SENZABARCODE
10 ANNI 2012 2022
Non siamo tutti uguali !



Cronaca

# Intelligenza artificiale: Etica, Rischi e Governance

24 Dicembre 2022 SenzaBarcode Redazione intelligenza artificiale, Pierluigi Testa, Think Tank Trinità dei Monti

## Dalla più famosa partita di scacchi all'esigenza della spiegabilità del risultato. A.I. Etica, Rischi e Governance di una Rivoluzione in corso.

Martedì 20 dicembre 2022, il Think Tank Trinità dei Monti, fondato e diretto da **Pierluigi Testa**, ha organizzato nella storica sede dell'Hotel Palazzetto in Roma, nel cuore di Piazza di Spagna, un incontro di approfondimento sull'Intelligenza Artificiale, con ospite **Germana Lo Sapio**, Magistrato amministrativo con qualifica di Consigliere presso il TAR per la Campania.

Nell'introduzione il Presidente del Think Tank Trinità dei Monti Pierluigi Testa ha ricordato come l'intelligenza artificiale è stata equiparata all'elettricità. Entrambe hanno determinato cambiamenti epocali nella storia dell'umanità.

L'elettricità ha caratterizzato la Seconda rivoluzione industriale a metà dell'800; l'Intelligenza artificiale (IA) è considerata una delle tecnologie trainanti della Quarta rivoluzione industriale. Come parte delle scienze computazionali, l'IA è oggetto di studi e applicazioni e i primi investimenti per la ricerca si sono visti negli Stati Uniti a partire dagli anni Cinquanta.

## Negli ultimi dieci anni la ricerca ha fatto passi da gigante

grazie, da un lato, alla enorme disponibilità dei dati, per effetto della pervasiva digitalizzazione della società attraverso la rete, i social network, l'interoperabilità tecnologica e semantica tra i sistemi informativi, la tracciabilità di tutti gli aspetti delle interazioni umane; dall'altro, grazie allo sviluppo delle capacità di calcolo ad alte prestazioni (High Performance Computing HPC) indispensabile per elaborarli.

Vi sono però almeno due caratteri che distinguono l'IA dall'elettricità e la rendono fonte di nuove sfide etiche e giuridiche. Il primo è la diversa velocità dello sviluppo, cosicché ogni regola o categoria giuridica, che pare al momento adeguata a "governare" il fenomeno, rischia di diventare presto obsoleta. Il secondo è la diffusa consapevolezza dei rischi connessi agli usi dell'IA per i meccanismi di funzionamento della democrazia, rischi che dipendono dalle caratteristiche tecniche di questo sistema di tecnologie.

L'evento è proseguito con il contributo della special guest **Germana Lo Sapio, Magistrato amministrativo con qualifica di Consigliere presso il TAR per la Campania**, su:

## 1. La definizione dell'Intelligenza Artificiale

In ambito giuridico è talmente importante l'esigenza di definire l'intelligenza artificiale che le istituzioni europee si occupano del tema almeno dal 2017 e in ogni documento non compare ancora un atto normativo vincolante. Si sta lavorando al regolamento sull'intelligenza artificiale. In tutta la produzione documentale dal 2017 ad oggi, è stato profuso un enorme sforzo per fornire alcune definizioni giuridiche.

Si è arrivati all'ultima definizione – che è inclusa nella bozza di regolamento – in cui si afferma che l'**intelligenza artificiale** si riferisce ai **sistemi che, elaborando dati e utilizzando una notevole capacità di calcolo, raggiungono obiettivi assegnati con un certo grado di autonomia rispetto alle istruzioni iniziali**.

In quest'ultima definizione non c'è più il riferimento alla simulazione dell'intelligenza umana.

L'intelligenza artificiale abbraccia due grandi categorie di approcci:

1. il *machine learning*, che include anche il *deep learning*.
2. I sistemi simbolici esperti, più "*old fashioned*", in cui si danno alla macchina un insieme di conoscenze e la macchina elabora delle deduzioni ragionate, imitando il ragionamento umano.

