

DIPARTIMENTO DI GIURISPRUDENZA

Cattedra di Metodologia della scienza giuridica

Intelligenza artificiale e regolamentazione:
un'analisi comparatistica alla ricerca di un
equilibrio tra regole e innovazione

Chiar.mo Prof. Antonio Punzi

RELATORE

Chiar.mo Prof. Filiberto Brozzetti

CORRELATORE

Bianca Stella Bruschi
Mat. 155003

CANDIDATA

Anno accademico 2022/2023

Indice

INTRODUZIONE	4
CAPITOLO 1: L'INTELLIGENZA ARTIFICIALE E IL SUO UTILIZZO	6
1. <i>STORIA DEL PENSIERO ARTIFICIALE: GLI ALBORI DELL'IA</i>	6
2. <i>LA COMPLESSITÀ NEL DEFINIRE IL CONCETTO DI IA</i>	10
2.1 LA DEFINIZIONE FILOSOFICA (IA INGEGNERISTICA, IA COGNITIVA).....	13
2.2 LE DEFINIZIONI TECNICHE.....	16
2.2.1 <i>Le modalità di apprendimento: il Machine Learning nelle sue varie declinazioni e il c.d. “black box effect”</i>	19
2.2.2 <i>Il carburante dell'IA: i dati (e i loro potenziali “bias”)</i>	25
2.3 LA DEFINIZIONE NORMATIVA.....	35
3. <i>LE OPPORTUNITÀ E I RISCHI DELL'INTELLIGENZA ARTIFICIALE</i>	41
3.1 LE OPPORTUNITÀ: APPLICAZIONI NEL SETTORE DELLA MEDICINA, DEL LAVORO, E DELLA CRISI CLIMATICA.....	44
3.2 LE CATEGORIE DEI RISCHI “ATTUALI” DELL'IA: GLI ELEMENTI CRITICI INTRINSECI E QUELLI ELEMENTI DERIVANTI DA PRECISE APPLICAZIONI DEI SISTEMI.....	53
3.3 UN CENNO AI RISCHI VIRTUALI: LA SINGOLARITÀ TECNOLOGICA.....	63
CAPITOLO 2: LA REGOLAMENTAZIONE DELL'IA NELL'UNIONE EUROPEA	66
1. <i>PREMESSE: LA CORSA ALLA REGOLAMENTAZIONE DELL'IA E IL MODELLO REGOLAMENTARE EUROPEO NEL CAMPO DELL'ECONOMIA DIGITALE</i>	66
2. <i>LA STRADA PER LA PROPOSTA DI REGOLAMENTO (2017-2021); LA SCELTA DELLO STRUMENTO LEGISLATIVO DA ADOTTARE E LA BASE GIURIDICA IDONEA</i>	73
3. <i>L'AI ACT (2021)</i>	81
3.1 IL PERCORSO NORMATIVO DELL'AI ACT.....	81
3.2 GLI OBIETTIVI PERSEGUITI DALLA PROPOSTA DI REGOLAMENTO: QUELLI ESPLICITI, IL “FIRST MOVER ADVANTAGE,” E LA RICERCA DI UN “EFFETTO BRUXELLES”.....	84
3.3 IL CONTENUTO DELLA PROPOSTA DI REGOLAMENTO.....	89
3.3.1 <i>L'ambito di applicazione e definizioni rilevanti</i>	89
3.3.2 <i>L'approccio risk-based: pratiche vietate, sistemi ad alto rischio e a rischio minimo</i> ..	91
3.3.3. <i>Gli obblighi di trasparenza per sistemi intermedi e regolamentazione dell'IA generativa</i>	99
3.3.4 <i>Le misure a sostegno dell'innovazione e governance dell'IA</i>	101
3.4 SPUNTI DI CRITICA DELL'AI ACT.....	103
3.5 CONCLUSIONI SULL'AI ACT: IL RAPPORTO TRA AI ACT ED INNOVAZIONE; LA CONCRETA POSSIBILITÀ DI UN EFFETTO BRUXELLES.....	108
4. <i>LE DIRETTIVE SULLA RESPONSABILITÀ CIVILE DEI SISTEMI DI IA</i>	112
CAPITOLO 3: LE ALTRE POTENZE MONDIALI E LA REGOLAMENTAZIONE IDEALE DELL'IA	117
<i>PARTE I – GLI STATI UNITI D'AMERICA</i>	120
I.1 LA POSIZIONE STATUNITENSE NEL MERCATO DELL'IA E L'APPROCCIO ALLA REGOLAMENTAZIONE.....	120
I.2 LE INIZIATIVE DURANTE LA PRESIDENZA DONALD TRUMP (2016-2020).....	123
I.3 LA BIDEN ADMINISTRATION: DALL'AI BILL OF RIGHTS ALL'EXECUTIVE ORDER No. 14110.....	125
<i>PARTE II – LA REPUBBLICA POPOLARE CINESE</i>	135
II.1 IL MODELLO “STATE-DRIVEN” CINESE NELL'ECONOMIA DIGITALE E LO STATO DEL MERCATO DELL'IA IN CINA.....	135
II.2 LA REGOLAMENTAZIONE DELL'IA IN CINA.....	140
<i>PARTE III – COMPARAZIONE, SPUNTI DI REGOLAMENTAZIONE E CONCLUSIONI</i>	146
III. 1 COMPARAZIONE, CONVERGENZE A LIVELLO INTERNAZIONALE E STANDARDIZZAZIONE... ..	146

III. 2 ALLA RICERCA DI UN EQUILIBRIO REGOLAMENTARE OTTIMALE: DALLE VALUTAZIONI D'IMPATTO SUI DIRITTI FONDAMENTALI E SULL'ETICA (HRESIA) ALL'AI LITERACY	151
III. 3 CONCLUSIONI	156
CONCLUSIONE	161
BIBLIOGRAFIA.....	163

Introduzione

L'intelligenza artificiale (IA) è una tecnologia rivoluzionaria che sta rapidamente trasformando la società globale. Le sue applicazioni sono innumerevoli e promettono di migliorare numerosi aspetti della vita quotidiana nei settori più svariati: dalla medicina all'industria, dalla sostenibilità all'istruzione. Essa consiste in uno strumento tecnologico assolutamente polifunzionale che può essere implementato con successo in un'ampia gamma di processi, con straordinarie possibilità di progresso ed innovazione. Tuttavia, l'ascesa dell'IA pone anche nuove sfide e interrogativi di natura giuridica, soprattutto in relazione ai rischi che essa è in grado di porre a livello individuale e collettivo, sia nei confronti dei diritti e delle libertà dei singoli, che anche nei confronti dello stato di diritto e della democrazia a livello globale.

Il presente elaborato mira a fornire un quadro complessivo della regolamentazione in materia di IA, tramite un'analisi comparativa tra i vari approcci assunti da una serie di entità sovranazionali. In *primis*, il Capitolo 1 si propone un'analisi della natura dell'IA il più multidisciplinare possibile, comprendendo elementi di natura prettamente filosofica, nonché relativa all'informatica, e, in ultimo, ai risvolti di natura normativa. A ciò segue la trattazione delle opportunità e dei rischi attuali e virtuali che l'IA pone in una serie di settori; la corretta focalizzazione di questi elementi è prodromica e fondamentale per procedere ad una trattazione informata in tema degli obbiettivi stessi della regolamentazione.

Il Capitolo 2, di seguito, propone un'analisi dell'approccio dell'Unione Europea alla regolamentazione dell'IA. L'UE ha redatto la prima bozza di regolamento europeo indirizzato alla disciplina dell'IA in maniera orizzontale e sostanzialmente omnicomprensiva nel 2021 (*AI Act*). Sebbene la normativa europea in questo settore non sia ancora in vigore, le vicende legislative dell'*AI Act* sono di necessaria analisi per comprendere anche il contenuto del regolamento stesso. Il capitolo prosegue con una puntuale analisi della struttura normativa del regolamento, quest'ultimo modellato su un approccio in base al rischio posto dai vari sistemi di IA. Si prosegue poi con la valutazione di tutti quei meccanismi inseriti nel regolamento che hanno lo scopo di fornire una disciplina all'IA, controbilanciando questa necessità con la conservazione e la stimolazione di un ecosistema di innovazione e sviluppo a livello europeo, e che consenta all'UE di

essere riconosciuta contemporaneamente come pioniera nel campo della produzione e nella contestuale regolamentazione dell'IA. Per completare il quadro attinente alla regolamentazione europea, viene inoltre allargata la trattazione a quegli strumenti di diritto comunitario che si prefiggono l'obiettivo di disciplinare l'attività degli strumenti di IA *ex post*, cioè incidendo sul tema della responsabilità civile derivante dall'utilizzo degli strumenti di IA complessivamente considerati.

Il Capitolo 3, in ultimo, è suddiviso in tre parti: la prima attinente alla trattazione della regolamentazione dell'IA negli Stati Uniti, e il percorso da questi affrontato nell'arrivare alla determinazione di una cornice disciplinare, ancorché non di matrice legislativa ma piuttosto prettamente amministrativa, tramite l'emanazione dell'*Executive Order* No. 14110. La seconda parte, analizza la disciplina adottata finora dalla Repubblica Popolare Cinese, con lo scopo di evidenziare quali scelte siano state effettuate da un altro dei grandi *players* del mercato dell'IA, in maniera da operare un confronto trilaterale tra i tre modelli regolamentari proposti dalle entità sovranazionali in analisi.

La terza parte del Capitolo 3, infine, propone alcuni spunti di regolamentazione da implementare nelle varie soluzioni che sono state proposte dai singoli stati, nell'ottica di stimolare quanto più possibile un ambiente di innovazione in completo rispetto dei diritti fondamentali, nonché dei principi di giustizia sociale, democrazia e rispetto dello stato di diritto, alla ricerca di un equilibrio efficace tra regolamentazione e innovazione, non a scapito delle libertà e dei diritti dei singoli.

Il presente elaborato si avvale di un approccio assolutamente interdisciplinare, combinando l'analisi giuridica con quella di matrice sociologica, economica e tecnologica. Il risultato auspicato è quello di fornire una visione completa e aggiornata del tema della regolamentazione dell'IA, utile per futuri sviluppi normativi e per la costruzione di un futuro in cui l'IA sia una forza quanto più positiva per il bene dell'umanità complessivamente considerata.

Capitolo 1: l'Intelligenza Artificiale e il suo utilizzo

1. Storia del pensiero artificiale: gli albori dell'IA

Il concetto filosofico della possibile creazione di un'“Intelligenza Artificiale,” una *machina* in grado di performare attività con spiccate capacità tipicamente attribuibili agli esseri umani precede tutto ciò che viene cognitivamente associato al concetto di “IA” odierno. Volendo abbandonarsi in una retrocessione temporale significativa, nella Grecia ellenistica era già esistente un dibattito scientifico e filosofico sui c.d. *αὐτόματα*, le prime entità che svolgevano in maniera autonoma delle operazioni di calcolo sulla base di *input* esterni. D'altronde, il più antico “calcolatore” automatico risale proprio al I secolo a.C.: la cosiddetta Macchina di Anticitera, capace di prevedere eventi astronomici come il sorgere del sole, le fasi lunari, i movimenti planetari, gli equinozi e le date dei giochi olimpici nell'antica Grecia¹.

Prima ancora che la possibilità di costruire quelle tecnologie che oggi sbalordiscono per la loro capacità di eseguire operazioni ad alta complessità (si pensi ai sistemi di IA generativa²), prima ancora di costruire anche quei macchinari che sono passati alla cronaca come i primi sistemi propriamente “intelligenti,”³

¹ Carman, C. C., & Evans, J. (2014). On the epoch of the Antikythera mechanism and its eclipse predictor. *Archive for History of Exact Sciences*, 68(6), 693–774. <https://doi.org/10.1007/s00407-014-0145-5>.

² Ai fini di questa analisi introduttiva, è sufficiente far riferimento alla definizione di intelligenza artificiale generativa così come formulata dalle istituzioni europee nel gruppo di emendamenti presentati nel maggio 2023 alla Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce norme armonizzate in materia di intelligenza artificiale: essa è definita come l'insieme di “*sistemi di intelligenza artificiale specificamente destinati a generare, con vari livelli di autonomia, contenuti come testi complessi, immagini, audio o video.*” (Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale e modifica alcuni atti legislativi dell'Unione {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final, 21 aprile 2021 (d'ora in avanti “AI Act”), Art. 28b(4), così come strutturato nella Versione II del Parlamento Europeo. (per le varie versioni dell'*AI Act*, si veda il Capitolo 2, paragrafo 3.1).

