



**Istituto per le Applicazioni
del Calcolo**

Memoria Tecnica sul Disegno di Legge N 2316

(Disposizioni e deleghe al Governo in materia di intelligenza artificiale)

1. Premessa metodologica

La presente memoria offre un commento tecnico-scientifico al Disegno di Legge N 2316, che si propone di normare il settore dell'intelligenza artificiale in Italia. L'analisi adotta un approccio critico-costruttivo, valutando il testo dalla prospettiva dell'ingegneria, dell'informatica e delle scienze dei dati, con l'obiettivo di fornire un contributo basato sulla fattibilità tecnica, la precisione concettuale e l'impatto atteso sull'innovazione e sulla ricerca.

L'analisi si focalizzerà su:

Criticità tecniche e scientifiche: Verranno segnalate le imprecisioni, le ambiguità o le omissioni nelle definizioni, nei requisiti e nelle disposizioni che, dal punto di vista tecnico, potrebbero generare incertezze applicative, limitare l'innovazione, risultare difficilmente implementabili o non allineate con lo stato dell'arte e l'evoluzione tecnologica (es. definizione di modelli, requisiti di spiegabilità, gestione dei dati). Per ciascuna criticità, saranno proposte alternative o integrazioni basate su concetti scientifici e pratiche consolidate.

Raccomandazioni strategiche: Oltre alle osservazioni puntuali sul testo, saranno formulate raccomandazioni di carattere più ampio, orientate a promuovere un ecosistema dell'IA sano e innovativo in Italia, toccando temi come la necessità di competenze diffuse (alfabetizzazione sull'IA), la promozione di approcci tecnicamente efficienti e sostenibili, e la tutela della sovranità tecnologica intesa come capacità di sviluppo e gestione autonoma delle tecnologie chiave.

L'obiettivo è contribuire a definire un quadro normativo che sia non solo giuridicamente solido, ma anche tecnicamente informato e lungimirante, capace di governare efficacemente lo sviluppo e l'impiego dell'IA nel rispetto dei principi etici e dei diritti, senza ostacolare il progresso scientifico e tecnologico.



2. Analisi per Aree Critiche

Questa sezione analizza specifiche aree del Disegno di Legge, evidenziando criticità e suggerendo modifiche dalla prospettiva tecnico-scientifica.

A. Definizioni e Ambito di Applicazione (Art. 2, Art. 3)

Una normativa efficace sull'IA richiede definizioni precise che riflettano l'attuale panorama tecnologico e siano armonizzate con il quadro normativo internazionale, in particolare l'AI Act europeo. Si riscontrano alcune criticità tecniche legate alla terminologia e all'ambito di applicazione.

Art. 2 comma 1 lettera c e Art.3 comma c - Definizione di *modelli di intelligenza artificiale*. La definizione, citando l'articolo 3 punto 63 del c.d. AI Act rischia di non essere pienamente allineata con la terminologia e la portata del concetto di "general purpose AI model". Una possibilità sarebbe aggiungere *generale* come da definizione del c.d. AI Act. Una possibilità alternativa sarebbe anche cambiare all' articolo 3 comma 5 *modelli di intelligenza artificiale* per finalità generali con *modelli di intelligenza artificiale generali*.

Art. 3 commi 1, comma 2 - L'attuale formulazione degli articoli 3, commi 1 e 2, nel fare riferimento a "sistemi e modelli generali", crea un potenziale equivoco tecnico. Sembra implicare che le disposizioni si applichino prevalentemente o esclusivamente ai "modelli di IA generale" definiti all'Art. 2 comma 1 lettera c. Questo escluderebbe, o renderebbe incerta, l'applicazione della norma a sistemi e modelli di IA progettati ed addestrati per compiti specifici (come sistemi di classificazione di immagini, riconoscimento vocale per task specifici, modelli predittivi verticali). Per chiarire l'ambito e includere tecnicamente tali derivazioni, una possibile alternativa sarebbe aggiungere nell'Articolo 2 un nuovo comma che definisca i modelli *fine-tuned* o *adattati* come modelli generali successivamente addestrati per compiti specifici e/o per uno specifico problema.

