Paper

AVVOCATO ed IA

(*Competenze trasversali ed integrate dell'avvocato per procedere in relazione ad eventi "critici", localizzati in sistemi di IA)*

Indice

- Keywords
- Abstract
- Paper: Formazione dell'avvocato, Valutazione di Impatto, Leading cases
- Conclusione
- Fonti

Keywords

- Formazione
- Avvocato
- Conoscenza degli strumenti di IA sul mercato
- Valutazione di impatto; forensic investigation
- Provvedimento del Garante Privacy chatGPT
- Caso avvocato americano
- Caso italiano
- 360 c.p.p.
- *compliance*

Abstract

L'avvocato risulta essere sempre più "rapito" dall'utilizzo dell'IA come strumento di lavoro, affascinato, in un certo qual modo, dalla velocità di risposta del *tool de quo* e dalla certezza che ciò che viene affidato al "sistema info-telematico-computazionale-generativo" sia degno di essere "fidato strumento di lavoro" al pari di programmi con cui quotidianamente si interfaccia, come, ad esempio, le banche dati da cui trae la giurisprudenza utile per i casi che ha in difesa (anche queste, peraltro, non propriamente conosciute in punto *security*). Tuttavia ciò non è, allo stato dell'arte, ancora "perseguibile *in toto*" poiché il sistema non permette di riporre la propria "fiducia professionale" in determinati "applicativi" di cui non si ha sufficiente "controllo" (ad oggi non se ne ha quasi del tutto, a parere di chi scrive) in riferimento alla sicurezza della sua

programmazione ed alla presenza di talune vulnerabilità[1], nonché a quella dell'attendibilità dei dati inseriti utili "all'apprendimento automatico" e della conseguente, nonché relativa, elaborazione (*output*) ottenuta.

Paper

Questo scritto vuole prendere le mosse da un tema non troppo popolare. Evitando, quindi, ogni parafrasi sull'utilità (probabile, ad oggi) e sul progresso imperante dell'IA (concetto che ormai pare scontato), l'oggetto della tematica che qui si vuole trattare è la totale incertezza che, allo stato dell'arte, tale tecnologia oggi consegna (e non è facile dire ancora per quanto).

Ora, partendo dall'assunto che ad oggi l'IA trova il suo fattore determinante nella sua "capacità auto istruttiva", al netto del suo settore di impiego, quello che qui preoccupa è la mancanza di formazione, da parte della pletora dei suoi utilizzatori, in termini di sicurezza.

Formazione degli avvocati

Quello che oggi realmente "perplime" chi scrive è quanto noi avvocati, spesso, siamo più concentrati sui risultati ottenibili attraverso l'utilizzo dell'IA piuttosto che sui mezzi utilizzati per ottenere e conseguire quei risultati. Di fatto se non ci rendiamo conto di quale pericolo corriamo ogniqualvolta inviamo, ad esempio, a mezzo whatsapp, dati giudiziali (*rectius* dati particolari ai sensi del Reg. 679/2016 UE) dei nostri clienti (v. provvedimenti del giudice scambiati tra colleghi od inviati alle Cancellerie etc.), allora, forse, non siamo ancora pronti all'utilizzo di strumenti di IA.

Spieghiamoci meglio. Ogniqualvolta utilizziamo un software o un applicativo, come ad esempio quello di whatsapp, giusta anche la "caducazione" ormai perfezionatasi del Privacy Shield[2], noi non abbiamo alcuna certezza dell'allocazione (di fatto in America… o chissà dove) dei nostri dati e del loro relativo trattamento (e quanto sia sicuro o meno, nella fattispecie).

Posto ciò, trasliamo il tema sull'utilizzo dell'Intelligenza Artificiale e pensiamo a quale realmente sia "il dato" che introduciamo nel *tool* che andremo ad utilizzare, nonché ad impiegare. Ed ancora pensiamo al software in uso nello specifico e, soprattutto, alla casa produttrice di esso.

Ora, indipendentemente dalla circostanza per cui si verta di Open AI e dell'ormai "noto caso ChatGPT" (ad oggi nella sua versione -4) o di Google Bard (*ex multis*), ciò che desta preoccupazione, ad esempio, sono i codici sorgente delle case madri nonché, tanto per procedere *in medias*, la programmazione dei software *ad hoc* sotto il profilo della sicurezza.