Alcuni esempi di IA generativa meritevoli di essere menzionati sono l'IA conversazionale “ChatGPT” (generativa di testo) e “DALL-E” (generativa di immagini e contenuti multimediali a seconda del *prompt* fornito dall'utente), entrambi sviluppati dalla società OpenAI. Vi sono anche altri sistemi creati da società e start-up di dimensioni minori rispetto alle aziende c.d. *Big tech*, tra cui spiccano progetti come ad esempio “Pi AI”, bot conversazionale prodotto dalla società benefit statunitense Inflection AI, e noto nell'opinione pubblica come un'IA conversazionale “gentile.”

³ Si pensi ad esempio ad uno dei sistemi di intelligenza artificiale dall'impatto storico più noto, il computer Deep Blue, progettato dall'IBM e capace di vincere una partita a scacchi contro il campione del mondo Garry Kasparov nel 1997. Sebbene calcolatori come questo siano esempi di come la complessità applicata ai macchinari computazionali fosse già in grado di performare attività complesse come eccellere negli scacchi, i primi veri “sistemi intelligenti” immaginati dai pionieri di questo settore risalgono a più di quarant'anni prima. Studiosi ancorchè visionari nell'America

L'uomo moderno si è *in primis* interrogato sul rapporto tra corpo fisico e mente, sulla natura della cognizione umana e, successivamente, sulla capacità di replicare quest'ultima in entità separate, artificiali e automatizzate. Se le modalità di apprendimento e ragionamento tipicamente qualificate come umane sono infatti replicabili in sistemi che operino in tutto e per tutto alla stregua di “menti”, sarebbe possibile dunque realizzare veri e propri “cervelli artificiali” basati sull'utilizzo del silicio come elemento chimico dal quale scaturiscono le proprie funzionalità, anziché sul carbonio, come le forme di vita organiche.

Il metodo di costruzione e programmazione di sistemi intelligenti è radicalmente cambiato nel corso dei decenni. Alla fine degli anni '50, nel periodo di nascita effettiva di quella branca dell'informatica che è la disciplina dell'IA, il primo approccio degli informatici statunitensi fu quello di creare programmi in grado di elaborare “ragionamenti” sulla base di un'immensa serie di computazioni dettagliatamente fornite dal programmatore alla macchina⁴. Inoltre, questi modelli cercavano di emulare artificialmente il ragionamento umano senza tuttavia riprodurre le sottostanti strutture biologiche che ne consentono lo sviluppo⁵. L'obbiettivo era creare dei c.d. *expert systems*, sistemi informatici articolati che potessero operare a fianco degli esseri umani in campi di particolare complessità e rilevanza scientifica⁶. Sistemi come Deep Blue⁷, il supercomputer che conquistò la stampa mondiale dopo aver battuto il campione del mondo di scacchi Garri Kasparov, rientrano in questa classificazione di *expert systems*⁸.

Tuttavia, gli *expert systems* operanti sulla base di regole computazionali accuratamente inserite nel loro codice sorgente incontrarono un significativo

degli anni '50 già teorizzavano la nascita, secondo loro pressoché imminente, di tecnologie che erano fino ad ora solo descritte in romanzi di *science fiction*. Non a caso, il primo ad utilizzare la dicitura di “Intelligenza Artificiale” per denominare la disciplina consistente nella programmazione di sistemi capaci di produrre forme di ragionamento e simulare comportamenti umani fu l'informatico John McCarthy, alla nota conferenza che si tenne a Dartmouth nel 1956. Thompson, C. (11 maggio 2022). What the history of AI tells us about its future. *MIT Technology Review*. <https://www.technologyreview.com/2022/02/18/1044709/ibm-deep-blue-ai-history/>.

⁴ *Id.*

⁵ Contissa, G. (2017). *Information technology for the law.*, G. Giappichelli Editore, 173.

⁶ Un esempio di un primo “sistema esperto” nell'ambito medico fu MYCIN, prodotto all'Università di Stanford, Stati Uniti nell'ambito del Progetto Proteoma Umano (HPP). Questo è stato uno dei primi sistemi esperti di intelligenza artificiale sviluppato in campo medico nel '72, e utilizzato per diagnosticare le infezioni all'interno del sangue. Operando sulla base di sole 500 regole di produzione, poteva fornire diagnosi dei pazienti e suggerire trattamenti terapeutici a un livello di competenza simile a quello degli specialisti umani in questo campo. Copeland, B. (7 ottobre 2008). MYCIN | *Expert System, Medical Diagnosis & Treatment*. Encyclopedia Britannica. <https://www.britannica.com/technology/MYCIN>.

⁷ Vedi nota 3.

⁸ Thompson (2022).

ostacolo in situazioni in cui era necessario un qualsivoglia “adattamento” da parte della macchina⁹: essi non erano in grado di performare altrettanto bene quando la situazione con cui interagivano non rientrava nel set di leggi e regole sulla base delle quali erano stati programmati. Questa loro inflessibilità, in ultimo, determinò il fallimento nel progettare i modelli di IA come “sistemi esperti” integralmente fondati su attività di pura deduzione logica sulla base della grande quantità di regole impresse nel codice sorgente. Ebbene, ciò demarcò il necessario passaggio dell’IA da branca della logica matematica e computazionale, all’IA moderna, più connessionista¹⁰ e maggiormente affine a discipline quali la statistica¹¹.

La “nuova IA”¹² fa larghissimo uso di un modello tecnico di apprendimento che, seppur ideato anch’esso durante quel periodo di fioritura iniziale di questa disciplina negli anni ‘50¹³, ha trovato successo solo in tempi più recenti: le reti neurali (c.d. *neural networks*¹⁴). Infatti, sebbene queste siano state teorizzate oltre ottanta anni prima dei moderni sistemi di IA progettati al giorno d’oggi¹⁵, esse non sono state oggetto dello stesso entusiasmo promotore dell’innovazione che aveva interessato gli *expert systems*. Una delle cause fu certamente la limitata conoscenza del campo delle scienze cognitive e del funzionamento del cervello umano, delle sue funzioni e dell’intera architettura cerebrale¹⁶. Con lo sviluppo della conoscenza in questo settore, in parallelo con l’incremento delle capacità computazionali e una maggiore quantità di dati a disposizione degli sviluppatori, le reti neurali sono tornate ad interessare gli studiosi in questo settore per le loro potenzialità ancora inesplorate.

Esse (denominate anche con l’acronimo ANN, Artificial Neural Networks), consistono in modelli semplificati del cervello umano composti da un gran numero

⁹ *Id.*

¹⁰ L’approccio connessionista nel campo dell’IA mira a comprendere e riprodurre il funzionamento del cervello umano a livello neurale e, in particolare, il modo in cui gli esseri umani apprendono dalla propria esperienza. Secondo questa dottrina il comportamento intelligente dell’uomo richiede la riproduzione della struttura del cervello umano, composto da cellule organiche, e delle loro connessioni attraverso le sinapsi. Contissa (2017), 149.

¹¹ Floridi, L. (2022). *Etica dell’intelligenza artificiale: Sviluppo, opportunità, sfide*. Raffaello Cortina Editore, 25.

¹² *Id.*

¹³ La prima rete neurale, seppur semplice, fu il sistema SNARC (*Stochastic Neural Analog Reinforcement Calculator*) costruito da Marvin Lee Minsky ed Dean Edmonds nel 1951 presso l’Air Force Research Laboratory in Ohio. Contissa (2017), 173.

¹⁴ Floridi (2022), 24.

¹⁵ L’idea di un modello computazionale che potesse simulare parzialmente l’attività neurale umana fu coniata per la prima volta dagli statunitensi Pitts e McCulloch nel 1943. Contissa (2017), 174.

¹⁶ *Id.*

di nodi (neuroni artificiali¹⁷) e da “sinapsi” che collegano un neurone artificiale all’altro. Ogni collegamento tra neuroni ha un peso che misura la forza della connessione tra i nodi, mentre ad ogni nodo della rete corrisponde un algoritmo specifico che regola le sue reazioni e il suo *output* a seconda dell’immissione di dati fornitogli¹⁸.

La rete così configurata interagisce con un ecosistema di *input* volti ad allenarne le potenzialità: come con l’apprendimento dei bambini, l’ANN impara grazie ad un continuo processo di *trial and error*, memorizzando i progressi e i fallimenti, nell’ottica di una riduzione al minimo dei tassi di errore¹⁹.

L’IA basata sulle *neural networks* ha trovato un terreno fertile in cui evolversi fino ai giorni nostri. Queste costituiscono il caposaldo di algoritmi allenati attraverso il *machine learning*²⁰, ma ancor di più di quella particolare categoria di “*deep*” *learning*, algoritmi che ad oggi costituiscono le fondamenta della tecnologia emergente più importante del secolo.

A tal proposito, questo elaborato inizia dunque a muovere i primi passi idealmente da quell’interrogativo che venne inquadrato da Alan Turing in uno degli articoli scientifici che hanno fatto la storia della scienza informatica²¹: “*Possiamo pensare le macchine?*”²² Il quesito, sebbene compaia in una pubblicazione risalente al 1950, risuona sotto diverse forme in numerosi dibattiti odierni, convegni, e seminari a livello globale, essendo rimasto comunque assolutamente attuale.

¹⁷ I primi neuroni artificiali furono il prodotto dello psicologo statunitense Frank Rosenblatt, che, nel 1950, progettò un primo *software* neurale, chiamato Perceptron; in seguito, creò un sistema di riconoscimento delle immagini a partire da veri e propri neuroni artificiali collegati tra di loro, sistema denominato Mark1 Perceptron. *Id.*

¹⁸ *Id.*, 149, 179.

¹⁹ *Id.*, 180.

²⁰ Il machine learning (ML) è un processo che consente la creazione di algoritmi in grado di vedere e riconoscere modelli da enormi quantità di dati e di applicarli per fare previsioni. Gli algoritmi di ML sono al centro di qualsiasi sistema di IA. Di solito questi prevedono una fase di “addestramento” della macchina e una successiva fase di “test”. Tuttavia, sebbene esistano diversi tipi di ML (supervisionato, non supervisionato o rinforzato), tutti i tipi di addestramento e tutte le fasi necessitano di enormi quantità di dati aggregati sia per “alimentare” che per “testare” questi sistemi. Per approfondire la nozione e il funzionamento effettivo del ML, vedi il sottoparagrafo 2.2.1. Hao, K. (5 aprile 2021). What is machine learning? *MIT Technology Review*. <https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/>.

²¹ Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, *LIX* (236), 433–460. <https://doi.org/10.1093/mind/lix.236.433>.

²² Turing (1950), 433.

2. La complessità nel definire il concetto di IA

La nozione di IA è tutt'altro che univoca o di semplice definizione: nella sua complessità definitoria, vi è il rischio di coniare spiegazioni meramente tautologiche, che non chiariscono la vera natura di questo strumento tecnologico²³. Anche negli ambienti di studio di questo settore, sono state proposte diverse definizioni, ciascuna delle quali corretta in sé e per sé. Tuttavia, nessuna è risultata sufficientemente convincente da superare in modo inequivocabile le altre. In maniera estremamente semplicistica, alcuni hanno descritto l'IA come un sistema informatico in grado di:

- ricevere dei compiti da portare a termine (delle c.d. *tasks*);
- imparare come portare a termine questi compiti attraverso modalità di apprendimento artificiale che coinvolgono una grande quantità di dati (da alcuni denominati “*Big data*”²⁴);
- adattare il proprio *output* durante tutto il periodo di utilizzo, non limitandosi ad apprendere unicamente nella fase di allenamento iniziale, ma migliorando la propria *performance* nel tempo²⁵.

La problematica principale di una definizione di tal genere è che questa serie di condizioni necessarie (ma talvolta non sufficienti) possono essere soddisfatte mediante modalità pressoché infinite²⁶. Vi sono inoltre ulteriori ragioni che contribuiscono a spiegare le sfide affrontate nel delineare con esattezza la natura dell'Intelligenza Artificiale. Prima fra tutti l'estrema polivalenza e multiformità dei sistemi *AI-powered*, che si presentano sotto forme diverse ed estremamente

²³ Ad esempio, si prenda in considerazione la definizione assolutamente tautologica che, nel 2020, compariva sull'Enciclopedia online Wikipedia alla voce “Intelligenza Artificiale”: “È l'intelligenza mostrata dalle macchine, in contrasto con l'intelligenza naturale mostrata dagli esseri umani.” Floridi, (2022), 40.

²⁴ Stabilire la definizione di *big data* non è un compito facile. Il professor Johnathan Stuart Ward e il professor Adam Barker dell'Università di St. Andrews (Edimburgo, Scozia) hanno intervistato i rappresentanti delle più influenti aziende tecnologiche (tra cui Microsoft, Oracle e Meta), chiedendo la loro definizione di *big data*. Il risultato dello studio porta alla seguente descrizione, ancorché in parte tautologica: “*Big data* è un termine che descrive l'archiviazione e l'analisi di insiemi di dati grandi e o complessi utilizzando una serie di tecniche tra cui, ma non solo, *NoSQL*, *MapReduce* e *machine learning*”.

