Intelligenza Artificiale e Responsabilità di sviluppatori, fornitori e utilizzatori

Autore: Dott. Giacomo Conti

Revisori: Dott. Antonio Vetrò
Prof. Maurizio Borghi

Dicembre 2023

Studying the Internet, exploring its potential & experimenting new ideas

Via Pier Carlo Boggio 65/A, 10129 Torino, Italia

(http://nexa.polito.it/contacts-en)

+39 011 090 7217 (Telephone)

+39 011 090 7216 (Fax)

info@nexa.polito.it

Mailing address:

Nexa Center for Internet & Society

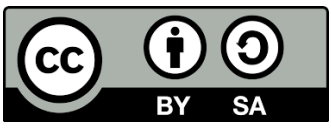
Politecnico di Torino - DAUIN

Corso Duca degli Abruzzi, 24

10129 TORINO

ITALY

The Nexa Center for Internet & Society is a research center affiliated to the Department of Control and Computer Engineering of Politecnico di Torino (http://dauin.polito.it) and the Department of Law of the University of Turin (https://www.law.unito.it).



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Indice

Introduzione 5

Software come bene o come servizio? 6
L'approccio europeo all'intelligenza Artificiale: l'*IA Act*, la *Liability Directive*. 10
IA ACT 12
IA LIABILITY DIRECTIVE 17
Approcci extra UE alla normazione dell'IA: gli Stati Uniti 20
Approcci extra UE alla normazione dell'IA: la Cina. 23
Conclusioni 26

Introduzione

I modelli generativi di intelligenza artificiale, stanno rapidamente prendendo piede nei più disparati ambiti della società, modificando i precedenti paradigmi comunicativi, di lavoro, di efficienza e di automazione. La loro influenza si estende a diversi settori: dalle strategie aziendali all'assistenza sanitaria; dalle pratiche pedagogiche ed educative alla ricerca universitaria; dallo sviluppo di software all'intrattenimento e alle arti. Sebbene le loro capacità sembrino ed effettivamente siano rivoluzionarie, questi sistemi generativi di I.A. presentano sfide complesse.

Modelli come GPT-4, ChatGPT, Luminous, Bard e Bing nel dominio del testo, Stable Diffusion e DALL-E 2 per le immagini, Synthesia per la creazione di video e MusicLM per l'audio si sono presto diffusi con grande successo tra gli utenti. All'orizzonte si intravedono ancora più applicazioni, con la prospettiva di snellire i compiti, specialmente quelli più oberanti e ripetitivi, permettendo di dare priorità alle attività principali, d'ingegno e di creatività, portando potenzialmente a una distribuzione delle risorse più efficiente ed equa.

Tale diffusione e tali potenzialità sono tanto grandi quanto grandi sono i possibili costi degli errori e dei danni causati da un qualsiasi malfunzionamento di questi sistemi. Sono chiaramente possibili situazioni in cui algoritmi sbagliati e pregiudizievoli, o *dataset* viziati, o ancora usi impropri od illeciti dei sistemi di IA generativa possano causare gravi danni ai privati ed alla società. Come ogni tecnologia che vuole dirsi rivoluzionaria, il potere di questi modelli di I.A. può rivelarsi essere un'arma a doppio taglio[1].

Da ciò prende le mosse quello che forse è la madre di tutte le questioni: chi si assume la responsabilità quando, a partire dall'uso di una IA generativa, è stato arrecato un danno? Gli sviluppatori che hanno addestrato il modello, gli utenti che lo hanno utilizzato o le piattaforme che lo hanno ospitato? La causa del danno deriva da dati errati su cui si è basata la decisione di IA, da un malfunzionamento dell'algoritmo, da un uso improprio del sistema o da una sua evoluzione non prevista? Quando un sistema di IA generativa genera testo, audio o immagini a partire da elementi protetti da diritto d'autore si tratta di semplice ispirazione o di violazione di copyright?

Le peculiarità dei sistemi di Intelligenza Artificiale complicano notevolmente l'identificazione delle cause di un danno, la determinazione delle responsabilità e l'applicazione delle normative vigenti in materia. Questa complessità può portare le vittime di tali danni a incontrare ostacoli nel rintracciare l'origine del problema e nel fornire le prove necessarie per ottenere un risarcimento adeguato.

E' fondamentale, quindi, mettere in atto garanzie solide in caso di problemi o danni correlati all'utilizzo di questa tecnologia. Questa precauzione non solo tutelerebbe i danneggiati, ma incentiverebbe anche investimenti nel settore e favorirebbe l'adozione dell'I.A. da parte di cittadini e aziende. Questa esigenza di consolidare la fiducia del pubblico è stata peraltro chiaramente sottolineata anche dalla Commissione Europea, già nel 2020, in una relazione al Parlamento e al Consiglio sulle implicazioni dell'I.A.[2].

[1] P. Hacker, A. Engel, M. Mauer, *Regulating ChatGPT and other Large Generative AI Models,* 2023 ACM Conference on Fairness, Accountability and Transparency, 12-15 Giugno 2023, https://dl.acm.org/doi/pdf/10.1145/3593013.3594067

[2] Commissione Europea, *Relazione della Commissione al Parlamento Europeo, al Consiglio e al Comitato Economico e Sociale Europeo, "sulle implicazioni dell'intelligenza artificiale, dell'internet delle cose e della robotica in materia di*

Almeno per ciò che concerne il mondo occidentale, all'orizzonte appaiono due fondamentali sistemi per risolvere questo dilemma, a partire dagli indirizzi dei due più influenti attori internazionali, cioè gli Stati Uniti e l'Unione Europea:

1) Vi è un approccio **ex-ante** e **centralizzato**, di ispirazione europea e basato su una legislazione univoca che regoli e definisca le prerogative, i requisiti ed i limiti dei sistemi di intelligenza artificiale all'interno di un determinato mercato, agendo preventivamente rispetto al verificarsi dei problemi ed mirando ad arginare i rischi, ma al prezzo di limitare la portata innovativa e la "corsa all'oro" tipica di una nuova tecnologia;
2) Ed un approccio **ex-post** e **decentralizzato**, di ispirazione statunitense, in cui si predilige una regolamentazione minima a livello centrale (o federale, in questo caso), demandando ai singoli stati di normare la questione e limitando gli interventi a quelli di singole agenzie federali, come la FTC (Federal Trade Commission). Invece di stabilire norme preventive, la FTC interviene a posteriori, sanzionando quelle imprese che di volta in volta sono determinate come responsabili di un danno causato dai loro sistemi di IA. Questo approccio permette una maggiore libertà nell'innovazione, ma rischia di rendere insufficiente o inesperibile la tutela per i soggetti lesi.

Prima di addentrarsi nelle moderne questioni giuridiche e tecniche riguardanti la responsabilità dell'IA nello specifico, ed i piani legislativi adottati per farne fronte, è necessario tuttavia avere chiara la genesi storica che ha portato ad essi. Il problema principale del software, indipendentemente dal fatto che sia esso un'Intelligenza Artificiale o un tradizionale programma non dotato di autonomo apprendimento, è ancora oggi legato alla sua stessa natura.

Software come bene o come servizio?

Si diceva che quale che sia l'approccio regolativo migliore o quello di volta in volta ritenuto più efficace, i sistemi di Intelligenza Artificiale sono prima di tutto dei *software*, ed in quanto tali è difficile ancora oggi dare una precisa definizione della loro natura. In particolare, non è mai sopito né risolto il dibattito fondamentale: se il software cioè sia un bene oppure un servizio.

Nell'attuale era digitale, la concezione del software come merce fisica è stata radicalmente rimodellata. In precedenza, il software era commercializzato su supporti tangibili, quali cassette, floppy disk, CD e DVD, un formato che ne facilitava la percezione come bene vendibile e trasferibile. Tuttavia, l'espansione di Internet ha progressivamente spostato l'acquisto e la distribuzione del software verso un modello di servizio basato su licenze, con la vendita che si concretizza non più nel trasferimento di un oggetto fisico, ma nell'assegnazione di diritti d'uso attraverso account digitali. E' il sistema della cosiddetta Digital Distribution, che viene definita come "un metodo di distribuzione nel quale il contenuto è distribuito senza l'uso di un medium fisico, solitamente con un download da internet verso la casa del privato consumatore"[3].

sicurezza e di responsabilità", 19 febbraio 2020, https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:52020DC0064

[3]*Digital Distribution Law & Legal Definition*, USLegal, 25 Aprile 2012

Questo cambiamento ha sollevato importanti questioni legali e di responsabilità. I produttori di software, in particolare quelli che operano nel campo dell'intelligenza artificiale, devono navigare in un contesto regolamentare complesso, dove la determinazione della responsabilità per software difettoso non è più immediata. Inoltre, come si è visto, l'IA, con le sue capacità di apprendimento e di decisione autonoma, potrebbe causare danni non previsti dalla programmazione iniziale, ampliando così la portata e la natura dei danni rispetto al software convenzionale.

Il primo obiettivo da raggiungere è stato considerato dal mercato di natura prettamente economica: occorreva evitare che, tramite la dematerializzazione del software, si facesse un uso illecito ed incontrollato dello stesso. Questo significava sia evitare la violazione del copyright (il software viene utilizzato cioè da chi non ne ha il diritto, non avendolo acquistato), sia controllare con precisione quali utenti avessero a loro disposizione quali software. I principali attori del mercato, hanno quindi deciso di puntare ampiamente sui cosiddetti sistemi di DRM, o Digital Rights Management.

Questi sistemi di protezione solitamente oggi subordinano la fruizione di un software ad una chiave alfanumerica, detta "licenza", che viene legata ad uno specifico account, a sua volta collegato ad una persona fisica, e certificato tramite vari sistemi (come una doppia conferma e-mail o un login nel quale è ulteriormente richiesta una chiave generata sul cellulare dell'utente). Risalire ai soggetti responsabili di aver utilizzato in modo illecito un certo software sembrerebbe dunque essere astrattamente possibile. Tuttavia, nella pratica, queste protezioni vengono spesso aggirate tramite il mascheramento degli indirizzi IP (utilizzando le sempre più diffuse reti private virtuali, o VPN), la creazione di *fake* (cioè di identità fittizie), o la retro-ingegnerizzazione dei software, che permette l'eliminazione dei loro sistemi di protezione (il cosiddetto *cracking*).

La volontà economicistica dietro alla base di questo sistema, che tende a tutelare gli sviluppatori contro le violazioni del copyright, determina al contempo una situazione sfavorevole per l'utente, perchè egli può vedersi bloccato il proprio *account* e non poter più fruire del prodotto che ha precedentemente acquistato. Non solo: qualsiasi interruzione nella catena che porta l'acquirente a utilizzare il software porta inevitabilmente all'impossibilità di utilizzarlo. Da parte sua può non avere una connessione ad internet con la quale autenticarsi; da parte degli sviluppatori può verificarsi, in futuro, un fallimento o un riassestamento delle risorse che elimini i *server* dai quali i dati sono presi; e questa stessa evenienza può verificarsi per i sistemi di protezione, situati presso terzi che non sono neppure gli sviluppatori del software protetto[4]. Inoltre gli account possono essere sospesi, cancellati o comunque dismessi spesso senza che l'utente possa opporsi.

