ASSOCIAZIONE ITALIANA PER L'INTELLIGENZA ARTIFICIALE

c/o CNR-ISTC
Via S. Martino della Battaglia 44, 00185, Roma
C.F. 96111060586 – P.IVA 14789761005

email: segretario@aixia.it
Web: http://www.aixia.it

AI REGULATION AND 2021 COORDINATED PLAN ON AI

FEEDBACK DEL CONSIGLIO DIRETTIVO DELL'AIXIA

Nel ringraziare la IX Commissione Trasporti e la X Commissione Attività produttive della Camera dei deputati per il coinvolgimento nel processo di audizione finalizzato ad acquisire utili elementi di conoscenza e di valutazione sull'atto dell'Unione europea COM (2021) 206 final, recante "Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'intelligenza artificiale", il Consiglio Direttivo dell'Associazione Italiana per l'Intelligenza Artificiale ritiene di dover evidenziare e stimolare una più approfondita riflessione su tre specifici articoli che saranno analizzati nel seguito di questa breve nota.

ART. 3 - DEFINITIONS

Article 3 Definitions

(1) 'artificial intelligence system' (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with.

ANNEX I ARTIFICIAL INTELLIGENCE TECHNIQUES AND APPROACHES referred to in Article 3, point 1

(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning;
(b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems;
(c) Statistical approaches, Bayesian estimation, search and optimization methods.

Il primo punto di riflessione attiene all'articolo 3 sopra riportato, che evidentemente riveste un ruolo essenziale nella proposta della Commissione Europea, in quanto esso aspira a definire con precisione ciò che deve esser considerato come sistema di AI e, di conseguenza, ciò che è da ritenersi assoggettato alla regolamentazione stessa. Non a caso, tale articolo è da tempo al centro del dibattito pubblico e ha, invero, attirato già non poche critiche e perplessità. Da più parti, infatti, è stato ribadito che la definizione di sistema di AI fornita dalla Commissione Europea presenta alcune marcate criticità, in particolar modo legate all'ampiezza della definizione stessa che sembra includere anche tecnologie che, oggigiorno, non sono considerate dalla comunità scientifica come strettamente collegate all'AI, quali ad esempio quelle dei metodi di ricerca e ottimizzazione (*search and optimization methods*). Nella sostanza, con la lista di tecnologie elencate all'Annex I, la quasi totalità delle attuali applicazioni industriali finirebbero per essere incluse nella definizione di

sistema di AI. L'Associazione Italiana per l'Intelligenza Artificiale fa dunque proprie queste osservazioni, raccomandando di definire in maniera più precisa i contorni delle tecnologie di interesse, onde evitare di abbracciare sistemi e approcci utilizzati già da decenni in numerosi contesti applicativi ben lontani da ciò che oggi si intende come intelligenza artificiale.

E, tuttavia, nell'analizzare l'articolo 3, l'Associazione ritiene altresì fondamentale sottolineare che la sua problematicità non può risolversi in una semplice modifica all'elenco proposto all'Annex I, magari con l'espunzione di qualche voce relativa a una specifica tecnologia. Ciò che preoccupa l'Associazione – al pari di altri importanti attori dell'ecosistema Europeo dell'AI che hanno già preso specifica posizione su questo tema – è, infatti, l'impianto stesso della definizione che risulta ben lontano da ciò che la comunità internazionale largamente intende come intelligenza artificiale. In primo luogo, l'Associazione vuole rimarcare che l'Intelligenza Artificiale è una *disciplina* radicata nell'informatica, e pertanto essa non deve essere confusa con le tecnologie e applicazioni che essa genera: la definizione dovrebbe risolvere questa ambiguità di fondo, nell'ottica di una visione prospettiva in cui lo studio sugli aspetti fondazionali sia portato avanti in modo libero e non condizionato, mentre le applicazioni siano indirizzate verso obiettivi benefici per l'umanità e il pianeta evitando effetti distopici. Anche risolvendo tale ambiguità di fondo, tuttavia, la scelta di discriminare ciò che possa essere incluso nel novero dell'AI sulla base delle specifiche tecniche e tecnologie adottate per lo sviluppo dei sistemi risulterebbe comunque piuttosto atipica e problematica, in quanto pone una eccessiva enfasi sul "mezzo" utilizzato dai sistemi AI, ignorando del tutto le loro finalità e le loro peculiarità.

