

IL PUNTO SU...

Popper, “PRODIGIT” e giustizia predittiva

di Enrico Marello - 24 ottobre 2022

A. Il progetto PRODIGIT sta suscitando un grande interesse nella comunità scientifica e nella comunità applicativa del diritto tributario. Su questa rivista, per esempio, analizzano diversi profili della questione Francesco Farri (in questa Rivista, 12 ottobre 2022) e Alberto Marcheselli (in questa Rivista, 20 ottobre 2022).

Il progetto PRODIGIT «*si propone di attuare un importante processo di innovazione della Giustizia Tributaria, con il supporto della tecnologia digitale e della intelligenza artificiale, ed è volto a realizzare un miglioramento della trasparenza e della efficienza della stessa, mediante la realizzazione di cinque linee di intervento: 1. digitalizzazione dei processi interni del CPGT; 2. reingegnerizzazione del sito internet istituzionale del CPGT; 3. implementazione della Banca dati nazionale di giurisprudenza di merito accessibile gratuitamente e pubblicamente dal sito internet istituzionale del CPGT; 4. realizzazione di un modello di c.d. giustizia predittiva; 5. creazione sperimentale, in otto Regioni italiane (Emilia Romagna, Lazio, Liguria, Marche, Puglia, Sardegna, Sicilia e Veneto), e presso la Corte di Giustizia tributaria di I grado di Brescia, del laboratorio digitale del giudice tributario denominato “TRIBHUB”*» (cfr. il bando sul sito /oppure cfr. il bando di selezione esperti per la disciplina ed il funzionamento del laboratorio digitale del giudice tributario, pubblicato sul sito del Consiglio di Presidenza della Giustizia Tributaria www.giustizia-tributaria.it, 22 settembre 2022).

Ad oggi, il progetto è in fase di avvio e i tratti realizzativi concreti del progetto non sono conosciuti.

B. Dalle prime informazioni, risulta che il progetto, sfruttando anche l’esperienza pregressa (e *ongoing*) di altri progetti di Analytics for Decisions of Legal Cases, si indirizzerà, in questo primo momento, verso la massimazione di un numero rilevante di sentenze tributarie (si dice un milione).

Al momento, i metodi di questa operazione di *summarization* non sono noti, così come non è noto il modo in cui lo strumento prodotto sarà messo a disposizione dei giudici tributari e degli operatori tutti.

Si tratta di un momento cruciale e fondativo di un nuovo modo di intendere la giustizia tributaria. Siamo dinanzi all’introduzione di strumenti che cambieranno il modo di decidere e difendere.

Non è il caso di indulgere in ciechi tecno-ottimismo e neppure di brontolare sulla superiorità assiologica di strumenti e procedimenti decisionali che saranno presto marginalizzati, come lo è stato il cavallo, inteso come mezzo di locomozione.

E', invece, il momento di mettere in chiaro alcuni principi fondanti la logica e la filosofia di questa operazione.

La comprensione, la condivisione e, soprattutto, l'applicazione di questi principi possono rendere possibile l'utilizzo ragionato dei nuovi strumenti da parte della comunità di applicazione e di studio.

Mi soffermo su due di questi principi, che regolano la fase di costruzione del nuovo sistema, per come possono essere brevemente formalizzati.

(I) PARZIALITÀ

Ogni strumento tecnologico ha, per il profano, l'aura dell'oggettività e della necessità: lo strumento proposto appare come il frutto necessario ed unico di un progresso che non poteva muoversi altrimenti. Non appena si acquisisce qualche consapevolezza dello strumento, si comprende, al contrario, la sua parzialità e la sua dipendenza da un progetto e da ipotesi teoriche (e realizzazioni pratiche) fortemente soggettive.

Anche nel caso di PRODIGIT deve essere chiaro che il prodotto finale (qualunque esso sia) dipenderà strettamente: (1) dalle persone che hanno realizzato il progetto e che hanno accettato alcune ipotesi teoriche e non altre e (2) dagli strumenti tecnici adoperati per realizzare le ipotesi.

Supponiamo che la prima operazione di PRODIGIT sia davvero solo la creazione di un grande massimario (semi-automatizzato nella sua creazione?). Bene, deve essere chiaro a tutti noi che altre persone e altri strumenti avrebbe creato risultati differenti. Quindi, se il nuovo massimario, in relazione ad una certa ipotesi, ci offrirà un certo grado di informazioni (per esempio che la stessa fattispecie è stata decisa un numero x di volte, con un certo risultato α , con certe frequenti ricorrenze testuali – le massime), dobbiamo sempre tenere a mente che questa è una delle molte ricostruzioni possibili e che progetti diversi avrebbero proposto come prodotto finale informazioni diverse.

Diverse ipotesi di partenza, diverse persone e diverse tecnologie possono raggiungere risultati differenti: lo ripeto, perché va cristallizzato.

Ed è nella logica di base della scienza che questi diversi risultati si confrontino, per valutare l'accuratezza, la correttezza, la fondatezza delle ipotesi assunte.

