Diritto e Giustizia
IL QUOTIDIANO DI INFORMAZIONE GIURIDICA

La proposta di Regolamento UE sull'Intelligenza Artificiale: i profili operativi del nuovo quadro normativo europeo – Parte Seconda

Prof. Avv. Alessandro del Ninno
Studio Legale Tonucci & Partners

§ 1. Introduzione: le tematiche oggetto di analisi

Nel primo dei cinque saggi – pubblicato su *Diritto e Giustizia* il 28 aprile scorso - in cui si è scelto di suddividere l'analisi pratica della proposta di *Regolamento del Parlamento europeo e del Consiglio che stabilisce norme armonizzate in materia di intelligenza artificiale (l'Artificial Intelligence Act) e che modifica talune norme legislative dell'Unione* (di seguito, per brevità, il "*Regolamento IA*") abbiamo illustrato e commentato operativamente i profili generali del Regolamento IA, l'oggetto e le finalità delle nuove regole, l'ambito di applicazione, le principali definizioni giuridiche dei concetti e – infine- abbiamo approfondito le scelte legislative circa i divieti imposti a particolari "*pratiche di intelligenza artificiale*".

Nel presente contributo sarà approfondita l'analisi pratica in merito all'approccio *risk based* che emerge dalla proposta di Regolamento IA (anche comparativamente con l'approccio basato sul rischio adottato dal Regolamento 679/2016), per poi procedere all'esame della classificazione dei sistemi di Intelligenza Artificiale *ad alto rischio* e all'avvio della descrizione – per tali sistemi - degli obblighi pratici in capo a chi li immette sul mercato o li mette in uso. I fornitori, difatti, devono "*strutturare, attuare, documentare e mantenere*" (in base a quanto previsto dall'articolo 9 della proposta di Regolamento IA) un complessivo sistema di *risk management* per i sistemi IA ad alto rischio, fatto di requisiti legali, tecnici e documentali che dunque cominceremo ad esaminare.

§ 2. La classificazione dei sistemi di Intelligenza Artificiale ad alto rischio: l'approccio risk based come scelta legislativa per la regolamentazione dei sistemi IA

Il Titolo III della proposta di Regolamento IA contiene le regole specifiche per i sistemi di IA che creano un *rischio elevato* per la salute e la sicurezza o diritti fondamentali delle persone fisiche. Intanto: non bisogna confondere le regole su tali sistemi con i divieti applicabili alle "*pratiche di intelligenza artificiale*" (cfr. art. 5 Reg. IA) che sono state commentate nel precedente contributo. I sistemi IA *alto rischio* sono difatti perfettamente leciti – anche se assai "rischiosi" – purchè conformi ai requisiti regolamentari previsti. In linea con un approccio basato sul rischio, i sistemi IA ad alto rischio sono difatti ammessi sul mercato dell'Unione Europea subordinatamente al rispetto di rigorosi requisiti obbligatori e alla preventiva valutazione della loro conformità.

La classificazione di un sistema IA "*ad alto rischio*" si basa sul suo "*scopo previsto*" (si veda la specifica definizione dell'art. 3), in linea con il criterio già adottato nella vigente

legislazione sulla sicurezza dei prodotti. Pertanto, l'approccio *risk based* del legislatore europeo e l'assegnazione di un sistema IA alla categoria "*ad alto rischio*" dipende non solo dalla funzione svolta, ma anche dalle sue finalità e dalle specifiche modalità di operatività del sistema.

L'articolo 6 della proposta di Regolamento IA introduce i criteri per la classificazione dei sistemi IA e identifica due categorie principali di sistemi di IA ad alto rischio:

1. i sistemi IA destinati ad essere utilizzati come componenti di sicurezza di prodotti soggetti alla valutazione di conformità *ex ante* da parte di terzi;
2. altri sistemi di IA *stand-alone* che hanno un impatto sui diritti e le libertà fondamentali degli individui, come elencati nell'Allegato III alla proposta di Regolamento (tale elenco di cui all'Allegato III contiene un numero limitato di sistemi di IA i cui i rischi si sono già concretizzati o è probabile che si concretizzino nel prossimo futuro).

Più in particolare, e senza distinzione se un sistema di IA venga *immesso sul mercato* oppure *messo in servizio* (si ricordino le definizioni di cui si è ampiamente dato conto nel precedente contributo), e indipendentemente dai prodotti che di seguito vengono menzionati alle lettere (a) e (b), un sistema di IA deve essere considerato *ad alto rischio* se sono soddisfatte entrambe le seguenti condizioni:

a) il sistema di intelligenza artificiale è destinato a essere utilizzato come componente di sicurezza di un prodotto, oppure è esso stesso un prodotto a cui si applica la normativa di armonizzazione dell'Unione europea elencata nell'allegato II al Regolamento (cioè tutta una serie di direttive e regolamenti tecnici elencati in detto Allegato che fanno riferimento alla disciplina su immissione sul mercato e sicurezza di macchine, veicoli, trasporti aerei, stradali e ferroviari, prodotti, apparati radio, ascensori, apparati di diagnostica medica, apparati di sicurezza personale, etc);

b) il prodotto il cui componente di sicurezza è il sistema di intelligenza artificiale o il sistema di intelligenza artificiale stesso come un prodotto, è tenuto a subire una valutazione di conformità condotta da una terza parte allo scopo di essere immesso sul mercato o messo in servizio, ai sensi della normativa di armonizzazione dell'Unione elencata sempre nell'allegato II.