Ciò che distingue profondamente i sistemi AI dagli altri approcci informatici e algoritmici è l'idea che essi siano sviluppati con l'obiettivo di replicare funzioni, abilità e comportamenti che sono propri delle persone e della loro intelligenza. Questa idea è uno dei principali lasciti di Alan Turing che nel suo lavoro *Computer Machinery and Intelligence* del 1950, affrontando la domanda "Can machine think?" e scontrandosi inevitabilmente con la difficoltà di caratterizzare e definire i contorni dell'intelligenza umana prima ancora che di quella artificiale, propone un approccio molto pragmatico basato sul principio di "imitazione": macchine intelligenti sono macchine in grado di agire simulando il modo con cui una persona agirebbe. Da qui prende corpo il famoso gioco dell'imitazione: una persona interagisce con un agente - persona o macchina che sia - senza vederlo, trattando cioè l'agente come una black-box e focalizzando pertanto il proprio interesse solo a ciò che esso è in grado di produrre operando in un certo ambiente. Ecco, dunque, che il punto essenziale della definizione dell'AI non è tanto su quale sia la specifica tecnologia che tali macchine utilizzano per esibire un comportamento intelligence; non si tratta di capire se un certo algoritmo sia o meno utilizzato in un sistema; tali questioni rientrano pienamente nel "velo dell'ignoranza" che definisce il gioco dell'imitazione. Turing, in sostanza, ci ha già ammonito sull'importanza di caratterizzare l'AI in termini di come essa si rivela alle persone piuttosto che in termini di una qualche tecnologia.

In aggiunta ai punti sopra delineati, giova altresì rimarcare alcune argomentazioni di natura più tecnica che dovrebbero indurre a maggiore prudenza rispetto all'idea di impostare la definizione di sistema AI esclusivamente sulla base di una lista di tecnologie:

- In primo luogo, la tecnologia è in continua evoluzione e un qualsivoglia elenco non potrà essere esaustivo, specie nell'ottica del medio-lungo periodo. Peraltro, già nella sua attuale formulazione, numerose tecnologie oggi disponibili e riconducibili alla disciplina dell'AI sono ignorate nella definizione di cui all'articolo 3. A puro titolo esemplificativo, basti osservare che, al momento, sono del tutto esclusi dalla definizione i sistemi hardware e le metodologie software che si basano sull'interazione di agenti intelligenti.
- In aggiunta, vi è poi da considerare che, di fronte a una regolamentazione che definisca con precisione una lista di tecnologie "attenzionate", è tutt'altro remota l'ipotesi che gli attori dell'ecosistema inizino a sviluppare sistemi in cui tali tecnologie vengono "nascoste" o simulate con altre tecnologie che al momento non sono riportate nella regolamentazione. Nel mondo informatico è ben noto quanto sia agevole "mascherare" un componente software o una certa tecnologia; e certamente si tratta di dispiegare approcci e soluzioni ben alla portata dei maggiori player internazionali. Ecco, quindi, che insistere su una elencazione di tecnologie potrebbe finire per avere come paradossale effetto avverso quello di rendere del tutto vana la definizione stessa, portando a uno scenario che avvantaggia chi avrà le competenze tecniche e la massa critica per aggirare maliziosamente, ma del tutto lecitamente le previsioni regolamentari.

Tanto premesso, la raccomandazione dell'Associazione Italiana per l'Intelligenza Artificiale è di riformulare l'attuale definizione dei sistemi AI, enfatizzando le peculiarità generali che caratterizzano tali sistemi piuttosto che le tecnologie con le quali essi sono implementati. L'attuale lista potrebbe comunque essere riportata a titolo esemplificativo, espungendo le tecnologie chiaramente non AI e ribadendo che la lista non è esaustiva in quanto a essa sono da ricondursi tutte le forme di implementazione software e hardware che replicano il comportamento umano e abilitano nuovi livelli di automazione e delegazione.

ART. 10 – DATA AND DATA GOVERNANCE

> **Article 10 Data and data governance**
> 3. Training, validation and testing data sets shall be relevant, representative, free of errors and complete. They shall have the appropriate statistical properties, including, where applicable, as regards the persons or groups of persons on which the high-risk AI system is intended to be used. These characteristics of the data sets may be met at the level of individual data sets or a combination thereof.