MIT Technology Review, *The Big Data Conundrum: How to Define It?* (23 ottobre 2013) <https://www.technologyreview.com/2013/10/03/82990/the-big-data-conundrum-how-to-define-it/>.

²⁵ Per una definizione di IA così strutturata, ancorché semplicistica, vedi Ward, J. (2022). *The loop: How Technology Is Creating a World Without Choices and How to Fight Back*. Hachette Books., 130.

²⁶ *Id.*

composite. L'IA è intrinsecamente uno strumento *general-purpose*²⁷: con ciò, si intende il fatto che lo stesso identico sistema può essere utilizzato per scopi e obiettivi molteplici. Si prenda come esempio uno strumento di riconoscimento facciale che impiega IA nel trovare corrispondenza tra i dati biometrici di un individuo e la propria identità: esso potrà essere impiegato sia per compiere attività più semplici quali ad esempio aprire il cancello di un appartamento tramite l'autenticazione facciale di chi vi abita, ma anche come strumento di sorveglianza biometrica (o addirittura comportamentale) dei detenuti all'interno di un carcere²⁸. Questo dato, oltre a causare l'elusività della nozione di IA stessa, rende anche molto complicata la valutazione di impatto concernente i rischi connessi all'uso di un certo strumento sul pubblico, potendo implicare rischi ed effetti diversi nei confronti di diverse classi di individui²⁹.

In secondo luogo, l'Intelligenza Artificiale è una tecnologia dinamica, che non conosce stasi: essa muta e adatta i suoi parametri di funzionamento fin dal primo momento in cui viene iniziato l'allenamento dei propri algoritmi, così come durante la sua interazione costante con chi le utilizza. Il comportamento di questi sistemi può cambiare considerevolmente quando gli vengono forniti nuovi dati, o quando vengono previsti nuovi usi o nuove integrazioni³⁰. A ciò si ricollega anche la capacità, sempre più attuale, dell'IA di essere oggetto di processi di innovazione e aggiornamento estremamente rapidi, che si muovono di pari passo con la popolarità di queste tecnologie, cresciuta esponenzialmente negli ultimi due anni³¹. A questo proposito basti pensare che nel 2022 gli investimenti nel settore dell'IA (sia privati che pubblici complessivamente considerati) sono aumentati di tredici

²⁷Edwards, L. (2022) Regulating AI in Europe: Four Problems and Four Solutions, *Ada Lovelace Institute*, 2, 6 <https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/>.

²⁸ Si prendano come esempi di questo tipo di applicazione le società come l'Americana *Corsight*, che produce sistemi di IA da integrare sia nel pubblico che nel privato. Uno dei servizi offerti è proprio l'utilizzo di riconoscimento facciale per incrementare la sicurezza nella gestione delle carceri. *Enhancing Jail Management and Prisoner Management System with Facial Recognition | Corsight*. (18 maggio 2023). Corsight. <https://www.corsight.ai/prisons/>.

²⁹ Edwards (2022), 6-7.

³⁰ *Id.*

³¹ Uno dei fattori che ha generato questo aumento repentino di popolarità è stato il rilascio al pubblico di GPT-3.5 nel 2021 ad opera dell'azienda statunitense OpenAI, che ha indubbiamente segnato un punto di svolta nel campo dell'IA, determinando chiaramente una fase precedente e una successiva. Heaven, W. D. (31 agosto 2023). ChatGPT is everywhere. Here's where it came from. *MIT Technology Review*. <https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/>.

volte rispetto al 2013, registrando un incremento esponenziale soprattutto nel triennio tra il 2019 e il 2022³².

In terzo luogo, la presenza di IA è talvolta non evidente al pubblico dei consumatori che interagiscono con essa, o per motivi di scarsa trasparenza degli sviluppatori, o, più frequentemente, a causa della stretta sinergia che si crea fra lo strumento di IA in questione e la piattaforma sulla quale esso è integrato³³. Un esempio è costituito dalle piattaforme social come Facebook o Instagram (entrambe di proprietà di Meta, Inc.): sebbene siano conosciute come “semplici” *social media*, esse incorporano una molteplicità di algoritmi *AI-powered* nell'erogazione dei propri servizi sia di proprietà di Meta, Inc. che servizi posseduti da terze parti³⁴ (tra cui, ad esempio, algoritmi di moderazione di contenuti, di raccomandazione di post sponsorizzati e non, di affinamento delle ricerche interne alla piattaforma ecc.). L'utente Facebook medio non è usualmente a conoscenza della presenza di questi sistemi, poiché non riceve una contestuale informazione sulla piattaforma stessa, ma solo attraverso complesse *policies* non alla portata dell'utente comune non qualificato³⁵.

Per i motivi così esposti, si comprende come affidarsi ad un'unica definizione di Intelligenza Artificiale possa risultare non solo difficile, ma anche potenzialmente fuorviante per chi si trovi, in ultimo, nella posizione di dover definire la regolamentazione che la riguarda. Tuttavia, dal momento che la complessità di circoscrivere un concetto non dovrebbe diminuirne l'interesse ad esplorarne la sua natura³⁶, per fornire una complessiva immagine di ciò che può costituire l'IA, si è voluto fare riferimento a due diversi impianti definitori, per poi approdare alla tematica della sua complessa definizione normativa. Verranno quindi analizzate una nozione prettamente filosofica di IA, e a seguire, una serie di nozioni ad essa afferenti, di natura strettamente tecnica.

Questa scelta deriva dal fatto che, la stessa dicitura “Intelligenza Artificiale” è stata considerata da alcuni una silloge, ovvero di ricorrenza di una stessa parola

³² *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 4: the economy, 184 (2023) https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.

³³ Edwards (2022), 7.

³⁴ Edwards (2022), 7.

³⁵ Columbro, D. (2022). *Dentro l'algoritmo: Le formule che regolano il nostro tempo*. Effequ, 23-25.

³⁶ Floridi (2022), 41-44.

in due sensi diversi, in senso letterale e in senso simbolico³⁷. Questa figura retorica suggerisce che il concetto di IA possa essere interpretato sia in modo figurato, come un intero campo scientifico, sia letterale, come uno strumento fisico o una formalizzazione tecnica³⁸. L'analisi di entrambe queste sue caratteristiche, due facce della stessa medaglia, consentirà di procedere al dibattito estremamente attuale concernente i confini definitivi che si intende dare all'IA nell'ottica dell'assetto regolamentare e normativo che la riguarda.

2.1 La definizione filosofica (IA ingegneristica, IA cognitiva)

I padri fondatori della disciplina dell'Intelligenza Artificiale, che si riunirono nella celebre Conferenza di Dartmouth del 1956, inquadrarono così lo scopo dei sistemi di IA, agli albori di un'epoca in cui tali concetti stavano iniziando a prendere idealmente una forma: *“far sì che una macchina agisca con modalità che sarebbero definite intelligenti se un essere umano si comportasse allo stesso modo.”*³⁹ Questa inquadratura pragmatica dell'IA, tramite la focalizzazione non di ciò che essa “è,” ma di ciò che essa “fa”, consente di svincolarsi dal rispondere a quell'interrogativo formulato da Turing⁴⁰ che si è richiamato nel paragrafo 1, sull'intrinseca capacità delle macchine di *“pensare.”*⁴¹ A rilevare non è il *come* la macchina agisca per arrivare ad un dato risultato (se in maniera “intelligente” o meno), ma piuttosto il risultato tangibile ed evidente che da essa scaturisce⁴².

Se è possibile dunque accettare questa nozione, seppur controfattuale⁴³, dell'IA, si può comprendere anche il motivo per il quale Turing stesso concluse che tale quesito era *“troppo insensato per meritare di essere discusso,”*⁴⁴ poiché entrambi i termini (“Macchina” e “pensare”) sono privi di definizioni chiare e univoche.

Delimitare l'IA attraverso una nozione controfattuale e pragmatica però consente di cogliere una bipartizione concettuale rilevante: l'IA può essere infatti

³⁷ Benbouzid, B., Meneceur, Y., & Smuha, N. A. (2022). Four shades of AI regulation: a cartography of normative and definitional conflicts, *Réseaux*, N° 232-233(2), 29–64, 5 <https://doi.org/10.3917/res.232.0029>.

³⁸ *Id.*

³⁹ *Id.*

⁴⁰ Turing (1950), 433.

⁴¹ Floridi (2022), 44-48.

⁴² Floridi (2022), 45.

⁴³ *Id.*, 44.

⁴⁴ Turing (1950), 442. *Vedi anche* Floridi (2022), 273.

concepita come un sistema a due anime, una “ingegneristica” e una “cognitiva.”⁴⁵ L’IA cosiddetta “ingegneristica”, altresì nota come *Weak AI* (IA “debole”⁴⁶) è quella interessata alla riproduzione dei risultati positivi che scaturiscono da comportamenti “intelligenti.”⁴⁷ La si riconosce nei dispositivi *smart* ormai presenti nella quotidianità, e che consentono di portare a termine compiti che vanno di gran lunga oltre le potenzialità cognitive e di calcolo di un singolo cervello umano. Questo tipo di IA è, tuttavia, meramente riproduttiva dei comportamenti e degli approcci che vengono giustamente riconosciuti dall’esterno come comportamenti “intelligenti:” non è tuttavia foriera di una qualsivoglia capacità intellettuale endogena⁴⁸.

Dall’altro lato figura quella IA “cognitiva,” anche denominata *Strong AI*: per essa si intende quel particolare e delimitato settore delle scienze cognitive che mirano a produrre un’intelligenza umana endogena a partire da sistemi artificiali, con strutture che ricalcano quelle biologiche umane. Questa seconda categoria non sarebbe *riproduttiva*, come quella ingegneristica, bensì *produttiva* di intelligenza⁴⁹. Sebbene la *Strong AI* costituisca uno di quei temi che ha generato il fascino e lo stupore dell’opinione pubblica con il vertiginoso aumento di popolarità di questa tecnologia⁵⁰, essa non è stata mai realizzata (né, secondo diversi studiosi, può essere fisicamente sviluppata⁵¹). L’idea di un’applicazione dell’IA che vada oltre i limiti di ciò che è artificiale e ciò che è umano, con una suggestiva sfocatura dei contorni di entrambi, è certamente affascinante: ma rimane, allo stato attuale delle cose, pura fantascienza⁵².

Dunque, si ritorna alla *Weak AI*: quel modello che, in grado di eseguire operazioni, valutazioni e predizioni incredibilmente più complesse di quelle che un cervello umano sia in grado di gestire, paradossalmente non rappresenta nulla di più che un calcolatore estremamente complesso, dotato di innumerevoli parametri,

⁴⁵ Floridi (2022), 48.

⁴⁶ La tradizionale bipartizione tra *Weak AI/Strong AI* non consiste tanto in una vera e propria dicotomia tra due diverse concezioni di questa tecnologia, ma rispecchia piuttosto un differente approccio nell’immaginare o meno la possibilità di progettare un’intelligenza “completa” a mezzo di *hardware* non necessariamente organici, e non solo meramente riproduttiva dei comportamenti “intelligenti” tipici degli esseri umani. Contissa (2017), 145.

⁴⁷ Floridi (2022), 48.

⁴⁸ *Id.*

⁴⁹ *Id.*, 49.

⁵⁰ *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 8: public opinion, 332 (2023).

⁵¹ Contissa (2017), 145. Per un approfondimento sul tema del potenziale sviluppo di un’IA in grado di “produrre” (e non “riprodurre”) l’intelligenza umana a tal punto da sfuggire dal “controllo” umano, si veda il paragrafo 3.3.

⁵² Floridi (2022), 49.

in grado solamente di emulare l'intelligenza⁵³. Anche nel mondo della ricerca vi è stato un mutamento delle priorità di sviluppatori e programmatori proprio perché si sono realizzati i limiti nel cercare di produrre un' intelligenza "endogena": ad esempio, il progetto AlphaGO⁵⁴ del noto laboratorio di IA DeepMind, di Alphabet (Google), è stato interrotto a favore dello sviluppo di algoritmi da utilizzare nel settore scientifico e medico, senza alcuna pretesa di creare un dispositivo che mostrasse né dimostrasse "intelligenza" di alcun tipo, ma che semplicemente fosse in grado di coadiuvare gli studiosi nella ricerca⁵⁵. Qualche anno dopo il cambio di rotta di DeepMind, è stato sviluppato un progetto di successo chiamato AlphaFold, un algoritmo in grado di prevedere le strutture "di ripiegamento" delle proteine che ha determinato un significativo avanzamento nel settore medico e della ricerca scientifica⁵⁶. Questo a dimostrazione del fatto che l'impostazione di progresso nel settore della ricerca sull'IA sta sempre più tendendo alla programmazione di algoritmi complessi da utilizzare in campo scientifico, glissando sulla capacità di questi di essere forieri di capacità intellettuali effettive, tangibili, ed endogene.