Inoltre, in aggiunta ai due criteri sopra riportati e in generale applicabili per applicare interpretativamente a un sistema IA il requisito dell'*alto rischio*, l'Allegato III della proposta di Regolamento elenca una serie di sistemi di Intelligenza Artificiale che già di per sé stessi sono stati valutati dal Legislatore comunitario come sistemi "*ad alto rischio*" (in quanto hanno già concretamente dimostrato l'*alea* nelle loro modalità operative) e dunque la loro mera inclusione nell'elenco dell'Allegato III (che la Commissione Ue rivedrà e aggiornerà o integrerà ogni tre anni, in base allo sviluppo tecnologico) solleva

l'interprete dal valutare se un sistema sia o meno ad altro rischio: basta confrontare il sistema IA di volta in volta in questione con quelli di cui all'elenco dell'Allegato III.

§ 2.1. Segue. L'elenco dei sistemi di IA ad alto rischio nell'Allegato III al Regolamento: l'analisi pratica

Quali sono allora i sistemi di Intelligenza Artificiale ad alto rischio elencati nell'Allegato III alla proposta di Regolamento?

Eccoli di seguito:

1. Identificazione biometrica e classificazione delle persone fisiche:
(a) Sistemi di IA destinati a essere utilizzati come sistemi di identificazione biometrica a distanza in tempo reale ed ex post.

2. Gestione e funzionamento operativo delle infrastrutture critiche:
(a) Sistemi di IA destinati ad essere utilizzati come componenti di sicurezza nella gestione e operatività della circolazione stradale e nella fornitura di acqua, gas, riscaldamento ed elettricità.

3. Istruzione e formazione professionale:
(a) sistemi di IA destinati a essere utilizzati allo scopo di determinare l'accesso o assegnazione di persone fisiche a istituti di istruzione e formazione professionale;
(b) Sistemi di intelligenza artificiale destinati ad essere utilizzati allo scopo di valutare gli studenti in istituti di istruzione e formazione professionale e per valutare i partecipanti a nell'ambito dello svolgimento di test comunemente richiesti per l'ammissione agli istituti scolastici.

4. Occupazione, gestione dei lavoratori e accesso al lavoro autonomo:
a) sistemi di IA destinati ad essere utilizzati per il reclutamento o la selezione di persone fisiche, in particolare per pubblicizzare offerte di lavoro, effettuare lo screening o il filtraggio delle candidature, valutare candidati durante colloqui o prove selettive;
(b) Intelligenza Artificiale destinata a essere utilizzata per prendere decisioni sulla promozione e la cessazione di rapporti contrattuali di lavoro, per l'assegnazione di mansioni e per il monitoraggio e valutazione delle prestazioni lavorative e del comportamento delle persone nell'ambito di rapporti di lavoro.

5. Accesso e fruizione di servizi essenziali nel comparto privato, di servizi pubblici e fruizione di benefici:

(a) Sistemi di IA destinati ad essere utilizzati dalle autorità pubbliche – oppure per conto di autorità pubbliche - per valutare l'idoneità e il diritto delle persone fisiche di ricevere assistenza pubblica, benefici o altri servizi pubblici, nonché per gestire ove del caso i procedimenti di concessione, riduzione, revoca, contestazione di tali benefici o servizi;
(b) sistemi di IA destinati ad essere utilizzati per valutare il merito creditizio delle persone fisiche o per stabilire un punteggio (*credit scoring*), ad eccezione dei sistemi di IA messi in servizio da fornitori su piccola scala per uso proprio;
(c) Sistemi di intelligenza artificiale destinati ad essere utilizzati per assegnare o per stabilire la priorità nella assegnazione e risposta per l'invio di servizi di pronto intervento di emergenza, anche da parte dei vigili del fuoco o di personale sanitario di pronto soccorso medico;