Il risultato di questo è che ci si chiede oggi se ad acquistare licenze su internet si sia veramente proprietari di qualcosa, o se si fruisca solo di un servizio che arbitrariamente può finire[5].

Questo discorso, che nasce per il software tradizionale, è ancora oggi assai applicabile alle Intelligenze Artificiali, che vengono fatte fruire dagli utenti dietro un abbonamento ed appoggiandosi su servizi terzi (come server o database) necessari perché il programma di I.A. possa funzionare.

Una licenza, in questo senso, non è quindi che un diritto limitato di godimento del software, subordinato all'accettazione di determinati termini e condizioni, che spesso sono sfavorevoli all'utente e vengono iscritti nella cosiddetta EULA, o *End User License Agreement*, con evidenti sgravi

[4] *Denuvo,* su www.denuvo.com

[5] John Walker, *Thought: Do We Own our Steam Games?,* Rock Paper Shotgun, 1 Febbraio 2012

di responsabilità da parte del produttore o del rivenditore del software[6].

In Europa, la c.d. sentenza Oracle[7], pronunciata nel 2012[8], ha riguardato la rivendita di licenze software usate, sfidando la nozione che il software, una volta dematerializzato e legato a un account specifico, non possa essere rivenduto come un bene fisico. Oracle sosteneva di offrire un servizio anziché un prodotto fisico, ma la Corte di Giustizia stabilì che non vi è distinzione tra il momento dell'acquisto e quello del download, riconoscendo di fatto il software come un bene e permettendo peraltro la rivendita di licenze usate sotto il principio dell'esaurimento comunitario. Malgrado gli iniziali titoli sensazionalistici di taluni giornali[9], a distanza di anni si rileva una scarsa incisività nella stessa nelle pratiche commerciali delle attuali *software house*.

Il caso Oracle è stato fondamentale per sancire la possibilità di avere un mercato dell'usato[10] per il software di oggi, sempre più dematerializzato.

Seppure la giurisprudenza europea appaia più vicina alla tradizionale concezione di software come di un bene, le problematiche sono onnipresenti, e le insidie dietro l'angolo.

Si immagini la situazione di un *software* che viene venduto ma che per essere fruito deve appoggiarsi ad un'architettura esterna: ad un servizio, dunque, di terzi. Ad esempio, ad un sito di autenticazione in cui è necessario registrarsi dopo aver acquistato il programma per dimostrarne l'autenticità e poterlo quindi fruire. Viene da sé che la fruizione di quello che all'apparenza è un bene viene subordinata al corretto funzionamento del servizio che garantisce la stessa fruibilità. Che succede, dunque, se il bene diventa infruibile per via di un servizio lacunoso?

E' chiaro che qui il bene/software di per sé è funzionante: certamente la tutela giuridica non si concentrerà su di esso, ma sul servizio interrotto.

Nel caso delle IA, si pensi a come il server su cui è basata l'interfaccia testuale con la quale si interagisce con ChatGPT vada offline, o sia irraggiungibile in talune parti del mondo: non è per forza detto che sia la stessa impresa autrice dell'IA o del software in generale (in questo caso OpenIA), ad occuparsi del sito web e dell'interfaccia utente, fornendo i servizi al di là del software in sé richiesti per poterlo fruire. L'I.A. in sé funziona, ma non si può utilizzare.

Si potrebbe sostenere che in questi casi si crei un contratto di prestazione di servizi avente come oggetto la disponibilità dei server o comunque degli strumenti del produttore che consentono al compratore del software la fruibilità di quest'ultimo; come corrispettivo il prezzo già pagato per

---

[6] John Walker, *Thought: Do We Own our Steam Games?,* op. cit.

[7] Osborne Clarke, *The end of the UsedSoft case and its implications for "used" software licences,* Osborneclarke.com, 11 maggio 2015.

[8] C-128/11, disponibile per intero su https://curia.europa.eu/juris/document/document.jsf?docid=124564&doclang=IT

[9] Elio Cogno, *Il software usato si può rivendere: la Corte europea dà torto a Oracle,* Il Fatto Quotidiano, 5 luglio 2012

[10] Se si prende a piene mani dalla concezione di software acquistato attraverso l'ottenimento di una licenza di uso, come di un servizio dunque e non di un bene, tutto ciò che ha a che fare con il "mercato secondario", o "dell'usato", non è concepibile, e nascono anche quesiti sull'effettiva proprietà che l'utente ha sul suo software . Come si potrebbe infatti rivendere il software, se esso fosse considerato un servizio intimamente connesso con il primo acquirente, l'unico dotato di un account legittimo e quindi l'unico in grado di fruirlo?
La chiave di licenza che permette all'acquirente di possedere il software è sua e solo sua, disponibile sul suo account nominale o su una piattaforma comunque privata e non può essere rivenduto o alienato in qualsiasi modo.

l'acquisto del software; come durata il tempo di supporto del software, e come metro di valutazione per l'adempimento il classico livello minimo di prestazione del servizio.

Un tale contratto, per quanto implicito, aprirebbe le porte ad una tutela giuridica avverso comportamenti lesivi della fruizione del servizio, non potendosi notare una qualche differenza tra questo tipo di fornitura e le più classiche, come quelle di interruzione della linea telefonica[11], soprattutto se in assenza di un <<congruo preavviso>>[12]. È naturale infatti che nel mondo del lavoro informatizzato di oggi l'impossibilità di utilizzare un software, specie se potenzialmente onnicomprensivo come un'I.A. generativa, possa dare origine a danni patrimoniali ingenti.

Questa concezione dubbia e confusa di quale disciplina dovrebbe essere applicabile al software non poteva di certo essere foriera di un futuro roseo per l'innovazione e lo sviluppo tecnologico, né per la tutela degli utenti e del pubblico, soprattutto a fronte di prorompenti novità quali quelle introdotte dai sistemi di Intelligenza Artificiale. L'Unione Europea ha pertanto deciso, dopo molti anni, di dotarsi di una legislazione complessiva sul tema dell'Intelligenza Artificiale: una sorta di compendio normativo unico nel quale scrivere ex-novo una disciplina innovativa e al passo coi tempi: nelle intenzioni dei suoi legislatori, l'AI Act, del quale ci si occuperà nel dettaglio in seguito.

Contemporaneamente ai dibattiti europei sulla natura del software, anche negli Stati Uniti le discussioni hanno abbondato. Specialmente, ci si è concentrati sul fatto che se il software è definito come un servizio, allora generalmente il produttore non risponderà di responsabilità extracontrattuale, a meno che non sussista qualche comportamento di negligenza o di dolo[13]. Se, al contrario, il software è un bene, allora il produttore sarà punibile a titolo di responsabilità oggettiva, o *strict liability*[14]. Esso sarà responsabile del difetto del prodotto anche se non avrà avuto colpa per il suo malfunzionamento. Dal momento che il software null'altro è se non un codice preciso che non lascia il minimo spazio al caso, è evidente che un minimo di "colpa", da parte dello sviluppatore, dovrà esserci per forza.

Su tale diatriba si scontrano due diverse concezioni. La prima ritiene che i produttori del *software* siano da ritenere *strictly liable*, o responsabili in via oggettiva, poiché il software è un bene e non un servizio e perchè da ciò deriverebbe una maggiore precauzione presa dagli stessi sviluppatori nel momento in cui il prodotto è messo in commercio. Questa tesi è interessante dal momento che verrà ripresa, come si vedrà, più avanti nel documento a proposito delle moderne proposte di responsabilità per software di intelligenza artificiale nello specifico. I sostenitori di questa tesi affermano che "gli sviluppatori sono quelli tra tutti nella migliore posizione per conoscere potenziali difetti e danni che potrebbero derivare dai loro prodotti". Pertanto, con una sorta di responsabilità oggettiva, essi dovrebbero fare più attenzione durante lo sviluppo degli stessi software[15].

Al contrario, coloro i quali si oppongono alla concezione del software come di un bene e negano pertanto la *strict liability* per i produttori motivano la loro tesi non solo, come faceva Oracle, puntando sulla concezione di "servizio", ma anche dicendo che una responsabilità oggettiva sarebbe

[11]Tribunale di Roma, sez. VII, sentenza 15/11/2014 , n° 22789

[12]Tribunale di Brindisi, sez. Ostuni, sentenza 30/12/2011, n° 17020

[13]Giuseppe Contissa, *Strict liability vs negligence: should software manufacturers be held strictly liable for software defects?* European University Institute, 13 aprile 2012

[14]Fred W. Morgan, *Strict Liability and the Marketing of Services vs. Goods: A Judical Review*, Vol 6, American Marketing Association, 1987

[15]Giuseppe Contissa, *Strict liability vs negligence: should software manufacturers be held strictly liable for software defects?*, Op. Cit.

troppo gravosa per i produttori di *software* e pertanto un enorme ostacolo al loro lavoro. Essi sarebbero costretti a prendere troppe precauzioni, e il bilanciamento di interessi tra l'eventuale risarcimento e il normale andamento dell'attività lavorativa sarebbe sbilanciato a favore del primo dei due concetti, col risultato che la *software house* dovrebbe aspettarsi cause civili ad ogni passo, e correre conseguentemente ai ripari. Ciò comporterebbe anche un'evidente diminuzione della capacità innovativa degli sviluppatori del software, perché, com'è ovvio, ogni innovazione porta con sé dei rischi e pertanto costoro sarebbero molto più inclini a mantenere lo status quo nei loro prodotti se a ciò si accompagnasse una maggiore sicurezza in sede risarcitoria[16]. Si noti che un tale dibattito si è ripetuto, praticamente uguale, durante gli studi della Commissione Europea per la proposta di una nuova Direttiva sulla responsabilità civile extracontrattuale dell'intelligenza artificiale[17]*, l'IA Liability Directive*, che si avrà modo di analizzare nel prossimo capitolo.

Vi è poi da considerare che, se si considera il software come un bene, negli Stati Uniti la sua vendita e diffusione è determinata dall'Uniform Commercial Code – una serie di regole che, non dissimilmente da quanto accade dei nostri codici, uniformano l'amministrazione e le transazioni commerciali nelle legislazioni dei vari Stati, ed anche la risoluzione delle controversie nate da queste tematiche.