A questo proposito, si è parlato di una scissione, di un "divorzio" che l'Intelligenza Artificiale ha causato tra la risoluzione efficace dei problemi complessi, (l'"*agire*") e il "*comportamento intelligente*."⁵⁷ L'IA è in grado di risolvere brillantemente i problemi che le vengono sottoposti senza rispecchiare quelle componenti di matrice prettamente e unicamente umana, di "*comprensione, consapevolezza, acume, sensibilità, [...] intuizioni, semantica, esperienza, bio-incorporazione, significato, persino saggezza*."⁵⁸ È a partire da questa concezione della stessa che si è definita l'IA come una "*capacità d'agire, interattiva,*

⁵³ *Id.*, 49-50.

⁵⁴ AlphaGo è un sistema di IA che utilizza diverse reti neurali con apprendimento "profondo" (vedi sottoparagrafo 2.2.1) sviluppato da DeepMind tra il 2014 e il 2017. Il laboratorio di Alphabet (Google) intendeva creare un "giocatore artificiale intelligente" in grado di sconfiggere i campioni mondiali del gioco "Go," dotato di un elevato grado di complessità a tal punto da far dubitare di poter essere insegnato ai dispositivi di apprendimento informatico. Tuttavia, nel 2017, AlphaGo ha battuto i *master* del gioco Go, e ha raggiunto il suo picco di complessità computazionale: a questo punto, il progetto fu abbandonato dalla casa madre DeepMind, per portare avanti progetti di diversa natura, principalmente scientifica, interrompendo quella tendenza a cercare di sviluppare sistemi che si misurassero in sfide con l'"intelligenza" umana, piuttosto orientandosi su strumenti che risolvessero delle questioni cruciali in campo scientifico. *Id.*, 50. *Vedi anche Google's AlphaGo "plays just like a human," says top ranked Go player.* (1° febbraio 2016). SiliconANGLE. <https://siliconangle.com/2016/02/01/googles-alphago-plays-just-like-a-human-says-top-ranked-go-player/>.

⁵⁵ *AlphaGo's next move.* (27 maggio 2017). Google DeepMind. <https://deepmind.google/discover/blog/alphagos-next-move/>.

⁵⁶ Floridi (2022), 51.

⁵⁷ *Id.*, 52-53.

⁵⁸ *Id.*, 52.

autonoma, e autoapprendente:”⁵⁹ ciò, tuttavia, non ha nulla a che fare con l’“intelligenza” in quanto tale.

L’IA così definita ha guadagnato gran parte del suo successo grazie alla sua capacità di plasmare la realtà attorno ad essa in modo tale da rendere l’ambiente circostante sempre più favorevole al suo utilizzo⁶⁰. Questo può essere il motivo per il quale, ad oggi, il futuro dell’umanità che prende forma è sempre più caratterizzato da sistemi di IA ubiqui e sostanzialmente invisibili che svolgeranno calcoli complessi, valutazioni, previsioni e scelte per conto degli esseri umani. Va invece sfumando quella concezione futuristica di macchine fisiche e robot antropomorfi dotati di intelligenza e altre capacità intellettuali tipicamente umane che, secondo alcuni, avrebbero condiviso gli spazi quotidiani con l’umanità negli anni a venire. Quei robot umanoidi immaginati dagli autori di fantascienza che avrebbero affiancato gli esseri umani nella loro vita quotidiana, durante tutte le varie fasi della loro esistenza, fornendo assistenza e supporto costante, forse, non vedranno mai la luce del giorno.

2.2 Le definizioni tecniche

A questo punto, è necessario spostare l’attenzione di questa analisi su quella seconda interpretazione del concetto di “Intelligenza Artificiale,” che la interpreta come un oggetto fisico o una vera e propria formalizzazione tecnica⁶¹, soffermandosi sugli aspetti più tecnologici la cui delineazione è necessaria, quasi propedeutica per esaminare successivamente la regolamentazione e le possibili strategie da utilizzare. Infatti, alcuni concetti prettamente tecnici devono essere correttamente inquadrati per comprendere le necessità e le implicazioni regolamentari, per operare un *drafting* normativo consapevole di tutti i risvolti specifici di queste tecnologie, anche se si volesse assumere un approccio alla normazione il più tecnologicamente neutrale possibile. Non è infatti possibile immaginare una legislazione in questo campo che, seppur dichiaratamente volta a fare leva su una “neutralità tecnica”⁶² per assicurare l’efficacia nel tempo di leggi

⁵⁹ *Id.*, 53.

⁶⁰ Floridi ha denominato questa capacità di influenza dell’IA sull’ambiente circostante e sulla modernità come capacità di “re-ontologizzazione” del mondo contemporaneo, consentito sempre più anche dall’allargarsi ad ogni angolo del globo di quella c.d. “infosfera”, cioè quell’atmosfera sempre più digitale, connessa, nella quale si ritrova immersa l’umanità intera. *Id.*, 25, 55.

⁶¹ Benbouzid *et al.* (2022), 5.

⁶² *Id.*, 21.

che regolano applicazioni in costante mutamento, prescindendo dagli elementi tecnologici che verranno di seguito analizzati.

Come è stato evidenziato nel paragrafo 2, la complessità di delineare un concetto univoco di IA si ripropone anche nell'ambito più tecnico e relativo alle applicazioni pratiche di questi sistemi. Il motivo risiede nel fatto che, come è stato anticipato, l'IA è uno strumento assolutamente *general-purpose*⁶³, un “*super-tool*”⁶⁴ che si presenta in diverse modalità, incorporata in dispositivi fisici, su piattaforme online, in *software* o siti web⁶⁵. Come precedentemente analizzato, le sue forme sono molteplici e non sempre riconoscibili agli occhi di coloro che sono meno esperti⁶⁶.

A questo proposito, può essere utile recuperare la generalissima descrizione fornita al paragrafo 2 di IA, e da lì procedere “a cascata” nell'analizzare quei componenti tecnici che afferiscono a questo campo. Tuttavia, non si intende solo fornire in maniera didascalica una cornice esplicativa per questi aspetti tecnici, ma si è scelto di affiancare a questi anche le relative problematiche applicative che questi possono generare.

L'IA è stata precedentemente delineata come un sistema in grado di ricevere dei compiti da svolgere (dei tasks), imparare come portarli a termine attraverso modalità di apprendimento artificiale che coinvolgono una grande quantità di dati, ed in ultimo di produrre un certo *output* che non è statico, ma continuamente aggiornato proprio dalle capacità di continuo apprendimento della macchina dalla propria esperienza⁶⁷. Da questa estrema semplificazione concettuale, emergono due questioni tecniche che necessitano di corretta individuazione e trattazione: l'apprendimento della macchina da un lato, e il materiale dalla quale questa è in

⁶³ Edwards (2022), 6-7.

⁶⁴ Questo termine molto calzante è stato utilizzato dal Prof. Andrea Amidei, Docente di Informatica e Diritto presso l'Alma Mater Studiorum Università di Bologna, in occasione di uno degli appuntamenti mensili del seminario permanente dell'area “Macchine Intelligenti e Diritto” tenutosi presso l'Università LUISS Guido Carli il giorno 25/10/2023, dal titolo “*Machina et Inventio: il Diritto e la Creatività dell'IA generativa.*”

⁶⁵ Alcuni esempi delle molteplici forme di manifestazione dell'IA: gli assistenti digitali sui moderni smartphones quali Siri o Google Assistant utilizzano Intelligenza Artificiale; gli algoritmi di moderazione dei contenuti su piattaforme quali Facebook, Youtube, Instagram, sono *AI-powered*; programmi online come Midjourney o DALL-E forniscono servizi di IA generativa in grado di generare immagini dai *prompt* di testo forniti dagli utenti; gran parte delle chat che forniscono assistenza clienti online utilizzano più o meno complessi algoritmi di IA per dare il migliore supporto possibile alla clientela e rendere le risposte della macchina il più rilevanti e accurate possibili.

⁶⁶ Columbro (2022), 23-25. Vedi paragrafo 2.

⁶⁷ Ward (2022), 130.

grado di apprendere dall'altro. Il *Machine Learning* (traducibile in italiano in “Apprendimento Automatico”) costituisce il metodo generale di apprendimento dei sistemi di IA; mentre il materiale che questi utilizzano come “carburante” che foraggia i propri processi decisionali sono i dati. Questi due temi saranno trattati a fianco di due fenomeni rispettivamente ad essi ricollegati che consentono di individuare quali possano essere gli eventuali risvolti negativi ricollegati ai fondamenti funzionali dei sistemi di IA.

L'esemplificazione dei potenziali aspetti problematici ricollegati rispettivamente al *Machine Learning* e alle grandi quantità di dati da esso utilizzati, funge da riprova al fatto che la tecnologia, per quanto frutto dell'affinamento della scienza informatica, non è intrinsecamente neutrale né rappresenta una soluzione la cui derivazione direttamente scientifica esclude ogni margine di errore⁶⁸. Essa non è né potrà mai costituire una panacea assoluta⁶⁹. Al contrario, essa può causare conseguenze pregiudizievoli direttamente collegate con il contesto sociale in cui i sistemi di IA sono impiegati. Basti pensare al fatto che gli algoritmi che sono normalmente percepiti come “oggettivi” e “neutrali” quali ad esempio quelli impiegati dai traduttori *online*, in realtà codificano il linguaggio da loro utilizzato per le traduzioni secondo modalità condizionate da vari *bias*, tra cui ad esempio quelli di genere⁷⁰.

Ancora, unicamente a titolo esemplificativo, il Libro Bianco sull'Intelligenza Artificiale pubblicato dalla Commissione Europea del 2020 ha riconosciuto delle macro-aree di danni che le macchine intelligenti potrebbero causare, proprio dal momento che queste possono avere dei risvolti non solo di utilità, ma anche potenzialmente nocivi: danni materiali (contro la sicurezza e la salute delle persone e danni alla vita o alla proprietà), ma anche immateriali (perdita della privacy, limitazione del diritto alla libertà di espressione e così via⁷¹).

Da ciò si evince il motivo per il quale la normazione in questo campo è ormai imprescindibile per limitare questi effetti potenzialmente dannosi che l'IA può causare, anche a causa delle sue caratteristiche intrinseche e di funzionamento che ci si appresta ad analizzare di seguito.

⁶⁸ O'Neil, C. (2016). *Weapons of math destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin UK, 218. Vedi anche Floridi (2022), 156-163.

⁶⁹ Floridi (2022), 226.

⁷⁰ *Id.*, 146.

⁷¹ *Libro Bianco della Commissione Europea sull'Intelligenza Artificiale – Un approccio europeo all'eccellenza e alla fiducia*, 11 COM (2020) 65 final (19 febbraio 2020).

2.2.1 Le modalità di apprendimento: il Machine Learning nelle sue varie declinazioni e il c.d. “black box effect”

Il *Machine Learning* (d’ora in avanti abbreviato in ML) consiste in quel settore della scienza dell’IA che raggruppa i vari metodi di affinamento e allenamento degli algoritmi tramite l’applicazione di specifici principi di matrice statistica. Esso consiste in una combinazione di tre elementi fondamentali: dati, modello e funzione obiettivo⁷². I metodi di apprendimento automatico combinano queste tre componenti all’interno di implementazioni efficienti dal punto di vista computazionale, che integrano quel principio scientifico di “*trial and error*” per adattare e in ultimo migliorare continuamente il proprio *output*⁷³.

Per la precisione, questi metodi di apprendimento leggono i dati generati all’interno di un dominio applicativo e attribuiscono loro “caratteristiche” ed “etichette.”⁷⁴ Le caratteristiche sono proprietà misurabili o calcolabili facilmente, mentre le etichette sono proprietà che richiedono l’intervento di esperti umani per determinarne le proprietà⁷⁵. L’obiettivo del ML consiste nel consentire alla macchina di “imparare” a prevedere l’etichetta di un determinato punto dati (*datapoint*) basandosi solo sulle sue caratteristiche. L’algoritmo che apprende tramite il ML non produce unicamente una predizione di risultato univoca, ma crea, sulla base dei dati fornitigli da chi lo allena e in base alle proprie capacità di riconoscere i *pattern*, un modello di informazioni strutturate⁷⁶, che può essere anche descritto come una sorta di “mappa di ipotesi.”⁷⁷ Per adattare o migliorare la mappa di ipotesi, viene utilizzata la suddetta funzione obiettivo o funzione di perdita⁷⁸, per quantificare la discrepanza tra le previsioni della macchina e i dati oggetto dell’allenamento⁷⁹. Per questo motivo il ML viene definito, più semplicemente, come un metodo di apprendimento applicato ad un sistema “*che migliora le proprie prestazioni su compiti futuri dopo aver effettuato un’osservazione.*”⁸⁰

⁷² Jung, A. (2022). *Machine learning: The Basics*. Springer., vii. Per la definizione di funzione obiettivo, vedi nota 78.