6. Applicazione della legge/attività di Forze dell'Ordine, autorità amministrative o giudiziarie.

(a) Sistemi di IA destinati ad essere utilizzati dalle autorità deputate all'applicazione della legge per lo svolgimento di valutazioni su base individuale del rischio connesso a una persona fisica (*risk assessment individuale*) al fine di valutare il rischio che una certa persona fisica commetta reati o sia recidiva o di valutare il rischio per altre persone fisiche di rimanere potenzialmente vittime di reato;
b) sistemi di intelligenza artificiale destinati ad essere utilizzati dalle forze dell'ordine come poligrafi e strumenti simili o per rilevare lo stato emotivo di una persona fisica;
c) Sistemi di intelligenza artificiale destinati ad essere utilizzati dalle autorità deputate all'applicazione della legge o dalle Forze dell'Ordine per rilevare i *deep fake*, secondo la procedura prevista dall'art. 52 della proposta di Regolamento;
d) sistemi di IA destinati ad essere utilizzati dalle autorità deputate all'applicazione della legge o dalle Forze dell'Ordine per la valutazione dell'attendibilità delle prove nel corso delle indagini o nell'ambito dei procedimenti penali;
e) Sistemi di intelligenza artificiale destinati ad essere utilizzati dalle autorità deputate all'applicazione della legge o dalle Forze dell'Ordine per la previsione del verificarsi o del ripetersi di un reato effettivo o potenziale basandosi sulla profilazione delle persone fisiche o per valutare i tratti e le caratteristiche della personalità o il passato comportamento criminale di persone fisiche o gruppi;
(f) sistemi di IA destinati ad essere utilizzati dalle autorità deputate all'applicazione della legge o dalle Forze dell'Ordine per la profilazione di persone fisiche nel corso di accertamento, indagine o perseguimento di reati;
(g) sistemi di IA destinati ad essere utilizzati per analisi dei crimini riguardanti persone fisiche, consentendo alle autorità deputate all'applicazione della legge o alle Forze dell'Ordine di cercare dataset complessi (correlati e non correlati), di grandi dimensioni disponibili da diverse fonti o in diversi formati al fine di identificare modelli (*patterns*) sconosciuti o scoprire relazioni nascoste nei dati.

7. Migrazione, asilo e gestione del controllo delle frontiere:

(a) Sistemi di IA destinati ad essere utilizzati dalle autorità pubbliche competenti come poligrafi e strumenti simili o per rilevare lo stato emotivo di una persona fisica;
(b) sistemi di IA destinati ad essere utilizzati dalle autorità pubbliche competenti per valutare un rischio, compreso un rischio per la sicurezza, un rischio di immigrazione irregolare o un rischio per la salute pubblica implicato dall'ingresso di una persona fisica che intende entrare o è entrata nel territorio di uno Stato membro;
(c) Sistemi di IA destinati ad essere utilizzati dalle autorità pubbliche competenti per la verifica dell'autenticità dei documenti di viaggio e della documentazione di supporto riferiti a persone fisiche e rilevare documenti non autentici controllando le loro caratteristiche di sicurezza;
(d) sistemi di IA destinati ad assistere le autorità pubbliche competenti per l'esame di domande di asilo, visti e permessi di soggiorno e associati reclami relativi all'idoneità delle persone fisiche che presentato istanze relative al riconoscimento di un particolare status.

8. Amministrazione della giustizia e processi democratici:

a) sistemi di IA destinati ad assistere un'autorità giudiziaria nella ricerca dei fatti, nella loro interpretazione e nell'interpretazione della legge al fine di applicare la legge a un insieme concreto di fatti.

L'analisi delle otto categorie di sistemi IA ad alto rischio (e della loro ulteriore articolazione di dettaglio) fa agevolmente emergere una diversificata tipologia di rischi connessi ai sistemi IA che possiamo provare a classificare di seguito:

1. rischi oggettivi legati alla pericolosità operativa dei sistemi e alla gravità dei danni che in teoria potrebbero causare (cfr. classificazione n. 2 sulla gestione di infrastrutture critiche);
2. rischi connessi all'impatto sui diritti e sulle libertà fondamentali di *soggetti vulnerabili* come studenti minori, lavoratori, richiedenti asilo, etc [si vedano le classificazioni nn. 3, 4 e 7(d)];
3. rischi di discriminazione per i cittadini rispetto all'impiego da parte di autorità pubbliche di sistemi IA a supporto dello svolgimento di funzioni amministrative specifiche [si vedano le classificazioni 5(a) e 7(a),(b) e(c)];
4. rischi di discriminazione rispetto all'accesso di persone fisiche a servizi offerti da privati basati sulla profilazione [si veda la classificazione 5(b)];
5. rischi di discriminazione per i cittadini (e per la stessa democrazia) connessi all'impiego di sistemi IA a supporto dell'esercizio di poteri pubblici o costituzionali e dello svolgimento di funzioni di pubblica sicurezza a fini di prevenzione e accertamento dei reati o ai fini dell'amministrazione della giustizia [si vedano le classificazioni 6, 7(a) e 8];
6. rischi generali per le persone fisiche connessi alla classificazione biometrica e alla operatività di sistemi di riconoscimento biometrico che possono essere utilizzati per diversi scopi da autorità pubbliche o da privati (si veda la classificazione n. 1).