Per ciò che concerne la materia dei beni, la disciplina è certa: valgono le disposizioni dell'UCC, in particolare quelle dell'articolo 2 (a sua volta diviso in 105 paragrafi). Esse contengono, dapprima, una definizione di "beni" (*goods*): "tutto ciò che è mobile al momento dell'identificazione nel contratto di vendita, a parte il denaro con cui il prezzo viene pagato"[18]. Da ciò deriva un'uniformità legislativa che invece non si può riscontrare nella disciplina inerente ai servizi.

Tuttavia, nell'UCC il *software* non è mai citato tra i beni, né come esempio, né come condizione speciale (anche se da più voci si acclama un ripensamento di tutto l'articolo 2 dell'UCC aggiungendo la lettera B), che dovrebbe riguardare proprio la disciplina dei beni dematerializzati e del *software*[19]). Rimane pertanto una grandissima incertezza, anche negli Stati Uniti, sulla risposta alla domanda se il *software* sia un bene od un servizio, soprattutto ad opera della giurisprudenza.

L'approccio europeo all'intelligenza Artificiale: l'*IA Act*, la *Liability Directive*.

Il precedente capitolo dimostra come il diritto sia rimasto fondamentalmente indietro rispetto all'incedere sempre più repentino del progresso tecnologico applicato al software. La mancata definizione di regole precise sui beni digitali, o sulla somministrazione di servizi digitali, rende la tutela dei dritti per gli utenti e la corrispettiva enunciazione di obblighi per i produttori incerta ed inadeguata. Questo problema non poteva che acuirsi nell'ultimo periodo, nel quale l'incedere tecnologico si è fatto ancora più veloce, e la nascita e la diffusione degli strumenti di intelligenza

[16] Giuseppe Contissa, *Strict liability vs negligence: should software manufacturers be held strictly liable for software defects?*, Op. Cit.

[17] EU Commission, Liability Rules for Artificial Intelligence, 28 settembre 2022, https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en

[18] Uniform Commercial Code, Article 2, § 1, "Goods"

[19] Si veda, ad esempio, la proposta dell'American Law Institute insieme alla National Conference of Commissioners on Uniform State Laws durante un lungo *meeting*, Luglio 24-31, 1998, Cleveland, Ohio.

artificiale hanno posto questioni uniche e complesse per le quali almeno un minimo di regolamentazione è da ritenersi necessaria.

Vi sono infatti alcune caratteristiche intrinseche dei sistemi di Intelligenza Artificiale che rendono particolarmente complesse la normali questioni tipicamente riferite al software. Tra queste, ancora una volta la principale è con ogni probabilità l'attribuzione di responsabilità.

A causa dell'elevata capacità autonoma dei sistemi di I.A., che ne determina un'evoluzione costante ed una modifica delle loro funzionalità e della qualità del loro *output* nel corso del tempo, diventa complesso stabilire una catena precisa di responsabilità: lo sviluppatore del software di I.A. potrebbe essere inerme di fronte all'evoluzione del suo stesso sistema, che in qualche misura, addestrato su moli sempre maggiori di dati, "trascende" la sua originale programmazione e dà origina ad un output imprevisto e potenzialmente dannoso.

Se a ciò si aggiunge la questione della c.d. "opacità" degli algoritmi, cioè la marcata difficoltà da parte del pubblico (e degli stessi sviluppatori) di comprendere e prevedere fino in fondo con assoluta certezza gli *output* del sistema, si comprenderà come le vittime dei potenziali danni derivanti da una I.A. potrebbero andare incontro a difficoltà insormontabili nel dimostrare il nesso causale tra il danno subito ed il malfunzionamento del software, rischiando quindi di non essere adeguatamente risarcite.

La difficoltà nell'ottenimento del risarcimento è ulteriormente acuita dall'ormai consolidata tradizione dell'inserimento di clausole di esclusione della responsabilità da parte dei produttori o dei distributori del software: il riferimento è sempre all'*EULA* passata in rassegna nel precedente capitolo, che anche per i sistemi di IA tende ad escludere la responsabilità dei produttori e dei fornitori in caso di malfunzionamento.

A tali questioni, proprie di ogni sistema di I.A., se ne aggiungono poi altre nel momento in cui si passano in rassegna le Intelligenze Artificiali "generative": quelle che non solo ausiliano l'essere umano nel completamento di certe attività o prendono decisioni più o meno discrezionali, ma che producono un contenuto nuovo a partire da input dell'utente, attingendo ad enormi quantità di dati sui quali esse si sono formate. La generazione di nuovi dati da parte di queste IA a partire dai modelli sui quali è stato svolto il loro addestramento le rende distinte dalle altre, che solitamente hanno funzioni predittive o classificatorie, e non anche generative. Questo nuovo ambito applicativo rende possibile la creazione di nuovi dati e nuove relazioni tra i dati, non previste esplicitamente dagli sviluppatori umani a monte[20]. Non solo: si può anche dare il caso, che per il momento può sembrare fantascienza ma che indiscutibilmente si verificherà nel prossimo futuro, di I.A. generative che si addestrano su modelli i cui dati comprendono altre generazioni di ulteriori I.A. generative[21].

In primo luogo si pone la questione del diritto d'autore sui contenuti generati: se un'IA crea una composizione musicale o un testo letterario, chi ne detiene i diritti? Le attuali normative, evidentemente, attribuiscono i diritti d'autore agli esseri umani, lasciando quindi un possibile vuoto normativo (per quanto fantascientifico) quando si tratta di opere generate da IA.

[20] P. Hacker, A. Engel, M. Mauer, *Regulating ChatGPT and other Large Generative AI Models*, FACCT '23 ACM Conference, pp 1112-1123, 12 giugno 2023, https://dl.acm.org/doi/10.1145/3593013.3594067

[21] E. Strickland, *Are you still using Real Data to Train your AI,* IEEE Spectrum, febbraio 2022, https://spectrum.ieee.org/synthetic-data-ai

Ancor più che nelle IA tradizionali, il potenziale impatto sociale ed economico delle IA generative sembra ancora più dirompente: se un'IA è in grado di produrre lavori che tradizionalmente richiedono un input creativo umano, quali saranno le ripercussioni per i professionisti in questi campi? Come si regoleranno le professioni e le industrie se il lavoro creativo dovesse in larga parte essere assorbito da sistemi automatizzati?

Si pone anche la questione della sicurezza e della privacy: le Intelligenze Artificiali generative richiedono l'accesso a vasti database di informazioni, spesso contenenti dati personali. La protezione di questi dati è imperativa, ma la capacità delle IA di generare nuovi contenuti basati su tali dati introduce un livello di rischio che deve essere attentamente gestito[22].

Infine, il potenziale di credibilità degli output delle IA generative può portare alla creazione automatica di notizie false o di opere d'arte manipolative. Questo ha il potenziale di influenzare l'opinione pubblica e alterare il corso degli eventi sociali e politici. Le implicazioni etiche di questa capacità non sono soltanto teoriche ma si pongono con urgenza come problemi pratici che la società deve affrontare, specie in un momento di tensione globale qual è quello di oggi.

Nel tentativo di dare una regolamentazione precisa ed univoca a questi nuovi temi, per mitigarne gli evidenti rischi e indirizzare quanto più possibile l'I.A. verso un progresso non lesivo dei diritti dei cittadini, l'Unione Europea ha inteso promulgare atti normativi specificamente legati all'Intelligenza Artificiale, sorpassando quindi il tradizionale dibattito del software come bene o come servizio e decidendo di intraprendere una strada nuova, mai battuta e tendenzialmente inedita. Questa strada si sostanzia in due fondamentali atti: l'*AI Act* e l'*AI Liability Directive*.

IA ACT

L'AI Act si pone come il primo grande tentativo di regolamentare l'Intelligenza Artificiale all'interno dell'Unione Europea con un approccio olistico. L'obiettivo dell'AI Act è quello di classificare e regolamentare le applicazioni di intelligenza artificiale in base al loro concreto rischio di arrecare danno. L'atto, come si diceva, è innovativo, anche se ispirato a legislazioni precedenti, quali il GDPR, e, nella parte dedicata ai sistemi IA "ad alto rischio", il regolamento (UE) 2019/1020[23] che regolamenta la conformità dei prodotti disponibili nel mercato dell'Unione Europea. In questo senso, si può vedere una preferenza degli enti europei a considerare i software di Intelligenza Artificiale come dei prodotti (o almeno "parte" dei prodotti) anziché dei servizi, basandosi le nuove fonti normative su precedenti basi legislative relative ad essi.

Come nel caso del GDPR, anche l'AI Act si applica anche "ai fornitori e agli utenti di sistemi di IA situati in un paese terzo, laddove l'output prodotto dal sistema sia utilizzato nell'Unione". In altre parole, qualsiasi soggetto che esperisca effetti nel mercato dell'Unione tramite sistemi di IA è sottoposto alla normativa dell'IA Act. E, sempre come nel caso del GDPR, l'idea è quella di promulgare l'IA Act a fine 2023 e determinarne l'entrata in vigore due anni dopo, nel 2025. Questo dovrebbe consentire ai soggetti interessati sufficiente tempo per mettere in pratica le eventuali

[22] ESEMPIO REDDIT IMGUR

[23] Regolamento (Ue) 2019/1020 Del Parlamento Europeo e del Consiglio del 20 giugno 2019 sulla vigilanza del mercato e sulla conformità dei prodotti e che modifica la direttiva 2004/42/CE e i regolamenti (CE) n. 765/2008 e (UE) n. 305/2011, https://eur-lex.europa.eu/legal-content/IT/TXT/PDF/?uri=CELEX:32019R1020&rid=1

modifiche che l'IA Act imporrebbe ai loro prodotti, per arrivare al momento dell'entrata in vigore già preparati.

Nello specifico delle previsioni dell'IA Act, sono individuati diversi livelli di rischio riferiti a diverse modalità di utilizzo dei sistemi di IA. Alcune di queste pratiche sono determinate come troppo rischiose, e sono totalmente vietate: si includono in queste l'uso dell'IA per manipolare subliminalmente o sfruttare le vulnerabilità delle persone che possono causare danni fisici o psicologici, l'uso indiscriminato di identificazione biometrica in tempo reale in spazi pubblici[24] o l'uso di 'punteggi sociali'[25] derivati dall'IA da parte delle autorità per svantaggiare ingiustamente individui o gruppi.

Vi sono poi i sistemi ad **alto rischio**. In primo luogo, i sistemi di IA utilizzati in prodotti soggetti alla legislazione dell'UE sulla sicurezza dei prodotti[26], includendo con ciò i giocattoli, gli aeromobili, i dispositivi ad uso medico, e gli ascensori o montacarichi.