Il che ci conduce al secondo principio fondante questo momento di scoperta.

(II) RIPRODUCIBILITÀ

La logica e la filosofia della scienza ci insegnano alcune lezioni basilari, utili per mantenerci nell'ambito della razionalità critica.

Ci sono alcuni termini, come verificabilità, falsificabilità, comparabilità, con i quali dobbiamo iniziare a familiarizzare, vestendoli pur sempre di abiti giuridici.

Ecco alcune variazioni sul principio di riproducibilità:

II.1. Trasparenza delle ipotesi di partenza.

Ogni operazione di questo tipo dipende da una serie di ipotesi di studio che devono essere formulate dal gruppo di ricerca. Per fare un esempio cui siamo avvezzi: se massimazione sarà, si dovrà definire cosa intendiamo per massima (e quale il rapporto tra la massima e il *decisum*), il che non è operazione banale, come sa chiunque si sia fatto le ossa in quelle riviste dove la massimazione era operazione quasi sacra. Come è necessario chiarire l'ipotesi definitoria di massima, sarà opportuno che siano chiariti tutti gli assunti di partenza che indirizzeranno il lavoro: si tratta di ipotesi giuridiche, logiche, tecnologiche.

Queste dovranno essere rese pubbliche: alcune prima che inizi il lavoro, altre al termine, nessuna potrà essere taciuta.

La pubblicità di queste ipotesi consentirà il primo vaglio critico sulla correttezza dell'impostazione iniziale.

II.2. Disponibilità del dataset di lavorazione.

PRODIGIT lavorerà su alcuni dataset di sentenze: presumibilmente ci sarà un *training* dataset e un dataset di applicazione (il "milione di sentenze").

Questi dataset devono essere disponibili e trasparenti. Devono essere messi a disposizione di tutti coloro che, con finalità di studio e ricerca, intendono verificare l'integrità, la validità e l'utilizzabilità dei dati (come ha già detto bene Alberto Marcheselli: quante di quelle sentenze sono davvero utilizzabili agli occhi di un operatore esperto?) e che intendono testare con altre ipotesi e altri strumenti la producibilità di risultati differenti.

A scanso di equivoci, si può dire sin d'ora *apertis verbis* che il nostro gruppo di ricerca, finanziato dal Ministero dell'Università e della Ricerca (Progetti di ricerca di rilevante interesse nazionale – PRIN 2020 – La digitalizzazione dell'Amministrazione finanziaria), è interessato ad avere accesso al dataset di lavorazione, per studiare e verificare i dati (le sentenze utilizzate) e ricercare sistemi alternativi che possano essere comparati con quanto prodotto da PRODIGIT.

Il dataset di lavorazione dovrebbe essere messo a disposizione nel momento stesso in cui PRODIGIT inizierà l'elaborazione, per consentire la simultaneità nello sviluppo dei progetti (ed evitare capziosi e purtroppo non infrequenti adeguamenti del dataset ai risultati, sulla scorta dell'adagio secondo cui "se i fatti non rispecchiano la teoria, cambia i fatti").

Se non si consentirà l'accesso e l'utilizzo dei dataset, l'operazione non avrà alcun carattere scientifico, perché mancherà di uno dei caratteri fondanti: la riproducibilità.

III.3. Trasparenza degli strumenti adottati.

L'operazione farà uso di metodi statistici e strumenti informatici.

Tutti questi strumenti devono essere resi disponibili. Nella pubblicistica dell'intelligenza artificiale, i lavori dicono e non dicono, il che è comprensibile: sono spesso in ballo rilevanti interessi economici e segreti industriali.

Qui, invece, si sta creando uno strumento che influirà sul modo di assumere decisioni tributarie che, oggi come domani, sono potenzialmente distruttive dei risparmi, dei redditi e delle attività dei cittadini. Il tutto è parte di un procedimento democratico, come ha ben ricordato Alberto Marcheselli, notando che la giustizia, ieri oggi e domani, continua ad essere amministrata nel nome del popolo italiano. I margini di opacità non sono tollerabili.

C. Nella soluzione che personalmente preferisco, in un sistema pubblico, applicato alla decisione giudiziale, dovrebbe essere resa disponibile ogni singola riga di codice utilizzato. Se la soluzione appare estrema, se non ogni singola riga di codice, dovranno comunque essere sufficientemente descritte le componenti del codice, le variabili assunte ecc.

Il motivo di questa trasparenza è presto detto: se la razionalità e il fondamento della decisione creano uno *shifting* (almeno parziale) verso la fiducia nello strumento che ha creato l'informazione di base, una parte della critica alla decisione giudiziale si sposterà verso la critica al codice che ha creato l'informazione su cui la decisione si è fondata.

Realizzare uno strumento che non consenta di falsificare le ipotesi su cui si fonda e di criticare i mezzi adottati non è scienza o progresso, ma oscurantismo magico (“sembra il medioevo, più smart e più fashion”, come canta Willie) e noi tutti confidiamo che PRODIGIT resti nella luminosa logica della scoperta scientifica.