Art. 8 comma 1 - Riferimento a *modelli di base* senza definizione: L'Articolo 8, comma 1, fa riferimento ai *modelli di base*. Sembrerebbe corrispondere al concetto tecnico di "foundation models" per indicare modelli base di ampia scala addestrati su dati vasti e non etichettati, che servono come fondamento per lo sviluppo di numerosi modelli generali derivati. Tuttavia, il DDL non fornisce una definizione di *modello di base* nell'Articolo 2. Questa mancanza potrebbe generare ambiguità e rendere



difficile determinare con precisione a quali entità tecniche si applichino le disposizioni dell'Art. 8, creando incertezza sia per i ricercatori che per gli sviluppatori.

Un'alternativa per ovviare a tale criticità si otterrebbe nell'Articolo 2, tramite un comma aggiuntivo, la definizione di "modelli di base" (Foundation Models) come *modelli base di ampia scala addestrati su dati vasti e non etichettati, che servono come fondamento per lo sviluppo di modelli generali.*

B. Spiegabilità e Trasparenza (Art. 3, comma 3)

L'Articolo 3, comma 3, introduce il concetto di *spiegabilità*, fondamentale dal punto di vista etico e normativo, ma che necessita di maggiore precisione tecnica per garantirne l'effettiva implementazione e verifica.

La definizione di "spiegabilità" fornita nel comma è generica e non specifica i requisiti tecnici minimi necessari per considerare un sistema di IA realmente spiegabile in contesti applicativi. In particolare, mancano riferimenti espliciti alla riproducibilità dei risultati o dei processi decisionali e all'esplorabilità tecnica del sistema. Dal punto di vista scientifico e ingegneristico, la possibilità di riprodurre una decisione presa dal sistema (dato uno stesso input) e di poter "auditare" tecnicamente il suo funzionamento interno (comprendere, ad esempio, quali caratteristiche dei dati di input hanno influenzato maggiormente l'output) è cruciale non solo per la fiducia dell'utente, ma anche per l'identificazione e la correzione di bias (errori/pregiudizi/preconcetti), il debug del sistema e la verifica del rispetto dei requisiti. Una spiegabilità priva di questi elementi tecnici rischia di rimanere un concetto astratto e difficilmente verificabile nella pratica.

Una possibile alternativa risiede nell'estensione del comma 3 dell'Articolo 3 per integrare i requisiti tecnici di riproducibilità e auditabilità nella definizione di spiegabilità. Una possibile integrazione potrebbe essere: *"La spiegabilità" include la capacità del sistema o di chi lo fornisce, di offrire informazioni comprensibili sull'origine e sul processo che ha portato a una determinata decisione o output, assicurando ove tecnicamente fattibile la tracciabilità dei dati rilevanti utilizzati, la riproducibilità del processo decisionale dato un input specifico, e prevedendo meccanismi tecnici che consentano l'auditabilità algoritmica da parte delle autorità competenti e, nei limiti del necessario e proporzionato, degli utenti interessati o di loro delegati.*



C. Privacy e Sicurezza(Art. 4, Art. 6)

La protezione e la sicurezza dei dati personali è cruciale nello sviluppo e nell'utilizzo dell'IA. Gli articoli 4 e 6 affrontano questo tema, ma alcune considerazioni tecniche relative a metodologie avanzate e alla gestione dei risultati dell'addestramento possono migliorare l'efficacia della normativa in termini di privacy.

Si rileva l'assenza di riferimenti nell'**Articolo 4** (Trattamento dei dati) all'incoraggiamento all'uso di metodologie che offrano un robusto meccanismo nativo per la minimizzazione del trasferimento dati e quindi per la protezione della privacy (ad esempio "apprendimento federato"). Si rileva la necessità di un'integrazione dell'Articolo 4, nonostante il tema venga poi affrontato nelle disposizioni di settore successive, che espliciti la protezione della privacy nell'addestramento dei sistemi di intelligenza artificiale. Un esempio potrebbe essere: *Il trattamento di dati personali per finalità di sviluppo e utilizzo di sistemi di intelligenza artificiale è lecito se rispetta i principi del Regolamento (UE) 2016/679 e del decreto legislativo 30 giugno 2003, n. 196, anche attraverso l'adozione di tecniche che minimizzino la raccolta e il trasferimento dei dati originali, incluse metodologie di apprendimento decentralizzato che permettono l'elaborazione dei dati localmente.*

