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 Terza

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 secondo saggio è stata 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), e si è svolto l'esame della classificazione dei sistemi di Intelligenza Artificiale *ad alto rischio*. Il secondo contributo si è concluso con l'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.

In questo terzo contributo concluderemo l'esame del complessivo sistema di *risk management*, approfondendo gli obblighi di trasparenza, nonché l'impostazione cosiddetta *antropocentrica* dei sistemi IA, con l'obbligo di garantire la sorveglianza umana. Avvieremo poi l'analisi degli ulteriori obblighi organizzativi, tecnici, commerciali e di altro tipo in capo ai fornitori di sistemi IA ad alto rischio, con l'illustrazione – nello specifico – dei loro obblighi inerenti la c.d. *accountability* dell'Intelligenza Artificiale e degli aspetti pratici delle *procedure di conformità* a cui devono sottoporre i sistemi IA ad alto rischio prima di immetterli sul mercato.

§ 2. L'obbligo di un sistema di tracciamento, registrazione e conservazione dei logs/eventi generati da sistemi IA alto rischio

Come forma di obbligo di progettazione *by design*, l'articolo 12 del Regolamento IA prescrive - con riferimento ai sistemi IA *ad alto rischio* - di progettarli e svilupparli in modo tale che fin dall'inizio assicurino funzionalità di tracciamento e registrazione

automatica dei *logs* nel corso del loro funzionamento e in generale durante la loro operatività, in conformità agli standard tecnici del settore. La capacità di tracciamento e registrazione degli eventi/logs deve essere tarata in base allo "scopo previsto" e deve protrarsi per l'intero ciclo di vita del sistema IA ad alto rischio.

Con particolare ed esclusivo riferimento ai sistemi IA ad alto rischio rappresentati da quelli che consentono l'identificazione biometrica da remoto di persone fisiche – sa in tempo reale che *ex post* – il sistema di tracciamento e registrazione dei *logs* deve prevedere, quale nucleo minimo:

1. la registrazione di ciascun periodo di utilizzo del sistema (data e ora di avvio; data e ora di spegnimento; tempo complessivo di ciascuna sessione di operatività);
2. la banca dati di riferimento rispetto alla quale i dati di *input* sono statui controllati e comparati dal sistema;
3. i dati di *input* per i quali la ricerca ha condotto ad un *match*, consentendo cioè una comparazione tra dati contenuti nel database di riferimento e dati rilevati dal sistema di identificazione biometrica da remoto che ha condotto a una identificazione certa di una persona fisica;
4. i dati identificativi delle persone fisiche coinvolte nella comparazione e nella verifica dei risultati.

Come si può agevolmente intuire, gli obblighi tecnici di progettazione e sviluppo sopra illustrati hanno un impatto anche commerciale, dovendosi ad esempio per lo meno prevedere servizi (es: *cloud storage*) associati alla fornitura del sistema IA alto rischio per la conservazione prevista dei *logs*/eventi.

§ 3. La trasparenza verso gli utenti: le informazioni da fornire su tutti gli aspetti progettuali, operativi, tecnici e di sicurezza dei sistemi IA ad alto rischio

L'articolo 13 del Regolamento IA prescrive – sempre con riferimento ai sistemi IA *ad alto rischio* - di progettarli e svilupparli in modo tale che la loro operatività sia sufficientemente trasparente, nel senso di garantire all'utente che questi sia sempre in grado di interpretare i risultati (*output*) del sistema IA e di servirsene propriamente. In particolare, detto obbligo di trasparenza include la fornitura di istruzioni d'uso del sistema IA alto rischio in formato digitale (o in altro formato) che includano informazioni chiare, complete, corrette, concise, accessibili e comprensibili per gli utenti.

Fornire le istruzioni d'uso appena menzionate implica specificare:

(a) l'identità e i dettagli di contatto del fornitore e, ove applicabile, del suo rappresentante autorizzato;
(b) le caratteristiche, capacità e limitazioni delle prestazioni del sistema IA ad alto rischio, incluso:

(i) lo scopo previsto;
(ii) il livello di accuratezza, robustezza e sicurezza informatica rispetto al quale il sistema IA ad alto rischio è stato testato e validato e che può ragionevolmente essere atteso e qualsiasi circostanza nota o prevedibile che possa avere un impatto sul livello atteso di accuratezza, robustezza e sicurezza informatica;
(iii) qualsiasi circostanza nota o prevedibile relativa all'utilizzo del sistema IA ad alto rischio in conformità allo scopo previsto o tenendo presenti condizioni di utilizzo non conforme ragionevolmente prevedibile che può portare a rischi per la salute, la sicurezza o diritti fondamentali;
(iv) le prestazioni del sistema in rapporto a persone fisiche o gruppi di persone fisiche rispetto ai quali è previsto che il sistema IA debba essere operativo;
(v) quando appropriato, le specifiche sui dati di *input* o qualsiasi altra informazione rilevante in termini di addestramento del sistema, validazione, test, dataset impiegati, prendendo in considerazione lo scopo previsto del sistema IA ad alto rischio.