Il secondo punto che l'Associazione Italiana per l'Intelligenza Artificiale intende evidenziare è inerente all'articolo 10. Esso illustra alcuni requisiti che devono essere rispettati dai sistemi di AI

che vengono classificati come ad "alto rischio". In effetti, poiché la definizione di "alto rischio", dettagliata nell'Annex III, abbraccia un ventaglio estremamente ampio di possibili contesti applicativi, i requisiti definiti in tale articolo acquistano una particolare rilevanza. Essi potrebbero, infatti, configurarsi come uno dei principali ostacoli tecnici all'utilizzo effettivo e alla diffusione delle tecniche di intelligenza artificiale in Europa. A riguardo, l'Associazione intende in primo luogo evidenziare che, in assenza di un'accurata e profonda analisi, risulta difficile definire una classificazione dei rischi e suggerire una scala di valori degli stessi: alcune attività che giudichiamo a basso rischio potrebbero portare a effetti distopici nel medio-lungo termine, come è successo con molte tecnologie nel recente passato; altre che giudichiamo invece "rischiose" potrebbero invece risultare in futuro fondamentali per abilitare un modello di sviluppo pienamente sostenibile.

Ma anche volendo sorvolare sulla definizione di sistema ad "alto rischio", prima ancora che entrino in campo valutazioni di carattere economico legate ai costi che dovrebbero sobbarcarsi le imprese e le pubbliche amministrazioni per soddisfare i requisiti proposti dalla Commissione Europea per tali sistemi e prima ancora di avviare discussioni circa l'impatto che tali costi avrebbero nello sviluppo della competitività del nostro paese e dell'intera Europa, l'Associazione ritiene necessario essenziale ragionare più profondamente sul senso stesso di questi requisiti.

In effetti, quando la proposta di regolamentazione entra nel merito dell'utilizzo dei dati per lo sviluppo di sistemi di intelligenza artificiale, essa sembra delineare requisiti che hanno una valenza sostanzialmente teorica e speculativa: vari sono gli aggettivi utilizzati nell'articolo 10, e ciascuno di essi presenta un più o meno marcato grado di criticità nella sua applicazione pratica. In primo luogo, è opportuno sottolineare come un dataset non possa mai essere ritenuto *completo*, almeno non nell'accezione più propria ed estensiva che ha questo termine. Cosa significa, ad esempio, "completezza" in un dataset che registra valori clinici di pazienti? Forse ci si aspetta che il dataset includa tutti i possibili pazienti del mondo? Si parla poi di dataset *privi di errori,* di fatto ignorando che il concetto di errore è intrinseco alle misurazioni sperimentali: avrebbe senso parlare di grado di incertezza o di grado di approssimazione, ricordando che nella pratica nessuna misura (e dunque nessuna registrazione informatica della stessa) potrà essere avulsa da un possibile errore. E ancora: chi e come verifica la *rappresentatività* dei dati e la loro *rilevanza*? Tutti questi interrogativi restano e, inevitabilmente, resteranno senza risposta. La Comunità Europea, nell'interesse di garantire la più elevata possibile qualità del dato, ha finito infatti per assolutizzare i requisiti di qualità sui dati fornendo una chiave interpretativa di impossibile attuazione.

In aggiunta alle osservazioni tecniche sin qui delineate, l'articolo 10 è altresì problematico anche in termini concettuali in quanto sembra confliggere con l'attuale impostazione del GPDR che spinge, invece, per la rimozione delle informazioni sensibili nel trattamento automatizzato, e asserisce il diritto di ciascuna persona a non concedere l'autorizzazione per l'utilizzo dei propri dati. In sostanza, la proposta di regolamento si pone come obiettivo quello di lavorare con dati di elevata (teoricamente perfetta) qualità, mentre il GPDR definisce un contesto concreto che limita la qualità

dei dati a disposizione in considerazione delle esigenze di privacy. Tale dualismo rischia evidentemente di alimentare profonde incertezze, e rischia di creare un quadro normativo confuso che potrebbe arrecare nocumento al sistema produttivo e imprenditoriale. Di certo, esso non si muove nella direzione di facilitare l'adozione di soluzioni di AI basate sull'elaborazione automatizzata di dati, attraverso ad esempio tecniche di machine e deep learning.