L'**Articolo 6, comma 2**, focalizzandosi solo sulla localizzazione dei dati non affronta adeguatamente il rischio tecnico associato al trasferimento dei pesi o parametri di un modello di IA addestrato su dati sensibili localizzati. Sebbene i dati originali rimangano in Italia, i parametri (c.d. **pesi**) del modello derivati dall'addestramento su tali dati possono potenzialmente "memorizzare" o permettere la ricostruzione di informazioni sensibili sui dati di addestramento, specialmente in determinate condizioni o con modelli di grandi dimensioni. Il trasferimento all'estero di questi pesi, anche senza i dati originali, può quindi rappresentare un rischio per la privacy che la sola norma sulla localizzazione dati non copre tecnicamente.

Una possibile soluzione è rappresentata dall'aggiunta di un comma all'Articolo 6 per estendere il concetto di localizzazione anche ai derivati dell'addestramento su dati sensibili localizzati. Ad esempio: *I pesi, i parametri o altri derivati tecnici di un modello di intelligenza artificiale addestrato su dati sensibili, genetici o biometrici soggetti a obblighi di localizzazione ai sensi del comma 2, non possono essere trasferiti o resi accessibili al di fuori del territorio nazionale, se tale trasferimento o accesso comporta un rischio significativo di re-identificazione o inferenza sui dati sensibili sottostanti, salvo specifiche autorizzazioni basate su valutazioni tecniche del rischio e adeguate garanzie.*



D. Sovranità Tecnologica, Efficienza e Sostenibilità Ambientale (Art. 5, Art. 23)

Dalla prospettiva tecnico-scientifica, per promuovere una reale sovranità tecnologica, l'efficienza e la sostenibilità ambientale nell'ambito IA, si formulano le seguenti osservazioni e proposte relative agli Articoli 5 e 23:

Art. 5 (Principi Generali): Si rileva la mancanza di un principio esplicito che incoraggi lo sviluppo tecnico di modelli di IA computazionalmente efficienti, ottimizzati per hardware eterogeneo (non solo GPU ad alte prestazioni) e basati su dataset rappresentativi del contesto linguistico e culturale italiano. Per colmare questa lacuna, un'alternativa è rappresentata dall'introduzione di un nuovo comma nell'Articolo 5 che promuova attivamente la ricerca e lo sviluppo di tali modelli e l'uso di dati pertinenti al contesto nazionale, riconoscendone il valore per la sostenibilità, l'accessibilità e la pertinenza locale dell'IA.

Art. 23 (Investimenti e misure di sostegno): Si osserva che tra le aree di investimento prioritario non sono esplicitamente menzionate tecnologie abilitanti per l'efficienza e l'IA su dispositivi a risorse limitate, come i modelli lightweight. Tale criticità potrebbe essere affrontata aggiungendo specificamente queste aree nell'Articolo 23 per incentivare lo sviluppo tecnico in settori cruciali per il deployment diffuso e sostenibile dell'IA, in linea con gli obiettivi di efficienza, sovranità tecnologica e sostenibilità ambientale.

E. AI Literacy (Capo I, Artt. 11-15, 19, 24 comma e)

La capacità di comprendere, utilizzare e valutare criticamente i sistemi di intelligenza artificiale è un requisito tecnico e sociale fondamentale per lo sviluppo e l'impiego sicuro e responsabile dell'IA. L'**AI Literacy (alfabetizzazione sull'IA)** è riconosciuta come un pilastro normativo cruciale a livello europeo (cfr. Articolo 4 del c.d. AI Act). Per garantire che la legge promuova attivamente questa consapevolezza, è opportuno integrarla sia tra i principi fondanti che nelle disposizioni specifiche sulla formazione.