(c) le modifiche al sistema IA ad alto rischio e alle sue prestazioni che sono state predeterminate dal fornitore al momento dell'iniziale *assessment* della conformità del sistema, se tali modifiche sono previste;
(d) le misure di sorveglianza umana che devono essere previste, incluse le misure tecniche per facilitare l'interpretazione dei risultati (*outputs*) da parte degli utenti;
(e) il ciclo di vita atteso del sistema IA, inclusa l'indicazione di qualsiasi misura di manutenzione e cura atte ad assicurare il corretto funzionamento del sistema, come ad esempio gli aggiornamenti del software.

Quelli che precedono sono obblighi informativi variamente connotati (e che avranno anche un impatto sulle prassi commerciali e distributive): si va dalla trasparenza tecnica con le informazioni sulla *cybersicurezza*, alle informazioni – che in realtà sono vere e proprie "previsioni" – su conseguenze connesse all'utilizzo corretto del sistema IA ("scopo previsto") o non corretto (purchè detta non conformità sia ragionevolmente prevedibile); si passa inoltre da informazioni sulla operatività del sistema e dei risultati attesi variabili in base alla tipologia di utenti ("gruppi di persone") alle fondamentali informazioni (valevoli anche nella prospettiva della attivazione di responsabilità risarcitorie per eventuali danni) su come il sistema è stato sviluppato, addestrato e poi sottoposto a test e validato prima della sua immissione sul mercato o della sua messa in uso.

§ 4. Gli aspetti cybersecurity: gli obblighi di accuratezza, robustezza e resilienza dei sistemi IA ad alto rischio

Quanto alla specifica declinazione delle misure sulla *cybersicurezza* (su cui poi i fornitori devono dare le adeguate informazioni che abbiamo elencato più sopra), l'articolo 15 del Regolamento IA fissa obblighi tecnici specifici prevedendo che i sistemi IA ad alto rischio siano in generale fin dall'inizio (*by design*) progettati e sviluppati in modo tale che alla luce dello specifico "scopo previsto" operino per tutto il loro ciclo di vita con un

livello adeguato di accuratezza, robustezza e sicurezza informatica: E' tale attitudine che deve poi costituire oggetto di specifiche informazioni e istruzioni d'uso nei materiali e manuali di accompagnamento.

Altro obbligo tecnico è quello della *resilienza* dei sistemi IA ad alto rischio. Come è noto, gli obblighi di resilienza dei sistemi tecnologici sono stati introdotti da pochi anni nei testi normativi che disciplinano in senso lato le tecnologie (si pensi anche all'articolo 32 del GDPR che obbliga titolari e responsabili del trattamento a adottare come specifica *misura di sicurezza* nel trattamento dei dati personali "*sistemi e servizi di trattamento*" di cui deve essere garantita "*su base permanente la riservatezza, l'integrità, la disponibilità e la resilienza*").

La *resilienza informatica* è la capacità di reagire di fronte ad un evento tecnico avverso. Può ad esempio descrivere la capacità di un sistema di resistere ad un *cyberattacco* o ad un evento catastrofico (*disaster*), anche se un sistema tecnologico dovrebbe essere resiliente anche rispetto a molte altre tipologie di avversità (a maggior ragione agli eventi avversi che possono essere generati nel contesto operativo di sistemi IA autonomi, eventi che difatti devono essere ragionevolmente previsti quanto a possibili effetti dannosi su salute, ambiente o diritti fondamentali).

La resilienza dei sistemi IA ad alto rischio è declinata dal legislatore come capacità di reagire ad errori, *failure* nel perseguimento dello "scopo previsto" o a conflitti che si possono verificare nel contesto operativo del sistema IA oppure nell'ambiente in cui il sistema opera (si pensi al caso di auto a guida autonoma e ad errate rilevazioni di sensori, con rilevazione di informazioni ambientali errate). In particolare, gli errori rispetto ai quali il sistema deve essere a monte resiliente sono quelli generati dalla sua interazione con le persone o con altri sistemi.