⁷³ *Id.*

⁷⁴ Nella traduzione originale del testo sono denominate “*labels.*” *Id.*, 13.

⁷⁵ *Id.*

⁷⁶ Columbro (2022), 53.

⁷⁷ Jung (2022) vii.

⁷⁸ Per funzione obiettivo (*loss function*) si intende una particolare funzione di ottimizzazione matematica utilizzata per mappare un evento ed i costi ad essa associati, nell’ottica di ridurre al minimo gli stessi per garantire un risultato ottimale.

⁷⁹ Jung (2022), vii.

⁸⁰ Contissa (2017), 107.

In base alle modalità secondo le quali i metodi di ML valutano la qualità delle diverse mappe di ipotesi delineate, si possono distinguere tre tipi principali di ML: *Supervised*, *Unsupervised* e *Reinforcement Learning* (rispettivamente apprendimento supervisionato, non supervisionato e apprendimento per rinforzo⁸¹).

Nell'apprendimento supervisionato (*standard*), al sistema di IA vengono fornite una serie di coppie di *input-output*, una serie di esempi di dati e predizioni ad essi connessi, affinché il sistema possa estrapolare una serie di regole da essi⁸². Lo sviluppatore che allena il sistema, anche denominato come “*trainer*” successivamente fornisce un *input* che ha la funzione di testare il modello: se l'*output* generato dal sistema è errato, l'algoritmo alla base di questo si adatterà di conseguenza, nell'ottica di ridurre sempre al minimo il tasso di errore rispetto alle regole finora apprese⁸³. In questo modo i sistemi acquisiscono automaticamente il proprio insieme di regole che descrivono il modello. Questo tipo di apprendimento è utilizzato (parallelamente ad altri), ad esempio, dai complessi sistemi *AI-powered* che svolgono attività di moderazione dei contenuti online⁸⁴.

L'apprendimento rinforzato, o *Reinforcement Learning* differisce da quello supervisionato “*standard*” al sistema non vengono presentate le coppie di *input-output* per la definizione di regole primarie sulle quali il sistema baserà le proprie inferenze. Al contrario, il sistema di IA in questo caso apprenderà da un meccanismo di rinforzo basato su logiche premiali utilizzato dal *trainer*: se l'output generato è corretto, il *trainer* assegnerà un “premio” (ad esempio un “+1”) alla macchina, che sarà dunque attivamente incentivata ad ottenere più premi possibili, affinando contestualmente la propria capacità decisionale e riducendo al minimo il margine di errore⁸⁵. Questo tipo di apprendimento è particolarmente utilizzato con un certo successo, ad esempio, nel campo dei veicoli a guida autonoma⁸⁶.

L'apprendimento non supervisionato (o *Unsupervised*) invece, differisce dai precedenti, in quanto manca quella una strutturazione dei dati in *input-output* tipica del *Supervised Learning*, e non vi sono alcuni meccanismi premiali di rinforzo come nel *Reinforcement learning*, né altre tipologie di *feedback* da parte di chi

⁸¹ Jung (2022), 13.

⁸² Contissa (2017), 108.

⁸³ *Id.*

⁸⁴ Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1), 5-7 205395171989794. <https://doi.org/10.1177/2053951719897945>.

⁸⁵ Contissa (2017), 108.

⁸⁶ Jung (2022), 14-15.

allena la struttura di IA⁸⁷. Il metodo non supervisionato si basa esclusivamente sulla struttura intrinseca dei dati forniti alla macchina, non corredati di alcuna indicazione (senza quindi la c.d. “*label*”), che sono sufficienti all’algoritmo per formulare la propria mappa di ipotesi⁸⁸. Pertanto, i sistemi così allenati non necessitano di un *trainer* o di un esperto che fornisca le “etichette” per il set di dati che viene utilizzato nell’apprendimento iniziale⁸⁹.

Tuttavia, è da sottolineare come quelli descritti non siano le uniche tecniche che vengono utilizzate nel ML, che è un campo in continuo sviluppo in parallelo con l’innovazione tecnica del settore, l’aumento del volume e della qualità di dati disponibili e le capacità di calcolo dei computer che supportano i *software* utilizzati nella programmazione di questi complessi sistemi. Infatti, esistono ulteriori modalità di *training* dei sistemi di IA che, seppur rientranti nella più ampia categoria di ML, hanno un livello superiore di complessità intrinseca, oppure delle peculiarità applicative che consente di considerarli singolarmente.

Un tipo di ML dotato di una certa complessità strutturale è il cosiddetto “apprendimento profondo” o *Deep Learning*. I sistemi di apprendimento profondo utilizzano strutture di *Cloud Computing*⁹⁰ per addestrare modelli di grandi dimensioni su vaste quantità di dati. Quest’attività coinvolge una serie di numerose reti neurali più o meno nascoste, che operano a diversi livelli di profondità nel sistema. Queste sono utilizzate ad esempio nell’allenamento di tecnologie di elaborazione del linguaggio naturale (il c.d. *Natural Language Processing*), processo che si utilizza per “rilevare la struttura e il significato di un testo.”⁹¹ Questo tipo di metodo di apprendimento è ad esempio utilizzato nel contesto dei c.d. *Large Language Models*⁹² (LLMs) come il celebre ChatGPT.

⁸⁷ Contissa (2017), 108.

⁸⁸ *Id.*

⁸⁹ Jung (2022), 13-14.

⁹⁰ Per *Cloud Computing* si intende un’evoluzione nelle reti di computer che emerse nel 2000 con lo scopo di connettere una molteplicità di gruppi di *server* online che consentono l’accesso a un particolare servizio per l’appunto, online (*software*, piattaforme, *datasets* ecc.). Esempi di questi sistemi di *Cloud Computing* sono ad esempio i servizi mail, di e-commerce, di conservazione *file* offerti dalle grandi aziende tech. Contissa (2017), 34.

⁹¹ Columbro (2022), 31.

⁹² I Large Language Models (LLM) sono un sottotipo di modello di IA generativa specificamente progettato per comprendere e generare testi simili a quelli umani. Essi utilizzano metodi di *deep learning* e vasti set di dati per comprendere, condensare, e produrre testi di vario genere. L’espressione “IA generativa” (che verrà analizzata insieme alla potenziale normativa ad essa applicabile nel Capitolo 2, sottoparagrafo 3.3.3) è strettamente legata agli LLM. Kerner, S. M. (13 settembre 2023). *Large Language Models (LLMs)* <https://www.techtarget.com/whatis/definition/large-language-model-LLM>. Vedi anche Mészáros, J., Preuveneers, D., Marquet, E., Biasin, E., Peter, I. E., Santos, I. R., Vranckaert, K., Belkadi, L., &

Una diversa ipotesi di modalità di apprendimento dotato di particolarità applicative, frutto della ricerca più all'avanguardia degli ultimi anni nel campo della protezione dei dati personali, è il c.d. *Federated Learning*. Questo tipo di apprendimento non fa leva su grandi *dataset* centralizzati, ma piuttosto implementa l'addestramento di sotto-modelli di IA utilizzando *database* localmente archiviati in singoli dispositivi che utilizzano la tecnologia di IA⁹³. Senza entrare nelle specifiche caratteristiche di questo tipo di modello di apprendimento, basti sapere che ciò combina il ML con meccanismi di crittografia e computazione c.d. *multiparty*⁹⁴, che insieme consentono di raggiungere un elevato grado di protezione della privacy e della sicurezza delle informazioni utilizzate nel *training* di un modello di IA. Questa è solo un'altra possibile modalità di allenare un sistema di IA, nell'ottica di una tecnologia non solo più orientata alla protezione dei dati personali, ma anche ispirata a quel principio c.d. di *privacy by design*⁹⁵ e *by default*, per cui un determinata tecnologia è, strutturalmente, progettata in maniera tale da tutelare la privacy.

Così come descritte, le modalità di apprendimento utilizzate nell'allenamento dei sistemi di IA sono molteplici e varie, e sono impiegate diversamente a seconda del tipo di modello e di scopi che si intende raggiungere con la programmazione di un determinato strumento applicativo dell'IA. La precedente trattazione relativa al funzionamento dei modelli di apprendimento richiama consequenzialmente l'attenzione su una necessaria analisi critica di quali

Menéndez, N. (2023). Chatgpt: How Many Data Protection Principles Do You Comply with?, *Artificial Intelligence eJournal*. <https://doi.org/10.2139/ssrn.4647569>.

⁹³ Cheng, Y., Li, Y., Chen, T., & Yang, Q., (2020) Viewpoint Federated Learning for Privacy-Preserving AI, *Communication of the ACM*, vol. 63, 12, 33-37. <http://dx.doi.org/10.1145/3387107>

⁹⁴ La computazione c.d. *multiparty* (*Multi-Party Computing*) è una tecnica che consente a più parti di calcolare congiuntamente una funzione sui propri *input*, mantenendo tali *input* privati. Essa garantisce un elevato grado di sicurezza. Cheng *et al.* (2020), 33-34.

⁹⁵ La *privacy by design* è un approccio all'ingegneria dei sistemi informatici, ideato da Ann Cavoukian e formalizzato nel 1995 da un gruppo congiunto formato dall'*Information and Privacy Commissioner* dell'Ontario (Canada), dall'Autorità olandese per la protezione dei dati e dall'Organizzazione olandese per la ricerca scientifica applicata. Questo concetto prevede che la privacy degli individui e della collettività venga considerata in tutte le fasi del processo di progettazione di un certo sistema informatico, così da rappresentare un esempio di *design* sensibile ai valori umani di tutela della riservatezza e dei dati personali. Nonostante la complessità tecnica nell'applicazione e integrazione di tale principio nei sistemi informatici, sviluppi recenti nell'informatica e nell'ingegneria dei dati, insieme alla disponibilità e qualità delle tecnologie di miglioramento della privacy, contribuiscono a superare tali criticità, rendendo la *privacy by design* applicabile in contesti reali. Il principio di tecnologie *privacy by design* è anche confluito nella normativa europea sulla protezione dei dati personali. Art. 25, Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (d'ora in avanti abbreviato come "GDPR").

effetti (potenzialmente negativi) derivano da quei fattori indissolubilmente legati al *training* algoritmico.

I sistemi informatici come quelli di IA che, semplicisticamente, operano sulla base di un certo *input* dal quale dipende un determinato *output* che la macchina formula alla base delle regole impartitele o da essa apprese, possono essere paragonati a delle “scatole nere.”⁹⁶ La macchina che viene programmata con il ML non dispone di una trasparenza tecnica e di *default* tale da consentire di individuare la motivazione alla base di determinate associazioni o combinazioni di *input* e *output*. Infatti, le varie tecniche di ML che sono state descritte, comprese quelle che hanno delle particolarità applicative volte a dare priorità ai valori “umani” come la tutela della privacy, possono portare con sé questa problematica strutturale e intrinseca per la quale è estremamente complesso risalire alle motivazioni alla base di un certo *output*. Dunque, in base al c.d. “*black box effect*” (o “effetto scatola nera”), più gli algoritmi di IA sono complessi, maggiore sarà la complessità nel risalire precisamente a tutti i meccanismi decisionali e le computazioni interne che hanno portato ad una certa decisione della macchina.

Un dato (allarmante) di questo tipo poteva essere altresì tollerato in una fase assolutamente sperimentale dell’IA, ma ad oggi, con i sistemi *AI-powered* dai quali può dipendere, ad esempio, la possibilità di accendere un mutuo o di vedersi offerta una posizione lavorativa o una specifica mansione, o la scelta di un particolare farmaco per la cura di una malattia, diviene imprescindibile e necessario poter fornire dettagliate spiegazioni procedurali rispetto agli *output* ottenuti⁹⁷.

Inoltre, i progressi recenti nel campo del *Deep Learning* hanno portato alla diffusa convinzione che i modelli più complessi ed accurati debbano essere intrinsecamente non interpretabili, proprio perché volti a fornire risposte a problemi estremamente complicati. Secondo alcuni, la combinazione delle variabili per fare previsioni diventa così articolata che anche gli umani, inclusi gli stessi sviluppatori, non hanno modo di comprenderne precisamente l’*output*, né prevederlo⁹⁸. Questa credenza è inoltre avvalorata da quell’ uso, storicamente fondato,

⁹⁶ Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. *Conference on Fairness, Accountability, and Transparency (FAT* '19)*. 59-60, <https://doi.org/10.1145/3287560.3287598>.

⁹⁷ Ward (2022), 134-135. Vedi anche Donati, F., *Quale disciplina per l’intelligenza artificiale?*, 47, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.