Tanto premesso, la raccomandazione dell'Associazione Italiana per l'Intelligenza Artificiale è di riformulare l'attuale impianto dell'articolo 10, in particolare al comma 3, riducendo l'enfasi sui requisiti dei dati e spostandola, nell'ottica di aderire meglio ai principi del GPDR, sulla definizione delle proprietà – quali fairness e non discriminazione – che i sistemi di AI basati sul trattamento dei dati dovrebbero garantire.

ART. 17 – QUALITY MANAGEMENT SYSTEMS

Article 17 Quality management systems
1. Providers of high-risk AI systems shall put a quality management system in place that ensures compliance with this Regulation. […]

L'ultimo articolo su cui l'Associazione Italiana per l'Intelligenza Artificiale ritiene rilevante soffermare la propria attenzione è l'articolo 17, che assieme agli altri articoli dal Capitolo 3, delinea le responsabilità degli attori che realizzano sistemi di AI. In particolare, si vuole qui osservare che l'attuale impostazione del regolamento assegna ai *provider* un ruolo centrale nella gestione dei rischi dei sistemi di AI, probabilmente iper-semplificando il modello dello sviluppo e dell'utilizzo dei sistemi di AI. L'articolo 17 chiarisce, infatti, che tutta la complessità dell'attuale regolamentazione è in capo ai fornitori, che sono dunque tenuti a certificare le caratteristiche dei propri prodotti di AI e il rispetto dei requisiti definiti dalla Comunità Europea.

Tale approccio, evidentemente, ignora la complessità della catena del valore dell'AI interpretando queste tecnologie in un'ottica tradizionale di rapporto fornitore-consumatore. Nella pratica, invece, le soluzioni di AI solo raramente vengono acquisite già specificatamente verticalizzate e adattate su uno specifico contesto applicativo; più frequente è invece la definizione di framework e piattaforme applicative multisettoriali o comunque di ampio spettro che vengono verticalizzate e adattate a specifici domini con il ruolo attivo degli utilizzatori. Le piattaforme di AI hanno – peraltro con ovvi vantaggi di economia di scala – elevati gradi di flessibilità e offrono agli utilizzatori potenti leve di intervento. Ebbene, di queste dinamiche l'attuale regolamentazione non riporta alcuna traccia. In particolare, la proposta di regolamentazione non tiene conto del fatto che molti obblighi possono essere assunti solo dall'entità che gestisce il sistema di IA e il suo utilizzo concreto, in sostanza proprio dall'utente finale. Infatti, anche se un fornitore di un sistema di IA adottasse tutte le precauzioni dovute già in fase di design, non potrebbe comunque escludere alcuni

rischi che sono invece imputabili alla modalità di utilizzo e al contesto di esecuzione che l'utilizzatore definirà per quel prodotto. Banalizzando l'argomentazione nell'ottica di essere quanto più possibili espliciti, è come se addossassimo tutta la responsabilità per l'uso di un coltello da cucina al produttore del coltello stesso, anche nel caso in cui tale coltello venisse poi utilizzato non per i suoi "fini primari", ma per portare offesa a un altro essere vivente. Qualsiasi tecnologia sufficientemente "potente" porta con sé dei rischi; tali rischi devono essere analizzati, ponderati e mitigati attraverso un insieme di azioni implementative e culturali che riguardano tutta la catena di produzione e utilizzo, non solo una sua specifica parte.

L'effetto più evidente dell'attuale impianto normativo, invece, è quello di disincentivare i provider, essendo chiamati ad assumersi responsabilità su cui le proprie leve di intervento potrebbero essere parziali e limitate. I provider finirebbero, precauzionalmente, per classificare i propri prodotti come ad alto rischio e lo sviluppo di soluzioni realmente innovative e competitive su scala internazionale finirebbe per essere disincentivato. Questo è, tra tutti, l'aspetto più delicato: il clima di incertezza sulle responsabilità e il timore di possibili ripercussioni legali rischia di agire da freno allo sviluppo dell'intelligenza artificiale in Europa.

Tanto premesso, la raccomandazione dell'Associazione Italiana per l'Intelligenza Artificiale è di riformulare l'articolo 17 (e più in generale l'impianto del Capitolo 3), identificando meglio gli attori e le loro responsabilità, anche in relazione al ruolo da essi ricoperto nell'implementazione di adattamenti e personalizzazioni che modifichino in modo sostanziale le modalità di erogazione o finanche le finalità stesse del sistema per come concepite dai provider.