I sistemi IA ad alto rischio devono poi essere resilienti rispetto ai tentativi di terzi non autorizzati di alterarne l'utilizzo o le prestazioni sfruttando vulnerabilità. In che modo tali possibili attacchi possono essere condotti (e devono essere respinti da un sistema IA *resiliente* adottando misure adeguate per prevenirli e monitorare costantemente i tentativi di attacco)? Possibili attacchi sono ad esempio volti a manipolare, inquinare e alterare i *dataset* di addestramento (questo attacco si chiama "*data poisoning*" o "*inquinamento dei dati*"). Altro attacco è basato sui cc.dd. *adversarial examples,* input specializzati creati con lo scopo di confondere una rete neurale, con conseguente errata classificazione di un dato *input*.

Quanto al requisito della *robustezza,* (che in realtà riguarda non tanto il sistema IA quanto i suoi software o algoritmi, essendo la *robustezza* da intendersi come la capacità di software/algoritmi di comportarsi in modo ragionevole in situazioni impreviste, non contemplate dalle specifiche), il Regolamento IA indica tecnicamente come acquisirla: mediante soluzioni tecniche di *ridondanza* (secondo l'ingegneria del software, la *ridondanza* è uno strumento fondamentale per l'ottenimento dell'affidabilità del sistema dal momento che, qualora una parte del sistema non dovesse funzionare correttamente, è necessario vi sia una parte di riserva predisposta a svolgere la medesima funzione) o mediante la previsione dei cosiddetti *fail-safe plans* (un *fail-safe plan* è una pratica di progettazione che in caso di malfunzionamento o guasto o errore del sistema, consente di ridurre al minimo o addirittura rendere nulli i danni ad altre apparecchiature, all'ambiente o alle persone).

Di particolare interesse è poi la previsione normativa dedicata a quei sistemi IA ad alto rischio "*che continuano ad imparare e ad apprendere dopo la loro immissione sul mercato o la loro messa in uso*": difatti – in tali ipotesi – la misura di sicurezza specifica in sede di sviluppo del sistema è di assicurare che possibili *outputs* discriminatori che possano derivare dal fatto che il sistema impiega come *input* per future operazioni risultati della sua precedente operatività o del suo apprendimento (una volta messo in uso) siano gestititi appropriatamente con preventive misure specifiche di mitigazione. Gli sviluppatori devono cioè prevedere ciò che non è facilmente prevedibile: come apprenderà autonomamente un sistema, adottando contromisure specifiche. Il che – come si intuisce – cambia profondamente i paradigmi della sicurezza generale dei prodotti.

§ 5. La supervisione umana dei sistemi IA ad alto rischio. L'approccio antropocentrico all'Intelligenza Artificiale

Il Legislatore europeo ha voluto specificatamente costruire il Regolamento IA fondandolo sul c.d. *antropocentrismo dell'Intelligenza Artificiale* come presidio e baluardo dei diritti e delle libertà fondamentali degli individui/utenti messi in pericolo dagli specifici rischi dei sistemi IA (es: opacità algoritmica, complessità, dipendenza dai dati, apprendimento e comportamento sempre più autonomi, etc).

Il *Considerando* n. 48 del Regolamento IA evidenzia che i sistemi di intelligenza artificiale ad alto rischio dovrebbero essere progettati e sviluppati in modo che vi siano sempre persone fisiche che possono vigilare sul loro funzionamento. A tal fine, un'adeguata supervisione umana rientra tra le misure che dovrebbero essere identificate dal fornitore del sistema prima della sua immissione sul mercato o messa in servizio. In particolare, ed ove opportuno, tali misure dovrebbero garantire che il sistema sia soggetto a vincoli progettuali, funzionali e operativi integrati nel sistema (*in-built*) e che ilo sistema medesimo non possa bypassare, ad esempio mediante sovrascrittura. Corollario di tale opzione tecnica è l'obbligo progettuale di prevedere che il sistema IA ad alto rischio risponda sempre ad un essere umano, il quale – nello svolgimento dei compiti di supervisione – deve possedere specifici requisiti di competenza, esperienza e autorità.

