Privacy, sicurezza ed etica nell’utilizzo dei Big Data e dell’Intelligenza Artificiale

Presentazione di un quadro d’insieme degli aspetti di particolare rilievo in tema di privacy, sicurezza, ed etica per l’utilizzo di soluzioni basate sui Big Data e sull’Intelligenza Artificiale nella produzione, nella competizione dei mercati, nelle innovazioni di importanti settori pubblici, nella normalità della vita dei cittadini.

Premessa

I Big Data (BD) e l’intelligenza artificiale (IA) sono diventati un fattore strategico nella produzione, nella competizione dei mercati, nelle innovazioni di importanti settori pubblici, nella struttura funzionale delle abitazioni, nella normalità della vita quotidiana. L’enorme vantaggio offerto dai Big Data e dalla loro capacita di estrarre significato da masse di dati anche non strutturati ed il crescente potere di autoapprendimento delle macchine, è destinato a rafforzarsi.

Ma governare questi processi non è certamente un compito semplice; la capacita di estrarre dai dati informazioni che abbiano un significato e siano funzionali, richiede infatti lo sviluppo di sofisticate tecnologie e di competenze interdisciplinari che operino a stretto contatto. Inoltre, l’utilizzo di queste tecnologie e degli scopi con questi perseguiti mettono in discussione molti paradigmi consolidati del diritto ed espongono le persone a nuovi rischi, poiché i dati rappresentano la loro proiezione digitale.

Pertanto, l’attenzione alle soluzioni BD&IA non può riguardare soltanto le loro implicazioni scientifiche e tecniche o gli

effetti delle innovazioni sull'economia, ma deve concentrarsi anche sugli impatti per le persone fisiche, in quanto, possono creare le condizioni per cui vengono effettuate valutazioni o decisioni con un intervento umano reso via via più marginale, fino ad annullarsi, con effetti dirompenti sul modo di vivere e articolare le esistenze e relazioni, in termini individuali ma anche sociali e politici.

Le Autorità europee in materia di protezione dei dati hanno più volte sottolineato il bisogno di accompagnare questi fenomeni attraverso un più rigoroso approccio etico e di generale responsabilità. In particolare, è necessario che le organizzazioni pubbliche e private adottino misure di garanzia per la trasparenza dei processi basati su tali tecnologie, anche per la progressiva difficolta a mantenere un effettivo controllo sui dati, derivate dall'opacità delle modalità di raccolta, dai luoghi di conservazione, dai criteri di selezione e di analisi. Inoltre, è necessario che sia effettuata una valutazione del potenziale discriminatorio che deriva dal loro utilizzo, che può nascere per effetto di profilazioni sempre più puntuali ed analitiche, in un modo tale che può portare ad annullare l'unicità della persona, il suo valore, la sua eccezionalità.

Elementi di fiducia

Per analizzare i Big Data, le organizzazioni utilizzano sempre più tecniche di analisi avanzate, predittive e prescrittive, che utilizzano spesso l'intelligenza artificiale e il Machine Learning (ML) in particolare, per comprendere e raccomandare azioni basate sull'analisi di elevati volumi di dati da più fonti, interne o esterne. I casi d'uso tipici includono, ad esempio, l'onboarding dei clienti, il rilevamento delle frodi e l'automazione dei processi di back office, ecc.

Al riguardo, le soluzioni di BD&IA offrono l'opportunità di riutilizzare i dati dall'uso previsto originalmente e di fornire approfondimenti e trovare correlazioni tra set di dati

che altrimenti non sarebbero stati previsti. Nuovi canali per la raccolta di dati derivanti da sviluppi tecnologici comportano una costante evoluzione delle fonti di dati e dei metodi per ottenere i dati (ad esempio raccolti, derivati, dedotti e forniti).

Come indicato dalle linee guida etiche per l'IA affidabile del Gruppo di Esperti ad Alto Livello della Commissione europea sull'IA [1], l'affidabilità e fiducia nelle soluzioni IA è essenziale per consentire il massimo beneficio dalle potenziali opportunità offerte e allo stesso tempo garantire l'uso corretto, sicuro e responsabile di tali soluzioni. Un quadro per un uso responsabile e affidabile delle soluzioni BD&IA potrebbe essere basato su una serie di elementi fondamentali, quali:

- **Etica**: lo sviluppo, la diffusione e l'uso di qualsiasi soluzione di IA dovrebbe aderire ad alcuni principi etici fondamentali, che possono essere integrati fin dall'inizio in qualsiasi progetto AI, in una sorta di approccio "etico *by design*". Ciò significa anche il business case presentato all'inizio di un progetto di IA può includere un'analisi di alto livello della conformità ai principi etici e rifiutare soluzioni non rispondenti.
- **Spiegabilità e interpretabilità:** un modello di analisi è spiegabile quando il suo comportamento interno può essere compreso direttamente dall'uomo (interpretabilità) o quando possono essere fornite spiegazioni (giustificazioni) per i principali fattori che hanno portato alla sua produzione.

La Spiegabilità è solo un elemento di trasparenza, che consiste nel rendere disponibili dati, funzionalità, algoritmi e metodi di formazione per eventuali verifiche esterne e costituisce una base per la costruzione di modelli affidabili. La necessità di spiegabilità è maggiore ogni volta che le decisioni hanno un impatto diretto sulle persone e dipendono

dal contesto particolare e dal livello di automazione implementato.

Infatti, i modelli IA possono diventare rapidamente "scatole nere", sistemi opachi per i quali il comportamento interno non può essere facilmente compreso e per i quali, pertanto, non è facile verificare come un modello di analisi abbia raggiunto una certa conclusione o previsione. L'opacità di una soluzione AI può variare a seconda della complessità del modello sottostante e della modalità di apprendimento. Ad esempio, le reti neurali tendono ad essere più opache, a causa della complessità intrinseca dell'algoritmo sottostante, rispetto agli alberi delle decisioni, il cui funzionamento interno può essere più facilmente compreso dagli umani. Questa opacità tecnica è direttamente collegata al concetto opposto di spiegabilità.

