

LEGAL ANALYTICS

Giustizia predittiva, ecco i tre pilastri per capire gli impatti della tecnologia sul diritto

[Home](#) > [Documenti Digitali](#) > [Giustizia Digitale](#)

Le tecnologie digitali impattano sul rapporto tra diritto e sistema democratico, in particolare riguardo alla possibilità di decisioni libere e informate: in questo contesto è utile approfondire tre aspetti legati tra di loro come predizione giuridica, accesso ai dati e libertà dei giudici

2 ore fa

Amedeo Santosuoso

IUSS Pavia e Dipartimento giurisprudenza UNIPV



Predizione giuridica, accesso ai dati e libertà del giudicare sono tre facce del rapporto tra **diritto e sistema democratico** nell'era delle tecnologie digitali. Sono necessari **dati giudiziari ampi e di buona qualità** (come sentenze, atti degli avvocati e atti processuali) per sviluppare **la predizione giuridica quantitativa**, che altro non è che lo studio del diritto con le modalità tecniche oggi possibili.

È necessaria una conoscenza appropriata dell'esperienza giuridica perché vi siano decisioni libere, cioè consapevoli e informate. Qui di seguito approfondiamo perché questi tre aspetti devono essere affrontati in connessione uno con l'altro.

Indice degli argomenti

Predizione giuridica quantitativa, come funziona

Gli obiettivi

Lo scenario secondo tre studi

Giustizia predittiva in Italia

Il framework giuridico di riferimento

Gli aspetti privacy

La libertà del giudice

Perché tenere alta l'attenzione critica

Conclusione

Note

Predizione giuridica quantitativa, come funziona

In un recente volume pubblicato presso MIT Press, due studiosi italiani della memoria, Tommaso Vecchi e Daniele Gatti, sostengono che la capacità di conservare informazioni (quello che comunemente è inteso come memoria del passato) è **una proprietà collegata al futuro**, tanto da potersi dire che “il vero scopo della memoria è fare predizioni”. Secondo autorevoli studi in materia “il sistema cognitivo anticipa continuamente e attivamente ciò che sta per accadere sulla base della memoria”. Insomma, l'informazione sul passato è intrecciata con i processi predittivi, che siano umani o strettamente biologici o ambientali.

Il **sistema cognitivo** mette in pratica i risultati dell'elaborazione delle relazioni causali, e dei rapporti tra eventi, basati sull'apprendimento precedente, cioè sulla memoria. Tale apprendimento può avvenire estraendo regolarità statistiche (costruendo sequenze o modelli di eventi) o facendo un uso più o meno esplicito della legge di inferenza e dell'analogia tra gli eventi; questi modelli consentono, quindi, di fare delle previsioni.

Se questo è quello che accade nella nostra attività mentale in generale, qualcosa del tutto simile accade **nella mente del giurista dell'accademia o del pratico del diritto**, giudice o avvocato. Lo studio dei testi di dottrina o delle raccolte cartacee di giurisprudenza è il modo in cui sono state fatte le predizioni giuridiche dopo l'invenzione della stampa a caratteri mobili, usando carta, penna, annotazioni personali, memoria di studio e di esperienze pratiche.

Oggi, quella stessa attività la si può fare con gli strumenti di Legal Analytics (LA), un **ambiente tecnico nel quale confluiscono diverse discipline**, come Data Science, Intelligenza artificiale (AI), Machine Learning (ML), Natural Language Processing (NLP) e statistica. L'obiettivo della LA è di estrarre e rappresentare conoscenza giuridica, rinvenire correlazioni implicite e individuare tendenze ed effettuare previsioni relative alle decisioni, giudiziarie o anche circa la possibilità che un disegno di legge diventi legge realmente o circa le decisioni della PA o di soggetti privati.



Gli obiettivi

La differenza rispetto al passato **sta nelle tecniche usate**, che ora consentono di raggiungere i medesimi risultati in un tempo nettamente inferiore (fattore tempo) e di far emergere connessioni o correlazioni non visibili anche a un occhio esperto e che, una volta emerse, si offrono alla riflessione giuridica (estrazione di conoscenza).