La *sorveglianza umana* è una parte fondamentale del meccanismo di *risk management* dei sistemi IA ad alto rischio che i fornitori devono implementare e che abbiamo visto essere complessivamente connotato da obblighi di trasparenza, conservazione di informazioni, etc. L'obbligo di *human oversight* ruota intorno alla prescrizione principale di cui all'art. 14.1 del Regolamento IA di progettare e sviluppare i sistemi IA ad alto rischio in modo tale che questi possano *effettivamente* essere supervisionati, monitorati e controllati da esseri umani durante il periodo in cui il sistema IA è operativo e in uso.
La sorveglianza umana è immaginata come misura legislativa (e pratica) atta a costituire un presidio per prevenire o minimizzare i rischi alla salute, all'ambiente o ai diritti e alle

libertà fondamentali degli utenti quando un sistema è utilizzato in conformità allo scopo previsto (il che la dice lunga sui rischi connessi a tali sistemi, ben presenti al legislatore anche in caso di utilizzo del tutto conforme alle specifiche tecniche e operative e agli scopi d'uso), oppure quando viene utilizzato in modo difforme, ancorché ragionevolmente prevedibile. Ma la sorveglianza umana si pone anche come baluardo ultimo: l'intervento dell'essere umano quando le altre misure di gestione/contenimento del rischio - pure previste dal Regolamento IA - si rivelano inefficaci. Riecheggiano nelle previsioni normative quei rischi che in molti film fantascientifici (ma che oggi avrebbero una trama del tutto scientifica e attuale) vedono i sistemi IA acquisire la cosiddetta *singolarità* (consapevolezza di sé stessi e capacità di autodeterminazione) liberandosi dal controllo umano…

La sorveglianza umana come presidio va garantita da sviluppatori e fornitori adottando misure che - ove possibile – vanno integrate nel sistema all'atto della sua progettazione e sviluppo o devono comunque essere implementate dall'utente. Sono misure che devono consentire all'essere umano a cui è assegnato il compito di sorvegliare e monitorare il sistema IA durante la sua operatività:

a) di comprendere appieno prestazioni e limiti del sistema IA ad alto rischio, monitorandone debitamente le operazioni in modo tale da poter immediatamente rilevare e gestire/risolvere nel più breve tempo possibile segni di anomalie, disfunzioni, malfunzionamenti o prestazioni inattese (ovviamente rispetto allo "scopo previsto" di ciascun sistema IA);
b) di rimanere consapevole della possibile tendenza a fare affidamento automaticamente o a fare eccessivo affidamento sull'output prodotto da un sistema di intelligenza artificiale ad alto rischio, in particolare per i sistemi di IA ad alto rischio utilizzati per fornire informazioni o raccomandazioni per le decisioni che devono essere prese dalle persone fisiche;
c) di essere in grado di interpretare correttamente l'output del sistema di IA ad alto rischio, prendendo in considerazione in particolare le caratteristiche del sistema nonché i mezzi e i metodi di interpretazione disponibili;
d) di essere in grado di decidere, in qualsiasi situazione particolare, di non utilizzare l'output di un sistema di IA ad alto rischio o altrimenti di ignorarlo, andare oltre o invertirlo;
e) sistema;
f) di essere in grado di intervenire sul funzionamento del sistema di IA ad alto rischio o di interrompere l'operatività del sistema tramite un pulsante "*stop*" o una procedura simile.

Chi è dunque il "*sorvegliante umano*"? In base alle misure elencate, deve essere in primo luogo un soggetto dotato di elevata competenza ingegneristica e funzionale circa gli aspetti operativi del sistema IA ad alto rischio che deve consentirgli di *prevedere* o *interpretare* possibili segni di un funzionamento del sistema IA che non è quello atteso. Ma anche la competenza più sofisticata potrebbe non bastare: l'opacità dei sistemi IA e la loro

sempre crescente autonomia potrebbe ingannare l'essere umano sotto forma, ad esempio, di un disfunzionamento "strisciante" e non percepibile dal sorvegliante. Questi deve poi saper interpretare i risultati del sistema, il suo *output* (circostanza non sempre possibile stante le innumerevoli combinazioni o anche significati che è possibile associare a tali *output*) e soprattutto deve poter intervenire con il più classico dei sistemi e delle soluzioni: *spegnere il sistema*. Questa misura dello *switch off* appare a chi scrive un po' ingenua, soprattutto se rapportata a scenari non troppo futuri dove i sistemi IA potrebbero come prima cosa intervenire sugli apparati di erogazione dell'energia, impedendo al povero essere umano di spegnere o di staccare l'alimentazione al sistema IA….
Che vi sia uno specifico obbligo di prevedere lo *switch off* come misura di sicurezza nell'ambito della sorveglianza umana la dice comunque assai lunga sui rischi connessi all'Intelligenza Artificiale.