- **Equità ed eliminazione di fattori di discriminazione**: un altro aspetto importante di un modello IA affidabile è la sua correttezza. Il principio di correttezza richiede che il modello garantisca la protezione dei gruppi di persone contro fattori di discriminazione (diretta o indiretta). La discriminazione può interessare, in particolare, le popolazioni più piccole e i gruppi vulnerabili (ad esempio discriminazione basata sull'età, disabilità, riassegnazione di genere, matrimonio o unione civile, gravidanza o maternità, razza, religione o convinzioni personali, sesso, orientamento sessuale, ecc.). Per garantire l'equità (non discriminazione), il modello dovrebbe essere privo di pregiudizi; generalmente, i dati contengono distorsioni quando non sono rappresentativi della popolazione in questione. Ciò può portare a discriminazioni, ad esempio quando una classe di persone meno rappresentate nel set di dati di *training* riceve risultati più o meno favorevoli semplicemente perché il sistema ha appreso solo da alcuni esempi e non è in grado di generalizzare

correttamente. Tuttavia, la discriminazione può esistere senza pregiudizi o discriminazione diretta; può derivare da attributi sensibili che fungono da variabili di input, indipendentemente dalla distorsione.

- **Tracciabilità e verificabilità**: l'uso di soluzioni di tracciabilità aiuta a monitorare tutte le fasi, i criteri e le scelte effettuate durante tutto il processo, che consente la ripetizione dei processi risultanti nelle decisioni prese dal modello e aiuta a garantire la verificabilità del sistema. I passaggi coinvolti in una decisione presa da un modello BD&IA possono essere tracciati dalla raccolta dei dati (anche da fonti di dati di terze parti) al momento in cui viene presa la decisione e anche oltre, poiché le organizzazioni potrebbero ancora essere in dovere di spiegare come sono stati prodotti i risultati e perché sono state prese le decisioni.
- **Protezione dei dati personali**: i dati devono essere adeguatamente protetti con un sistema BD&IA affidabile e conforme alla vigente normativa sulla protezione dei dati. Le organizzazioni dovrebbero fondare il trattamento di dati personali con soluzioni BD&IA su una base legale[1] e rispondere agli obblighi previsti dalla normativa comunitaria e nazionale in materia di protezione dei dati. L'organizzazione (o "titolare del trattamento dei dati") è responsabile del trattamento lecito, equo e trasparente dei dati personali e della garanzia che i dati personali siano stati raccolti per finalità determinate, esplicite e legittime e non ulteriormente trattati in modo incompatibile con tali finalità[2].
- **Sicurezza**: l'utilizzo di nuove tecnologiche introducono anche nuove tecniche di attacco che sfruttano le vulnerabilità dei sistemi. È importante mantenere un controllo tecnico sugli ultimi attacchi alla sicurezza delle informazioni e sulle relative tecniche di difesa e garantire che la governance, la supervisione e

l’infrastruttura tecnica siano in atto per un’efficace gestione del rischio ICT.

- **Protezione dei consumatori**: un sistema BD&IA affidabile dovrebbe rispettare i diritti dei consumatori e proteggere i loro interessi. I consumatori hanno il diritto di presentare un reclamo e ricevere una risposta in un linguaggio chiaro che può essere chiaramente compreso. Il principio di spiegabilità è la chiave per far fronte a questo obbligo.

Protezione dei dati personali

L’utilizzo di sistemi di BD&IA è da considerarsi in crescita grazie al fatto che la tecnologia e il processo decisionale basato sui dati sono considerati fonti di soluzioni sempre più accettate, anche se non esiste ancora una metodologia che consenta di effettuare una valutazione basata su riscontri oggettivi dell’impatto complessivo, ma esistono elementi indicanti che la loro analisi può avere un impatto orizzontale significativo sia sul settore pubblico che su quello privato. Al riguardo, la Commissione europea riconosce il potenziale delle tecnologie e dei servizi basati sui dati, nonché dei Big Data, quali catalizzatori di crescita economica, innovazione e digitalizzazione nell’UE.

E’ da tener conto però che i progressi delle tecnologie di comunicazione e l’uso massiccio di dispositivi elettronici e di monitoraggio, dei social media, delle interazioni e delle reti web, compresi i dispositivi che comunicano informazioni senza intervento umano, hanno portato allo sviluppo di enormi insiemi di dati in costante crescita che, attraverso l’analisi e tecniche avanzate di trattamento, tracciano un quadro senza precedenti del comportamento umano, della vita privata e delle società. La “*Risoluzione del Parlamento europeo del 14 marzo 2017 sulle implicazioni dei Big Data per i diritti fondamentali*”[2] (di seguito Risoluzione del Parlamento europeo) sottolinea che la conformità con la vigente

legislazione in materia di protezione dei dati, unitamente a solide norme scientifiche ed etiche sono fondamentali per creare fiducia nelle soluzioni dei Big Data e considerarle affidabili.

Pertanto, anche in assenza di norme nazionali e internazionali che disciplinino puntualmente la materia dei Big Data, è rintracciabile nel Regolamento UE 2016/679 (di seguito Regolamento) la normativa di riferimento per eccellenza in materia di protezione dei dati. Infatti, molti casi di applicazione di BD&IA riguardano il trattamento di dati personali, che possono riferirsi ad attività di marketing, all'efficientamento dei servizi produttivi, al miglioramento dei servizi ai cittadini, alla prevenzione di attività fraudolente, ecc. Nella fattispecie, il Considerando 91 del Regolamento sancisce l'applicazione delle disposizioni anche ai trattamenti su larga scala, definiti come trattamenti riguardanti una notevole quantità di dati personali a livello regionale, nazionale e sovranazionale che potrebbero incidere su un vasto numero di interessati e presentare potenzialmente un rischio elevato per i diritti delle persone fisiche. I Big Data rientrano, quindi, nella definizione di larga scala e, pertanto, devono essere trattati nel rispetto delle disposizioni previste dal Regolamento.

Il rispetto delle disposizioni previste dal Regolamento pone nuove sfide alle organizzazioni, richiedendo una maggiore attenzione alle strategie di governo dei dati che devono essere attuate. In particolare, il principio di *accountability* (*ex* articolo 5, comma 2 del Regolamento) richiede che le organizzazioni siano in grado di garantire e dimostrare il rispetto dei principi relativi al trattamento dei dati personali.