---

[1] Tra le tante… il cybercriminale è alla ricerca continua delle cosiddette vulnerabilità *zero-day*.

[2] Trattasi dell'accordo tra EU ed USA *illo tempore* in vigore - ad oggi "decaduto"- che disciplinava il regime di trasferimento dei dati personali tra l'Unione Europea e gli Stati Uniti.

Invero solo una conoscenza puntuale dei prodotti, delle sue funzionalità, dei suoi codici sorgente, delle API e dei fornitori di esse (ove questi siano altri, come quasi sempre accade, rispetto al "creatore" del "programma BOT") e del conseguente trattamento dati, potrebbe mettere al sicuro il giurista che ne fa uso.

Saper distinguere tra una BOT ed un'altra, saper valutare la credibilità dell'azienda produttrice di essa o meno, saper leggere le "informative privacy e protezione dei dati" degli strumenti acquistati (*ahinoi*, spesso scaricati in modalità *free* a seguito di una ricerca superficiale sul nostro browser "preferito"), saper leggere le linee-guida operative, saper distinguere un codice sorgente sicuro perché magari in *open source* ed ancora saper distinguere ciò che è utile che un sistema di IA possa "consegnare" all'avvocato o meno (avvocato il quale esercita una professione il cui pericolo di violazione dei diritti e delle libertà fondamentali dell'individuo è sempre "dietro l'angolo"), "fa un'enorme o, meglio, una VITALE differenza"!

Valutazione di Impatto

Ma quali sono le modalità da scegliere per la valutazione dei sistemi di IA?

La risposta alla domanda che precede, per quanto possa apparire quasi filosofica e di astratta concezione per i più, invero non lo è affatto! Il GDPR (Reg. UE 679/2016), tra i vari apporti tecnico-normativi - in punto Data Protection/Cybersecurity - utili al tema qui oggetto di trattazione, offre una soluzione interessante (fra le diverse percorribili) e lo fa con il suo "*risk based thinking*" che trova la propria massima espressione all'art. 35, meglio conosciuto come Valutazione di Impatto[3]. Tale misurazione/valutazione del rischio della tecnologia sfruttata (approccio da "*Risk assessment*")[4] diviene strumento fondamentale se ben conosciuto e se davvero si è in grado di utilizzarlo (onde evitare, ad esempio, di fare valutazioni di impatto al pari di quei DVR aziendali che in determinate realtà sono, spesso, esclusivamente dei "copia-incolla tratti dalla rete").

Altresì, a tutto questo, va aggiunta una necessaria analisi della tecnologia in uso da parte dell'avvocato e/o del *team* specializzato che lo affianca (*team* che, in una società in piena *digital trasformation*, pare sempre più necessario avere accanto), tecnologia che nel caso dell'IA è fondata su un sistema di *Machine*

---

[3] Si riporta qui il solo comma 1 seppur non pienamente esaustivo:
"*Quando un tipo di trattamento, allorché prevede in particolare l'uso di nuove tecnologie, considerati la natura, l'oggetto, il contesto e le finalità del trattamento, può presentare un rischio elevato per i diritti e le libertà delle persone fisiche, il titolare del trattamento effettua, prima di procedere al trattamento, una valutazione dell'impatto dei trattamenti previsti sulla protezione dei dati personali. Una singola valutazione può esaminare un insieme di trattamenti simili che presentano rischi elevati analoghi*".

[4] "*La valutazione di impatto del trattamento (D.P.I.A., cioè Data Protection Impact Assessment) è un processo volto a descrivere il trattamento, valutarne la necessità e la proporzionalità e a gestire gli eventuali rischi* - Risk Assessment sta proprio ad indicare la valutazione del rischio - *per i diritti e le libertà delle persone derivanti dal trattamento dei dati personali*". Ove si ritenessero essere "*sussistenti rischi per le libertà e i diritti degli interessati, si dovranno individuare le misure specifiche richieste per attenuare o eliminare tali rischi*" (protezionedatipersonali.it).