L'obiettivo rimane largamente simile a quello del passato: **rivolgersi alla memoria pregressa per formulare predizioni**. Con un'interessante differenza metodologica (che riguarda non solo il diritto ma tutto il mondo della ricerca scientifica): mentre nell'impostazione usata tradizionalmente in approcci sperimentali o quasi-sperimentali si andava dalla costruzione di un modello verso (*forward*) i dati osservabili, dove si cercavano conferme o smentite, nell'apprendimento automatico si usano, invece, i dati osservabili per costruire il modello (*inverse*)^[1]. In altri termini, il rapporto ipotesi-test, alla base del metodo scientifico moderno, è invertito ed è l'analisi con tecnologie *machine learning* di un appropriato insieme di dati a far emergere un possibile modello da poi testare.

In sintesi, si può dire che **quanto migliore sarà il livello di analisi dei dati del passato**, tanto maggiore sarà **l'accuratezza della predizione di decisioni future**, sapendo però che vi è un limite connaturato a qualsiasi predizione quantitativa, costituito dal fatto di basarsi sul presupposto implicito che le condizioni di partenza siano stabili (isomorfismo), mentre ciò non accade nella realtà e, meno che mai, nel diritto.

Lo scenario secondo tre studi

Tutto ciò si potrebbe fare e lo si è cominciato a fare a un primo livello sperimentale nell'ultima decade (se si escludono alcuni sporadici tentativi del passato). Mi riferisco agli studi di Nikos Aletras e altri sulle decisioni della Corte Europea dei Diritti dell'uomo^[2], a quello di Daniel Katz e altri sulle decisioni della Corte Suprema degli USA^[3] e a quello di Nay sui disegni di legge presentati negli USA e alla previsione della loro trasformazione in leggi approvate dal potere legislativo^[4].

★ LIVE STREAMING, 11 MARZO

6 aziende raccontano: AI e Cloud ready, l'esperienza di oggi per costruire il futuro

Big Data # Cloud



Scopri di più

Iscriviti subito!

Questi tre studi adottano tecniche diverse (analisi testuale quello sulla CEDU, analisi extratestuale quello sulla Corte USA, mentre l'ultimo adotta un criterio misto che combina i due precedenti) e si basano su insiemi di dati (*dataset*) di diversa natura e raccolti con modalità diverse, tanto da potersi dire che difficilmente essi costituiscano un vero e proprio stato dell'arte.

Giustizia predittiva in Italia

In Italia, limitandosi alle sole realtà in cui si fa uso di tecniche di *machine learning*, si ha notizia di uno studio in corso presso la Scuola Superiore Sant'Anna di Pisa, in collaborazione con i Tribunali di Genova e di Pisa, avente come oggetto la giurisprudenza in materia di alimenti per la separazione e di alimenti per il divorzio in Italia e di uno studio condotto dall'**Università di Bologna e dall'Università di Pavia** all'interno del progetto PRIN "LAILA", avente ad oggetto l'analisi (e relativa predizione) delle decisioni in materia di proprietà intellettuale. È possibile che altro sia in movimento e mi scuso con chi non si vedesse citato. Ma per procedere su questa strada è necessaria la disponibilità ampia dei dati.

Le tecnologie di *machine learning* devono essere applicate a insiemi di dati (*dataset*) che devono essere appropriati, per quantità e qualità, come si diceva sopra. La Cina, tanto per fare un esempio, rende milioni di dati liberamente accessibili via web al sito, che, affermano con vanto, "è diventato il più grande database di sentenze del mondo", dal quale si può arrivare ad accedere a *China Judgments Online*, dove "chiunque può leggere quasi tutti i documenti delle sentenze pronunciate dai tribunali cinesi dal 2013". Ora, senza voler competere con questo grande Paese asiatico per quantità di dati e per qualità democratica, **in Italia vi è certamente una gran quantità di dati disponibili**. Per avere idea della dimensione basti pensare alla storia dell'informatica giuridica in Italia (di cui il lavoro del

Centro Elaborazione Dati -CED- della Cassazione è chiaro esempio) e all'accumulo quotidiano delle decisioni e degli atti di parte che il processo civile telematico (PCT) produce, e altro ancora. A dispetto di ciò, però, non abbiamo niente di paragonabile al libero accesso che il governo cinese consente.