L'elenco di cui all'Allegato III (soggetto a revisione triennale da parte della Commissione) è inoltre emendabile e soggetto a future integrazioni da parte ella Commissione UE non solo per i fisiologici sviluppi della tecnologia: la Commissione UE dovrà procedere alla valutazione se inserire nell'elenco altri sistemi IA ad alto rischio sulla base di stringenti criteri che (anche per non impattare sul mercato) sono rigorosamente elencati all'articolo 7 della proposta di Regolamento IA, tra cui: lo scopo previsto del sistema IA; il tipo di impiego del sistema (o la probabilità di un suo impiego); se il sistema ha già causato un danno alla sicurezza o alla salute che non è reversibile o ha in altro modo impattato pregiudizialmente sui diritti e sulle libertà fondamentali di una pluralità di soggetti, in base a precise allegazioni documentali trasmesse alla Commissione UE dalle autorità nazionali di controllo), etc.

§ 2.2. Il Regolamento IA e il GDPR e il rispettivo approccio risk based: "IA ad alto rischio" vs. "rischi aventi probabilità è gravità diverse", "rischio" e "rischio elevato" nel GDPR.

I due fondamentali testi legislativi rappresentati dalla proposta di Regolamento sull'Intelligenza Artificiale (ovviamente non in vigore e appena all'inizio del percorso di approvazione) e dal Regolamento Generale sulla protezione dei dati personali 679/2016 appaiono avere su molti aspetti una stretta connessione, e – anzi – è il GDPR a porsi come modello di riferimento in più punti addirittura ripreso dal Regolamento IA quanto a meccanismi, contenuti regolatori o organismi appositamente istituiti: si pensi non solo alla definizione di "*dati biometrici*" che l'articolo 3(33) della proposta di Regolamento IA riprende dall'articolo 4(14) del GDPR, ma anche alla grande rilevanza data alle valutazioni di impatto preventive, oppure alla istituzione di un *Comitato europeo sull'Intelligenza Artificiale* (sul modello dello *European Data Protection Board*) fino all'impianto sanzionatorio che nella proposta di Regolamento IA ricalca il sistema del GDPR (con sanzioni fino a 10, 20 o 30 milioni di Euro o fino al 6% del fatturato mondiale).

Appare poi del tutto ovvio come i due testi normativi siano strettamente legati anche rispetto alla circostanza che i dati (personali e non personali) sono il presupposto della operatività e il "motore" – diremmo il "carburante" vitale – di ogni sistema di intelligenza artificiale che attraverso quantità sempre più vaste di dati (*dataset*) si addestra, si perfeziona, impara e persegue con sempre maggiore efficienza ed efficacia lo "scopo previsto". Tanto che uno degli aspetti pratici di maggiore rilevanza è appunto l'istituzione di un sistema di *data governance*, prescritto dall'articolo 10 del Regolamento IA come parte di un più ampio sistema di gestione del rischio.

Ma è appunto *l'approccio basato sul rischio* il terreno sul quale le due normative possono essere comparate per trarne spunti interessanti, come in tale sede – sia pure sinteticamente – si proverà a fare.

Il GDPR incentra l'approccio *risk based* sui rischi per i diritti e le libertà delle persone fisiche, *aventi probabilità e gravità diverse,* che possono derivare da *trattamenti di dati*

personali suscettibili di cagionare un danno fisico, materiale o immateriale, pregiudizio alla reputazione; discriminazioni, furto o usurpazione d'identità; perdite finanziarie o qualsiasi altro danno economico o sociale significativo; perdita della riservatezza dei dati personali protetti da segreto professionale, etc.

Vi sono poi rischi connessi al trattamento che il GDPR individua – potremmo dire oggettivamente:

- *in base alla natura dei dati* (si pensi al trattamento di dati di particolare natura che rivelano l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, l'appartenenza sindacale, nonché dati genetici, dati relativi alla salute o i dati relativi alla vita sessuale o a condanne penali e a reati o alle relative misure di sicurezza);
- *in base alla natura del trattamento e agli impatti che questo può determinare sulla sfera giuridica delle persone fisiche* (si pensi ai trattamenti di valutazione di aspetti personali, in particolare mediante profilazioni o analisi o previsione di aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze o gli interessi personali, l'affidabilità o il comportamento, l'ubicazione o gli spostamenti, al fine di creare o utilizzare profili personali; ma si pensi anche a tutti quei casi in cui gli interessati rischiano di essere privati dei loro diritti e delle loro libertà o venga loro impedito l'esercizio del controllo sui dati personali che li riguardano);
- *in base alla natura delle persone interessate* (si pensi alle categorie dei cosiddetti *soggetti vulnerabili* come i minori, gli anziani, i diversamente abili, i richiedenti asilo, ma anche i consumatori o i lavoratori);
- *in base a parametri quantitativi* (si pensi al trattamento che riguarda una notevole quantità di dati personali oppure un vasto numero di interessati).