Learning che va "trattata" anche con strumenti di analisi del rischio afferenti alla *Forensic Investigation*[5], utili a scovare quali criticità si possono, per esempio, annidare nel sistema IA in uso, tra cui *Malware*, *Warm*, *Virus* etc… (parlare invece di *Deep Learning* pare, a chi scrive, ad oggi non sempre corretto vista la difficoltà di impostare una vera e propria rete neurale artificiale. Si parla di DL, infatti, in casi quali: "*la computer vision per le auto senza conducente, i robot droni impiegati per la consegna di pacchi o l'assistenza in emergenza (per esempio per la consegna di cibo o per trasfusioni in zone di crisi); il riconoscimento e la sintesi vocale e linguistica per chatbot e robot di servizio; il riconoscimento facciale per sorveglianza in paesi come la Cina; il riconoscimento immagini radiologiche o l'individuazione di sequenze genetiche o di molecole farmaceutiche; sistemi di analisi per la manutenzione predittiva su una infrastruttura o un impianto analizzando i dati dei sensori dell'IoT*)[6].

D'altro canto l'IA porta con sé anche il fenomeno dei "*data poisoning*" (avvelenamento dei dati) ovvero quello che Luigi Carrozzi, funzionario del Garante per la protezione dei dati, così definisce per Federprivacy: "*L'ampia "superficie d'attacco" offerta dai sistemi di Intelligenza Artificiale a causa della loro complessità espone tali sistemi a varie tipologie di minacce, molte riconducibili a quelle tipiche dei sistemi ICT, ma altre con caratteristiche specifiche, che è necessario conoscere e saper gestire adeguatamente.*

Il data poisoning è una delle più concrete minacce per i sistemi di Machine Learning che ne compromette le funzionalità. Tali minacce devono essere efficacemente gestite, in particolare quando ai sistemi è richiesto un elevato livello di affidabilità visti i potenziali impatti sull'uomo (ad es. applicazioni per infrastrutture critiche, sistemi a guida autonoma, settore sanitario, etc.).

Il data poisoning può essere contrastato innanzitutto includendo le applicazioni di ML ed i dataset utilizzati per l'addestramento nei processi di sicurezza dell'organizzazione proteggendo tali sistemi da possibili intrusioni di soggetti non autorizzati e da inserimenti di codice malevolo. È inoltre raccomandata l'adozione di specifici approcci di data quality (con particolare riferimento alle dimensioni dell'accuratezza e della completezza dei dati) a cominciare delle catene di fornitura dei dati (supply chain), *attraverso opportuni controlli dei dati in ingresso acquisiti da terze parti e la* "*sanificazione*" *dei dati, al fine di identifi-*

---

[5] A mezzo della investigazione forense, condotta con software *ad hoc* e attraverso metodiche investigative, appunto, specifiche, possiamo attivare l'analisi dati - *data mining* - onde ottenere informazioni utili alla consistenza o meno della criticità di cui si vuole ottenere contezza. Invero tali investigazioni ricoprono sia la fase della ricerca di tali criticità - al pari di quanto si farebbe in fase giudiziale ove tale attività è volta all'individuazione ed acquisizione dei mezzi di prova - sia quella dell'individuazione di vulnerabilità nel sistema info-telematico, eventualmente lesive dei diritti e delle libertà degli individui. Non vada taciuta la circostanza secondo cui tale attività va condotta anche attraverso l'utilizzo di procedure attuabili grazie a metodi specifici di *intelligence* condotta su fonti aperte (*OSINT*), nonché attraverso investigazioni condotte in presenza (analisi documentali, interviste, *audit*, investigazioni difensive dell'avvocato e dei suoi eventuali consulenti esperti in materia etc…).

[6] www.ai4business.it.

care ed escludere possibili dati corrotti nei data set di addestramento. La complessità e la tumultuosa evoluzione della tematica richiede un approccio corale".