Si tratta, quindi, di regolare ed assicurare l'*accountability* dei sistemi BD&IA, applicando il principio di trasparenza delle scelte che vengono fatte dagli algoritmi (ad esempio, con appositi diritti come il "*right of explanation*"

citato nel considerando 71 del Regolamento e poi non disciplinato nel testo normativo), l'istituzione di sistemi di tracciatura di ogni scelta effettuata dal sistema, in modo da poter renderlo trasparente, la costruzione di modelli statistici, su cui opera l'intelligenza artificiale, che assumano quale componente fondamentale l'acquisizione del *feedback* sulle scelte compiute, così da poter effettuare migliori decisioni in futuro.

Oltre al tema della trasparenza del sistema anche quello dei dati che sono trattati dallo stesso assume un'importanza primaria. È noto che in informatica vige il detto "*garbage in, garbage out*" ed in quest'ottica la qualità e pertinenza del dato su cui un sistema di intelligenza artificiale opera le proprie scelte incide in maniera fondamentale sull'esito delle stesse. È così necessario che tali sistemi, specialmente se utilizzati per assicurare l'ordine pubblico e la sicurezza sociale, evitino di fondare le proprie valutazioni su dati non attinenti alla singola persona e che possano avere conseguenze sulla stessa (come valutazioni in base al luogo di residenza, al comportamento della comunità, etc.).

In aggiunta ai temi sopra citati, molti specialisti di settore ritengono che l'articolo 22 del Regolamento sia particolarmente attinente alle applicazioni IA che elaborano dati personali. L'articolo limita il diritto dei titolari del trattamento dei dati di prendere decisioni su un soggetto "basato esclusivamente sul trattamento automatizzato", laddove tali decisioni abbiano effetti giuridici o altrettanto significativi sull'individuo. Tale processo decisionale individuale automatizzato include, a titolo esemplificativo, la profilazione, definita dall'articolo 4 del Regolamento come:

"*qualsiasi forma di trattamento automatizzato di dati personali consistente nell'utilizzo di tali dati per valutare determinati aspetti personali relativi a una persona fisica, in particolare per analizzare o prevedere aspetti riguardanti*

il rendimento professionale, la situazione economica, la salute, le preferenze personali, gli interessi, l'affidabilità, il comportamento, l'ubicazione o gli spostamenti di detta persona fisica"

Il Comitato Europeo per la Protezione dei Dati (in precedenza denominato Article 29 Working Party) sta lavorando attivamente, in particolare emettendo pareri e linee guida, per supportare le organizzazioni nell'adempimento degli obblighi previsti dal Regolamento come, ad esempio, il documento "Guidelines on Automated individual decision-making and Profiling" [3] che rappresenta una linea guida essenziale per le organizzazioni che utilizzano o pianificano di utilizzare soluzioni IA.

Sicurezza delle Informazioni

La sicurezza delle informazioni è definita come la protezione delle informazioni e dei sistemi da accesso, uso, divulgazione, interruzione, modifica o distruzione non autorizzati al fine di garantire la riservatezza, l'integrità e la disponibilità. Nel contesto BD&IA, la sicurezza dei dati e la sicurezza dei modelli sono di particolare importanza in quanto entrambi sono aspetti della sicurezza delle informazioni che sono essenziali per il corretto funzionamento degli algoritmi di BD&IA. Mentre la sicurezza dei dati si concentra sulla protezione della riservatezza, integrità e disponibilità dei dati, la sicurezza dei modelli indirizza gli attacchi e le corrispondenti misure di protezione specifiche dei modelli IA.

La rapida espansione dell'intelligenza artificiale in nuovi settori con nuove parti interessate, unita a un panorama delle minacce in evoluzione e all'enorme crescita degli strumenti BD&IA, presenta sfide difficili per la sicurezza delle informazioni.

Alcuni dei principali tipi di attacco che colpiscono, in

particolare, le soluzioni IA includono:

- furto del modello IA;
- attacchi di avvelenamento (*poisoning*);
- attacchi avversari (compresi gli attacchi di evasione).

Gli attacchi di furto / estrazione dei modelli vengono utilizzati per "rubare" i modelli replicandone il funzionamento interno. Questo viene fatto semplicemente sondando il modello target con un elevato numero di query di previsione e utilizzando la risposta ricevuta (la previsione) per addestrare un altro modello. Il modello clonato può raggiungere un alto livello di precisione, anche superiore al 99,9%.

In attacchi di avvelenamento, gli aggressori influenzano deliberatamente i dati di *training* per manipolare i risultati di un modello predittivo. Questo tipo di attacco è particolarmente valido se il modello è esposto a Internet in modalità online, ovvero il modello viene continuamente aggiornato imparando da nuovi dati.

Un attacco "avversario" consiste nel fornire un campione di dati di input che è stato leggermente alterato per indurre il modello a classificarlo erroneamente. Nella maggior parte dei casi, una perturbazione così piccola (che sostanzialmente rappresenta un "rumore") può essere così sottile che un essere umano non se ne accorge nemmeno (ad es. Quando il problema riguarda una classificazione dell'immagine e l'immagine in ingresso è stata alterata da un rumore non evidente a l'occhio umano).

Un esempio particolare di un attacco avversario è un attacco di evasione, che consiste nel creare l'input per fare in modo che il modello eviti il rilevamento di un particolare oggetto / elemento.

I ricercatori hanno recentemente sviluppato alcune nuove tecniche di difesa per sconfiggere questi attacchi, ad esempio

aggiungendo un de-noiser prima della fase di input. Inoltre, sono in fase di sviluppo librerie open source che offrono tecniche per testare la solidità dei modelli contro questo tipo di attacco.