Il GDPR, inoltre, non richiama parametri fissi nella individuazione dei rischi e nella loro classificazione: sostanzialmente, introduce una prima e generica distinzione tra "*rischi aventi probabilità e gravità diverse*" ancorando poi l'approccio *risk based* alle due categorie del "*rischio*" e del "*rischio elevato*", tipologie a cui il Legislatore ricollega obblighi e adempimenti diversificati. Per esempio la Valutazione di impatto (DPIA) ai sensi dell'articolo 35 GDPR è obbligatoria solo un caso di trattamenti che comportano un "*rischio elevato*" (come quelli derivanti ad esempio dalla introduzione di tecnologie nuove o innovative) mentre basta un ordinario "*rischio*" connesso al trattamento per obbligare alla tenuta del Registro delle attività di trattamento anche se per i requisiti dimensionali (meno di 250 dipendenti) non si sarebbe soggetti a tale obbligo documentale (cfr. art. 30, comma 5 GDPR). O – ancora – nell'ambito della disciplina introdotta in merito alla violazione dei dati personali (*data breach*) l'intensità del rischio è il presupposto di obblighi diversi: se il rischio è *elevato*, il titolare dovrà procedere a comunicare la violazione agli interessati (oltre a dove effettuare la notifica all'Autorità di controllo); mentre non dovrà procedere né alla notifica né alla comunicazione della *data breach* ove il titolare valuti che

l'incidente di sicurezza che ha comportato la violazione dei dati personali è improbabile che rappresenti un rischio per i diritti e le libertà degli interessati.

Ma in base a quali criteri i Titolari e i Responsabili del trattamento procedono a valutare probabilità e gravità del rischio e a classificare – anche ai fini delle contromisure da documentare – i rischi ordinari ed elevati? L'unica indicazione pratica fornita dal Legislatore nel GDPR la ritroviamo al Considerando 76 ove si legge che "*la probabilità e la gravità del rischio per i diritti e le libertà dell'interessato dovrebbero essere determinate con riguardo alla natura, all'ambito di applicazione, al contesto e alle finalità del trattamento. Il rischio dovrebbe essere considerato in base a una valutazione oggettiva mediante cui si stabilisce se i trattamenti di dati comportano un rischio o un rischio elevato*".

Infine, in questa brevissima rassegna sul concetto di rischio nel GDPR, va ricordato che l'adozione di misure tecniche e organizzative adeguate per dimostrare la conformità da parte del Titolare o del Responsabile del trattamento per quanto riguarda l'individuazione del rischio connesso al trattamento, la sua valutazione in termini di origine, natura, probabilità e gravità, passa anche attraverso la strutturazione di politiche interne e processi per attenuare il rischio fin dalla progettazione di un trattamento (si pensi ai principi di *privacy by design* e di *privacy by default*).

L'approccio *risk based* che ritroviamo invece nella proposta di Regolamento UE sull'Intelligenza Artificiale si connota per una maggiore chiarezza pratica che sottrae a valutazioni soggettive la corretta individuazione del rischio connesso alla operatività di un sistema IA immesso sul mercato o semplicemente messo in uso. Si intende dire che la individuazione di un "*alto rischio*" connesso ai sistemi IA non è lasciata agli operatori e alle loro valutazioni soggettive da documentare (come – per i trattamenti di dati personali – accade per Titolari e Responsabili), ma è fatta direttamente dal Legislatore, che:

1. elenca addirittura tali sistemi nello specifico Allegato III;
2. fissa un duplice criterio generale per cui sono da considerarsi ad alto rischio in ogni caso i sistemi IA che sono componenti di sicurezza di apparati o sono essi stessi prodotti regolamentati dalla normativa di armonizzazione elencata nell'Allegato II.

Quindi non c'è spazio interpretativo per la valutazione del rischio, con il Legislatore del Regolamento IA che in un certo senso avoca a sé la individuazione della tipologia di rischio e anche l'aggiornamento periodico dei parametri e dei criteri legislativi e tecnologici applicati (si pensi alla revisione triennale dell'Allegato III cui procederà la Commissione o all'articolo 7 che prescrive – alla Commissione – i criteri interpretativi per gli aggiornamenti periodici dell'elenco dei sistemi IA ad alto rischio).

In pratica: per individuare se un sistema IA sia o meno ad alto rischio, all'operatore non resta che fare una mera comparazione formale tra il suo sistema, i sistemi menzionati dalla legislazione di armonizzazione di cui all'Allegato II o i sistemi elencati per categorie in Allegato III.