⁹⁸ Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don’t Need To? A Lesson from An Explainable AI Competition. *Harvard Data Science Review*, 1(2), 3. <https://doi.org/10.1162/99608f92.5a8a3a3d>.

dell'apprendimento automatico in contesti definiti “a basso rischio”, come la pubblicità sui siti e *social networks* o la ricerca sul web, settori in cui, fortunatamente, le decisioni non hanno un impatto significativo sulle vite umane⁹⁹.

Una diversa conclusione è stata, tuttavia, raggiunta da alcuni studiosi che hanno partecipato alla *Explainable Machine Learning Challenge* organizzata da Google, insieme alla *Fair Isaac Corporation* (FICO) e cinque università statunitensi nel 2019. Essi, infatti, nel costruire una sistema di *credit scoring* completamente trasparente e comprensibile sia rispetto all'*output* che rispetto al *training*, hanno completamente sfatato il mito della dicotomia tra *interpretability* (intesa come capacità di comprensione ed interpretazione delle decisioni algoritmiche) ed *accuracy* (accuratezza) dei modelli di IA¹⁰⁰.

Secondo la loro analisi, gli algoritmi *black-box*, progettati in quegli ambiti in cui le variabili per determinare un certo *output* sono molteplici e ad alto livello di complessità, potrebbero essere positivamente sostituiti da algoritmi più trasparenti, conservando lo stesso livello di accuratezza, e, in ultimo di utilità sociale¹⁰¹. Essi hanno inoltre rilevato che il motivo alla base di questa mancanza generale di esplorazione di algoritmi di IA incentrati maggiormente sulla trasparenza fin dalla loro iniziale progettazione è da ritrovarsi principalmente in una ritrosia dettata dall'*industry* dell'IA, e non da motivazioni tecniche che impediscano, di fatto, la costruzione di modelli più comprensibili¹⁰².

Alcuni hanno concluso, in modo riduttivo ed essenzializzante, che gli algoritmi *black box* sono semplicemente meno onerosi da produrre: infatti, costruire IA “interpretabile” necessita l'osservazione di una serie di limitazioni e accortezze, che aumentano i costi di produzione¹⁰³. Sebbene dunque, sia meno oneroso progettare sistemi di IA le cui decisioni sono sostanzialmente opache, è ugualmente molto oneroso operare un *reverse-engineering*¹⁰⁴ dell'*output* algoritmico di modelli

⁹⁹ Tuttavia, si rammenta che, sebbene l'utilizzo di algoritmi che utilizzano IA in settori simili non comporti diretti rischi per le vite umane, essi potrebbero mettere a rischio l'equilibrio della democrazia, aumentando il rischio di eccessiva polarizzazione, in ultimo, la “tribalizzazione” del discorso politico. Bartlett, J. (2018). *The People vs Tech: How the Internet is killing democracy (and how we save it)*. Penguin Random House UK. Chapter 2: *the Global Village*, 41-68.

¹⁰⁰ Rudin & Radin (2019), 3. Vedi anche Ward, (2022) 134-138.

¹⁰¹ Rudin & Radin (2019), 7-8.

¹⁰² *Id.*

¹⁰³ Ward (2022), 135-136.

¹⁰⁴ Per *reverse-engineering* in questo contesto si intende la possibilità di comprendere le motivazioni e i fattori alla base di una certa decisione algoritmica a partire dalla decisione stessa, operano un percorso all'inverso per capire quali meccanismi decisionali abbiano portato ad una certa conclusione. *Id.*, 136.

*black box*¹⁰⁵. Questi fattori, entrambi due facce della stessa medaglia, spiegano il maggiore successo degli algoritmi a scatola nera rispetto a quelli dotati di maggiore trasparenza, che hanno un valore etico nettamente superiore. In conclusione, sebbene “*l’assenza di trasparenza sia una caratteristica intrinseca degli algoritmi di autoapprendimento, [...] [e sebbene sia] certamente reale la difficoltà di spiegare l’output degli algoritmi di ML, è al contempo importante non lasciare che questa difficoltà incentivi le organizzazioni a sviluppare sistemi complessi per sottrarsi alle [proprie] responsabilità.*”¹⁰⁶

2.2.2 Il carburante dell’IA: i dati (e i loro potenziali “bias”)

Nel 1974, in quel periodo storico in cui si registrava uno dei numerosi c.d. “inverni dell’IA,”¹⁰⁷ il filosofo americano Hubert L. Dreyfus si interrogava sulla possibilità di concepire un’Intelligenza Artificiale che riproducesse efficacemente in tutto e per tutto il pensiero umano. A questo interrogativo, il filosofo rispondeva negativamente, identificando due aree di tematiche in cui l’IA di quel tempo, dotata di quei caratteri che sono stati identificati nei paragrafi 1 e 2.1, non era riuscita ad eccellere e aveva conseguentemente disatteso le aspettative dei più ottimisti del settore¹⁰⁸. Queste due aree erano la “risoluzione di problemi” (*problem solving*) e il “riconoscimento dei *pattern*” (*pattern recognition*). Lo stato dell’arte di quel periodo, in cui le macchine intelligenti avevano performato in maniera deludente rispetto le speranze, portò Dreyfus ad immaginare che “[*a causa*] della struttura aperta del *pattern recognition* e dell’infinità di fatti che possono essere rilevanti nel *problem solving*, [...] [*anche se*] qualsiasi prestazione specifica dell’intelligenza umana potrebbe essere simulata su un computer, [...] un comportamento pienamente intelligente sarebbe impossibile, in linea di principio, per una macchina digitale.”¹⁰⁹

¹⁰⁵ *Id.*

¹⁰⁶ Floridi (2022), 154.

¹⁰⁷ L’inverno dell’IA consiste in una fase di temperamento delle aspettative, durante la quale tecnologia, *business* e *media* escono da una “bolla” ottimistica per affrontare in modo realistico le capacità effettive dell’IA. Questi periodi sono contrassegnati da investimenti più cauti e dalla transizione dei giornalisti verso nuovi argomenti di tendenza, temporaneamente lasciando da parte l’entusiasmo iniziale per questa tecnologia. Storicamente, l’IA ha attraversato diversi inverni, con rilevanti fasi di riflessione alla fine degli anni ‘70 e tra gli anni ‘80 e ‘90. Floridi, L. (2020). AI and Its New Winter: from Myths to Realities. *Philos. Technol.* 33, 1 <https://doi.org/10.1007/s13347-020-00396-6>.

¹⁰⁸ Dreyfus, H. L. (1974). Artificial Intelligence. *The Annals of the American Academy of Political and Social Science*, 412, 21–33. <http://www.jstor.org/stable/1040396>

¹⁰⁹ *Id.*, 32.

La disillusione che Dreyfus manifesta, è però attenuata comunque da un senso di fiducia nei confronti del futuro sviluppo dell'IA. Se infatti la possibilità di riprodurre l'intelligenza umana *tout court* artificialmente non è fisicamente possibile, rimaneva aperta, a quel tempo, la possibilità di immaginare fin dove le macchine si sarebbero spinte nel “simulare” la *performance* intellettuale degli umani. Secondo il filosofo americano, per riprodurre un comportamento umano, le macchine “*richiederebbe[ro], come minimo, l'archiviazione e il facile accesso a grandi quantità di dati. Solo in questo modo il computer potrebbe iniziare a simulare l'uso dell'esperienza passata e del contesto attuale per strutturare l'esperienza presente, un processo che sembra essere necessario per qualsiasi riconoscimento di modelli complessi o per la risoluzione di problemi.*”¹¹⁰ Nel 1974, l'archiviazione e l'accesso agevole a vaste quantità di dati erano solo un concetto astratto meramente potenziale; ad oggi, questa visione si è concretizzata nella realtà dei *Big Data* impiegati nel ML con l'esponenziale aumento delle capacità di computazione e archiviazione dei computer attuali.

Il *training* algoritmico così come descritto nel sottoparagrafo 2.2.1 non può essere analizzato in maniera avulsa da un altro aspetto cruciale per comprendere il corretto funzionamento dei sistemi che utilizzano l'IA: i dati. Essi sono la materia prima che foraggia gli algoritmi di IA: è solo grazie l'attuale presenza di ingenti quantità di dati che è possibile allenare sistemi così complessi come quelli che sono emersi recentemente e hanno catturato l'attenzione dell'opinione pubblica e dei media. La centralità dei dati nel campo dell'IA risiede principalmente nella loro natura aggregata: i c.d. *Big data*, cioè enormi masse di informazioni collegate, attribuibili (o meno) a dei soggetti individuati o individuabili¹¹¹, che costituiscono il vero e proprio “carburante” dell'IA¹¹². Non a caso, quell'attività tipica di ML con i quali i sistemi *AI-powered* sono in grado di evidenziare *pattern*, trovare correlazioni su cui basare successivamente le decisioni algoritmiche è chiamato “*data mining*.”¹¹³

¹¹⁰ Dreyfus (1974), 32-33.

¹¹¹ I *Big Data* possono contenere anche dati personali (secondo quella definizione di dato personale contenuta all'art. 4(1) del GDPR); ma possono anche essere costituiti in gran parte da dati non necessariamente riconducibili ad un individuo, ad esempio dati scientifici, di rilevamento dell'ambiente circostante, meteorologici.

¹¹² Per una definizione accurata di *Big Data*, vedi nota 24. La definizione “*Big*” non si riferisce necessariamente alla loro intrinseca caratteristica di enormi quantità di informazioni, ma nella loro natura aggregata.

¹¹³ Traducibile in italiano come “estrazione di dati/attività estrattiva di dati.” Contissa (2017), 106.

Ancora, la centralità dei dati come risorsa fondamentale per l'economia digitale in generale è ancora più enfatizzata nel contesto europeo: già nel 2018, la Commissione Europea aveva indicato i dati come “*la materia prima del mercato unico digitale*,”¹¹⁴ segnando che questi avrebbero ricoperto un ruolo centrale non solo dal punto di vista del mercato unico, ma anche, dal punto di vista normativo. Proprio l'Unione Europea, nella “strategia europea per i dati” (2020) ha chiaramente identificato una linea direttiva comunitaria per la condivisione e l'utilizzo dei dati in un contesto conforme ai principi di matrice europea, sempre e necessariamente in linea con la tutela della privacy, secondo quei principi cardine stabiliti dal GDPR¹¹⁵. I successivi interventi normativi del c.d. “*Data Package*”¹¹⁶, contenenti il *Data Governance Act* (entrato in vigore quest'anno dopo un periodo di *vacatio*¹¹⁷) e la successiva proposta di Regolamento che denominata c.d. *Data Act*¹¹⁸, sono l'esplicitazione a livello legislativo della necessità per l'UE di mettere al centro una delle risorse più importanti di cui essa dispone sfruttarne il valore economico *in primis* a livello interno. Ciò è da valutare anche nell'ottica del raggiungimento di un certo grado di “sovranià digitale” dell'UE sia dal punto di vista economico che normativo, nei confronti delle altre grandi potenze mondiali¹¹⁹. Ed infatti, non a caso si è ritenuto che, rispetto ad altri (come, ad esempio, gli Stati Uniti d'America o la Repubblica Popolare Cinese, che dispongono rispettivamente di un fortissimo potere digitale produttivo combinato con una spinta per il progresso gli uni, e una forte capacità di fornire infrastrutture e disponibilità di materie prime l'altra¹²⁰), l'Unione abbia come risorsa principale proprio i dati europei, grazie ai quali è capace di fare leva nei confronti degli altri *big players* a livello globale¹²¹.

¹¹⁴ Comunicazione della Commissione Europea “Verso uno spazio comune europeo dei dati” COM 232 final, 25 aprile 2018.

¹¹⁵ Comunicazione della Commissione Europea, “Una strategia europea per i dati”, COM 66 final, 19 febbraio 2020.

¹¹⁶ Così denominato in Smuha, N. A. (2023). Digital Sovereignty in the European Union: Five Challenges from a Normative Perspectives. *Working Paper - ERA Conference Proceedings*, 3.

¹¹⁷ Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio, del 30 maggio 2022, relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724 (*Data Governance Act*), 30 maggio 2022.

¹¹⁸ Proposta di Regolamento del Parlamento Europeo e del Consiglio riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo (*Data Act*), COM 68 final, 23 febbraio 2022.

¹¹⁹ Smuha (2023), 3, 6-9.

¹²⁰ Bradford, A. (2023). *Digital empires: The Global Battle to Regulate Technology*. Oxford University Press.

¹²¹ *Id.*, 354-355.