In secondo luogo, sistemi di IA che rientrano in otto aree specifiche e che dovranno essere registrati in un database specifico nell'Unione Europea:
- Identificazione e categorizzazione biometrica di persone fisiche
- Gestione e funzionamento di infrastrutture critiche
- Educazione e formazione professionale
- Impiego e gestione dei lavoratori
- Accesso e fruizione di servizi privati essenziali e servizi e benefici pubblici
- Applicazione della legge
- Gestione dell'immigrazione, delle richieste di asilo e controllo delle frontiere
- Assistenza nell'interpretazione e applicazione della legge

I sistemi ad alto rischio richiedono una valutazione di conformità obbligatoria prima di essere immessi sul mercato. L'atto prevede inoltre l'introduzione di un European Artificial Intelligence Board per promuovere la cooperazione nazionale e assicurare il rispetto del regolamento. Tale autorità avrà accesso a informazioni riservate, inclusi i codici sorgente dei sistemi di IA, e sarà soggetta a stringenti obblighi di riservatezza. Adotterà misure correttive per proibire, limitare, ritirare o richiamare sistemi di IA che non rispettano l'AI Act o che, pur essendo conformi, presentano un rischio eccessivo. In caso di persistente inosservanza, gli Stati membri dovranno prendere tutte le misure appropriate per limitare, proibire, ritirare o richiamare dal mercato il sistema di IA ad alto rischio in questione. Sono previste sanzioni amministrative variabili (fino a 30 milioni di euro o al 6% del fatturato annuale mondiale totale) a seconda della gravità dell'infrazione[27].

Tali stringenti obblighi potrebbero essere legittimamente considerati un pesante fardello per l'innovazione. Nel tentativo di limitare ciò, la Commissione propone che gli Stati membri o il Garante Europeo della Protezione dei Dati possano stabilire un "sandbox regolatorio", ossia un ambiente

[24] Questo utilizzo è l'unico ad avere previste alcune possibilità di eccezione: di fatto se è utile per ricercare potenziale vittime di reato o loro sospettati, e se è utile per prevenire "una minaccia specifica, sostanziale e imminente".
[25] Probabilmente un riferimento alla pratica del *social credit system* ormai diffusa in Cina e largamente portata come esempio totalitario di limitazione dei diritti tramite sistemi sempre più automatizzati.
[26] Il regolamento (UE) 2019/1020, citato in precedenza.
[27] Parlamento Europeo, *Briefing on Artificial Intelligence Act,* europarl.europa.eu, 2023, https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf

controllato che facilita lo sviluppo, il test e la validazione dei sistemi di IA innovativi per un periodo limitato di tempo prima che vengano immessi sul mercato. Il sandboxing permetterà ai partecipanti di utilizzare dati personali per promuovere l'innovazione nell'IA senza pregiudizio per i requisiti del GDPR. Altre misure sono specificamente orientate a fornitori di piccola scala e start-up.

Vi sono in seguito i sistemi a **rischio limitato**, o *limited risk*. Sono questi quelli nei quali esplicitamente figurano le I.A. che "interagiscono con gli esseri umani", e "che **generano o manipolano immagini**, audio o contenuto video"[28]. Per questo genere di sistemi di I.A., l'obbligo è fondamentalmente quello di trasparenza: di informare l'utente e quanti altri potenzialmente siano a contatto con il sistema o il prodotto del sistema di I.A. stesso (come un'immagine o un testo) che vi è o vi è stata un'interazione con una intelligenza artificiale.

A una prima visione la norma, che esplicitamente cita le I.A. generative come una sorta di sottocategoria delle I.A. a rischio limitato, sembrerebbe far pensare che tali I.A. siano unicamente sottoposte a questi obblighi di trasparenza. In effetti, l'introduzione delle I.A. generative nell'AI Act sembra un po' un'aggiunta posticcia e poco organica nel resto del testo, nonché foriera di alcuni problemi. Va infatti intanto sottolinato come l'originale proposta di AI Act proveniente dalla Commissione Europea non prendeva in esame le I.A. generative nello specifico. Questa mancanza è stata presa in esame dal Parlamento Europeo in fase di negoziazione, ed è ora al centro del dibattito sull'A.I. Act. E' stato infatti il Parlamento Europeo a ritenere che i modelli fondativi di IA generativa, come ChatGPT, usati per generare arte, musica e altri contenuti, sarebbero dovuti essere soggetti a obblighi di trasparenza stringenti, al pari di quelli dei sistemi *low risk* individuati precedentemente dalla Commissione Europea.

Ma nel concreto, la situazione non è così semplice. Intanto, c'è un problema di metodologie concrete per raggiungere tale trasparenza: i produttori di tali modelli e di contenuti generativi dovrebbero rivelare che il contenuto è stato generato da IA e non da esseri umani, addestrare e progettare i loro modelli per prevenire la generazione di contenuti illegali e pubblicare informazioni sull'uso di dati di addestramento protetti dal diritto d'autore. Infine, tutti i modelli fondativi dovrebbero fornire tutte le informazioni necessarie affinché i distributori e gli utenti a valle possano adempiere ai loro obblighi ai sensi dell'atto sull'IA.

Tale richiesta non è irragionevole in astratto, ma rischia di diventare estremamente difficile da rispettare nel concreto, specie per ciò che concerne l'obbligo di trasparenza riguardante i contenuti generati da IA: proprio lo stesso paragrafo precedente di questo scritto è stato generato tramite ChatGPT, ma è stato anche opportunamente rivisto e modificato da un essere umano. Il risultato è un testo fondamentalmente indistinguibile da uno realizzato interamente a mano, e non rilevabile come artificiale da alcuno dei (carenti, si deve dire) sistemi di controllo autoriale. In che termini dunque dovrebbe porsi la questione quando l'I.A. generativa non è altro che un ausilio ad un'opera umana? Allo stesso modo, un'immagine generata da una I.A. potrebbe essere ampiamente modificata da un essere umano: sarebbe lo stesso obbligatorio da parte dell'autore dichiarare la sua provenienza artificiale? In questo senso, seppure la terminologia dell'A.I. Act, opportunamente rivisto dal Parlamento Europeo per includere le I.A. generative, parli genericamente di *providers* e *distributors*, il destinatario finale delle prescrizioni è in realtà spesso l'*utente*, che utilizza sistemi di I.A. per i suoi scopi e deve in seguito dichiarare la provenienza artificiale dei suoi lavori.

[28] Parlamento Europeo, *Briefing on Artificial Intelligence Act,* Op. Cit.

Si pensi al caso appunto di ChatGPT che generi un testo: chi si dovrebbe considerare come *provider*, o, in italiano, *fornitore*? Intanto, esiste un problema tecnico: ChatGPT è innanzi tutto un servizio web, che tramite API è collegato al modello di I.A. vero e proprio, GPT-4. ChatGPT *non è* l'intelligenza artificiale vera e propria, ma, per dirla in modo estremamente semplificato, è un servizio che si *collega* a questa I.A. e ne permette un utilizzo comprensibile per l'essere umano, così fa come qualsiasi altro *software* che si lega alle API e ottiene gli stessi risultati. Essendo le API pubblicamente disponibili, ognuno potrebbe a questo punto diventare un produttore, ed essere soggetto agli obblighi dell'I.A. Act[29].

La bozza attuale dell'Atto sull'IA sembra presupporre che il mercato funzioni in modo estremamente più lineare: un venditore di chatbot assembla un'IA e poi rivende un sistema più complesso basato su di essa.
Ma come si è visto, le API stanno diventando sempre più accessibili e configurabili in modo intuitivo anche dagli utenti finali. Di conseguenza, per riflettere meglio la realtà attuale, la bozza dell'Atto sull'IA potrebbe dover superare la distinzione tradizionale tra 'utente' e 'fornitore' per essere realmente efficace[30]. Non solo: perché si superi quel requisito di "trasparenza" e che appaia chiaro dunque che un certo prodotto sia stato generato dall'Intelligenza Artificiale generativa, cosa dovrebbe accadere nel concreto? Che, per esempio, ChatGPT produca ogni volta una stringa di testo che ne segnali l'artificialità, come ad esempio un tag, o un "#bot"? Un'immagine generata dall'IA dovrebbe contenere un *watermark*, cioè una piccola scritta o un simbolo sovrimpresso?

Tutte queste soluzioni sono facilmente eludibili: i watermark possono essere eliminati dall'utente finale, così come i tag o le scritte apposite. Non solo: già dalle prime iterazioni delle IA generative (in primo luogo ChatGPT) proliferano su internet le metodologie con le quali aggirare i blocchi e imporre all'I.A. di produrre output illeciti o sconvenienti[31]. L'unico sistema realmente efficace dovrebbe essere per forza di cose subdolo: uno "stile" immediatamente riconoscibile che dimostri l'artificialità di un certo prodotto. Un modo di esprimersi particolare per i servizi di chat, ad esempio, o certe caratteristiche specifiche in un'immagine generata. Eppure queste caratteristiche sarebbero intanto facilmente intuibili e quindi eludibili, e richiederebbero pertanto un costante aggiornamento da parte degli sviluppatori. Secondariamente, questo imporrebbe alla tecnologia certe caratteristiche in qualche misura politiche, evidentemente non rispettando i requisiti di neutralità tecnologica che dovrebbero essere alla base di una regolamentazione normativa.

Vi è poi un complesso discorso riguardante il potenziale delle I.A. generative. Queste infatti stanno assumendo ruoli sempre maggiori anche all'interno di applicazioni che più probabilmente sarebbero da ricondurre a rischi elevati. In questi casi, le I.A. generative non sono gli strumenti esclusivi con i quali vengono svolti compiti "ad alto rischio", ma possono coadiuvare nel loro svolgimento, piano piano assurgendo alla loro progettazione in astratto onnicomprensiva.

Il punto centrale dell'IA generativa è che può rapidamente affermarsi come sistema di IA generale, pensato cioè per un numero potenzialmente infinito di scopi diversi. L'IA Act prende in considerazione questo tema, e regolamenta, all'articolo 4b la concreta applicazione cui di volta in

[29] In teoria, alcune API, tra le quali la più famosa, cioè quella di OpenAI da cui deriva GPT e ChatGPT, sono a pagamento. Ma non tutte, ed esistono in ogni caso numerosi sistemi per ottenere prove temporanee o utilizzi gratuiti delle stesse. Si veda https://github.com/ayaka14732/ChatGPTAPIFree
[30] M. Veale, F. Zuiderveen Borgesius, *Demystifying the Draft EU Artificial Intelligence Act*, 6 luglio 2021, https://doi.org/10.9785/cri-2021-220402
[31] P. Hacker, A. Engel, M. Mauer, *Regulating ChatGPT and other Large Generative AI Models,* Op. Cit.

volta viene applicata una certa I.A. generativa. I critici dell'AI Act sottolinano come anche per l'IA generativa debbano essere costituite regole apposite[32].