Poi – certamente – i diversi approcci *risk based* del GDPR e del Regolamento IA hanno anche dei rilevanti punti di contatto, se non di sovrapposizione: ad esempio, i *trattamenti* di dati personali implicati dai sistemi IA alto rischio elencati all'Allegato III del Regolamento comporteranno sempre il "*rischio elevato*" per i diritti e le libertà fondamentali degli interessati di cui al GDPR; oppure, laddove nella fase di addestramento o di operatività di un sistema IA vengano in considerazione volumi considerevoli di dati (dataset) o di persone interessate, ecco che questa altro non è se uno dei parametri che il GDPR fissa per l'individuazione del "*rischio elevato*".

Infine, può osservarsi che – in un certo senso – le prescrizioni e gli obblighi pratici per la *gestione del rischio* sono ben più articolati nel Regolamento IA rispetto agli istituti analoghi del GDPR (DPIA, *privacy by design* e *privacy by default*, etc): proprio di tali requisiti di risk assessment dobbiamo ora parlare nel paragrafo che segue.

§ 3. I requisiti di conformità di un sistema di Intelligenza Artificiale ad alto rischio quale presupposto per la sua immissione sul mercato e messa in servizio: strutturare, attuare, documentare e manutenere un sistema di risk assessment

Il Capo II del Titolo III del Regolamento IA stabilisce i requisiti legali per i sistemi di IA ad alto rischio in relazione a: caratteristiche dei dati, predisposizione della *data governance*, tenuta di documentazione e registri, obblighi di trasparenza da parte dei fornitori di sistemi IA e fornitura di informazioni agli utenti, supervisione umana, robustezza, precisione e sicurezza dei sistemi. Il *set* minimo di requisiti di conformità è il risultato delle indicazioni operative e tecniche di oltre 350 organizzazioni che hanno partecipato – su impulso della Commissione UE - a lavori biennali onde predisporre un nucleo regolatorio e tecnico che fosse altresì coerente con raccomandazioni, principi e standard internazionali onde garantire che il *framework IA* proposto sia compatibile con le regole adottate dai partner commerciali internazionali dell'UE.

Le due prospettive di partenza per la valutazione di conformità dei sistemi IA ad altro rischio sono:

- l'esame dello "*scopo previsto*" del sistema;
- l'avvenuta implementazione di un *sistema di risk management*.

Con particolare riferimento al *sistema di risk management*, l'articolo 9 della proposta di Regolamento IA prescrive che ai fini della immissione sul mercato o della messa in servizio di un sistema IA, deve essere "*strutturato, attuato, documentato e manutenuto*" uno specifico sistema di gestione del rischio.

Il *sistema di risk management* dei sistemi di IA alto rischio consiste in un processo ripetuto e continuo durante l'intero ciclo di vita di tali sistemi, soggetto a regolare

aggiornamento. Il Legislatore europeo – con approccio pratico – ne individua direttamente all'interno della norma le sue fasi qualificanti:

a) identificazione e analisi dei rischi noti e prevedibili associati a ciascun sistema di IA ad alto rischio;
(b) stima e valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio viene sia utilizzato in conformità allo scopo previsto che in modalità improprie che sono però ragionevolmente prevedibili;

(c) valutazione di altri rischi che possono sorgere sulla base dell'analisi dei dati raccolti dal sistema di monitoraggio successivo all'immissione sul mercato (il sistema di monitoraggio successivo all'immissione sul mercato di un sistema di IA è uno specifico obbligo normativo che l'articolo 61 del Regolamento indirizza ai fornitori, prevedendo che questi ultimi – in modo proporzionato alla natura della tecnologia IA e ai rischi implicati – istituiscano e documentino un sistema successivo all'immissione sul mercato di un sistema IA al alto rischio per la raccolta attiva e sistematica di dati rilevanti sulle prestazioni dell'intero ciclo di vita dei sistemi. I dati raccolti - da documentare e analizzare - sono forniti direttamente dagli utenti o raccolti tramite altre fonti e devono consentire al fornitore di valutare la conformità dei sistemi IA propri ai requisiti qui elencati;

(d) adozione di adeguate misure di gestione del rischio, in base allo stato dell'arte, affinché qualsiasi rischio residuo associato a ciascun pericolo nonché il rischio residuo complessivo (che va spiegato e comunicato all'utente) del sistema di IA ad alto rischio siano giudicati accettabili, a condizione che il sistema di IA ad alto rischio venga utilizzato in conformità allo scopo previsto o in modalità improprie che sono però ragionevolmente prevedibili.

Come individuare le contromisure più appropriate (che tra l'altro, in caso di accesso di minori ai sistemi IA ad alto rischio o in caso di impatto sui minori devono essere particolarmente rigorose) per la gestione dei rischi di un sistema IA ad alto rischio? Il Legislatore europeo fissa alcuni parametri quali:

a) eliminazione o riduzione dei rischi per quanto possibile mediante un'adeguata progettazione e sviluppo;
b) se del caso, attuazione di adeguate misure di mitigazione e controllo in relazione ai rischi che non possono essere eliminati;
c) fornitura di adeguate informazioni agli utenti, in particolare relativamente a stima e valutazione dei rischi che possono emergere quando il sistema di IA ad alto rischio viene sia utilizzato in conformità allo scopo previsto che in modalità improprie che sono però ragionevolmente prevedibili e, dove appropriato, formazione per gli utenti.