Pertanto, è importante per le organizzazioni che adottano soluzioni IA mantengano un controllo tecnico sulle vulnerabilità dei sistemi utilizzati e siano regolarmente aggiornate sui progressi degli attacchi alla sicurezza e delle relative tecniche di difesa. In generale, le funzioni aziendali che gestiscono la sicurezza delle informazioni devono avere il compito di individuare e valutare i rischi associati, definire le strategie di gestione del rischio più opportune e garantire che l'azienda abbia le capacità e le risorse necessarie per affrontare e ridurre tali rischi durante tutto il ciclo di vita dei processi IA. A seguito di ciò, è necessario definire e attuare garanzie adeguate per la sicurezza dei dati, che ricomprendano misure tecniche e organizzative per garantire un livello di sicurezza adeguato al rischio. Tali misure potrebbero essere affrontate nell'ambito di una strategia globale di gestione della sicurezza delle informazioni o, in alternativa, stabilendo ruoli dedicati e un quadro di gestione della sicurezza specifico per la sicurezza dei dati in relazione a soluzioni BD&IA, consentendo all'organizzazione di sviluppare una strategia e procedure per monitorare, rilevare e rispondere agli incidenti di sicurezza dei dati relativi ai sistemi BD&IA, alle loro fonti di dati e ai fornitori di tecnologia di terze parti.

In alcuni casi questa può essere una sfida importante, a causa della scarsa esperienza pregressa relativa alle tecnologie da parte di molte organizzazioni. La politica e l'industria devono perseguire con coerenza e rapidità l'obiettivo di determinare un quadro di riferimento chiaro relativo alla protezione della privacy e della sicurezza dei cittadini, definendo norme e standard di riferimento circa le

caratteristiche di protezione dei sistemi BD&IA destinati ad un largo uso.

Al riguardo, è da segnalare il lavoro della ETSI Industry Specification Group on Securing Artificial Intelligence (ISG SAI), che ha un ruolo chiave in Europa per il miglioramento della sicurezza dell'IA attraverso la produzione di standard tecnici. Il gruppo ISG SAI ha l'obiettivo di creare standard per preservare e migliorare la sicurezza delle nuove tecnologie IA. In particolare, L'Istituto Europeo per gli Standard nelle Telecomunicazioni (in inglese European Telecommunications Standards Institute, acronimo ETSI) è un organismo internazionale, indipendente e senza fini di lucro ufficialmente responsabile della definizione e dell'emissione di standard nel campo delle telecomunicazioni in Europa. Il gruppo di ISG SAI sviluppa specifiche tecniche e relazioni per affrontare aspetti importanti per la sicurezza dell'intelligenza artificiale:

- Proteggere l'IA dagli attacchi: dove l'IA è un componente di un sistema che necessita di protezione;
- Mitigare gli attacchi che sfruttano l'IA: dove l'intelligenza artificiale viene utilizzata per migliorare i vettori di attacco convenzionali o creare nuovi vettori di attacco;
- Usare l'IA per migliorare le misure di sicurezza: proteggere i sistemi dagli attacchi laddove l'utilizzo dell'IA fa parte della "soluzione" o viene utilizzato per migliorare le contromisure più convenzionali.

Le parti interessate dall'attività del gruppo ETSI ISG SAI comprendono utenti finali, produttori, operatori e governi. Per maggiori informazioni sul lavoro di ETSI ISG SAI è possibile riferirsi al sito istituzionale al seguente link: https://www.etsi.org/technologies/securing-artificial-intelligence#mytoc3

Un altro istituto di riferimento per la sicurezza delle

soluzioni BD&IA è il NIST (National Institute of Standards and Technology), che è un'agenzia del governo degli Stati Uniti d'America con una reputazione consolidata negli anni nello sviluppo di standard per rendere le tecnologie più sicure, utilizzabili, interoperabili e affidabili. La ricerca NIST si concentra su come misurare e migliorare la sicurezza e l'affidabilità dei sistemi IA. Ciò include la partecipazione allo sviluppo di standard internazionali che garantiscono innovazione, fiducia del pubblico e fiducia nei sistemi che utilizzano le tecnologie IA. Inoltre, il NIST sta applicando l'IA ai problemi di misurazione per ottenere una visione più approfondita della ricerca stessa e per comprendere meglio le capacità e i limiti dell'IA.

L'11 febbraio 2019, l'Executive Order on Maintaining American Leadership in Artificial Intelligence ha incaricato il NIST di sviluppare "un piano per l'impegno federale nello sviluppo di standard tecnici e strumenti correlati a sostegno di sistemi affidabili, robusti e affidabili che utilizzano tecnologie di intelligenza artificiale".

In particolare, il NIST parteciperà allo sviluppo di standard di intelligenza artificiale, insieme al settore privato e al mondo accademico, che affrontano questioni e principi sociali ed etici, governance e politiche e principi sulla privacy. L'impegno relativo agli standard IA include:

- Supportare e condurre ricerca e sviluppo dell'IA
- Sviluppare attivamente standard IA
- Sviluppare e attuare politiche di supporto, comprese le politiche normative ove necessario.

Mentre la comunità dell'intelligenza artificiale ha concordato sul fatto che questi problemi devono tener conto degli standard dell'intelligenza artificiale, molte decisioni devono ancora essere prese sul fatto che vi siano ancora abbastanza basi scientifiche e tecniche per sviluppare tali standard.

Al riguardo, l'ENISA[3] ha pubblicato il 15 Dicembre 2020 il rapporto denominato "AI Threat Landscape", che sottolinea l'importanza della sicurezza informatica e della protezione dei dati in ogni parte dell'ecosistema IA, al fine di creare tecnologie affidabili per gli utenti. In particolare, lo studio dell'agenzia europea, oltre a presentare le minacce applicabili alle soluzioni IA, pone le basi per le prossime iniziative politiche sulla sicurezza informatica e le relative linee guida tecniche, sottolineando quelle che saranno le sfide rilevanti.

I principali punti trattati nel rapporto possono essere riassunti in:

- Definizione del campo di applicazione dell'IA nel contesto della sicurezza informatica secondo un approccio basato sul ciclo di vita, delineando l'ecosistema dei sistemi e delle applicazioni dell'IA.
- Identificazione degli elementi che costituiscono un ecosistema IA, in modo da individuare cosa deve essere protetto e le vulnerabilità che possono essere esposte per la sicurezza dell'ecosistema IA.
- Classificazione delle minacce per i diversi elementi dell'ecosistema IA durante il suo ciclo di vita.
- Mappatura delle minacce IA mediante una tassonomia dettagliata, al fine di identificare eventuali scenari di attacco e definire i termini nelle prossime valutazioni dei rischi e dei controlli di sicurezza necessari per mitigarli.