Il framework giuridico di riferimento

Eppure, nel quadro giuridico internazionale, europeo e italiano vi è più di un fondamento giuridico della libertà di accesso alle decisioni. In un interessante contributo, al quale rinvio per il dettaglio, Enrico Consolandi^[5] ricorda come l'**art. 6 CEDU**, quando stabilisce che "la sentenza deve essere resa pubblicamente", pone una regola che è rivolta sia alla tutela delle parti di quel processo, sia a garantire l'interesse generale alla correttezza della giurisdizione e alla conoscibilità dell'ordinamento: una sentenza resa pubblicamente è una decisione presa in modo legale in nome dell'intera comunità. Non va infatti dimenticato che la giurisdizione è attività pubblica ed è sotto il controllo dell'opinione pubblica.

Un'analisi OSCE del 2008 sull'accesso alle decisioni giudiziarie afferma che "il principale fine del diritto a un giudizio pubblico è di assicurare che l'amministrazione della giustizia sia accessibile e aperta alla valutazione pubblica. Perciò il diritto ad avere un giudizio che si svolga pubblicamente può essere rivendicato da chiunque, incluse persone che non sono parti del processo"^[6]. Inoltre, passando a una questione di dettaglio, ma assai significativa, **la necessità di avere una banca dati delle decisioni**, che sia completa e accessibile, è diventata una stretta necessità da quando i giudici possono motivare le sentenze "anche con riferimento a precedenti conformi" (art. 118 disp. att. c.p.c.).

Ora, o il precedente è conoscibile dalle parti in giudizio, e allora potrà dirsi garantita la condizione di parità tra le parti (art. 111, comma 2 Cost.), oppure la non piena e paritaria accessibilità ai precedenti porta a decisioni potenzialmente prese in violazione di quella garanzia costituzionale di parità e, alla fine, a una violazione del diritto di difesa (art. 24 Cost.), specie se il giudice fa riferimento a un precedente non versato in giudizio.

Gli aspetti privacy

E la privacy? Non è questa la sede per affrontare i dettagli del problema (per il quale rinvio ancora all'articolo di Consolandi), si può solo dire che il vecchio *codice privacy* (l'art. 52 d.lgs. n. 196/2003, non modificato dal successivo regolamento) **dava facoltà alla parte che ritenga di avere "motivi legittimi"** di chiedere tempestivamente (e cioè prima del deposito della decisione che definisce il grado di giudizio) che sia apposta "sull'originale della sentenza o del provvedimento, un'annotazione volta a precludere, in caso di riproduzione della sentenza o provvedimento in qualsiasi forma, per finalità di informazione giuridica su riviste giuridiche, supporti elettronici o mediante reti di comunicazione elettronica, l'indicazione delle generalità e di altri dati identificativi del medesimo interessato riportati sulla sentenza o provvedimento".

Leggendo la norma in controluce (e in linea con il comma 7, secondo il quale "fuori dei casi indicati nel presente articolo è ammessa la diffusione in ogni forma del contenuto anche integrale di sentenze e di altri provvedimenti giurisdizionali") si può affermare che, quando il diretto interessato non esercita quella facoltà, **si espanda l'interesse generale di rilievo costituzionale** alla piena conoscibilità delle decisioni giudiziarie. Il controllo sulla circolazione dei dati e delle notizie personali costituisce l'ambito della privacy o riservatezza delle persone fisiche. Lo stesso ordine di idee, però, se applicato a istituzioni deputate allo svolgimento di un'attività essenzialmente di pubblico interesse, cambia di segno, perché la privacy delle istituzioni pubbliche è il "segreto" e l'arbitrio del segreto.

Dunque, ferma restando la facoltà prevista dal nostro ordinamento che una parte chieda tempestivamente l'anonimizzazione (che potrebbe risolversi in una pseudonimizzazione, magari più coerente con le finalità del trattamento), la regola generale non può che essere quella della piena conoscibilità delle decisioni giudiziarie. E oggi **la pubblicazione e la conoscenza dell'intero patrimonio delle decisioni**

adottate ben può considerarsi qualcosa che si realizza attraverso la processabilità e l'utilizzabilità con tecniche di *machine learning* di tutto il patrimonio delle decisioni pubbliche.