Infine, il Regolamento IA specifica che tutti i sistemi IA ad alto rischio rappresentati da apparati di identificazione o classificazione biometrica da remoto in tempo reale o *ex post* devono prevedere come misura aggiuntiva di sorveglianza umana che nessuna azione o decisione possa essere presa dall'utilizzatore sulla base della identificazione biometrica come risultante da un sistema IA, se l'identificazione medesima non è verificata e confermata da almeno *due diverse persone fisiche*. Il presidio umano previsto dal Regolamento IA per i sistemi di identificazione e classificazione biometrica pare allora riecheggiare – e rafforzare – una garanzia già vigente come diritto generale: ci si riferisce al diritto di non essere sottoposti a una decisione completamente automatizzata che produca effetti giuridici o che incida in modo analogo significativamente sulla sfera giuridica di una persona introdotto dall'articolo 22 del Regolamento 679/2016 sulla protezione dei dati personali. E anche in questo caso, lo specifico presidio (che si basa su un divieto assoluto derogabile solo in caso di consenso dell'interessato, necessità di adempiere a un contratto o di adempiere a un obbligo legale) è rappresentato dal diritto dell'interessato che dovrebbe subire gli effetti di una decisione completamente automatizzata di ottenere l'intervento umano, di esprimere la propria opinione e di contestare la decisione automatica eventualmente assunta.

§ 6. Gli obblighi orizzontali per i fornitori di sistemi di IA ad alto rischio: implementare il "sistema di gestione della qualità dei sistemi IA ad alto rischio" quale forma di accountability dell'Intelligenza Artificiale

Il Capo III del Regolamento IA (artt. 16-29) introduce una serie di obblighi orizzontali per i fornitori di sistemi di IA ad alto rischio e prevede altresì obblighi proporzionati in capo sostanzialmente a tutti gli attori della catena del valore dell'Intelligenza Artificiale: dagli importatori ai distributori, dai rappresentanti autorizzati agli utilizzatori.

I fornitori di sistemi IA ad alto rischio devono intanto rispettare i requisiti e attuare gli adempimenti che abbiamo analizzato in precedenza con riferimento alla

predisposizione di un sistema di *risk assessment* fatto di trasparenza, obblighi documentali, tracciamento dei *logs* e così via. Ma il *sistema di gestione del rischio* è per il Legislatore europeo solo una delle componenti del più ampio *sistema di gestione della qualità* i cui contenuti sono dettagliati nell'articolo 17 e che si sostanzia in una sorta di *accountability* dell'Intelligenza Artificiale, sul modello del GDPR, che ancora una volta appare fungere da paradigma di riferimento. Difatti i fornitori di sistemi IA ad alto rischio devono implementare un *sistema di gestione della qualità* inteso come insieme di misure che consentano al fornitore di dimostrare e documentare la complessiva conformità a quanto previsto dal Regolamento IA. Tale sistema deve essere proporzionato alle dimensioni della realtà organizzativa del fornitore, dunque non vi sono requisiti dimensionali specifici da cui discende l'obbligo (come accade ad esempio nel GDPR con alcune previsioni legate a specifiche dimensioni del Titolare, si pensi ai 250 dipendenti come presupposto dell'obbligo di redazione e tenuta del Registro delle attività del trattamento): qualsiasi fornitore, indipendentemente dalle dimensioni organizzative (ma proporzionatamente a queste) deve implementare e gestire il *sistema di gestione della qualità* per i sistemi IA ad alti rischio.