In particolare, il rapporto è diretto ai responsabili delle politiche quando sviluppano linee guida future sulle implementazioni sicure dell'IA, agli esperti tecnici per supportare valutazioni del rischio personalizzate e agli organismi di standardizzazione per supportare i futuri standard di sicurezza dell'IA.

In conclusione, il panorama delle minacce dell'IA dell'ENISA

identifica le sfide e le opportunità per implementare sistemi e servizi di IA sicuri in tutta l'Unione, sottolineando la necessità di misure di sicurezza più mirate e proporzionate per mitigare le minacce individuate.

Quadro normativo di riferimento

L'Unione Europea (di seguito UE) ha istituito un quadro giuridico rigoroso, i cui obiettivi sono, tra gli altri, quelli di garantire la tutela dei consumatori, far fronte alle pratiche commerciali sleali e proteggere i dati personali e la privacy dei cittadini. L'acquis comunitario prevede, inoltre, norme settoriali specifiche (ad esempio per il settore dell'assistenza sanitaria o dei trasporti). Tali disposizioni vigenti del diritto dell'UE continueranno ad applicarsi in relazione anche ai processi di BD&IA, ma potrebbero rendersi necessari determinati adeguamenti di tale quadro giuridico per tenere conto della trasformazione digitale e dell'uso delle specifiche tecnologie. Di conseguenza, gli aspetti già trattati dalla legislazione orizzontale o settoriale in vigore (ad esempio per quanto riguarda i dispositivi medici o i sistemi di trasporto) continueranno ad essere disciplinati da tale legislazione.

Esiste un vasto corpus di norme vigenti dell'UE in materia di sicurezza dei prodotti e di responsabilità per danno da prodotti difettosi, comprendente norme settoriali ed integrate dalle legislazioni nazionali; tale corpus di norme è pertinente e potenzialmente applicabile a una serie di applicazioni di BD&IA emergenti.

Per quanto riguarda la protezione dei diritti fondamentali e dei diritti dei consumatori, il quadro legislativo dell'UE comprende atti normativi come la direttiva sull'uguaglianza razziale (Direttiva 2000/43/CE), la direttiva sulla parità di trattamento in materia di occupazione e di condizioni di lavoro (Direttiva 2000/78/CE), le direttive sulla parità di trattamento tra uomini e donne per quanto riguarda l'accesso a

beni e servizi e in materia di occupazione (Direttiva 2004/113/CE; Direttiva 2006/54/CE), una serie di norme sulla tutela dei consumatori (es.: Direttiva sulle pratiche commerciali sleali n. 2005/29/CE e Direttiva sui diritti dei consumatori n. 2011/83/CE), sulla protezione dei dati personali e sulla privacy, in particolare il Regolamento generale sulla protezione dei dati (Regolamento UE 2016/679), nonché altre normative settoriali riguardanti la protezione dei dati personali, come la direttiva sulla protezione dei dati nelle attività di polizia e giudiziarie (Direttiva (UE) 2016/680). Inoltre, a partire dal 2025 si applicheranno le norme sui requisiti di accessibilità dei beni e dei servizi previste dall'atto europeo sull'accessibilità (Direttiva (UE) 2019/882).

I diritti fondamentali devono, inoltre, essere rispettati in sede di attuazione di altre normative dell'UE, anche nel settore dei servizi finanziari, della migrazione o della responsabilità degli intermediari online.

Se da un lato la legislazione dell'UE rimane in linea di principio pienamente applicabile, indipendentemente dall'utilizzo di tecnologie BD&IA, dall'altro è importante valutare se tale normativa possa essere adeguatamente applicata per far fronte ai rischi derivanti dall'impiego di tali sistemi o se sia necessario apportare adeguamenti a determinati strumenti giuridici.

Ad esempio, gli operatori economici rimangono pienamente responsabili della conformità dei processi che presuppongono l'utilizzo di tecnologie BD&IA alle norme vigenti a tutela dei consumatori: è vietato utilizzare algoritmi che sfruttano il comportamento dei consumatori in violazione delle norme vigenti, e qualunque infrazione a tale divieto è punita di conseguenza.

Pertanto, è da prevedere nei prossimi anni un'evoluzione del quadro normativo di riferimento, al fine di disciplinare il

corretto utilizzo delle tecnologie BD&IA per la tutela dei consumatori e dei diritti fondamentali delle persone.

Norme sull'Intelligenza Artificiale

La definizione del quadro normativo per l'utilizzo di tecnologie AI è rappresentata dallo sviluppo di un percorso teso a realizzare una legislazione europea in materia. Tale percorso è stato iniziato ad aprile 2019, quando la Commissione europea ha pubblicato, in allegato alla sua Comunicazione al Parlamento europeo, al Consiglio, al Comitato economico e sociale europeo e al Comitato delle Regioni dal titolo "Creare fiducia nell'intelligenza artificiale antropocentrica", le linee guida etiche per un'intelligenza artificiale affidabile redatte da un gruppo di esperti di alto livello da essa istituito [1]. Lo stesso gruppo di esperti fece seguire nel giugno 2019 le sue raccomandazioni politiche e di investimento per un'intelligenza affidabile.

Pochi mesi dopo il suo insediamento, la Presidente della Commissione, Ursula Von der Leyen, fissò tra le sue priorità in ambito legislativo quella di guidare l'Europa attraverso la transizione digitale per rispondere alle sfide relative alle nuove tecnologie. La Presidente ha così dichiarato di voler introdurre al più presto una legislazione specifica sull'intelligenza artificiale, simile al Regolamento sulla protezione dei dati, sulla base dei rapporti realizzati dagli esperti innanzi richiamati.

Nel Febbraio del 2020 è stato realizzato un primo passo verso la realizzazione di una legislazione europea, attraverso la pubblicazione del *c.d.* **Libro Bianco sull'IA** [4]. I libri bianchi della Commissione (in inglese white paper) "sono documenti che contengono proposte di azione dell'UE in un settore specifico. Essi costituiscono, talvolta, il prolungamento dei libri verdi, il cui scopo è invece quello di avviare un processo di consultazione a livello dell'UE. Lo scopo dei libri bianchi è di avviare una discussione con il

pubblico, le parti interessate, il Parlamento europeo e il Consiglio allo scopo di raggiungere il consenso politico.