Circostanze queste (tutte) che presuppongono di adottare dei sistemi di analisi del rischio, sempre a parere di chi scrive, anche in ottica di attività irripetibili, ai sensi dell'art. 360 c.p.p. e, quindi, ad esempio, attraverso attività tecniche messe in atto così come si farebbe in quella fase del processo penale in cui si agisce con immediatezza, onde evitare che i mezzi di prova possano "deperire" vista la loro natura mutevole e, talvolta, "autodistruttiva". In tal senso si pensi, ad esempio, alla ricerca di punti specifici in cui l'attaccante (hacker) ha inoculato l'eventuale malevolo e che solo la più veloce analisi possibile - oltre, ove necessaria, la cristallizzazione di quanto eseguito su copia di lavoro *ad hoc* predisposta (c.d. copia forense) - può portare in termini di comprensione dello stato di effettiva sicurezza/insicurezza dello strumento di IA che si è ad utilizzare. Infatti l'immediato intervento (o meglio un intervento preventivo in casi simili) potrebbe essere di fondamentale importanza al fine di evitare che l'utilizzatore del suddetto strumento di IA, assolutamente ignaro della nascosta operatività dello stesso malevolo, possa "inficiare" il proprio sistema info-telematico sovrascrivendo su quelle che possono essere "prove fondamentali" per evidenziare le "cattive azioni di altri" e divenendo, peraltro, veicolo di infezione per terzi, anche ottenendo, magari, risultati di ricerca falsati e non veritieri. Ed ancora si pensi a quanto vi è da fare, in punto investigazioni forensi digitali e fisiche, per accertare l'attendibilità delle fonti/dati ricavati dall'uso dell'IA (sui quali dati essa si auto-addestra) onde evitare danni di immagine, lesioni della "*brand-reputation*" ed affini criticità, le quali possono, addirittura, limitare o bloccare l'attività professionale del suo utilizzatore, con tutta una serie di ripercussioni giudiziali e disciplinari a proprio debito.

Altresì vada ancora evidenziato come l'utilizzatore di IA dovrebbe informare, nel più performante modo possibile, il proprio cliente (*interessato* ai sensi della normativa sulla Protezione dei dati) circa l'utilizzo di tali strumenti, in virtù del sacrosanto principio di trasparenza richiesto, giustamente, dalla normativa di settore e in ragione della possibilità, in capo al suddetto interessato, di opporsi a tale utilizzo.

Leading Cases

Vale la pena, a questo punto, considerare il caso di Chat-GPT, la cui società produttrice veniva messa "in stato di allerta", da parte del nostro Garante per la protezione dei Dati, causa, *ex multis*, di "*trattamento dati non sicuro, non protezione dei minori e pericolo di outputs (risultati di ricerca ed utilizzo) falsati e non attendibili*".

In tal senso pensiamo ulteriormente al caso del collega americano Steve Schwartz (per cui vale la pena spendere qualche parola in più, perché costituisce fatto meno noto), il quale "*si era presentato in aula convinto di poter vincere, forte di un alleato solido e ben informato: l'intelligenza artificiale di ChatGpt… Egli rappresentava un cliente in una causa contro una compagnia aerea ed ha finito con il rimediare, però,*

una brutta figura: sei sentenze di cause precedenti (da questi richiamate) *si sono rivelate fasulle, basate su casi mai esistiti*"[7] … ottenendo così anche una "non felice ripercussione disciplinare".

Last but not least, la situazione di "chi", a mezzo di un sistema vulnerabile proprio e di un applicativo di IA "*downloaded free*", ha ottenuto "in combinato disposto" un "crack" del proprio sistema info-telematico con conseguente "*data breach*" e correlati.

Conclusione

L'IA è, e certamente sarà, uno strumento utilissimo, di potenza straordinaria, ma se fin da subito non ne scopriamo ed apprendiamo le funzionalità, soprattutto quelle nascoste (potenzialmente anche molto pericolose sotto il profilo delle superfici di attacco cibernetiche), ci troveremo innanzi ad una turba di *netizen* (*rectius cybercitizen*) alla guida di una "Ferrari" pur sprovvisti di patente. L'auspicio di chi scrive è quello di assistere, partecipandovi attivamente, ad una crescita della formazione trasversale, integrata e multidisciplinare del giurista e di tutti gli addetti ai lavori (e non solo), così come richiesto dalla realistica *compliance* sul tema, visto il cammino che tutti noi siamo "costretti" a percorrere in direzione di "sconosciute realtà artificiali".

Fonti: **- guerradirete.com** **- federprivacy.org**
- protezionedatipersonali.it **- www.ai4business.it**
- www.affaritaliani.it **- garanteprivacy.it**

12.06.2023

Avv. Paolo PISANO

Team.O.ne.T
CyberTech Forensic
Security & Investigation

- Cofondatore del "Team.O.ne. T. Cyber Tech Forensic - Torino"
- Membro del CTS "Data Security" di ALFAFORM
- Socio Federprivacy
- Docente in "data protection, cybersecurity e cybercrime"
- Socio aderente presso LAIC: "Laboratorio Avvocati Investigatori Criminologi"

[7] www.affaritaliani.it