La libertà del giudice

La libertà dei giudici è, innanzitutto libertà dalle pressioni che il sistema politico o istituzionale possa esercitare nei loro confronti. Ma, inutile dire, non è libertà di decidere indipendentemente dal diritto o di non rispondere alle questioni proposte in giudizio dalle parti. La questione **in qualche modo nuova** che si pone si può così formulare: se oggi la conoscenza del diritto passa attraverso sistemi di analisi quantitativa dei precedenti, quali sono i rischi che la libertà dei giudici sia compromessa, nel senso di essere condizionata?

Un problema sul quale vi è grande sensibilità è quello della profilazione dei giudici e anche di interi uffici. Se il problema è reale, il rischio è a mio avviso attenuabile con relativa facilità, con alcuni interventi regolatori, come linee guida, alcuni espliciti divieti, accorgimenti tecnici e altro. Detto questo non credo sia giusto alimentare la falsa idea che l'orientamento culturale di un giudice o di un ufficio debbano essere al riparo dalla pubblica conoscibilità. **Una sorta di pretesa di opacità del giudicante** non mi pare che sia conforme ai principi costituzionali e di civiltà della funzione giurisdizionale e, alla fine, che sia realisticamente esigibile, visto che (esclusi atteggiamenti intimidatori o diffamatori) è scrupolo professionale di un avvocato, prima di proporre una lite, informarsi sugli orientamenti esistenti in una determinata sede e parlarne con il cliente.

È, piuttosto, utile focalizzare una questione diversa, che può **incidere grandemente sulla qualità delle decisioni**. L'introduzione di supporti decisionali basati su elaborazioni quantitative (*data-driven*) quali cambiamenti induce nel lavoro dei giudici (come di qualsiasi decisore), e quali cambiamenti ciò produce nella logica della decisione e della motivazione? E, poi, come possono coesistere la natura intrinseca degli schemi emergenti dall'analisi giuridica (e, in alcuni casi, la loro limitata spiegabilità) e il diritto alla spiegazione delle decisioni pubbliche, che disposizioni costituzionali fondamentali riconoscono?

Sono questioni complesse che possono essere affrontate solo con la diffusione della conoscenza **di queste tecniche tra i giudici** e lo sviluppo di una cultura aperta e critica al tempo stesso. In altra sede ho espresso la mia opinione sull'eccesso di preoccupazione che vi è a proposito delle scatole nere (*black boxes*) nei processi decisionali e individuato alcune soluzioni^[7].

Perché tenere alta l'attenzione critica

Ma, accanto alla crescita culturale dei giudici, vi è da chiedersi in quale modo si possa contrastare lo sviluppo di forme di passivizzazione conformistica verso le proposte delle macchine e se **qualche dispositivo tecnico possa aiutare a tenere sveglia l'attenzione critica**. Si pensi a cosa potrà accadere quando i giudici italiani verranno forniti di sistemi di costruzione documentale (*document builder*) che possono essere interrogati attraverso una maschera nella quale il giudice inserisce i dati essenziali del caso da decidere (materia, domande delle parti, riferimenti normativi essenziali, epoca dei fatti, e altro) e ricevere come risposta una "proposta" di possibile percorso decisionale, i cui passaggi sono corredati da riferimenti giurisprudenziali e normativi. Un giudice pressato da un ruolo pesante e di fronte a casi seriali potrebbe essere tentato di adagiarsi acriticamente su quella "proposta".

Qui può essere di aiuto il sistema di *prediction poisoning* che Stefano Quintarelli ha proposto di utilizzare per le decisioni mediche^[8] e che, mi pare, possa essere utile anche nel contesto della decisione giudiziaria. Quintarelli parte dal presupposto che quello medico-macchina è un "sistema a due opinioni" in cui l'opinione della macchina tenderebbe a prevalere. Fa, quindi, **l'ipotesi che la macchina faccia la previsione sbagliata** e che il medico, pur intimamente non concordando, si conformi alla previsione della macchina, provocando o contribuendo a provocare un danno al paziente.