Talvolta, alcuni hanno definito i dati come il nuovo “petrolio,”¹²² contrapponendo la vitalità di questa risorsa all’importanza che essi ricoprono nel contesto digitale attuale, ancor di più con riguardo ai sistemi di IA. I dati, infatti, condividono con il combustibile fossile la necessità di essere sottoposti ad un processo di raffinazione accurato, prima di essere utilizzati e prima di poter estrarre il loro valore nel contesto economico¹²³. Ancora, il paragone è certamente calzante nel caso dei dati che fungono per l’appunto da “carburante” ai sistemi di IA, poiché questa “*ha bisogno di dati per essere addestrata e pertanto di dati per applicare il suo addestramento.*”¹²⁴ D’altronde l’industria del carbone fossile e dei dati hanno in comune la natura prettamente “estrattiva”: questi, collezionati allo stato grezzo, devono essere correttamente rielaborati per essere sfruttati, sia per fini commerciali¹²⁵ che per i fini di *training* algoritmico. È dunque certamente efficace la contrapposizione tra dati e petrolio come risorse le quali, se adeguatamente oggetto di raffinazione, hanno entrambe un’elevata attitudine ad essere oggetto di sfruttamento economico.

Tuttavia, questa analogia non è più sostenibile oltre ai punti in cui è stata appena esaminata. Il paragone non risulta essere più così appropriato quando si prendono in considerazione altre caratteristiche dei dati: essi, a differenza di una risorsa come il petrolio, sono durevoli, non esclusivi, facilmente trasportabili (sia in termini di velocità che di praticità), condivisibili, e in ultimo, virtualmente infiniti. Infatti, essi sono oggetto di una crescita esponenziale inesorabile¹²⁶: si pensi solamente che, se nel 2018 la quantità di dati “*catturati, creati o riprodotti*” presenti nella c.d. “Infosfera”¹²⁷ era pari a 18 zettabyte¹²⁸, si prevede che entro il 2025 questi aumentino a circa 175 zettabyte¹²⁹. Un aumento esponenziale di questo tipo difficilmente si concilia con la natura intrinsecamente limitata e in costante diminuzione delle risorse petrolifere, che certamente non si generano con la stessa

¹²² Alongi, A., & Pompei, F. (2019). *Il valore dei dati nell’ecosistema digitale: etica e diritti nell’era del web*. EditorialeNovanta, 102.

¹²³ Floridi (2022), 67.

¹²⁴ *Id.*, 68.

¹²⁵ Alongi & Pompei (2018), 102.

¹²⁶ Floridi (2022), 67.

¹²⁷ Per la definizione di “Infosfera,” vedi nota 60.

¹²⁸ Un zettabyte è una unità di misura dell’informazione ed è equivalente a 10^{21} bytes. Quindi, uno zettabyte corrisponde a un miliardo di terabyte o mille miliardi di byte.

¹²⁹ Floridi (2022), 26. Vedi anche Libro Bianco della Commissione Europea sull’Intelligenza Artificiale – *Un approccio europeo all’eccellenza e alla fiducia*, 11 COM (2020) 65 final (19 febbraio 2020), 4.

facilità¹³⁰. Ancora, questi due risorse differiscono in base al prezzo che ad esse viene attribuito: mentre il greggio ha un prezzo ben preciso, la monetizzazione dei dati dipende dal contesto in cui questi vengono impiegati e dall'utilizzo che si intende farne¹³¹.

Il valore economico dei dati è ad oggi un aspetto certamente indiscutibile e noto a chiunque viva immerso nell'Infosfera, e dunque non limitato solo ai tecnici del settore. Questo perché, alla nozione di "dato" a sé stante, va affiancata anche la nozione, ben più articolata, di dato "personale." I dati non riconducibili ad un individuo specifico, un soggetto c.d. "interessato,"¹³² hanno un'utilità e un valore commerciale sostanzialmente diversi dai dati personali, che contengono al loro interno informazioni personali dalle molteplici implicazioni commerciali e no.

Innanzitutto, i dati personali sono quelli *"riguardant[i] una persona fisica identificata o identificabile [...] si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale."*¹³³ La presenza di grandi quantità di dati personali ha, a suo tempo, dato origine una modifica del mercato per ricavare un nuovo tipo di valore economico, secondo un processo che ha modificato sensibilmente i rapporti tra stati, cittadini, governi, mercati e organizzazioni¹³⁴.

Un esempio noto di come il pubblico è venuto a conoscenza di queste nuove metodologie di estrazione del valore dei dati è il celebre caso Cambridge Analytica, con cui l'omonima società inglese, che si occupava di ricerca sociologica nel settore politico, ha ottenuto i dati personali di milioni di utenti Facebook senza il loro consenso per tracciare mappature psicografiche di elettori da utilizzare nel fornire pubblicità targettizzata agli utenti delle piattaforme social¹³⁵. Le conseguenze nocive di un utilizzo dei dati personali di questo tipo, volto a manipolare l'opinione pubblica (a propria insaputa) e in ultimo i processi democratici, ha gettato luce sulla

¹³⁰ Floridi (2022), 67.

¹³¹ *Id.*

¹³² La normativa europea sui dati personali (GDPR) denomina i soggetti a cui i dati sono riconducibili i cosiddetti "interessati al trattamento" (o in inglese, "*data subjects*"). Art. 4(1) GDPR.

¹³³ *Id.*

¹³⁴ Frosini, T. E., *L'Ordine Giuridico del digitale*, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli, 28.

¹³⁵ Frosini, 28. Vedi anche Wylie, C. (2020). *Il mercato del consenso: come ho creato e poi distrutto Cambridge Analytica*. Longanesi.

portata economica e sociale dell'utilizzo di dati personali degli utenti sulle piattaforme social, "auto schedature volontarie di massa"¹³⁶ nella disponibilità delle *Big tech* intenzionate a sfruttarne le potenzialità generative di profitto fino all'ultima goccia.

Da ciò emerge un particolare rilevante: sebbene siano i singoli individui che interagiscono quotidianamente con i sistemi informatici a generare i propri dati personali, altri beneficiano dalla capitalizzazione di questi ultimi¹³⁷. Sono i grandi attori nel campo del digitale delle piattaforme online a giovare *in primis* e in abbondanza di quel plusvalore derivante dai dati personali: dati in grado di fornire un quadro altamente dettagliato dell'interessato, delle sue caratteristiche, abitudini, inclinazioni e necessità, collettivamente identificati con una nozione brillantemente coniata dalla Prof.ssa Shoshana Zuboff, celebre autrice dell'opera "*Il capitalismo della sorveglianza*"¹³⁸: il c.d. "*surplus comportamentale*."¹³⁹ Lo scandalo di Cambridge Analytica ha permesso l'inizio di un processo di maturazione della coscienza collettiva relativa ai dati personali e al loro sfruttamento. Nel campo dell'IA, infatti, questo sfruttamento non può che presumersi assolutamente presente poiché materialmente funzionale allo sviluppo dei sistemi che utilizzano questa tecnologia.

È d'uopo sottolineare come la convergenza tra temi di tutela dei dati personali e IA sia sotto la lente d'ingrandimento delle autorità indipendenti europee: prima fra tutte, l'Autorità Garante della Protezione dei Dati italiana che, con il provvedimento n. 112 del 30 marzo 2023, ha temporaneamente limitato l'utilizzo del noto chatbot di OpenAI, ChatGPT, a causa della sua *non-compliance* rispetto la normativa europea ed italiana di tutela dei dati personali¹⁴⁰. Si tornerà su questo tema nel paragrafo 3.2, in cui verrà discusso, *inter alia*, il complesso equilibrio tra tutela della privacy e utilizzo dei dati personali per il *training* algoritmico.

Così come è stato fatto per il Machine Learning (sottoparagrafo 2.2.1.), è opportuno evidenziare in questa sede le criticità che "il carburante" dell'IA, cioè i dati, possono generare nel momento in cui i sistemi *AI-powered* vengono utilizzati.

¹³⁶ Frosini, (2023), 28.

¹³⁷ Alongi & Pompei (2018), 107.

¹³⁸ Zuboff, S. (2019) *Il capitalismo della sorveglianza: il futuro dell'umanità nell'era dei nuovi poteri*, LUISS University Press.

¹³⁹ Il "surplus comportamentale" viene definito da Zuboff come il valore generato dall'esperienza complessiva degli utenti sulle piattaforme Internet, che viene successivamente utilizzato per la profilazione degli utenti a fini di marketing. Zuboff, (2019) Capitolo 3: il surplus comportamentale.

¹⁴⁰ Provvedimento n. 112 del 30 marzo 2023, Autorità Garante per la protezione dei dati personali.

Questo poiché, in via assolutamente generale, la tecnologia dell'IA, in qualsiasi settore o contesto venga dispiegata, non è mai neutrale, né va percepita come tale sulla scorta di un pensiero “tecnosoluzionista” che tende a considerare la tecnologia come una sorta di panacea¹⁴¹. Nel campo dell'IA, quest'attitudine tecnosoluzionista si traduce in una percezione umana di oggettività meccanicistica associata dall'analisi generata dagli algoritmi di IA. Ciò può comportare che i decisori umani trascurino le proprie valutazioni per l'appunto “umane”, poiché ripongono una maggior fiducia nelle macchine in base ad un c.d. “pregiudizio dell'automazione.”¹⁴²

Fare affidamento su un modello di IA “a scatola nera,” i cui meccanismi decisionali non sono né trasparenti né comunque comprensibili ai più, significa non solo fare affidamento sulle computazioni utilizzate da suddetto modello, ma anche sull'intero *database* su cui esso è stato allenato¹⁴³. Ogni serie di dati ragionevolmente complessa (come gli immensi *datasets* utilizzati nel *training* algoritmico) contiene inevitabilmente delle imperfezioni di natura variabile: da enormi quantità di dati mancanti (che non sono casuali), a errori sistematici nel set di dati, a problemi di raccolta dei dati che causano una distribuzione dei dati diversa da quella originariamente prevista¹⁴⁴. Tutte queste ipotesi di *databases* in qualche modo “inquinati” generano dei pregiudizi o polarizzazioni, i c.d. *bias*.

Ancora, un fattore strutturale che talvolta può risultare in decisioni automatizzate che portano con sé delle problematiche per coloro che ne sono soggetti, consiste nel fatto che qualsiasi modello di algoritmo matematico che viene progettato, viene addestrato sulla base di dati storici, appartenenti al passato¹⁴⁵. Questi costituiranno quell'insieme di informazioni sulla base delle quali i sistemi di IA formuleranno le proprie decisioni: per quanto aggiornati siano, essi risalgono, inevitabilmente al passato. Pertanto, se sono presenti delle polarizzazioni di qualsiasi sorta, queste riverbereranno nelle stesse decisioni algoritmiche, amplificandone l'incisività (potenzialmente negativa) sul contesto sociale in cui viene impiegato l'algoritmo.

Un esempio, seppure ormai risalente ad una decade fa, può essere presentato per meglio comprendere il tipo di conseguenze a cui l'utilizzo di un certo *dataset*

¹⁴¹ Floridi (2022), 86-88, 146-147.

¹⁴² *Id.*

¹⁴³ Rudin & Radin (2019), 5-6.

¹⁴⁴ *Id.*

¹⁴⁵ O'Neil (2016), 38.

contente *bias* può portare. Nel 2014, Amazon.com, Inc. sviluppò un algoritmo di IA da implementare nel sistema di *recruiting* di figure tecniche (ingegneri, *developers*, ecc.).¹⁴⁶ L'algoritmo era in grado di esaminare migliaia di domande di lavoro contemporaneamente, automatizzando e ottimizzando considerevolmente i processi di assunzione. Ciò avveniva tramite una classificazione dei *curricula* dei vari candidati, per invitare ai colloqui di selezione solo coloro che avessero ricevuto i punteggi più alti secondo la graduatoria prodotta dall'algoritmo. Tuttavia, il modello continuava a penalizzare i *curricula* che contenevano la parola "donna," o "femmina," conferendo a questi punteggi più bassi. Infatti, l'algoritmo aveva ricevuto come dati di *training* (e continuava a ricevere durante il suo utilizzo) candidature prevalentemente provenienti da uomini, semplicemente perché, in proporzione, un numero maggiore di uomini si candidava a ruoli tecnici quali *software engineer* o sviluppatori. Il *bias*, indubbiamente già presente nel mondo del lavoro del settore STEM¹⁴⁷ tradizionalmente a prevalenza maschile, si è sostanzialmente riprodotto nell'algoritmo, creando un circolo vizioso interno alla *black box* decisionale: l'algoritmo, sebbene astrattamente idoneo a massimizzare l'efficienza nelle selezioni di nuovi dipendenti Amazon così come era stato progettato, ha in realtà efficientemente amplificato a livello esponenziale un pregiudizio di genere che, in assenza di automatizzazione, sarebbe rimasto (eventualmente) come *bias* umano della sola figura preposta ad esaminare le candidature e a valutarle.