Non solo: i medesimi critici sottolineano come secondo l'attuale versione del testo gli incentivi per i fornitori di I.A. di dotarsi di sistemi di gestione e limitazione del rischio siano fondamentalmente zero. L'AI Act consente al fornitore di "rinunciare" all'applicazione delle disposizioni in materia di rischio elevato escludendo esplicitamente tutti gli usi ad alto rischio nelle istruzioni per l'uso[33]. Non è peraltro ancora del tutto chiaro quali siano gli obblighi degli utenti professionali o finali in tale situazione, e in che modo potrebbero avvenire i controlli che determinino se un certo utente sta utilizzando un'IA generativa per scopi ad alto rischio.

In un simile vuoto normativo, che non rende possibile applicare all'IA generativa usata per scopi ad alto rischio le stesse disposizioni che si rinvengono, all'articolo 4b), per l'IA generale, non può che porsi l'ipotesi alternativa ed estrema che tutti i sistemi di IA generativa rientrino nella categoria ad alto rischio, perché non si può escludere che possano essere utilizzati anche in un'area ad alto rischio.

Sono questi i motivi che spingono numerosi autori a ritenere che sia più opportuno considerare l'I.A. generativa come una categoria a sé stante, e di assoggettarli a obblighi e requisiti legali adatti alle loro caratteristiche[34], considerando anche gli equilibri tra diversi soggetti che utilizzano o subiscono l'uso di tali sistemi: i fornitori, gli utenti professionali e gli utenti comuni. A questo proposito si suggerisce di trarre ispirazione dai già formulati requisiti del Digital Services Act: la predisposizione di meccanismi di segnalazione[35], l'indicazione dei comportamenti ritenuti compatibili con l'uso dello strumento[36], l'obbligo di informare tempestivamente le autorità sulle misure adottate a seguito di contenziosi giudiziali[37], nonchè la predisposizione di sistemi interni di gestione dei reclami[38], e la sospensione o l'allontanamento permanente del contravventore dal servizio[39].

---

[32] C. Djeffal, *The EU AI Act at a crossroads: generative AI as a challenge for regulation,* European Law Blog, 24 luglio 2023, https://europeanlawblog.eu/2023/07/24/the-eu-ai-act-at-a-crossroads-generative-ai-as-a-challenge-for-regulation/

[33] L'Art. 4 (c) dell'IA Act, "eccezioni all'articolo 4b", sottolinea infatti come "l'articolo 4b non applica nel momento in cui il fornitore ha espressamente escluso tutti gli usi ad alto rischio nel manuale di istruzioni o nelle informazioni che accompagnano il sistema di IA generale".

[34] N. Helberger, N. Diakopoulos, *ChatGPT and the AI Act,* Internet Policy Review, Vol 12 n°1, 16 febbraio 2023, https://policyreview.info/pdf/policyreview-2023-1-1682.pdf

[35] Art. 14 DSA, ancora

[36] Art. 12 DSA: "I prestatori di servizi intermediari includono nelle loro condizioni generali informazioni sulle restrizioni che impongono in relazione all'uso dei loro servizi per quanto riguarda le informazioni fornite dai destinatari del servizio. Tali informazioni riguardano tra l'altro le politiche, le procedure, le misure e gli strumenti utilizzati ai fini della moderazione dei contenuti, compresi il processo decisionale algoritmico e la verifica umana. Sono redatte in un linguaggio chiaro e privo di ambiguità e sono disponibili al pubblico in un formato facilmente accessibile."

[37] Art. 8 DSA. Si noti che gli ordini delle autorità nazionali devono essere estremamente specifici, dovendo indicare tra l'altro le motivazioni della decisione, gli URL esatti, eventuali informazioni aggiuntive e la determinazione dei mezzi di ricorso a disposizione del prestatore del servizio e del destinatario del servizio che ha caricato i contenuti illeciti.

[38] Art. 17 DSA, già in parte toccato dal 15 laddove si è parlato di meccanismi di tutela e garanzie interni al servizio. In questo caso dev'essere obbligatoriamente previsto l'accesso ad un sistema interno di gestione dei reclami efficace "che consenta di presentare per via elettronica e gratuitamente reclami" contro le decisioni di sospendere l'accesso alle informazioni, di rimozione delle informazioni, di rimozione o sospensione dell'account dei destinatari (il famoso *ban*).

[39] Art. 20 DSA

È del resto chiaro anche alle voci più critiche dell'IA Act, comunque, che occorrerà un lungo e parallelo studio del fenomeno prima di giungere alla formulazione di norme: in tale senso, non si rilevano ancora proposte concrete.

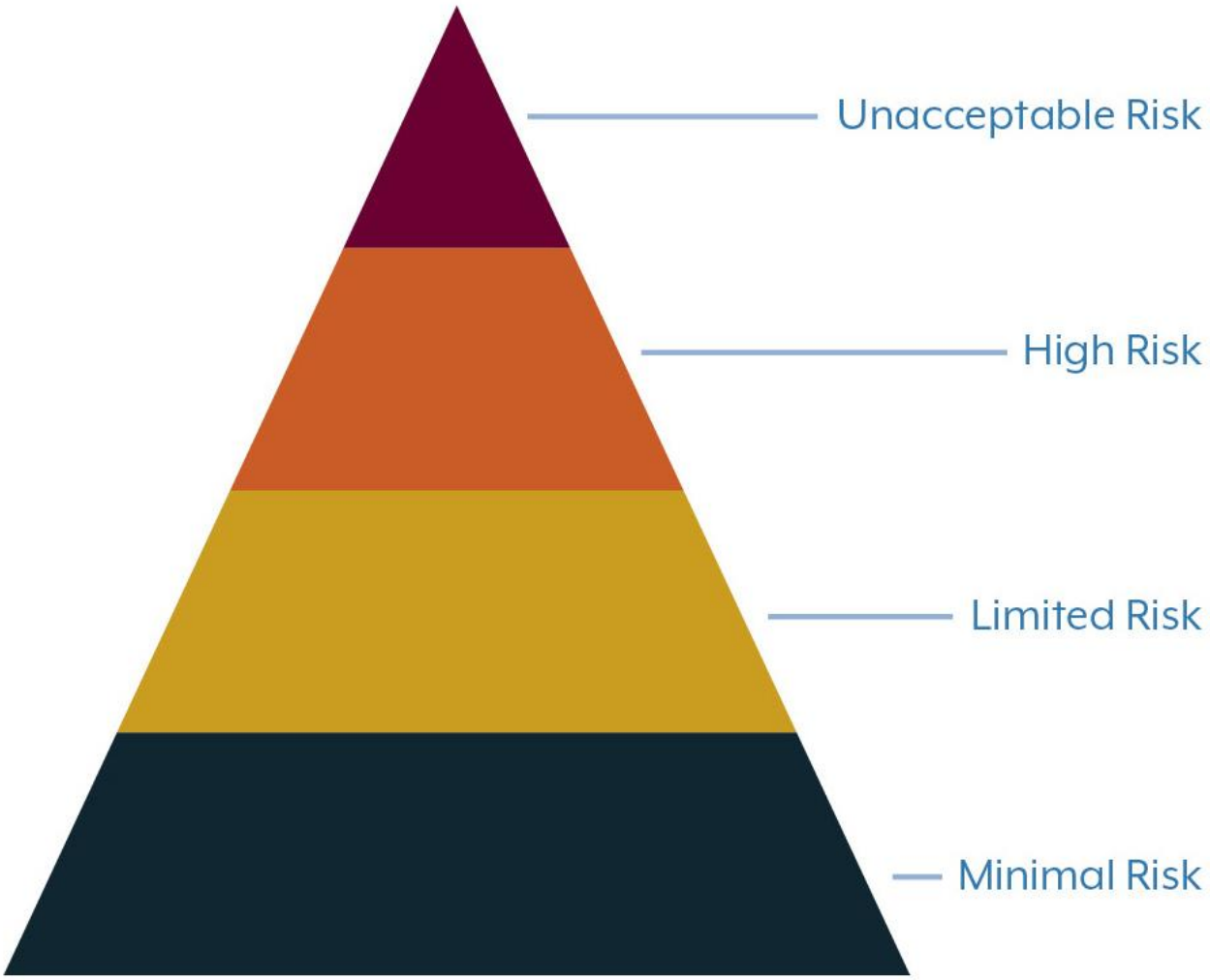


Figura 1: la piramide dell'IA secondo l'AI act. Non è prevista una sezione appositamente dedicata all'IA generativa.

IA LIABILITY DIRECTIVE

Con l'obiettivo di "adattare il diritto privato alle esigenze della transizione digitale"[40], la Commissione Europea l'AI Liability Directive mira a modernizzare il quadro di responsabilità dell'UE introducendo norme specifiche per i danni causati dai sistemi di IA[41]. Questa iniziativa, che si

[40] European Parliament, *Briefing: Artificial Intelligence liability directive,* febbraio 2023, https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/739342/EPRS_BRI(2023)739342_EN.pdf
[41] EU Commission, *Liability Rules for Artificial Intelligence*, 28 settembre 2022, https://commission.europa.eu/business-economy-euro/doing-business-eu/contract-rules/digital-contracts/liability-rules-artificial-intelligence_en

inserisce nell'ambito di un più ampio contesto di transizione verso l'economia digitale, punta a una riforma delle norme sulla responsabilità civile extracontrattuale per rispondere alle sfide poste dall'autonomia, complessità e opacità dei sistemi IA, fattori che rendono inadeguati gli attuali regimi di responsabilità.

Rispondendo agli ostacoli emersi da un'indagine del 2020 e agli impegni presi dalla Commissione von der Leyen a partire dal Libro bianco sull'IA del febbraio 2020, la direttiva mira a garantire benefici economici significativi, stimando un potenziale incremento del valore del mercato dell'IA tra i 500 milioni e 1,1 miliardi di euro. Socialmente, la direttiva si prefigge di rafforzare la fiducia nell'IA attraverso un regime di responsabilità efficiente, che assicuri il rispetto delle richieste di risarcimento, mentre politicamente si propone di ridurre la frammentazione legislativa e incrementare l'efficienza legale nell'UE.

E' effettivamente comprensibile come l'autonomia di queste tecnologie, che permette ai sistemi IA di operare e prendere decisioni senza intervento umano diretto, la loro complessità, data dalla vasta rete di algoritmi e dati che ne costituiscono il fondamento, e l'opacità, ossia la difficoltà di interpretare e comprendere il processo decisionale interno e le modalità di funzionamento dei sistemi IA, pongono questioni giuridiche inedite e complesse che vanno al di là delle capacità delle normative tradizionali di affrontare adeguatamente queste nuove realtà tecnologiche.

In particolare, i regimi di responsabilità attuali, che presuppongono una chiara attribuzione di colpa e una catena di causalità ben definita, si trovano a disagio di fronte a sistemi che possono apprendere, evolvere e operare in modi non completamente prevedibili o spiegabili. In questo contesto, definire chi sia responsabile per i danni causati da un sistema IA, e su quali basi, diventa una questione complessa che può ostacolare lo sviluppo e l'adozione dell'IA stessa, nonchè creare, come teme la Commissione Europea, un generale clima di sfiducia nei confronti della tecnologia.