Altro specifico parametro per la eliminazione o riduzione dei rischi legati all'uso del sistema AI ad alto rischio, è il seguente. Vanno tenute nel debito conto la conoscenza tecnica, l'esperienza, l'addestramento e la formazione che possono ragionevolmente

attendersi da parte dell'utente, così come l'ambiente in cui il sistema IA ad alto rischio è destinato ad essere operativo.

I sistemi di IA ad alto rischio devono inoltre essere testati allo scopo di identificare le più appropriate misure di gestione del rischio ed i test (che ovviamente da un punto di vista procedurale e contenutistico devono essere parametrati allo scopo previsto del sistema IA alto rischio) devono garantire che tali sistemi funzionino in modo coerente rispetto allo scopo previsto e in conformità ai requisiti che stiamo analizzando.

In che momento della catena di sviluppo del sistema IA devono essere condotti i test? Il Legislatore UE non individua una particolare fase tecnica nella catena di sviluppo: i test possono essere condotti in qualsiasi momento appropriato purchè prima della immissione sul mercato o della messa in servizio.

§ 3.1 Segue: risk assessment e data governance: gli adempimenti pratici.

La capacità di un sistema IA di raggiungere lo "scopo previsto" dipende – tra gli altri fattori preponderanti – dalla *qualità* dei *dataset* impiegati per l'addestramento, lo sviluppo, i test e la convalida di un sistema IA, in termini di completezza, ampiezza e rappresentatività, contestualizzazione, disponibilità ed esattezza dei dati. Più ampio, completo e rappresentativo sarà il *dataset*, minori saranno i rischi di risultati (*output*) discriminatori o errati forniti dal sistema IA nel perseguimento dello "scopo previsto". Ecco perché la proposta di Regolamento IA obbliga i fornitori - come parte del più ampio sistema di *risk assessment* per i sistemi IA ad alto rischio – ad istituire uno specifica sistema di *governance dei dati*, prevedendo che i dati impiegati per l'apprendimento/addestramento i test e la convalida dei risultati rispetto a un dato sistema IA rispondano a specifici requisiti di qualità.

I requisiti normativi di qualità dei *dataset* sono legati anche a buone prassi gestionali caratterizzanti la *data governance.* Ai sensi dell'articolo 10 della proposta di Regolamento le corrette pratiche gestionali in merito alle fasi di addestramento, test e convalida di un sistema IA riguardano in particolare:

a) corrette scelte di *design* del sistema di elaborazione dei dati;
b) l'impostazione della raccolta di dati in base al principio di pertinenza;
c) l'impostazione corretta delle operazioni di trattamento dei dati, per i quali devono prevedersi fasi quali annotazione, etichettatura, pulizia, arricchimento e aggregazione;
d) la formulazione di ipotesi e assunzioni pertinenti, in particolare per quanto riguarda la capacità dei dati di essere aderenti, pertinenti e rappresentativi rispetto a tali ipotesi e assunzioni;
e) informazioni che i dati dovrebbero misurare e rappresentare;
f) la valutazione preventiva della disponibilità, della quantità e dell'adeguatezza dei *dataset* necessari;

g) la valutazione di possibili pregiudizi o effetti discriminatori connessi ai dati (*bias*);
h) l'identificazione di eventuali lacune o carenze nei dati e come tali lacune e carenze (*data gap*) possono essere colmate.

In ogni caso, il Legislatore fissa degli obblighi stringenti circa le caratteristiche che devono avere i dati di addestramento, convalida e test (sia come *dataset* individuali che come combinazione di *dataset*), i quali:

- devono essere pertinenti, rappresentativi, privi di errori e completi;
- devono avere proprietà statistiche appropriate con riferimento – ad esempio, ed ove applicabile - alle persone o ai gruppi di persone rispetto ai quali si prevede di impiegare il sistema di IA ad alto rischio;
- devono prendere nella dovuta considerazione (o meglio: essere rappresentativi di) tutte le caratteristiche e gli elementi che risultano peculiari rispetto a specifici contesti comportamentali, geografici o funzionali all'interno dei quali il sistema IA ad alto rischio è destinato ad operare.