Il Libro Bianco ribadisce la necessità di un approccio comune europeo all'IA per raggiungere dimensioni sufficienti ed evitare la frammentazione del mercato unico. L'introduzione di iniziative nazionali rischia di compromettere la certezza del diritto, di indebolire la fiducia dei cittadini e di ostacolare l'emergere di un'industria europea dinamica. La costituzione di un futuro quadro normativo per l'IA in Europa consentirà, pertanto, di creare un "ecosistema di fiducia" unico. A tal fine, dovrà garantire il rispetto delle norme dell'UE, comprese le norme a tutela dei diritti fondamentali e dei diritti dei consumatori, in particolare per i sistemi di IA ad alto rischio gestiti nell'UE. La costruzione di un ecosistema di fiducia è un obiettivo strategico in sé e dovrebbe dare ai cittadini la fiducia di adottare applicazioni di IA e alle imprese e alle organizzazioni la certezza del diritto necessaria per innovare utilizzando l'AI.

La Commissione ritiene che, per stabilire in generale se una determinata applicazione di IA debba essere ritenuta ad alto rischio, occorre valutare gli interessi in gioco e considerare se il settore interessato e l'uso previsto per tale applicazione implichino rischi significativi, in particolare per quanto concerne la protezione della sicurezza, dei diritti dei consumatori e dei diritti fondamentali. Più specificamente, un'applicazione di IA dovrebbe essere considerata ad alto rischio se soddisfa i due criteri cumulativi descritti di seguito:

- In primo luogo, l'applicazione di IA è utilizzata in un settore in cui i rischi sono generalmente ritenuti più probabili. I settori interessati dovrebbero essere elencati in maniera specifica ed esaustiva nel nuovo quadro normativo. Ad esempio, settori dell'assistenza sanitaria, dei trasporti, dell'energia e parti del settore pubblico. L'elenco dovrebbe essere

periodicamente rivisto e modificato, ove necessario, in funzione dei pertinenti sviluppi nella pratica.

- In secondo luogo, l'applicazione dell'IA nel settore in questione è, inoltre, utilizzata in modo tale da poter generare rischi significativi. Questo secondo criterio riconosce il fatto che non tutti gli usi dell'IA nei settori selezionati comportano necessariamente rischi significativi. La valutazione del livello di rischio derivante da un determinato uso potrebbe basarsi sull'impatto per i soggetti interessati. Ad esempio, usi delle applicazioni di IA che producono effetti giuridici, o effetti altrettanto rilevanti, sui diritti di una persona o di una società; usi che presentano il rischio danni materiali o immateriali significativi; usi che producono effetti non ragionevolmente evitabili dalle persone fisiche o giuridiche.

L'applicazione dei due criteri cumulativi garantirebbe un ambito di applicazione del quadro normativo mirato e atto a garantire la certezza del diritto. Le prescrizioni obbligatorie contenute nel nuovo quadro normativo sull'IA si applicherebbero in linea di principio unicamente a quelle che siano state identificate come applicazioni ad alto rischio utilizzando i due criteri cumulativi sopra descritti.

Fatto salvo quanto precede, possono esistere anche casi eccezionali in cui, a causa dei rischi in gioco, l'uso di applicazioni di IA per determinati scopi deve essere considerato ad alto rischio di per sé, ossia indipendentemente dal settore interessato, per cui si applicherebbero comunque prescrizioni puntuali. A titolo esemplificativo, si possono citare in particolare i seguenti usi:

- Alla luce dell'importanza che riveste per le persone, e tenendo conto dell'acquis dell'UE sulla parità in materia di occupazione, l'uso delle applicazioni di IA nei processi di selezione del personale e nelle situazioni che incidono sui diritti dei lavoratori

sarebbe sempre considerato “ad alto rischio”.
- L’uso delle applicazioni di IA a fini di identificazione biometrica remota e l’impiego di altre tecnologie di sorveglianza intrusive sarebbero sempre considerati “ad alto rischio”.

In conclusione, gli orientamenti esposti dalla Commissione per un’intelligenza artificiale affidabile si basano su un principio di IA antropocentrica (*cfr.* “Creare fiducia nell’intelligenza artificiale antropocentrica” [5]) volto a garantire che i valori umani rivestano un ruolo centrale nelle modalità di sviluppo, distribuzione, utilizzo e monitoraggio dei sistemi di IA, garantendo il rispetto dei diritti fondamentali, tra cui quelli sanciti nei trattati dell’Unione europea e nella Carta dei diritti fondamentali dell’Unione europea, accomunati dal riferimento a un fondamento condiviso radicato nel rispetto della dignità umana, nei quali l’essere umano gode di uno status morale unico e inalienabile.

Le linee guida tracciate dalla Commissione per la definizione di un quadro normativo comune nell’UE si prefiggono di garantire una costante sorveglianza umana dei sistemi IA, la robustezza tecnica e la sicurezza degli algoritmi, la riservatezza dei dati nonché il rispetto del principio di accountability stabilito dal Regolamento UE 2016/679, la trasparenza e il rispetto della diversità, la non-discriminazione e l’equità.

Norme sui Big Data

I Big Data, secondo la “*Risoluzione del Parlamento europeo del 14 marzo 2017 sulle implicazioni dei Big Data per i diritti fondamentali*” [2] (di seguito Risoluzione del Parlamento europeo), si riferiscono alla raccolta, all’analisi e all’accumulo ricorrente di ingenti quantità di dati, compresi i dati personali, provenienti da una serie di fonti diverse, che sono oggetto di un trattamento automatizzato mediante algoritmi informatici e tecniche avanzate di trattamento dei

dati, che usano sia informazioni memorizzate sia in streaming, al fine di individuare determinate correlazioni, tendenze e modelli (analisi dei Big Data).