Rispetto al GDPR, l'*accountability*/documentabilità/responsabilizzazione appare dettagliata in maniera ancor più precisa, poiché il fornitore deve dotarsi di un "*sistema che deve essere documentato in modo sistematico e ordinato sotto forma di politiche, procedure scritte e istruzioni*" e deve includere *almeno* quanto segue (dunque da assumersi quale nucleo meramente minimo degli obblighi):

1. una vera e propria *strategia per la conformità normativa*, compresa la conformità alle procedure di valutazione e la conformità delle procedure per la gestione delle modifiche ai sistemi IA ad alto rischio;
2. tecniche, procedure e azioni sistematiche che devono essere impiegate nelle fasi di progettazione, controllo della progettazione e verifica della progettazione dei sistemi IA ad alto rischio;
3. tecniche, procedure e azioni sistematiche che devono essere impiegate nelle fasi di sviluppo, controllo qualità e garanzia di qualità dei sistemi IA ad alto rischio;
4. procedure di esame, test e validazione da attuare prima, durante e dopo lo sviluppo dei sistemi IA ad alto rischio, ivi inclusa la previsione della specifica periodicità con la quale tali azioni devono essere ripetute;
5. specifiche tecniche, inclusi standard, da applicarsi in assenza (o in caso di inapplicabilità o non completa applicabilità) dei rilevanti standard armonizzati a livello UE;
6. i mezzi per garantire la conformità dei sistemi IA ad alto rischio al sistema di risk assessment previsto al Capo II del Regolamento IA;
7. sistemi e procedure di *data management*, incluse le fasi di raccolta dei dati, analisi dei dati, etichettatura dei dati, conservazione dei dati, filtraggio dei dati, *data mining*, aggregazione dei dati, *data retention* e qualsiasi altra operazione riguardante i dati che sia eseguita prima e allo scopo della immissione sul mercato o della messa in servizio di un sistema IA ad alto rischio (probabilmente è questo il punto di più stretta interrelazione tra

GDPR e Regolamento IA, nella prospettiva degli obblighi per i fornitori/Titolari del trattamento);

8. la strutturazione, implementazione e il mantenimento di un sistema di monitoraggio dei sistemi IA ad alto rischio dopo la loro immissione sul mercato ai sensi di quanto previsto dall'articolo 61 del Regolamento (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 di cui ai sistemi di *risk assessment* e *quality assessment*);

9. procedure per censire incidenti gravi o malfunzionamenti gravi dei sistemi IA ad alto rischio ai sensi di quanto prescritto dall'articolo 62 del Regolamento IA. L'articolo 62 del Regolamento IA prescrive difatti ai fornitori di sistemi IA ad alto rischio di notificare qualsiasi incidente o malfunzionamento grave che costituisca violazione delle leggi UE a tutela dei diritti e delle libertà fondamentali alle autorità nazionali di sorveglianza dei mercati dello Stato Membro in cui si è verificato l'incidente o il malfunzionamento grave. Ancora una volta è individuabile un paradigma del GDPR: quello della notifica della *data breach*, qui declinato con l'obbligo di *notifica immediata*, nel senso che va effettuata subito dopo che il fornitore ha accertato (o ragionevolmente ravvisa) un collegamento causale tra il sistema IA e l'incidente o il malfunzionamento, oppure entro al massimo 15 giorni dal momento in cui il fornitore diviene consapevole dell'incidente o del malfunzionamento del sistema IA. A loro volta le autorità nazionali di sorveglianza dei mercati – ricevuta la notifica – dovranno informare le autorità pubbliche o agenzie dello Stato Membro incaricate a livello nazionale dell'*enforcement* di leggi e regolamenti a tutela dei diritti e delle libertà fondamentali, in base a specifiche Linee Guida che la Commissione UE redigerà appositamente entro i 12 mesi successivi all'entrata in vigore del Regolamento IA, per disciplinare tale procedura);

10. un sistema per la gestione delle comunicazioni con: (i) le autorità nazionali competenti nello Stato Membro, ivi incluse procedure specifiche che consentano o facilitino l'accesso ai dati da parte di tali autorità; (ii) gli enti di notifica; (iii) altri operatori, clienti o altri soggetti interessati;

11. sistemi e procedure per la tenuta delle registrazioni e di tutta la documentazione, anche informativa, rilevante;

12. un sistema che chiarisca il quadro complessivo delle responsabilità di chiunque sia chiamato a giocare un ruolo nell'ambito del sistema di *gestione della qualità* qui descritto che definisca e distingua ruoli e funzioni, dalla direzione a tutti quelli chiamati ad operarvi (e anche questa previsione riecheggia le "persone autorizzate" degli articoli 29 e 32 del GDPR).

§ 6.1 Segue. Gli altri obblighi dei fornitori di sistemi di IA ad alto rischio: come attuare in pratica le procedure di valutazione della conformità dei sistemi IA

Oltre alla messa in atto di un sistema di *quality assessment* che include anche l'importante sistema di *risk assessment*, come visto, i fornitori di sistemi IA ad alto rischio sono chiamati ad ulteriori adempimenti, come ad esempio quello inerente la predisposizione di un sistema di tracciamento dei *logs* generati e della documentazione tecnica del sistema IA, in base alle indicazioni contenute nell'Allegato IV al Regolamento (ne abbiamo parlato nella Parte Seconda, a cui si rinvia).