### Quello che conta, tuttavia, è comprendere in prospettiva evolutiva cosa includa veramente l'intelligenza artificiale, che oggi risulta composta da tre elementi fondamentali:

1. **I dati:** l'exploit dell'intelligenza artificiale c'è stato negli ultimi anni grazie alla progressiva digitalizzazione della realtà, dei fenomeni reali. In questo processo si include la tracciabilità di tutti gli aspetti della nostra vita quotidiana, professionale e personale.
2. **La potenza di calcolo:** questa è una narrativa che spesso viene esclusa anche dagli stessi studi giuridici, come se fosse un mondo a parte. In realtà è lì che risiede la potenza dell'intelligenza artificiale.
3. **Gli algoritmi** che sono formule matematiche pensate dai matematici proprio per far funzionare l'intelligenza artificiale.

Sui dati si è prima partiti negli anni '90 con il *web*, poi si è affermato il *social web*, infine oggi abbiamo il *semantic web* ovvero il *web 3.0*, in cui i dati vengono trasmessi da macchina a macchina, includendo in questa categoria anche l'*Internet of things.*

Sui dati la Commissione Europea ha fatto una previsione affermando che nel 2025 si arriverà a 175 zettabyte di dati (1 zettabyte equivale a 180 milioni di volte le informazioni contenute in tutti i documenti della biblioteca del congresso di Washington).

Ancor più impressionante è che **il 90% della rappresentazione dei dati virtuali è stata prodotta negli ultimi due anni**; infatti, ogni minuto:

- su Google vengono fatte 6 milioni di domande;
- su whatsapp ci sono 69 milioni di messaggi;
- vengono inviate 197 milioni di e-mail.

Sono tutti i dati che entrano nel grande mondo del web, alimentando in modo invisibile i sistemi di l'intelligenza artificiale. **Questi dati sono considerati il nuovo il nuovo petrolio**.

La differenza tra la quarta rivoluzione industriale, quella in corso, e la seconda rivoluzione industriale, che si fondava sull'elettricità, è che la quarta è anche una rivoluzione ambientale, ecologica, nel senso che cambia l'ambiente in cui viviamo. Infatti, la Commissione Europea nel definire l'intelligenza artificiale parla di "eco-sistema".

## 2. Eventi che hanno che ha fatto sì che l'intelligenza artificiale raggiungesse le masse

Tra più importanti eventi che hanno proiettato l'intelligenza artificiale nella coscienza di massa ci sono due occasioni particolari che si riferiscono in realtà al tema del gioco.

Ci sono eventi che hanno colpito la generazione nata prima degli anni '90 e che hanno come paradigma comune il gioco.

Il gioco in realtà è una sorta di pilastro della rivoluzione digitale: se pensiamo ad esempio a tutte le piattaforme mondiali a cui accediamo quotidianamente, tutte sono caratterizzate da una *user experience* che richiama il gioco.

Qualunque applicazione di successo che utilizza l'intelligenza artificiale nella sua fruibilità deve dare l'idea di un gioco.

Questo implica una sorta di contraddizione – che crea anche grandi equivoci dal punto di vista della comprensione – tra quello che appare e che superficialmente sembra semplice, *user friendly*, (facilmente utilizzabile), e il sottostante che è dato da una struttura molto complessa, tendenzialmente instabile e soggetta ad un cambiamento continuo, che deve essere manutenuta e continuamente aggiornata.

L'aggiornamento, infatti, è parte del DNA dell'ambiente digitale.

Nel **1997** ci fu una famosissima serie di partite di scacchi, passata alla storia, tra **Kasparov e *"Deep Blue"* di IBM**, uno scontro serrato tra l'uomo e la macchina in cui ha sempre prevalso l'uomo, il campione russo, fino all'ultima partita in cui ha vinto la macchina.

## Fu la prima volta che una macchina sconfisse l'uomo in un gioco considerato espressione di intelligenza

Questo evento marcò un grande passo per l'umanità, così come riconosciuto anche da Garry Kasparov dopo trent'anni che ebbe modo di affermare di essere stato il "*primo lavoratore intellettuale ad essere messo in discussione da una macchina*".

Fu la prima grande vittoria dell'intelligenza artificiale contro l'uomo, una macchina che elaborava qualche milione di possibili risposte al secondo.