L'enorme disponibilità di ogni sorta di informazione è una prospettiva allettante e apre innumerevoli scenari di ricerca; ad esempio, il progetto 4cE, descritto dal direttore generale dell'Oms Tedros Adhanom, può spiegare come i big data e le soluzioni IA possano trasformate i dati informatici al fine di salvare la vita ai malati Covid. In particolare, il progetto ha coinvolto 350 ospedali di 6 nazioni e riguardato più di 35mila ricoverati, dei quali sono state digitalizzate e confrontate le cartelle cliniche in una prima pubblicazione nell'aprile scorso. Grazie all'utilizzo dell'IA, che ha elaborato questi parametri nelle cartelle cliniche esaminate, è stato possibile definire come e quando intervenire sui malati gravi per evitare che peggiorino sino alla morte.

Pur avendo benefici incontestabili, però, si presenta la possibilità di risalire alle fonti di questa conoscenza e ai dati delle persone fisiche, minacciando i loro diritti sulla protezione dei dati che li riguardano. Infatti, è molto comune che un data set includa, nella sua forma originale, informazioni che sono riconducibili a specifici individui – un nome, una data di nascita, un indirizzo IP. Ciò diventa potenzialmente pericoloso nel momento in cui si dispone di una chiave per collegare dati su uno stesso individuo provenienti da fonti diverse – informazioni sanitarie, preferenze e opinioni espresse sui social network, dati collegati alle spese online e ai metodi di pagamento.

Pertanto, il rispetto degli obblighi previsti dal Regolamento UE 2016/679 in materia di protezione dei dati risulta come logica conseguenza. Ciò nonostante, sono molte le difficoltà di applicazione delle disposizioni del Regolamento a questa tipologia di trattamento. Alcune di esse fanno riferimento all'oggetto della tutela. Secondo l'art. 2 del Regolamento, l'ambito di applicazione materiale riguarda il trattamento di

dati personali, che permettano l'identificazione dell'interessato. Nei Big Data il trattamento riguarda dati personali e anche dati non personali. Tuttavia, a causa dell'ingente quantità e della varietà di algoritmi che possono essere ad essi applicati, in alcuni casi, distinguere i dati personali da quelli non personali diventa estremamente difficile. La natura dei dati da analizzare, in effetti, sarà definibile solo a seguito dell'elaborazione, per cui un dato non personale potrà condurre ad un dato personale, ed un dato anonimo, invece, potrà, a seguito dell'analisi, condurre all'identificazione dell'interessato. La natura propria dei Big Data e questa difficoltà di distinzione dei dati personali dai dati non personali, in essi connaturata, genera ulteriori criticità, rendendo ardua l'applicazione di alcuni principi previsti nel Regolamento eppure ritenuti essenziali per il raggiungimento dell'*accountability* del titolare del trattamento.

In primis, la varietà e la quantità dei dati analizzati impediscono di verificare sin dall'origine l'esattezza degli stessi (principio di esattezza). In secondo luogo, l'impossibilità di stabilire dall'inizio le finalità della raccolta e dell'analisi delle informazioni a causa da un lato delle potenzialità intrinseche del dato e, dall'altro, dalla varietà delle elaborazioni che possono condurre a nuovi dati (principio di minimizzazione dei dati e di principio di limitazione delle finalità).

L'apparente inesauribilità dei dati, trattabili per qualsiasi finalità, fa sorgere criticità anche per il principio di limitazione della conservazione. Infine, merita attenzione anche il consenso dell'interessato. Nei Big Data non sempre vi è un consenso informato dell'interessato per le ulteriori finalità perseguite. Ciò deriva proprio dal gap strutturale e conoscitivo tra l'utente ed il titolare del trattamento. L'interessato-utente non sempre ha tutte le informazioni per poter esprimere un consenso informato, ed inoltre, anche nel

caso in cui le avesse, non potrebbe comprenderle, poiché si fondano su algoritmi e software basati su un linguaggio complesso ai più incomprensibile.

Le considerazioni fin qui esposte mostrano come il ruolo assunto dai Big Data nel "funzionamento dei mercati, nel benessere dei consumatori, ma anche sotto il profilo sociale e democratico" è destinato a diventare sempre più rilevante e centrale. Tuttavia, la rilevanza e la centralità del fenomeno evidenziano, come sopra descritto, l'inadeguatezza del quadro normativo attuale, rendendo assolutamente necessario un intervento del legislatore. Anche l'Autorità Garante per la protezione dei dati (Garante Privacy), di concerto con l'Autorità Garante delle Comunicazioni (AGCOM) e l'Autorità Garante della Concorrenza e del Mercato (AGCM), si è fatta portatrice di questa esigenza.

Con il provvedimento del 10 luglio 2019, in effetti, è stata annunciata la fine dell'indagine conoscitiva ed interdisciplinare avviata nel maggio 2017 congiuntamente dalle tre Autorità con l'obiettivo di comprendere le implicazioni – per la privacy, la regolazione, la tutela della concorrenza e del consumatore – dello sviluppo di un'economia digitale fondata sulla raccolta e analisi di una mole sempre più ingente di dati.

Nel corso dell'indagine sono state svolte, da parte delle tre Autorità, circa quaranta audizioni, nell'ambito delle quali sono stati interpellati i principali operatori dell'economia dei dati, delle telecomunicazioni, dei settori finanziari e dell'editoria, nonché esperti e accademici. Sono state inviate richieste di informazioni ai grandi operatori digitali e sono pervenuti numerosi contributi. Il risultato dell'indagine conoscitiva ha portato alla costituzione delle Linee e Raccomandazioni di Policy [6] condivise relative al tema dei Big Data. Ciò che rende il documento interessante è il fatto che le Autorità hanno ritenuto opportuno anticipare, almeno in parte, le conclusioni delle analisi svolte, con la

pubblicazione di un documento che identificasse delle raccomandazioni destinate non tanto (e non solo) agli operatori economici coinvolti, quanto ai legislatori, al fine di indirizzare le future, necessarie regolamentazioni.