Per giungere all'attuale proposta di *Liability Directive*, la Commissione Europea ha svolto preliminarmente un'indagine, contrapponendo due blocchi dagli interessi distinti, ed interloquendo tra di loro: da una parte, le associazioni di consumatori; dall'altra, i produttori e le imprese del software che materialmente programmano i sistemi di IA.

Al centro del dibattito tra tali soggetti si è primariamente posta la questione dell'onere della prova in caso di danni causati da sistemi di intelligenza artificiale, cioè il problema di chi, e in quale modo, debba daterminare il collegamento, o nesso causale, tra un danno subito e l'utilizzo del sistema di IA.

Le associazioni di consumatori sostengono fortemente che i produttori, come creatori e utilizzatori di questi sistemi, dovrebbero assumersi una responsabilità oggettiva, simile a quella per le automobili. Questa posizione si basa sulla convinzione che chi beneficia dalla creazione e vendita dell'IA debba anche rispondere dei suoi difetti. D'altra parte, i produttori e le imprese di software ritengono che dovrebbe essere la vittima a provare la presenza di un malfunzionamento e il legame causale con il danno subito, evidenziando le sfide che l'attuale fase di innovazione e incertezza comporta nell'identificare i difetti dei sistemi IA.

Per risolvere questa contrapposizione, la Commissione ha proposto una strada intermedia che ruota sul perno del precedentemente descritto *AI Act,* e sull'introduzione di una presuzione di causalità (si darebbe cioè per assodato il collegamento tra l'uso del sistema di IA ed il danno subito).

Imponendo l'AI Act alcuni specifici requisiti per i sistemi di IA, e specialmente per quelli considerati "ad alto rischio", nel momento in cui un certo sistema non fosse rispettoso di quei requisiti si presumerebbe il nesso di causalità tra il danno e la non conformità. In altre parole, gli utenti che subissero un danno dall'utilizzo di sistemi di IA non conformi con l'AI Act sarebbero esentati dall'onere della prova.

Determinare tale non conformità potrebbe non essere semplice. Si sottolinea quindi l'importanza di considerare come specialmente importanti le "misure di gestione del rischio" adottate dai fornitori di sistemi di IA, riconoscendo pertanto la natura complessa ed in qualche modo parzialmente imprevedibile dei sistemi di IA, e dunque la potenziale difficoltà di individuarne i difetti.

Seppure l'IA Liability Directive non esprima norme specifiche legate all'IA generativa, al netto di questa metodologia può lasciare presumere che i produttori di IA generativa siano incentivati ad adottare misure di mitigazione dei rischi, ed imporre obblighi di tracciabilità durante lo sviluppo del software o durante il suo addestramento, così da ottenere una migliore tutela in caso di contenziosi rivolti contro di loro, oltre che una potenziale maggiore conformità con l'IA Act. Al contempo, qualora il sistema di IA generativa non fosse conforme con l'IA Act, sarebbero teoricamente le vittime a beneficiarne, attraverso la presunzione di causalità ed il conseguente alleggerimento dell'onere della prova.

La direttiva prevede infine un meccanismo di revisione periodica per aggiornare l'ambito di applicazione in base all'evoluzione tecnologica e normativa. Infine, per garantire un'armonizzazione normativa tra gli Stati membri, la direttiva stabilisce una scadenza specifica per il recepimento, allo scopo di istituire un approccio omogeneo e coordinato in tutta l'Unione Europea.

Ancora lontana dalla sua entrata in vigore, la IA liability directive sembra essere passata maggiormente sotto silenzio rispetto al più prorompente IA Act. Non si rileva, peraltro, alcuno studio o considerazione degna di nota per lo specifico campo dell'IA generativa e dell'attribuzione di responsabilità derivante da quella. In effetti, dato il forte collegamento con l'IA Act, non è per forza necessario che l'IA Liability Directive prenda esplicitamente in esame l'IA generativa, essendo sufficiente demandarne la normazione all'IA Act. Tenendo conto di ciò, potrebbe aggiungersi un'ulteriore motivo a sostegno della tesi che vorrebbe, in seno all'IA Act, una regolamentazione specifica e precisa per le IA generative, così da fugare eventuali dubbi potenzialmente derivanti in sede risarcitoria.

Approcci extra UE alla normazione dell'IA: gli Stati Uniti.

Se l'Unione Europea ha determinato da sempre che una regolamentazione *ex-ante* fosse il sistema più efficace per limitare i rischi (ed i conseguenti danni) di nuove tecnologie o pratiche, negli Stati Uniti il sistema di regole sull'I.A. è frammentario ed embrionale, e sembra più improntato ad una tutela *ex-post*, nell'ottica di limitare il minimo indispensabile il progresso tecnologico e la crescita delle industrie basate su queste nuote tecnologie.

Lo scenario normativo, legato in particolare alle norme sulla responsabilità nell'ambito dell'Intelligenza Artificiale, è multisfaccettato e influenzato da numerosi fattori: la volontà di non limitare l'avanzamento tecnologico, come si diceva, ma anche precedenti giurisprudenziali (di notevole importanza in un paese primariamente di Common Law) e linee guida provenienti da istituzioni non specificamente legate all'I.A., ma che si trovano sempre più spesso a dover fare i conti con essa.

Il mercato dell'IA negli Stati Uniti è in continua espansione, ed ha un valore che si presume raggiungerà i 270 miliardi di dollari nel 2026[42]. Rispetto a quanto avviene in Europa, dove come si è avuto modo di vedere l'I.A. Act imporrà alcune stringenti limitazioni rispetto a certi usi della tecnologia, negli Stati Uniti le porte sono spalancate per quasi ogni tipo di sperimentazione ed utilizzo: non per ultimo, quello della *facial recognition technology,* cioè i sistemi di riconoscimento facciale. Solo questo ambito, pare, varrà nel 2026 circa 12 miliardi di dollari, e i prodromi di un suo utilizzo diffuso sono già presenti, ed hanno già contribuito alla risoluzione di casi giudiziari, ma anche all'errata identificazione di soggetti.

Già nel 2015, ad esempio, la *Next Generation Identification Unit* dell'F.B.I. contribuì in modo decisivo al ritrovamento di Lynn Cozart, una guardia di sicurezza condannata per aver molestato tre bambini ed in stato di latitanza dal 1996. Fu un caso piuttosto importante per dimostrare le potenzialità dell'I.A. nella ricerca e nell'interpolazione di volti di persone invecchiati nel tempo. Si dice che il catalogo digitale di facce in mano all'FBI conti qualcosa come 548 milioni di immagini: un numero che non potrebbe essere efficacemente gestito soltanto da esseri umani, ma che raggiunge vette di efficacia infinitamente superiori se coniugato ad un sistema di I.A.[43].

Altra giurisprudenza nel corso del tempo ha contribuito a portare dubbi e specificazioni circa la regolamentazione della robotica computerizzata, anche a partire da casi del passato. Ad esempio, nel caso *Jones v. W + M Automation, Inc*., una denuncia contro un produttore e programmatore di un sistema di caricamento robotico per difetto di prodotto è stata respinta dalla Divisione d'Appello di New York[44]. Secondo il tribunale, gli imputati non erano responsabili per le lesioni del querelante nella fabbrica dove lavorava, poiché avevano dimostrato di aver "prodotto solo componenti non difettosi". Finché il robot e il software associato erano "ragionevolmente sicuri quando progettati e installati", gli imputati non erano responsabili per i danni del querelante. Tuttavia, l'utente finale, poteva essere ancora responsabile per aver modificato impropriamente l'hardware o il software. Questo tenderebbe ad implicare che i creatori di software o hardware AI non siano responsabili per eventuali danni, purché questi prodotti non fossero difettosi quando realizzati. A fronte delle norme attuali e dei precedenti giurisprudenziali, la valutazione se l'AI sia difettosa dipenderà in ultima istanza, come in altri casi di responsabilità del prodotto, dagli standard industriali prevalenti.

---

[42] R. E. Long, *Artificial Intelligence Liability: The Rules Are Changing,* Cyberlaw.stanford.edu, 17 marzo 2023, https://cyberlaw.stanford.edu/blog/2023/03/artificial-intelligence-liability-rules-are-changing-1

[43] E. Markowitz, *How the FBI uses facial recognition technology to fight crime,* Newsweek Magazine, 19 aprile 2016, https://www.newsweek.com/2016/04/29/fbi-biometrics-facial-recognition-next-generation-identification-449146.html

[44] *Thomas E. Jones v W & M Automation, Inc, https://law.justia.com/cases/new-york/appellate-division-fourth-department/2006/2006-05398.html*

E proprio a proposito di questi standard industriali, il *National Institute of Standards and Technology* (NIST), un'agenzia del Dipartimento del Commercio americano, ha rilasciato una guida volontaria per le compagnie che si occupano di tecnologie sensibili dal titolo *Artificial Intelligence Risk Management Framework*[45]. In esso si individuano tre punti fondamentali che, seppure siano ancora semplici consigli e raccomandazioni anziché obblighi propriamente legislativi, possono dare un'idea su ciò che il governo federale degli Stati Uniti potrebbe prendere in considerazione nel momento in cui volesse regolmentare l'I.A. basandosi sugli standard di settore:

1) **Affidabilità**: L'RMF enfatizza il miglioramento dell'affidabilità dei sistemi di IA per mitigare i potenziali danni. Definisce l'affidabilità con sette caratteristiche: sicurezza, protezione e resilienza, spiegabilità e interpretabilità, miglioramento della privacy, equità con gestione dei pregiudizi, responsabilità e trasparenza, validità e affidabilità.
2) **Gestione del rischio**: Il quadro riconosce i rischi unici associati all'IA, come i problemi di privacy dovuti alle capacità di aggregazione dei dati, i problemi di copyright sui set di dati di formazione, i problemi di qualità dei dati e la mancanza di consenso sui metodi di misurazione e sulle metriche.
3) **Gestione del ciclo di vita**: L'RMF delinea quattro funzioni chiave per gestire il rischio durante l'intero ciclo di vita del sistema di IA: mappatura del sistema di IA, misurazione attraverso varie valutazioni, governo con una cultura organizzativa incentrata sulla gestione del rischio di IA e gestione attiva dei rischi del sistema di IA.