La vittoria fu dovuta all'enorme capacità computazionale, e quindi di calcolo, della macchina, non c'era nulla di *machine learning*. I progettisti di "*Deep Blue*" avevano dato in pasto alla macchina tutte le migliori partite di scacchi della storia e quindi la macchina conosceva tutte le possibili risposte ad una singola mossa.

* In realtà, secondo **Antonio Ballarin**, **membro dell'International Neural Network Society e *Visiting Professor* alla University Canada West di Vancouver**, intervenuto all'evento come scienziato esperto, la sconfitta di Kasparov è più dovuta ad un errore di stanchezza: infatti lo scienziato afferma che dopo aver riguardato attentamente tutte le partite e tutte le mosse fatte dal campione sovietico, la sconfitta nell'ultima partita è dipesa da mosse che non risultavano le migliori possibili da parte di Kasparov ed evidentemente sbagliate.

Cosa diversa e ancor più globale fu la **vittoria di AlphaGo al gioco del *"Go"*** perché non solo sconfisse i più grandi campioni (tra cui Lee Sedol), ma soprattutto lo fece attraverso una serie di miglioramenti che essa stessa conseguì nel corso del gioco, evolvendo le sue capacità.

## 3. Etica e regolamentazione

Un problema che si pone a livello europeo con l'utilizzo dell'intelligenza artificiale è che talvolta questa produce dei risultati, ma non si riesce a spiegare il come sia riuscita a farlo e questo pone la macchina al pari di un oracolo a cui credere senza alcuna capacità critica, genera una sorta di pulsione alla fede nell'uomo sul fatto che il risultato ottenuto abbia eseguito il miglior processo possibile.

La spiegazione con cui si arriva ad un risultato costituisce una sorta di certificazione del processo ed è essa stessa parte del risultato. **Laddove non c'è una spiegazione si pone un problema etico**.

L'etica e la regolamentazione sono due tematiche connesse, la prima si cerca di risolverla attraverso la normativa.

**Uno dei problemi etici che pone l'intelligenza artificiale è proprio quello della spiegabilità.**

Come faccio ad utilizzare l'intelligenza artificiale che mi produce un risultato senza spiegarmi la motivazione, per orientare una decisione un giudizio amministrativo?

## Noi viviamo in un mondo fatto da regole e norme fondate sul principio della spiegazione.

a). L'**esigenza di spiegabilità** porta lo sviluppo dell'intelligenza artificiale ad un livello superiore.

L'ordinamento giuridico è fondato sulle motivazioni sottostanti alle decisioni che impattano nell'ordinamento giuridico sulla vita delle persone. Si parla di non comprensibilità del funzionamento.

b). Un secondo aspetto importante è la **perdita del controllo** da cui nasce l'esigenza di far intervenire un controllo umano sempre e comunque. Rispetto all'esigenza di controllo nella ultima bozza di regolamento che si chiama "AI Act" si prevede un articolo specifico che denota il **principio di sorveglianza umana**: chi sorveglia la macchina deve avere non solo la consapevolezza ma anche la competenza necessaria e la formazione adeguata e questo apre a molti fronti di interpretazione su come possano essere definite.

c). Un altro aspetto è quello dell'**autorità**, ovvero si deve avere la possibilità di esercitare la delega dell'autorità e quindi attraverso questa una persona deve essere in grado di spegnere la macchina, fare in modo che ci siano procedure che ricordino all'operatore che esiste anche il c.d. "*Automation bias*". Secondo questo *bias*, l'uomo tende ad affidarsi al risultato del calcolo, del processo, tende ad affidarsi, quindi, all'operato dalla macchina.

Un altro ***bias*** importante è quello **discriminatorio**, che nasce nel momento in cui gli operatori (gli esseri umani) alimentano le macchine con dati che all'origine sono o risultano fortemente discriminatori.

Eric Loomis fu condannato dalla Corte Suprema del Wisconsin a scontare una pena di 6 anni di reclusione, aggravata di ulteriori sei anni dipendenti dalla probabilità di recidiva. La probabilità è stata calcolata da Compas, un sistema di intelligenza artificiale utilizzato dalla giustizia statunitense.