Di particolare interesse pratico è poi l'obbligo per i fornitori di garantire che i sistemi IA ad alto rischio siano sottoposti a procedure di valutazione della loro conformità prima della loro immissione sul mercato o della loro messa in servizio: in base a quanto previsto dall'articolo 19 del Regolamento IA, solo dopo l'esito positivo della procedura di *conformity assessment* il fornitore potrà emettere – sotto la sua esclusiva responsabilità – la cosiddetta *dichiarazione di conformità* (con i contenuti previsti dall'articolo 48 e dall'Allegato V del Regolamento, che dovrà poi essere conservata per dieci anni dopo che il sistema IA ad alto rischio è stato messo sul mercato o è stato attivato) e potrà apporre il marchio CE (in modo visibile, leggibile e indelebile e inclusivo del numero di identificazione della competente autorità di notifica responsabile della procedura di valutazione di conformità) sul sistema IA, sul suo *packaging* o sulla documentazione di accompagnamento, a seconda dei casi, secondo quanto previsto dall'articolo 49.

Sottoporre il sistema IA ad alto rischio a una procedura di *conformity assessment* non è sempre obbligatorio: infatti – in base a quanto previsto dall'articolo 40 del Regolamento IA – se il sistema IA ad alto rischio è conforme a standard armonizzati (o a loro parti) come pubblicati sulla Gazzetta Ufficiale dell'Unione Europea (inclusi gli schemi di conformità sulla c*ybersecurity* ai sensi dello specifico Regolamento 2019/881), allora tale sistema si presumerà essere conforme a quanto richiesto dal Capo II (sistema di *risk assessment*, sistema di *quality management*, trasparenza, sicurezza, *data management*, etc). Resta sempre salva la possibilità che la Commissione UE ritenga che certi standard armonizzati – pur esistenti – siano comunque insufficienti rispetto alla tutela della salute, dell'ambiente o dei diritti fondamentali, potendo introdurre mediante atti delegati specifiche comuni e standard aggiuntivi (a cui i sistemi IA dovranno quindi conformarsi, cfr. art. 42 Reg. IA).

Venendo invece all'esame della procedura di conformità, l'articolo 43 del Regolamento IA prevede che per i sistemi IA ad alto rischio elencati al punto n. 1 dell'Allegato III (cioè i sistemi di identificazione biometrica da remoto delle persone fisiche) il fornitore deve attuare la procedura di valutazione della conformità secondo due diverse modalità: 1. attuare una procedura interna (laddove per dimostrare la conformità si avvalga di standard armonizzati come pubblicati nella GUCE o di specifiche comuni stabilite dalla Commissione) oppure 2. attuare una procedura esterna con il

coinvolgimento di un organismo notificato (ove non esistano o siano inapplicabili standard armonizzati o specifiche comuni).

Nel caso di valutazione interna, il fornitore dovrà procedere in base agli step elencati all'Allegato VI del Regolamento IA:

1. verificare che il *sistema di gestione della qualità* sia conforme a tutto quanto previsto all'articolo 17 (si vedano i punti da 1 a 12 che abbiamo illustrato al paragrafo 6 più sopra);
2. esaminare le informazioni contenute nella documentazione tecnica ai fini della verifica della conformità del sistema IA con i requisiti fondamentali stabiliti nell'ambito del sistema di *risk assessment*;
3. verificare che il processo di progettazione e di sviluppo del sistema IA e il suo monitoraggio successivo alla immissione sul mercato (secondo la procedura di raccolta delle informazioni e il processo che abbiamo illustrati più sopra richiamando l'articolo 61 del Reg. IA) sia coerente con la documentazione tecnica.

Nel caso di valutazione esterna, il fornitore dovrà presentare a un organismo notificato una specifica istanza di riconoscimento della conformità del sistema IA ad alto rischio nell'ambito di una complessa procedura articolata negli step elencati all'Allegato VII del Regolamento IA, che prevedono rigorosi controlli sul sistema di gestione della qualità, sulla documentazione tecnica presentata e sul sistema successivo di sorveglianza del sistema di qualità, una volta approvato (l'organismo di notifica prevede infatti *audit* periodici per la verifica che la conformità dei sistemi IA si mantenga nel tempo).