Al contempo, e forse con maggiore forza e rilevanza, la Federal Trade Commission (FTC), ha pubblicato numerosi post sul suo sito ufficiale delineando nel corso del tempo un *framework* accettagbile per l'uso e la costruzione dei sistemi di I.A.[46], avendo in questo caso riguardo più della tutela del consumatore[47] che degli interessi economici e tecnologici delle imprese[48]. Come nel caso degli standard del NIST, anche quanto formulato dall'FTC potrebbe essere un indizio su una possibile regolamentazione a livello federale dei sistemi di I.A. negli Stati Uniti. Tale regolamentazione, però, è opportuno ribadirlo, sembra ancora improbabile, ed è invece maggiormente possibile che ogni Stato dell'Unione provveda di per se stesso alla regolamentazione dei sistemi di I.A.

In ogni caso, la FTC enfatizza in particolar modo i seguenti temi:

1) **La Trasparenza**: La FTC sottolinea la necessità di trasparenza nell'utilizzo degli strumenti di IA, soprattutto quando si interagisce con i clienti, ad esempio attraverso i chatbot. Si consiglia alle aziende di non ingannare i consumatori sulle modalità di utilizzo degli strumenti automatizzati e di essere chiari sulla natura delle interazioni guidate dall'IA. Questo vale anche per la raccolta di dati sensibili, per la quale le aziende devono essere caute nel modo in cui ottengono i set di dati per i loro algoritmi. Ingannare i consumatori sui metodi di raccolta dei dati, come nel caso delle impostazioni di riconoscimento facciale di Facebook, può portare a un'azione della FTC.
2) **L'Equità nelle decisioni automatizzate**[49]: si richiede che le decisioni automatizzate, in particolare quelle che riguardano il credito al consumo, l'occupazione, l'assicurazione, l'alloggio o benefici simili,

[45] https://www.nist.gov/itl/ai-risk-management-framework
[46] FTC, *Using Artificial Intelligence and Algorithms,* 8 aprile 2020, https://www.ftc.gov/business-guidance/blog/2020/04/using-artificial-intelligence-and-algorithms
[47] FTC, *Chatbots, deepfakes, and voice clones: AI deception for sale,* 20 marzo 2023, https://www.ftc.gov/business-guidance/blog/2023/03/chatbots-deepfakes-voice-clones-ai-deception-sale
[48] FTC, *Keep your AI claims in check*, 27 febbraio 2023, https://www.ftc.gov/business-guidance/blog/2023/02/keep-your-ai-claims-check
[49] Un concetto piuttosto simile a quanto si rinviene in Europa all'art. 22 del GDPR, che pur non essendo esplicitamente basato sull'Intelligenza Artificiale dichiara che l'interessato ha il diritto di non essere sottoposto a decisioni basate unicamente sul trattamento automatizzato se tali decisioni producono effetti giuridici che lo riguardano o, analogamente, incidono sulla sua persona in modo significativo.

siano eque. Se una decisione viene presa sulla base di informazioni provenienti da un fornitore terzo, il consumatore deve ricevere uno specifico avviso che spieghi le ragioni della decisione. Le aziende devono essere in grado di spiegare in modo specifico perché a un consumatore è stato negato un servizio o un prodotto sulla base di un processo decisionale algoritmico.

3) **Il principio di non discriminazione:** L'IA non deve discriminare le classi protette. La FTC applica leggi come l'Equal Credit Opportunity Act (ECOA) e il Civil Rights Act del 1964 per evitare decisioni guidate dall'IA che potrebbero risultare discriminatorie. Le aziende sono incoraggiate a testare rigorosamente i loro algoritmi per garantire che non creino un impatto disparato sulle classi protette. La FTC esamina anche gli input e i risultati dei modelli di IA per valutare la discriminazione illegale, valutando se un modello di fatto neutro abbia un impatto disparato illegale sulle classi protette.
4) **L'Accuratezza e solidità empirica**: Le aziende che forniscono dati sui consumatori per decisioni basate sull'IA devono garantire l'accuratezza e l'integrità dei dati. Ai sensi del FCRA, devono implementare procedure per mantenere la massima accuratezza possibile delle relazioni sui consumatori e dare ai consumatori la possibilità di correggere le inesattezze. I modelli di IA devono essere convalidati e riconvalidati per garantire che funzionino come previsto e non facciano discriminazioni illegali, aderendo agli standard di "derivazione empirica, dimostrabile e statisticamente valida".
5) **Uso etico e responsabilità**: Le aziende devono considerare le implicazioni etiche e le questioni di equità nei loro algoritmi di IA. Prima di implementare i sistemi di IA, è necessario porsi domande sulla rappresentatività degli insiemi di dati, sulla considerazione dei pregiudizi, sull'accuratezza delle previsioni e sui potenziali problemi etici. Inoltre, le aziende che sviluppano l'IA dovrebbero proteggere i loro algoritmi dall'uso e dall'abuso non autorizzati e disporre di meccanismi di responsabilità, eventualmente coinvolgendo standard o competenze indipendenti.

Come si noterà, nello specifico di responsabilità si parla soltanto all'ultimo punto, e nell'ambito di un più ampio discorso inerente all'etica e ai possibili *bias* che i dati sui quali gli algoritmi di I.A. sono allenati possano portare.

In assenza di un quadro federale completo, i singoli Stati affrontano singolarmente temi regolamentari legati all'IA. Questo approccio si traduce in un panorama normativo diversificato e frammentato: proprio ciò che, al contrario, l'UE intende evitare. Ad esempio, la città di New York impone verifiche sui pregiudizi per gli strumenti automatizzati di decisione sull'impiego, mentre in Colorado vigono leggi che proteggono dalle pratiche assicurative sleali guidate dall'IA.

Come peraltro evidenziato in nota 48, è poi spesso la precedente legislazione sulla Privacy a dare spunti e strumenti per regolamentare l'Intelligenza Artificiale e le responsabilità derivanti da tali sistemi. Del resto, anche la legislazione sulla privacy si occupa spesso di proteggere i dati personali degli utenti, ed è ben noto che i sistemi di I.A. siano allenati proprio su ingenti quantità di questi stessi dati. Il CCPA in California, probabilmente la più avanzata legislazione degli Stati Uniti su questo tema, stabilisce ad esempio linee guida per la raccolta, l'uso e la condivisione di dati che sono applicabibli direttamente anche ai sistemi di I.A..

Lo stesso è opportuno dire per ciò che riguarda gli standard di trasparenza e di consenso, nonché di responsabilità per violazioni degli obblighi sulla buona tenuta (il "trattamento", diremmo in UE) dei dati. Si tratta di legislazioni parallele, più mature e comunque in buona parte compatibili con le nuove sfide derivanti dall'I.A., e pertanto si possono considerare, al pari delle raccomandazioni delle agenzie federali, come possibili spunti per una eventuale regolamentazione federale onnicomprensiva o per le frammentarie norme statali che inevitabilmente sorgeranno nel corso dei prossimi anni.

Si evidenzia infine come gli Stati Uniti, in ambito istituzionale, si stiano orientando ad uno sviluppo ed utilizzo responsabile dell'IA a tutela dei cittadini e contro ogni discriminazione[50] e si stiano mostrando inclini a lavorare in modo collaborativo sull'IA, con una tendenza crescente ad allinearsi agli approcci internazionali, in particolare dell'Unione Europea.

Approcci extra UE alla normazione dell'IA: la Cina.

Una terza via alla regolamentazione dell'I.A., e dell'I.A. generativa nello specifico, è data dalla Cina.

Già da molti anni l'attenzione della politica cinese sul tema è stata evidente. A partire dal 2017, con il *New Generation AI Development Plan*[51]*,* state promulgate numerose leggi e regolamenti che influenzano lo sviluppo dell'IA ed i requisiti che i prodotti o i servizi derivanti da essa debbono avere per poter essere immessi sul mercato.

In generale, fino al 2020 si rileva un approccio curiosamente simile a quello usualmente seguito in Occidente per la regolamentazione e l'investimento in queste tecnologie. I primi atti specificamente dettati per l'I.A. in Cina richiedono il rispetto della privacy e la necessità di essere sicuri e controllabili[52], nonché incoraggiano, tramite l'erogazione di fondi e lo sviluppo di piani decennali, lo sviluppo del settore. Si arriva anche alla raccomandazione di norme etiche che dovrebbero essere incorporate nello sviluppo e nella creazione stessa dei sistemi di IA, e alla previsione del principio secondo il quale gli esseri umani debbono mantenere il controllo sui sistemi di IA ed essere infine responsabili delle conseguenze del loro utilizzo[53].

Si rileva però anche una forte preoccupazione nel normare rapidamente situazioni di fatto, che concretamente incidono sul percepito interesse nazionale, o creano nel territorio della Cina allarme sociale. Alla prima categoria appartengono le regolamentazioni rapidamente emanate nel 2021 per controllare la diffusione di contenuti online tramite algoritmi, contenente numerosi vincoli di controllo del contenuto mostrabile ma anche tutele per potenziali conseguenze lavorative determinate dall'uso di algoritmi di IA[54], oppure quelle del 2022 legate all'utilizzo dell'IA in settori di particolare interesse, come la medicina e le scienze[55].

Alla seconda appartengono invece le norme più strettamente legate all'IA generativa. Nel 2022, la Cina ha inteso regolare le applicazioni IA utilizzate per generare testo, video e audio, correndo ai ripari dopo l'esplosione di popolarità nel paese di tali sistemi, ed il loro utilizzo illecito per la creazione di notizie false online, e per inganni perpetrati tramite i *deepfake*. A seguito del provvedimento emanato, il contenuto generato artificialmente in Cina deve essere obbligatoriamente indicato come tale, apponendo una

---

[50] Annuncio del 4 maggio 2023 dell'Amministrazione Biden-Harris su *New Actions to Promote Responsible AI Innovation that Protects Americans' Rights and Safety*, https://www.whitehouse.gov/briefing-room/statements-releases/2023/05/04/fact-sheet-biden-harris-administration-announces-new-actions-to-promote-responsible-ai-innovation-that-protects-americans-rights-and-safety/

[51] *China's New Generation Artificial Intelligence Development Plan (2017),* https://digichina.stanford.edu/work/full-translation-chinas-new-generation-artificial-intelligence-development-plan-2017/, traduzione del documento a cura dell'Università di Stanford.