### Questo sistema calcola la probabilità di recidiva di un reato partendo dai dati anagrafici e dal momento in cui viene effettuato un reato da parte di una persona

La valutazione del rischio è rappresentata da un grafico di tre barre che attestano da 1 a 10 il rischio di recidiva preprocessuale, il rischio di recidiva generale ed il rischio di recidiva violenta ed è volta a a predire la probabilità generale che gli individui con una storia criminosa simile siano più o meno propensi a commettere un nuovo reato una volta tornati in libertà.

Eric Loomis una volta condannato fece ricorso richiedendo la possibilità di accesso all'algoritmo utilizzato da Compas e la corte suprema rispose che comunque la decisione emersa sarebbe stata la stessa presa da un giudice in persona, così avvalorando la sentenza emessa dalla macchina.

Si è scoperto poi che il ***bias discriminatorio*** di questo sistema è che **aumenta del 50% la probabilità di recidiva quando si tratta di una persona di colore**.

d) Infine, esiste il problema di "stare al passo" dei regolatori (il c.d. dilemma di Collindrige) con la velocità di sviluppo e diffusione delle applicazioni che utilizzano sistemi di Intelligenza Artificiale.

## 4. La regolazione in UE e nel Regno Unito

**In Europa** a partire dal 2017 e si è cominciato ad approfondire il tema dell'intelligenza artificiale creando un gruppo di esperti che includesse non solo i giuristi, ma anche esperti tecnici perché la materia coinvolge entrambe le categorie.

Questo gruppo individuato ha delineato una serie di principi etici che dovevano orientare l'attività normativa.

Si era di fronte a due alternative: a) quella di lavorare sulle norme esistenti ed integrarle o b) quella di redigere un regolamento sull'Intelligenza Artificiale a parte.

Si è arrivati ad una proposta di una bozza di regolamento sull'intelligenza artificiale che ha trovato anche un orientamento favorevole anche da parte del Consiglio dell'Unione Europea.

Il documento è, ora, in fase di validazione da parte degli Stati membri.

L'approccio di questo regolamento è che non si riferisce ad un singolo settore, ma a tutti i settori che vengono toccati dall'intelligenza artificiale ed ha una **valenza extraterritoriale perché si applica a tutti i sistemi esterni all'Unione Europea che vogliono entrare nel mercato europeo**.

L'**approccio** di questo regolamento è **risk-based**, ovvero nel riconoscere l'intelligenza artificiale come un volano per l'economia, non possiamo bloccarne lo sviluppo, ma perseguire l'ecosistema di eccellenza e farlo in modo tale da garantire che i rischi non si traducano in danni effettivi.

Il regolamento prevede quattro livelli di rischio dal minimo a quello inaccettabile Sono previste delle procedure molto aggressive di tracciamento e di qualità dei dati che potrebbero tuttavia limitarne l'implementazione.

## In Europa in ogni caso è vietato, ad esempio, il "***social scoring***", utilizzato altrove (Cina).

* Secondo **Antonio Ballarin**, l'Intelligenza Artificiale è "materia viva" e come tale contiene sempre un errore implicito nel suo calcolo. Questo potrebbe rappresentare un limite all'AI Act, che non prevede la sanzionabilità all'errore nell'impostazione dell'Intelligenza Artificiale da parte di un data scientist che risulta quasi sempre guidato dall'esigenza di riprodurre un'area del cervello umano.

Il **Regno Unito**, che risulta molto avanti sull'innovazione, ha un approccio regolatorio più *light*, perché delega alle autorità di settore l'individuazione delle specifiche. Oltremanica il regolamento europeo è considerato talmente invasivo che rischia di bloccare lo sviluppo dell'economia.

Nel Regno Unito vengono identificati dei principi chiave che devono orientare la regolamentazione e tra questi c'è il principio della spiegabilità.

***

La governance etico-giuridica della rivoluzione digitale è una strada ancora in salita, tutta da costruire. Ma è un capitolo in cui è necessario che tutti prendano parte, ognuno con il suo ruolo.