Non si evince la necessità di una chiara e obbligatoria previsione formativa per le categorie (PA, professionisti, magistrati, clinici) che più interagiranno con sistemi IA in contesti critici. Dal punto di vista tecnico, non sancire l'AI literacy come principio indebolisce l'intero impianto normativo e non avere formazione obbligatoria mirata significa esporsi a rischi di errata implementazione, mancata identificazione di bias tecnici e vulnerabilità di sicurezza da parte dei decisori e operatori.



Istituto per le Applicazioni del Calcolo

Per allineare pienamente il DDL all'approccio europeo e rafforzare tecnicamente la sua efficacia, si potrebbero valutare le seguenti alternative:

Integrare l'AI Literacy tra i Principi Fondamentali (Capo I): Aggiungere l'AI Literacy come principio esplicito nel Capo I del Disegno di Legge. Questo sancirebbe a monte l'importanza strategica di una comprensione tecnica e critica dell'IA per tutti gli attori coinvolti, orientando l'interpretazione e l'applicazione di tutte le successive disposizioni. Esempio di principio proposto: *Lo sviluppo, la diffusione e l'utilizzo dei sistemi di intelligenza artificiale nel rispetto dei principi stabiliti dalla presente legge si fondano sulla promozione e sullo sviluppo dell'AI literacy come definita nell'articolo 4 del regolamento (UE) 2024/1689 a tutti i livelli della società, al fine di garantire la comprensione (tecnica e non) del funzionamento, delle capacità e dei rischi associati all'IA.*

Introdurre la formazione obbligatoria per alcune specifiche categorie(Art. 11-15): Coerentemente con il principio di AI literacy, si ribadisce la proposta di integrare gli articoli da 11 a 15 con l'obbligo di percorsi formativi, la cui attuazione e delega è demandata al Governo nel successivo Art. 24 comma e che dovrebbe contenere esplicito riferimento all'articolo sull'AI Literacy.

F. Diritto d'Autore e Generative AI (Art. 25)

L'emergere dell'IA Generativa e la sua dipendenza essenziale da vasti dataset per l'addestramento e la generazione sollevano questioni critiche relative al diritto d'autore. L'Articolo 25 affronta gli output dell'IA, ma è fondamentale considerare anche l'impatto tecnico-legale dell'utilizzo delle opere protette durante la fase di addestramento dei modelli.

Dal punto di vista tecnico, l'efficacia dei modelli di IA generativa nel produrre contenuti (testi, audio, immagini, video) di alta qualità, incluse, a titolo esemplificativo e non esaustivo, le capacità di replicare digitalmente lo stile distintivo di specifici autori o artisti, personaggi riconoscibili, o persino caratteristiche biometriche vocali di attori e cantanti, dipende in modo cruciale dalla presenza delle loro opere originali nei **dati di addestramento**. Sebbene l'opera generata possa essere nuova, il suo *stile* deriva direttamente dall'apprendimento dai dati di input. Si evidenzia una lacuna tecnica nel DDL: anche quando i dati di addestramento provengono da banche dati per le quali era stato fornito un consenso alla riproduzione, tale consenso faceva riferimento alla copia dell'opera in un contesto storico precedente l'avvento dell'IA generativa e non copriva esplicitamente l'utilizzo per **addestrare modelli algoritmici** capaci di apprendere stili e pattern. L'assenza di una disciplina chiara su come gestire, dal punto di vista tecnico e legale, l'uso di dati protetti (e degli stili in essi



Istituto per le Applicazioni del Calcolo

contenuti) per l'addestramento su larga scala, e la mancanza di previsioni su un'equa compensazione per gli autori, rappresentano un nodo critico che l'Articolo 25, focalizzato sull'output, non risolve per quanto riguarda l'input (i dati di training).

