Pietro Madaro, Maria del Amor De Ben Melgar, Ali Naserhelali
ISBN 978-88-548-9174-6, formato 17 × 24 cm, 364 pagine, 26 euro
30. Sandro BRIGNONE
Simulare la scuola. Costruzione di un modello ad agenti
Prefazione di Renato Grimaldi e Pietro Terna
ISBN 978-88-255-1178-9, formato 17 × 24 cm, 260 pagine, 20 euro
31. Lorenzo DENICOLAI, Alberto PAROLA (a cura di)
Digital Education. Ricerche, pratiche ed esperienze nei mondi mediali
Contributi di Francesca Alloatti, Marco Andreoli, Alberto Barbero, Fulvio Oscar Benussi, Alexandra Berndt, Barbara Bruschi, Daniel Burgos, Silvia Carbotti, Maria Carmela Catone, Stefania Cazzoli, Donatella Cesaroni, Angelo Chiarle, Riccarda Cristofanini, Mattia Davi, Alessandra De Nicola, Lorenzo Denicolai, Paolo Diana, Alice Di Leva, Francesco Fabbro, Rosanna Labalestra, Pietro Madaro, Michele Maranzana, Eugenio Melloni, Renata Metastasio, Elena Miele, Andrea Nardi, Fabio Nascimbeni, Gianluigi Nigro, Eleonora Pantò, Alberto Parola, Valentina Percivalle, Annamaria Poli, Maria Ranieri, Manuela Repetto, Valeria Ribani, Alessia Rosa, Barbara Rossi, Franca Rossi, Graziella Testaceni, Mattia Thibault, Annalisa Vio, Federica Viscusi, Franca Zuccoli
ISBN 978-88-255-2638-7, formato 17 × 24 cm, 396 pagine, 27 euro
32. Renato Grimaldi, Simona Maria Cavagnero, Maria Adelaide Gallina (a cura di)
Big data e tracce digitali. Una ricerca internazionale tra studenti universitari
Contributi di Francesca Sandro Brignone, Simona Maria Cavagnero, Maria Adelaide Gallina, Renato Grimaldi, Cristina Ispas, Stefano Poli, Francesco Mazzeo, Rinaldi Claudio Torrigiani
ISBN 978-88-255-2825-1, formato 17 × 24 cm, 100 pagine, 12 euro

Compilato il 8 ottobre 2019, ore 10:28
con il sistema tipografico L^AT_EX 2_ε

Finito di stampare nel mese di ottobre del 2019
dalla tipografia «The Factory S.r.l.»
00156 Roma – via Tiburtina, 912
per conto della «Gioacchino Onorati editore S.r.l. – unipersonale» di Canterano (RM)