Un altro dei (numerosi) punti di contatto tra la proposta di Regolamento IA e il GDPR sta nella previsione di cui all'articolo 10, comma 5 della proposta che autorizza i fornitori di sistemi IA a trattare nell'ambito del sistema di *data governance* e delle fasi di addestramento, convalida e test i dati che rivelano – ai sensi degli articoli 9, comma 1 del GDPR (e anche, come richiamati dalla norma, ai sensi degli artt. 10, comma 1 della Direttiva 2016/680 e 10, comma 1 del Regolamento 2018/725) l'origine razziale o etnica, le opinioni politiche, le convinzioni religiose o filosofiche, l'appartenenza sindacale, i dati genetici, i dati biometrici, i dati relativi alla salute o alla vita sessuale o all'orientamento sessuale della persona, purchè:

a) il trattamento sia strettamente necessario per assicurare il controllo, la individuazione e la correzione di errori e discriminazioni (*bias*) che possono verificarsi nelle dette fasi di addestramento, convalida e test di un sistema IA ad alto rischio;
b) il trattamento sia soggetto a garanzie adeguate per il diritti e le libertà fondamentali delle persone fisiche, come ad esempio la previsione di impostazioni tecniche che limitino il riuso; l'applicazione di misure tecniche parametrate allo stato dell'arte tecnologico per garantire sicurezza e riservatezza dei dati, come la pseudonimizzazione (altro concetto GDPR) o la crittografia.

In sostanza: il Regolamento IA individua una specifica deroga per superare il divieto di trattamento dei dati di particolare natura fissato in via generale all'articolo 9, comma 1 del GDPR, prescrivendo altresì specifiche misure tecniche (come le limitazioni al riuso) ad integrazione delle garanzie appropriate già note al GDPR.

§ 3.2 Segue: risk assessment e documentazione tecnica da predisporre prima della immissione sul mercato o messa in uso di sistemi IA ad alto rischio

Come vedremo nel prosieguo (nei successivi contributi a commento della proposta di Regolamento IA), uno degli obblighi principali in capo ai fornitori di sistemi IA ad alto rischio è quello della *accountability* e cioè della *documentabilità* del possesso dei requisiti legali e tecnici in capo ai sistemi IA ad alto rischio prima, durate e anche dopo la loro immissione sul mercato o della loro messa in uso.

L'articolo 11 della proposta di Regolamento si occupa nello specifico della predisposizione della documentazione tecnica di un sistema di IA ad alto rischio, che deve essere redatta e disponibile prima della sua immissione sul mercato o messa in uso. Per documentazione tecnica (che deve essere sempre mantenuta aggiornata) non si intende (o non si intende solamente) la manualistica contenente le istruzioni d'uso e le specifiche tecniche, funzionali e operative del sistema IA ad alto rischio, bensì trattasi di documentazione che deve essere redatta in modo tale:

- da comprovare la conformità del sistema IA ad alto rischio a tutti i requisiti legali e tecnici del *sistema di risk management* previsto dal Capo II del Regolamento IA (e dunque va documentato il complesso di: caratteristiche dei dati, attuazione della *data governance*, tenuta di documentazione e registri, assolvimento degli obblighi di trasparenza e di fornitura di informazioni agli utenti, le modalità con cui viene svolta la supervisione umana, le garanzie di robustezza, precisione e sicurezza dei sistemi);
- da mettere in condizione le autorità nazionali competenti e gli enti di regolazione o notifica di fruire di tutte le informazioni necessarie per verificare il rispetto di tutti i sopra citati requisiti.

L'Allegato IV del Regolamento fissa con precisione i seguenti contenuti che detta documentazione tecnica deve prevedere con riferimento a ciascun sistema IA ad alto rischio:

- una generale descrizione del sistema IA;
- una descrizione dettagliata degli elementi che compongono il sistema IA e delle fasi del suo sviluppo;
- informazioni dettagliate sul monitoraggio, il funzionamento e il controllo del sistema AI, in particolare per quanto riguarda: le sue capacità e limitazioni nelle prestazioni, il grado di accuratezza rispetto a persone o gruppi di persone rispetto alle quali il sistema è destinato ad operare; il generale e complessivo livello di accuratezza atteso rispetto allo scopo previsto; i risultati imprevisti che è prevedibile si possa verificare e le fonti di rischio per la salute e l sicurezza, i diritti fondamentali e le possibili discriminazioni rispetto allo scopo previsto; le modalità con le quali è condotta la supervisione umana, comprese le misure tecniche messe

in atto per facilitare l'interpretazione degli *output* dei sistemi di IA da parte degli utenti; informazioni specifiche sui dati di input, se applicabile;

- una descrizione dettagliata del sistema generale di *risk assessment*;
- una descrizione di qualsiasi modifica apportata al sistema IA durante il suo ciclo di vita;
- l'elenco degli standard e misure tecniche armonizzate in base alla legislazione applicabile;
- una copia della dichiarazione di conformità UE;
- una descrizione dettagliata del sistema in atto per valutare le prestazioni del sistema AI
- nella fase successiva all'immissione sul mercato, nell'ambito degli obblighi di monitoraggio post commercializzazione e distribuzione.