Per affrontare la criticità tecnica e legale dovuta all'uso di dati per l'addestramento, con particolare riferimento all'apprendimento dello stile, un'alternativa è rappresentata dalla modifica l'Articolo 25 (o aggiungere commi specifici) per disciplinare questa fase. Le soluzioni tecniche e normative possibili includono:

1. La necessità di ricostruire o aggiornare le banche dati di addestramento ottenendo specifici consensi che coprano esplicitamente l'uso delle opere per l'addestramento di modelli IA, distinguendolo dal mero consenso alla riproduzione.
2. L'implementazione di un sistema di compensazione, royalties o licenze collettive per gli autori le cui opere (e quindi stili) contribuiscono in modo significativo all'addestramento dei modelli generativi, riconoscendo il valore tecnico ed economico derivante da tale uso.

È fondamentale che l'Articolo 25 stabilisca un quadro chiaro che bilanci l'innovazione nell'IA generativa con la tutela dei diritti e del contributo creativo degli autori, considerando le specifiche tecniche dell'addestramento e dell'apprendimento dello stile.

3. Raccomandazioni Trasversali

Dalla prospettiva tecnico-scientifica, le seguenti raccomandazioni trasversali sono cruciali per rafforzare l'efficacia e la lungimiranza del quadro normativo sull'IA:

Promozione della Sostenibilità Tecnica e Ambientale: Considerato il significativo impatto ambientale derivante dall'addestramento e dall'esecuzione di modelli IA su larga scala, si raccomanda di integrare nel DDL principi o misure che promuovano attivamente lo sviluppo di modelli computazionalmente efficienti, l'ottimizzazione dell'uso delle risorse hardware e l'adozione di infrastrutture energeticamente sostenibili nel ciclo di vita dell'IA.

Meccanismo di Aggiornamento Periodico: Data la rapidissima evoluzione dello stato dell'arte nell'IA, è tecnicamente indispensabile prevedere una clausola di revisione periodica della legge o meccanismi che includano il contributo di esperti tecnici e scientifici per adeguare la normativa alle nuove sfide e opportunità tecnologiche emergenti.



4. Conclusione

La presente memoria tecnica ha esaminato il Disegno di Legge N. 2316 con un approccio critico-costruttivo, riconoscendo il suo valore come primo tentativo organico di normare l'intelligenza artificiale in Italia e di allinearsi al quadro europeo. L'analisi condotta dalla prospettiva tecnico-scientifica ha evidenziato come il testo offra una base importante, ma presenti aree suscettibili di miglioramento per garantire la sua efficacia pratica, la promozione di un'innovazione responsabile e la protezione dei diritti nell'attuale contesto tecnologico.

In sintesi, le principali aree che, da un punto di vista tecnico, richiedono un affinamento riguardano:

- La precisione delle definizioni e l'ambito di applicazione, per coprire in modo chiaro e tecnicamente accurato le diverse tipologie di modelli (generalisti, di base, fine-tuned) ed evitare ambiguità (cfr. Sezione A).
- I requisiti di spiegabilità e trasparenza, che necessitano di integrare metriche e concetti tecnici come la riproducibilità e l'auditabilità algoritmica per essere operativi (cfr. Sezione B).
- L'integrazione di tecniche avanzate per la tutela della privacy, e una disciplina più completa sulla gestione dei derivati tecnici dell'addestramento (pesi del modello) su dati sensibili (cfr. Sezione C).
- L'incentivo alla sostenibilità e all'efficienza computazionale, promuovendo lo sviluppo di modelli ottimizzati per hardware eterogeneo e l'uso di dataset pertinenti al contesto italiano, aspetti cruciali per una sovranità tecnologica tecnicamente solida (cfr. Sezione D).
- Il rafforzamento della AI Literacy come principio guida e l'introduzione di obblighi formativi con un esplicito contenuto tecnico per categorie professionali chiave (cfr. Sezione E).
- La gestione dell'intersezione tecnico-legale tra diritto d'autore e addestramento di IA generativa, per disciplinare in modo chiaro e compensato l'uso di opere protette nella fase di training (cfr. Sezione F).
- Le raccomandazioni trasversali (focus sulla sostenibilità ambientale e meccanismo di revisione periodica con esperti tecnici) completano il quadro degli interventi suggeriti per una legge che sia robusta, flessibile e in grado di accompagnare l'evoluzione tecnologica.

Redatta in data 08/05/2025