Per tenere desta l'attenzione critica del medico, la macchina potrebbe alterare consapevolmente alcune delle previsioni (tenendone traccia) e informare il medico che alcune delle previsioni che gli vengono comunicate (senza dire quali) sono in realtà l'opposto di ciò che

la macchina ha realmente elaborato. Queste previsioni sono *poisoned* (avvelenate). La macchina fa una previsione, ma al medico può essere comunicata una previsione *poisoned*, mentre quella autentica della macchina viene chiamata “sincera”. Le situazioni che si possono creare sono le seguenti:

- la previsione della macchina è “sincera” e quella del medico è concordante. In questo caso la diagnosi è confermata.
- La previsione della macchina è “sincera” e quella del medico è discordante. Il caso è evidenziato e viene richiesta una seconda opinione.
- La previsione della macchina è *poisoned* e quella del medico è discordante. In questo caso la diagnosi del medico è confermata.
- La previsione della macchina è *poisoned* e quella del medico è concordante. Il caso è evidenziato dal sistema e viene avviata la procedura di richiesta di una seconda opinione.

La “seconda opinione”, secondo Quintarelli, potrebbe essere fornita da un medico o potrebbe essere una seconda macchina, non *poisoned*, con diverso algoritmo e diverso addestramento. In sintesi, quando le opinioni del sistema binario (macchina-umano) sono in disaccordo, **deve essere generata una terza opinione** per risolvere la controversia. Questo è un semplice esempio di come il concetto di *redress by design* potrebbe essere implementato tramite il *prediction poisoning*.

Conclusione

A me sembra un’idea di **potenziale grande interesse anche per l’ambito giudiziario, sulla quale vale la pena di riflettere**. Qualcuno potrà dire che è un sistema troppo oneroso (avere un sistema doppio), ma il **preservare la capacità critica dell’umano mi pare avere un’importanza superiore**.

Non so naturalmente se e come possa essere implementato per le decisioni giudiziarie, ma mi sembra un modo intelligente di trarre tutti i benefici possibili dalle macchine, facendo salvo quello che ci sta più a cuore.

▶ WEBINAR

360On TV al centro della notizia. Nuovo Governo, la sfida: ricostruire l'economia italiana



Leggi l'informativa sulla privacy

Email

Consente l'invio di comunicazioni promozionali inerenti i prodotti e servizi di soggetti terzi rispetto alle Contitolari che appartengono al ramo manifatturiero, di servizi (in particolare ICT) e di commercio, con modalità di contatto automatizzate e tradizionali da parte dei terzi medesimi, a cui vengono comunicati i dati.

ISCRIVITI

Così si chiude il cerchio. La libertà del giudice non è libertà *contro* le macchine, ma *con* le macchine usate in modo intelligente, così da trarre ogni beneficio utile, dalla estrazione della conoscenza che i grandi dati contengono fino ad attuare sistemi di decisione informati, veloci e critici. ■

Note

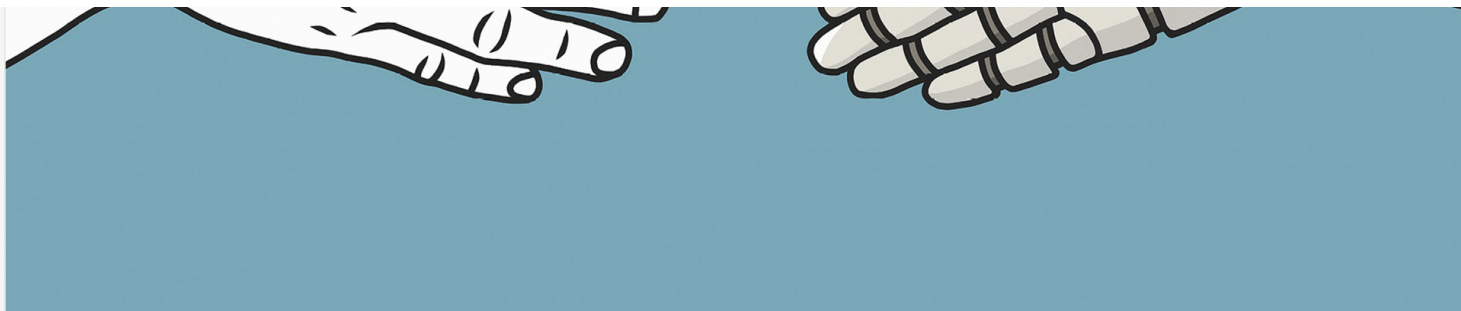
1. A. Santosuosso, *Intelligenza artificiale e diritto*, Mondadori Università, 2020, p. 99. ↑
2. Nikolaos Aletras – Dimitrios Tsarapatsanis – Daniel Preoțiu-Pietro – Vasileios Lampos, Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective, in «PeerJ Computer Science», 2, 2016. Disponibile all'indirizzo <https://doi.org/10.7717/peerj-cs.93>. ↑
3. Daniel Martin Katz, Bommarito, Blackman, A general approach for predicting the behavior of the Supreme Court of the United States, in «PLoS ONE» 12(4), 2017, disponibile all'indirizzo: <https://doi.org/10.1371/journal.pone.0174698>. ↑
4. J.J. Nay, Predicting and understanding law-making with word vectors and an ensemble model, in «PLoS ONE» 12(5): e0176999. <https://doi.org/10.1371/journal.pone.0176999>. ↑
5. Enrico Consolandi, *Conoscenza informatica della giurisprudenza e privacy*, in *Il libro dell'anno del diritto 2019 (2019)* ↑
6. *Access to Court Decisions, A legal Analysis of relevant international and national provisions*, OSCE 2008, p. 3. ↑
7. A. Santosuosso, *Intelligenza artificiale e diritto*, Mondadori Università, 2020. ↑
8. S. Quintarelli, *Il medico sarà succube dell'algoritmo? Le tecniche per ridurre il rischio*, in *Agenda Digitale*, 18 dicembre 2019 ↑