[52] *Governance Principles for New Generation AI: Develop Responsible Artificial Intelligence, https://digichina.stanford.edu/work/translation-chinese-expert-group-offers-governance-principles-for-responsible-ai/*

[53] *Ethical Norms for New Generation AI*, https://cset.georgetown.edu/publication/ethical-norms-for-new-generation-artificial-intelligence-released/ traduzione a cura del CSET – Center for Security and Emerging Technology

[54] *Provisions on the Management of Algorithmic Recommendations in Internet Information Services,* Traduzione del document a cura di *China Law Translate*, https://www.chinalawtranslate.com/en/algorithms/

[55] *Opinions on Strengthening the Ethical Governance of Science and Technology,* riassunto dal CARNEGIE-Endowment for International Peace, in M. Sheehan, *China's AI Regulations and How they get made,* Carnegie, 10 luglio 2023, https://carnegieendowment.org/2023/07/10/china-s-ai-regulations-and-how-they-get-made-pub-90117

"indicazione apposita" in un "luogo ragionevolmente posizionato" all'interno del contenuto generato[56]. Una soluzione, questa, che come si è visto è stata ritenuta possibile anche in Europa. In generale, comunque, l'atto normativo è interessante e merita una visione più approfondita, rappresentando un *framework* piuttosto completo per la regolamentazione dei servizi di IA generativa:

- È esplicitamente proibito l'utilizzo di tali sistemi per produrre o diffondere informazioni vietate dalla legge o dai regolamenti, e per trasmettere informazioni false (*fake news*).
- I provider dei servizi di IA generativa dovranno sempre assicurarsi di conoscere la catena di responsabilità che conduce all'utilizzatore illecito: pertanto, la registrazione degli utenti deve essere sottoposta alla verifica della reale identità dell'utente.
- I servizi di IA dovranno essere dotati di sistemi per identificare gli *input* degli utenti e poterli pertanto revisionare.
- È prevista una "responsabilità alla smentita" affidata ai provider dei servizi di IA: qualora una certa IA generativa fosse usata per generare un'informazione illecita, sarebbe compito del fornitore di quella IA smentire la notizia, confermandola come falsa e generata dal software.
- È fatto obbligo ai provider dei servizi di IA di dotarsi di sistemi di gestione dei reclami degli utenti.

Successivamente, in seguito di alcuni casi che nella tarda primavera del 2023 hanno causato un certo allarme sociale nel paese[57], la Cina ha deciso di agire ancora sull'I.A. generativa, dimostrando ormai di considerarla definitivamente come una categoria seperata dalla più ampia e generica categoria dell'IA, a differenza di quanto avviene in Europa[58].

La conseguente produzione normativa ha determinato l'entrata in vigore delle *Interim Measures for the Management of Generative AI Services*, il 15 agosto 2023.

Tali *Interim Measures* si applicano alle tecnologie di AI generativa utilizzate per fornire servizi al pubblico in Cina. Da ciò discende una doppia limitazione della normativa: essa non si applica a quante imprese sviluppino e svolgano ricerca sui sistemi di IA senza offrire servizi al pubblico (dimostrando in tal modo la preoccupazione del governo di Pechino di incentivare il settore il più possibile), e non si applica a quanti sistemi non siano esplicitamente definiti come di IA *Generativa*. Tuttavia, si segnala come la definizione di "IA generativa" data nelle *Interim Measures* sia piuttosto vaga: "quei modelli e tecnologie capaci di generare contenuti come testo, audio, immagini o video".

Le *Interim Measures* si applicano ai provider di servizi di IA generativa, cioè le "organizzazioni o individui che usano IA generativa per fornire servizi basati sulla stessa", inclusi quelli che forniscono tali servizi tramite interfacce programmabili (API). Inoltre, si applicano anche ai servizi forniti dall'estero che non rispettano le leggi cinesi, determinando la facoltà per le autorità cinesi di bloccare l'accesso a tali servizi nel paese.

Gli obblighi principali dei fornitori includono:

- Effettuare valutazioni di sicurezza secondo le normative esistenti e registrare informazioni sull'uso degli algoritmi presso gli enti designati;

---

[56] *Provisions on the Administration of Deep Synthesis Internet Information Services,* https://www.chinalawtranslate.com/en/deep-synthesis/

[57] E. Cao, E. Baptista, *Deepfake scam in China fans worries over AI-driven fraud*, Reuters, 2 maggio 2023, https://www.reuters.com/technology/deepfake-scam-china-fans-worries-over-ai-driven-fraud-2023-05-22/

[58] Liu, I.; D. Edmondson, *China: New Interim Measures to Regulate Generative AI*, Baker McKenzie, Agosto 2023, https://insightplus.bakermckenzie.com/bm/attachment_dw.action?attkey=FRbANEucS95NMLRN47z%2BeeOgEFCt8EGQJsWJiCH2WAWuU9AaVDeFglGa5oQkOMGl&nav=FRbANEucS95NMLRN47z%2BeeOgEFCt8EGQbuwypnpZjc4%3D&attdocparam=pB7HEsg%2FZ312Bk8OIuOIH1c%2BY4beLEAezirm3%2BK7wMU%3D&fromContentView=1#:~:text=The%20Interim%20Measures%20impose%20various,through%20programmable%20interfaces%20(APIs)%2C

- Assumere responsabilità di "elaboratore di informazioni personali" secondo la Legge sulla Protezione delle Informazioni Personali (PIPL), gestendo i dati personali in modo conforme alle leggi cinesi;
- Gestire i dati di addestramento e i modelli di base in conformità con la legge, rispettando i diritti di proprietà intellettuale e migliorando la qualità dei dati di addestramento;
- Rispettare i valori socialisti e non generare contenuti che minacciano la sicurezza nazionale o promuovono violenza, oscenità, falsità, ecc;
- Adottare misure contro contenuti illegali e situazioni di utilizzo illegale dei servizi AI da parte degli utenti;
- Etichettare i contenuti creati dall'AI generativa e prevenire la discriminazione nel design degli algoritmi e nella selezione dei dati di addestramento.

Come si vede, questi sono sostanzialmente degli obblighi rafforzati rispetto a quelli già previsti per i fornitori di servizi di IA non generativa. La differenza forse più importante, oltre all'ambito applicativo che attiene anche alle imprese estere che forniscono servizi in Cina, sta nell'ormai nota questione della responsabilità.

Essendo gli obblighi stabiliti a monte, gli sviluppatori di AI generativa potrebbero infatti essere ritenuti responsabili per problemi che emergono dalle azioni di aziende che utilizzano i loro strumenti AI. Questo significherebbe che se un'azienda cliente utilizza uno strumento AI generativo per fornire servizi e, a causa di quel servizio, si verificano problemi legali o danni, lo sviluppatore dello strumento AI potrebbe essere ritenuto parzialmente o totalmente responsabile. Gli stessi utenti finali possono avere il diritto di agire legalmente contro gli sviluppatori AI per eventuali danni o problemi causati dall'utilizzo degli strumenti AI.

Per risolvere questa attribuzione di responsabilità eccessivamente legata agli sviluppatori originari, si discute sul fatto che sarà necessario dotarsi di accordi chiari e specifici nei contratti tra gli sviluppatori di AI e le aziende che forniscono servizi al pubblico. Questi contratti dovrebbero delineare chi è responsabile in varie situazioni e come si gestiscono le eventuali problematiche legali.

Si può quindi osservare come anche in Cina permanga una situazione non precisamente lineare sull'attribuzione di responsabilità per uso di IA generativa, che può dare origine a incertezza e a squilibri sul mercato. Incertezza, questa, che è acuita dalla discrezionalità garantita all'autorità regolatrice cinese, la C.A.C., dalle *Interim Measures* in caso di vuoto normativo o di situazioni di dubbio. In quelle circostanze, l'autorità regolatrice può intervenire anche d'ufficio con un richiamo scritto, prescrivendo adempimenti alle imprese private.

In definitiva la Cina ha delineato un percorso regolamentare distintivo che evidenzia sia il suo impegno nel promuovere l'innovazione tecnologica sia la sua prontezza nel gestire i rischi e le sfide etiche connesse. Nonostante la volontà di stimolare lo sviluppo del settore dell'IA, il governo cinese ha mostrato, almeno sulla carta, un'attenzione particolare nel tutelare la sicurezza nazionale, la tutela dalle *fake news,* e la privacy (in quest'ordine), stabilendo normative che richiedono trasparenza e responsabilità dagli sviluppatori e fornitori di servizi IA[59].

L'approccio cinese, caratterizzato da una marcata responsabilità attribuita agli sviluppatori e una sorveglianza stringente sui contenuti generati, pone interrogativi significativi sul bilanciamento tra libertà di espressione, innovazione tecnologica e controllo statale. Inoltre, la complessità della catena di responsabilità in questo ambito solleva questioni legali e contrattuali rilevanti, specialmente per le aziende internazionali che mirano a operare nel mercato cinese: una incertezza acuita dalla discrezionalità lasciata alle autorità regolatrici, come la CAC, possono creare un ambiente di business imprevedibile, richiedendo alle aziende una maggiore attenzione e adattabilità nella navigazione di questo panorama legislativo in rapida evoluzione.

[59] M. Sheehan, *China's AI Regulations and How they get made,* Carnegie, 10 luglio 2023, https://carnegieendowment.org/2023/07/10/china-s-ai-regulations-and-how-they-get-made-pub-90117

Conclusioni

La regolamentazione dell'intelligenza artificiale (IA) generativa si colloca all'incrocio tra progresso tecnologico e imperativi legali, sfidando i confini tradizionali della responsabilità e della tutela dei diritti. L'Unione Europea, con il suo AI Act, ha stabilito un paradigma normativo che si concentra sulla classificazione basata sul rischio e impone obblighi stringenti per garantire sicurezza e trasparenza. Questo approccio riflette un impegno per una regolamentazione olistica che mira a bilanciare innovazione e protezione dei consumatori.

Negli Stati Uniti, la mancanza di un quadro normativo omogeneo per l'IA genera un panorama più frammentato. L'attenzione si concentra sulla tutela dei consumatori, sulla non discriminazione e sul rispetto della privacy, riflettendo un approccio più reattivo che proattivo. Questo lascia spazio a interventi normativi settoriali e statali, i quali, pur essendo efficaci in specifici contesti, potrebbero non fornire una coerenza e una prevedibilità necessarie a livello nazionale.

In Cina, la regolamentazione dell'IA generativa si distingue per il suo controllo incisivo e la sua focalizzazione sulla sicurezza nazionale e il controllo sociale. Le misure adottate riflettono la volontà di governare strettamente le tecnologie emergenti, pur promuovendo l'innovazione. Questo approccio potrebbe servire da modello per altri paesi che cercano di bilanciare la crescita tecnologica con la sicurezza e la stabilità sociale, ma potrebbe anche limitare la libertà di espressione e l'innovazione in determinati ambiti.

La varietà di questi approcci sottolinea la sfida di sviluppare un quadro normativo per l'IA generativa che sia efficace a livello globale. La questione della responsabilità nell'uso dell'intelligenza artificiale generativa si rivela essere un tema cruciale e complesso.

In conclusione, la regolamentazione dell'IA generativa rimane un terreno in continua evoluzione, che richiede una riflessione costante e un aggiornamento delle normative per tenere il passo con i rapidi sviluppi tecnologici. La responsabilità, intesa sia come conformità legale che come considerazione etica, diventa un fattore chiave per garantire che i benefici dell'IA generativa siano accessibili a tutti, minimizzando al contempo i rischi e proteggendo i diritti fondamentali.