Entrando nel merito delle Linee Guida, queste si sviluppano attraverso raccomandazioni ("Raccomandazioni") rivolte, come detto, eminentemente a legislatori e regolatori, e solo di riflesso (anche) agli operatori di mercato. In quest'ottica, particolare rilevanza assume la Raccomandazione 1, di carattere "programmatico", rivolta a Governo e Parlamento nell'ottica della promozione di "un appropriato quadro normativo che affronti la questione della piena ed effettiva trasparenza nell'uso delle informazioni personali (nei confronti dei singoli e della collettività)". Nel dettaglio, vengono brevemente considerate le disposizioni normative rilevanti in ambito privacy, concorrenza e comunicazioni, giungendo alla conclusione che sarebbe necessario un intervento di natura sostanziale soprattutto in materia di pluralismo informativo, considerato che la normativa è ancora oggi diretta a regolamentare un quadro più tradizionale, che non tiene in considerazione le peculiarità (e le possibili distorsioni) dell'economia digitale.

La Raccomandazione 5 è rilevante, invece, oltre che dal punto di vista contenutistico, anche sotto il profilo metodologico. Nel merito, è specificamente improntata su aspetti privacy, ed esprime due esigenze: da un lato, è necessario accertarsi della natura personale o non personale dei dati trattati (al fine di comprendere le disposizioni normative applicabili), dall'altro gli operatori dell'economia digitale dovrebbero essere chiamati a valutare se una persona fisica possa essere "ragionevolmente identificata" mediante l'utilizzo (e l'incrocio) di una molteplicità di dati anonimi. Sotto l'aspetto metodologico, la Raccomandazione in commento è rilevante dal momento che per la prima volta nel documento ci si rivolge non solo a legislatori e regolatori, ma anche agli

operatori del mercato.

Al riguardo, se è possibile individuare la norma che disciplina il trattamento di dati personali nel Regolamento UE 2016/679, quella relativa ai dati non personali risulta essere il Regolamento UE 2018/1807. Quest'ultimo si propone di eliminare le barriere che attualmente impediscono la libera circolazione transfrontaliera all'interno della UE dei dati che non riguardano persone fisiche identificate o identificabili. Le disposizioni contenute nel Regolamento UE 2018/1807 integrano quelle che prevedono la libera circolazione e la portabilità dei dati personali all'interno della UE contenute nel Regolamento UE 2016/679, ed assieme a queste ultime, nell'intenzione del legislatore europeo, contribuiscono alla creazione di quello "spazio comune europeo dei dati" auspicato più volte dalla stessa Commissione europea.

Nel Febbraio del 2020, è stato pubblicato il rapporto finale dell'indagine conoscitiva sui Big Data [7], che partendo da tre prospettive diverse e complementari, ha individuato le principali sfide per il loro utilizzo, tra cui:

- la centralità del dato, anche come bene economico e l'importanza della sua tutela come diritto fondamentale della persona;
- l'impatto della profilazione algoritmica e delle piattaforme on-line sul grado di concorrenza in vecchi e in nuovi mercati rilevanti;
- l'effetto del programmatic advertising sulla qualità dell'informazione e sulle modalità di diffusione e acquisizione della stessa;
- la tutela e la promozione del pluralismo on-line in un contesto informativo esposto a strategie di disinformazione e di hate speech;
- la necessità di garantire trasparenza e scelte effettive al consumatore, con particolare attenzione alla tutela dei minori, in relazione al consenso circa l'uso del

proprio dato;
- la protezione del dato personale anche in ambiti non attualmente coperti dal Regolamento UE 2016/679;
- la definizione di politiche di educazione in relazione all'uso del dato.

Riferimenti bibliografici

1. Ethics Guidelines for Trustworthy Artificial Intelligence, by High-Level Expert Group on Artificial Intelligence, published on 8 April 2019.
2. Risoluzione del Parlamento europeo del 14 marzo 2017 sulle implicazioni dei Big Data per i diritti fondamentali: privacy, protezione dei dati, non discriminazione, sicurezza e attività di contrasto.
3. Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, adopted on 3 October 2017 and revised on 6 February 2018 by Article 29 Working party
4. Libro Bianco sull'intelligenza artificiale -Un approccio europeo all'eccellenza e alla fiducia – pubblicato dalla Commissione Europea il 02 Febbraio 2020.
5. Comunicazione della Commissione al Parlamento Europeo, al Consiglio, al Comitato Economico e Sociale Europeo e al Comitato delle Regioni del 8 Aprile 2019.
6. Big Data. Linee guida e raccomandazioni di policy – Indagine conoscitiva congiunta di Agcom, Agcm e Garante privacy di Luglio 2019.
7. Indagine conoscitiva sui Big Data (rapporto finale) del 10 Febbraio 2020.

Note

[1] Articolo 6 del Regolamento UE 2016/679

[2] Articolo 5, comma 2 del Regolamento UE 2016/679

[3] ENISA – *European Union Agency for Network and Information*

Security è un centro di ricerca sulla *cyber security*, che contribuisce al mantenimento di elevati standard di sicurezza delle reti all'interno dell'Unione europea. Il lavoro dell'Agenzia si svolge in stretta connessione con i Paesi membri e con il settore privato.

Articolo a cura di **Ivano Pattelli**



Ivano Pattelli è Associate Manager presso Business Integration Partners S.p.A. e si occupa di sviluppo business e di servizi di consulenza in ambito cyber security e protezione dei dati personali, per clienti di grandi dimensioni nei diversi settori di appartenenza (pubblica amministrazione, sanità, telco, finance, luxury, ecc.). In precedenza, ha maturato un esperienza pluriennale in importanti società di consulenza ed in multinazionali del settore ICT (Ernst & Young, TTP, HPE), occupandosi principalmente di attività nell'ambito dell'IT Security Governance, Risk & Compliance, con particolare focus sugli aspetti di conformità normativa in materia di protezione dei dati personali. Nel corso degli anni ha conseguito diverse certificazioni, quali ISO/IEC 27001:2013 Lead Auditor, CISSP, CISA, CRISC, "Privacy Officer e Consulente della Privacy" di TUV Italia, European Privacy Auditor ISDP©10003:2015 e Lean 6 Sigma Green Belt.

https://www.ictsecuritymagazine.com/articoli/privacy-sicurezza-ed-etica-nellutilizzo-dei-big-data-e-dellintelligenza-artificiale/