@RIPRODUZIONE RISERVATA

WEBINAR

Intelligenza artificiale: Impatto sulle nostre vite, diritti e libertà





Il webcast è disponibile

GUARDA

EGUIDE

Richmond IT Director Forum: un confronto tra esperti per ripartire dall'innovazione

01 Mar 2021

Scaricalo gratis!

DOWNLOAD

Argomenti

mano nella società digitale



processo civile telematico

Canali

tà digitali



Documenti digitali



giustizia digitale

Articoli correlati





LO SCENARIO

L'IA leva per un nuovo modello socio-economico sostenibile: ecco come

16 Dic 2020

di Gianpiero Ruggiero e Piero Poccianti

Condividi 



LO SCENARIO

Intelligenza artificiale, gli usi pubblici in Italia: polizia, giustizia, sanità, fisco

07 Dic 2020

di Marina Rita Carbone

di Marina Rita Cardone

Condividi 



L'ANALISI

Gemello digitale fra innovazione e datacrazia: ecco perché serve riappropriarsi dei dati

19 Ott 2020

di Maria Pia Rossignaud

Condividi 

LO SCENARIO

L'IA leva per un nuovo modello socio-economico sostenibile: ecco come



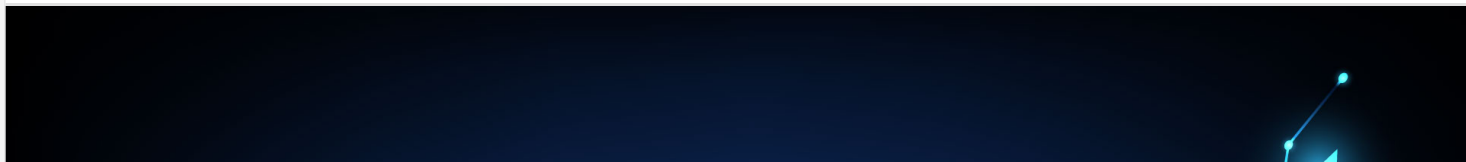
L'emergenza sanitaria ha imposto un cambiamento, in cui l'Intelligenza artificiale può rappresentare il fulcro di un nuovo modello di sostenibilità per economia e società: vediamo gli ambiti in cui l'applicazione della tecnologia potrebbe favorire questo processo

16 Dic 2020

[CONTINUA A LEGGERE](#)

WEBINAR

L'IT Financial Management per gestire la spesa tecnologica e finanziare l'innovazione e la trasformazione digitale





Inizia tra 16 gg 1 ore 9 min 54 sec

ISCRIVITI

WHITE PAPER

Come ottimizzare la gestione multi-cluster con Kubernetes

03 Feb 2021

Scaricalo gratis!

DOWNLOAD