Va specificato che l'obbligo di espletare la procedura di conformità del sistema IA ad alto rischio coinvolgendo un organismo notificato esterno non si applica invece a tutti gli altri sistemi IA ad alto rischio elencati ai punti da 2 a 8 dell'Allegato III del Regolamento IA, per i quali sarà sufficiente espletare la valutazione interna con gli step sopra illustrati previsti dall'Allegato VI. Resta tuttavia salva la possibilità per la Commissione UE di far sottoporre anche tali sistemi alla procedura che prevede il coinvolgimento di un organismo notificato se - mediante atti delegati – dovesse così decidere in futuro.

L'articolo 43.4 del Regolamento IA prevede inoltre che i sistemi IA ad alto rischio sono sottoposti a una nuova procedura di valutazione della conformità ogni qual volta vengano sostanzialmente modificati, indipendentemente dal fatto che il sistema modificato sia destinato ad essere ulteriormente distribuito o continui ad essere utilizzato dal medesimo utente. Interessante notare che non si considerano "modifiche sostanziali" che obbligano ad una nuova procedura di valutazione della conformità quelle modifiche che possono intervenire nell'ambito di taluni sistemi IA che continuano ad imparare anche dopo la loro immissione sul mercato, ove però lo sviluppatore abbia predeterminato (specificandole nella documentazione tecnica presentata) le possibili modifiche all'atto della originaria procedura di valutazione della conformità.

Una volta rilasciato, il certificato di conformità da parte dell'organismo notificato, questo avrà la validità (prorogabile) di volta in volta stabilita, ma che non può essere comunque superiore a 5 anni. Il certificato può essere sospeso, ritirato o limitato con decisione motivata dell'organismo che lo ha emesso laddove questo ravvisi che non vi è più rispondenza del sistema IA ai requisiti previsti e che non vi sono state azioni correttive da parte del fornitore.

L'articolo 47 del Reg. IA prevede poi una procedura derogatoria eccezionale in base alla quale una autorità nazionale di sorveglianza del mercato può autorizzare (notificandolo alla Commissione e agli altri Stati Membri, affinché possano eventualmente opporsi nei 15 giorni successivi) un sistema IA ad alto rischio ad essere immesso sul mercato anche in assenza o nelle more delle procedure sopra analizzate di valutazione della conformità, laddove vi siano eccezionali e giustificate ragioni di pubblica sicurezza, difesa della salute, dell'ambiente o di infrastrutture industriali critiche. L'autorizzazione in deroga avrà una validità temporanea (nell'attesa che si concludano le procedure di *conformity assessment* ordinarie) e deve comunque basarsi su motivate ragioni e sul convincimento da parte dell'autorità di vigilanza che la emette della conformità del sistema IA ai requisiti previsti.

Va infine ricordato che il fornitore (o il suo rappresentante autorizzato) deve:

- prima della immissione sul mercato, registrare tutti i sistemi IA ad alto rischio di cui all'Allegato III nel *database europeo dei sistemi IA stand-alone* che la Commissione istituirà e gestirà di intesa con gli Stati Membri e che sarà accessibile al pubblico. Il database conterrà le seguenti informazioni: nome, recapito geografico e dettagli per contattare il fornitore o del suo rappresentante autorizzato; il nome commerciale del sistema IA e qualsiasi ulteriore informazione chiara che consenta la sua identificazione e tracciabilità; descrizione specifica dello "scopo previsto" del sistema IA; status del sistema – ad esempio: sul mercato, operativo, ritirato, richiamato, etc; tipologia, numero e validità del certificato emesso dall'organismo notificato (che va reso disponibile in copia scansionata nel database) e dati identificativi dell'organismo notificato che lo ha emesso; Stati Membri nei quali il sistema IA è stato immesso sul mercato, reso operativo o in altro modo è stato reso disponibile; copia della dichiarazione di conformità ai sensi dell'art. 48; le istruzioni d'uso in formato elettronico, ad eccezione per quei sistemi IA ad uso pubblico, come quelli per il controllo delle frontiere; un indirizzo web ove reperire ulteriori informazioni (cfr. art. 60 e Allegato VIII del reg. IA);
- conservare per dieci anni tutta la documentazione: dalla documentazione tecnica , a quella inerente il sistema di gestione della qualità, fino alla dichiarazione di conformità ex art 48 del Regolamento.