A rendere ancora più incisive le conseguenze negative di decisioni algoritmiche di questo genere, vi è la natura prettamente *self-reinforcing* degli algoritmi di IA. Nel caso appena esaminato, le candidature provenienti da uomini, in misura sempre superiore rispetto alle donne, allenavano il sistema ad accettare più candidature maschili, secondo il *training* continuo dell'algoritmo che continua ad apprendere dalla propria esperienza. Più posizioni lavorative ottenute, in ultimo, da uomini, non facevano altro che confermare il *bias* della macchina, che riproponeva quello umano nelle proprie decisioni automatizzate. A ragione, uno dei caratteri strutturali di quei modelli di IA dotati di un certo grado di opacità e di una capacità di causare potenziali "danni" di qualsiasi natura, come l'algoritmo per le

¹⁴⁶ Dastin, J. (10 ottobre 2018). Insight - Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.

¹⁴⁷ L'acronimo indica, in inglese, le discipline quali *science, technology, engineering and mathematics*.

valutazioni dei *curricula* impiegato da Amazon, è proprio l'attitudine ad amplificare gli stessi, che spesso si verificano su larghissima scala¹⁴⁸.

Dunque, un errore o un *bias* nel *dataset* dato in pasto all'algoritmo può determinare un funzionamento che incide negativamente sul tessuto sociale su larga scala. A ben vedere, questo dato preso singolarmente, potrebbe, sul lungo periodo, disincentivare lo sviluppo di sistemi di IA, a causa delle conseguenze sociali pregiudizievoli che potrebbero insorgere. Al momento però, la capacità di utilizzare algoritmi come quello di impiegato da Amazon, ha determinato principalmente un grande vantaggio per le aziende, che beneficiano di risparmi significativi per i processi ottimizzati tramite l'automazione¹⁴⁹.

Ammesso che sia anche possibile osservare la presenza di pregiudizi algoritmici, l'eradicazione di *bias* presenti nei dati non è tuttavia cosa facile, soprattutto quando questi hanno matrice di genere, razziale, o di classe: alcuni infatti sostengono che “*le diseguglianze strutturali di vecchia data sono spesso profondamente radicate negli insiemi di dati degli algoritmi e raramente, se non mai, emendate.*”¹⁵⁰ A ragione, è stato osservato che “*la profondità computazionale senza la profondità storica o sociologica è soltanto un apprendimento superficiale e non un apprendimento profondo* (è evidente il gioco di parole tra l'apprendimento profondo, cioè il *deep learning*, e quello denominato “superficiale” poiché tralascia gli aspetti storici e sociali da applicare nelle valutazioni dello stesso algoritmo).”¹⁵¹

Può tuttavia essere menzionata una tendenza del mondo accademico relativo al riconoscimento dei potenziali *bias* presenti negli algoritmi che fa ben sperare. Infatti, lo studio dei pregiudizi algoritmici è oggetto di analisi da parte degli studiosi in proporzione sempre crescente: il tasso di pubblicazioni scientifiche che si occupano di questo tema nell'ambito di conferenze internazionali quali la *Conference on Fairness, Accountability, and Transparency* (FAccT) o la

¹⁴⁸ La matematica e scienziata dei dati Cathy O'Neil ha coniato un'etichetta evocativa e auto esplicativa per i sistemi algoritmici che combinano scalabilità, danno potenziale e opacità, in una sorta di arma a tre tagli: queste sono le c.d. *Weapons of Math Destruction* o Armi di distruzione matematica; evidente è il gioco di parole con l'espressione inglese *weapons of mass destruction*. Vedi O'Neil (2016), 29-31.

¹⁴⁹ O'Neil (2016), 118.

¹⁵⁰ Floridi (2022), 152.

¹⁵¹ *Id.*, citando Ruha Benjamin, sociologa e professoressa all'Università di Princeton che studia il rapporto tra innovazione ed equità, in particolare razza, giustizia e tecnologia *Ruha Benjamin on deep learning: Computational depth without sociological depth is 'superficial learning.'* (n.d.). <https://www.techregister.co.uk/ruha-benjamin-on-deep-learning-computational-depth-without-sociological-depth-is-superficial-learning/>.

Conference on Neural Information Processing Systems (NeurIPS) è aumentato considerevolmente negli ultimi due anni¹⁵².

Per concludere la presente sezione sui dati, va fatta menzione di alcune innovazioni che, se adeguatamente sfruttate in futuro, possono portare alla minimizzazione dei rischi inerenti ai dati utilizzati nel *training* algoritmico, aprendo le prospettive per un futuro differente per questi sistemi. Fino ad oggi, è pacificamente valso il principio secondo il quale maggiore la quantità di dati, tanto migliore è il risultato raggiunto dal sistema di IA, in termini di accuratezza¹⁵³. Tuttavia, questa associazione “*quantità dei dati* → *qualità dei risultati*” potrebbe starsi ribaltando proprio grazie allo sviluppo dell’IA stessa: un insieme di dati, ben curati e accuratamente selezionati, possono portare a risultati ancor più precisi e con un’alta probabilità di essere sostanzialmente privi di *bias* o gravi errori, in quanto soggetti ad un controllo di qualità più certosino¹⁵⁴. In futuro, dunque, potrebbe essere capovolta la narrativa dei *Big data* da impiegare nel *training* di algoritmi, e preferire i dati “piccoli,” adottando una strategia di “*quality over quantity*.”

In ultimo, un’altra tendenza che attiene alla qualità dei dati utilizzati nei *training* degli algoritmi è l’utilizzo dei dati c.d. “sintetici,” cioè quelli integralmente creati dalla stessa IA per porre una soluzione permanente al susseguirsi di preoccupazioni nel campo della protezione dei dati personali impiegati come base di addestramento dei modelli, specialmente quelli utilizzati in ambito medico, che utilizzano dati sanitari dotati di un alto livello di sensibilità¹⁵⁵. I dati sintetici sono autogenerati dall’IA stessa, strutturati in modo da conservare le caratteristiche di base dei dati personali di individui direttamente identificati o identificabili; tuttavia, essi non sono realmente riconducibili ad alcun soggetto interessato, in quanto sono “sintetizzati” dalla macchina stessa, senza per l’appunto incorrere in violazioni di

¹⁵² *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 3: Technical AI Ethics, 158, 164. (2023)

¹⁵³ Floridi (2022), 68.

¹⁵⁴ Floridi, per spiegare la differenza tra un *training* basato su enormi quantità di dati e dati in minori quantità ma molto più “curati,” mette a confronto due sistemi di IA entrambi sviluppati nel settore sanitario: il primo, sviluppato in ambito pediatrico, è stato allenato sulla base di 101,6 milioni di *datapoints* derivanti dalle cartelle cliniche dei piccoli assistiti di un centro medico a Guangzhou, in Cina, affinché questo fornisca un supporto diagnostico ai pediatri; il secondo, invece, sviluppato dall’azienda Deepmind di Google con lo scopo di fornire un supporto ai professionisti dell’oftalmologia nella diagnosi di malattie retiniche, è stato addestrato solamente su 14.886 scansioni di retine dei pazienti. Al di là dei diversi scopi perseguiti da questi due sistemi di IA, il divario nella quantità di dati utilizzati per il *training* non si è riflesso sulla qualità dell’*output* ottenuto, in quanto questi hanno entrambi performato con ottimi risultati. *Id.*, 68-69.

¹⁵⁵ *Id.*, 71.

privacy o riservatezza¹⁵⁶. Con questo genere di dati “l’IA non è mai costretta a lasciare il suo spazio digitale, dove può esercitare il controllo completo su qualsiasi input e output dei suoi processi.”¹⁵⁷ Se utilizzati con accortezza e con un elevato controllo sui rendimenti delle macchine che utilizzano in tutto o in parte i dati sintetici per il proprio allenamento, anch’essi possono costituire una strategia vincente di *privacy by design* che minimizza i rischi per la privacy dei soggetti interessati nel *training* algoritmico.

2.3 La definizione normativa

È d’uopo a questo punto soffermarsi sulla definizione normativa dell’Intelligenza Artificiale. La scelta di offrire una visione multilaterale e sfaccettata dell’IA, piuttosto che un’unica nozione identificativa di natura tecnica, è anche sintomo dell’incremento, a livello globale, di testi normativi che introducono il concetto di IA all’interno degli ordinamenti. Secondo l’Artificial Intelligence Index Report 2023 dello *Stanford Institute for Human-Centered Artificial Intelligence* (HAI), dal 2016 al 2023 si è registrato un aumento dell’utilizzo del termine “IA” nei disegni legislativi a livello nazionale e internazionale: se nel 2016 l’IA era stata menzionata solo una volta dai disegni di legge, ad oggi si registrano più di 37 proposte di legge attive, sia in corso di approvazione o anche solo semplicemente proposte di fronte agli organi legislativi di varie nazioni ed organi internazionali¹⁵⁸ (tra cui figura anche l’Unione Europea con il c.d. *AI Act*, che sarà estensivamente trattato nel capitolo 2).

Non è dunque più possibile prescindere dal fornire un quadro sulle nozioni adottate nel campo legislativo. Queste, infatti, tendono a soddisfare un fine completamente diverso dalla definizione tecnico-scientifica, volta ad evidenziare i requisiti necessari e sufficienti di cui i sistemi devono disporre per essere considerati *AI-powered*¹⁵⁹. Al contrario, le definizioni contenute dalle fonti normative hanno come scopo primario e principale quello di rendere precisi i limiti di ambito, applicabilità e implementazione delle regole che esse stesse contengono¹⁶⁰.

¹⁵⁶ *Id.*

¹⁵⁷ *Id.*, 72.

¹⁵⁸ *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 6: policy and governance, 266-284 (2023).

¹⁵⁹ Floridi, L. (2023) L. On the Brussels-Washington consensus about the legal definition of Artificial Intelligence. *Philos. Technol.* 1-14, <https://doi.org/10.1007/s13347-020-00396-6>.

¹⁶⁰ Floridi (2023), 12.

A ragione, redigere una definizione per l'IA è un compito certamente complesso che ha richiesto (e tuttora richiede) il raggiungimento di compromessi a livello politico. Un esempio è quello raggiunto dall'UE, seppure la normativa sull'IA sia ancora in via di approvazione nel momento in cui si scrive, per cui il compromesso raggiunto dalle forze politiche può considerarsi ancora non scolpito nella pietra, né certamente avente alcuna forza di legge. Infatti, negli stadi di genesi del regolamento denominato *AI Act* presso il Parlamento Europeo, ma anche nei successivi “triloghi” che hanno avuto luogo tra le istituzioni dell'UE dopo l'approvazione del testo da parte dell'organo rappresentativo comunitario¹⁶¹, si è osservata una divergenza di opinioni tra le correnti più conservatrici, che preferivano una definizione più ristretta di IA, rispetto l'ala dei progressisti (compresi gli stessi correlatori dell'*AI Act*, tra cui figura il parlamentare italiano Brando Benifei), più inclini ad una nozione omnicomprensiva di questa tecnologia¹⁶².

Alla base della complessità di delineare normativamente i tratti fondamentali di questa tecnologia, ormai più che semplicemente emergente, la cui necessità di regolamentazione è sempre più evidente, vi sono diversi fattori che meritano un'analisi preliminare e separata rispetto a quelle generali questioni esaminate *supra*, nel paragrafo 2.

In *primis*, vi è la sfida rappresentata dal cristallizzare nella lettera della legge una tecnologia che è soggetta ad un processo di innovazione ad una velocità esponenzialmente crescente¹⁶³. Questa difficoltà, si è tradotta e in una delle critiche che più frequentemente viene mossa nei confronti di normative come l'*AI Act*, cioè che questa, seppure non ancora approvata né entrata in vigore, sarebbe già obsoleta o potrebbe esserlo molto presto. Riducendo come obiettivo di tale critica solo la definizione normativa dell'IA, questa corrente di pensiero intende sottolineare la possibilità che, una volta scolpito nella lettera della legge il concetto di IA, questa tecnologia possa evolversi in maniera imprevedibile, anche in modo tale da eludere

¹⁶¹ Per una completa cronologia delle vicende che hanno interessato e correntemente interessano l'Unione Europea, vedi il paragrafo 3.1 del Capitolo 2.

¹⁶² Bertuzzi, L. (8 novembre 2022). *Artificial Intelligence definition, governance on MEPs' menu*. [www.euractiv.com](https://euractiv.com). <https://euractiv.com/section/digital/news/artificial-intelligence-definition-governance-on-meps-menu/>. Vedi anche Bertuzzi, L. (7 marzo 2023). *EU lawmakers set to settle on OECD definition for Artificial Intelligence*. www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/>.

¹⁶³ Schepisi, C., *Brevi note sulla “dimensione europea” della regolamentazione dell'intelligenza artificiale*, 57, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.

le disposizioni di legge e gli standard da esse predisposti. Tentare di circoscrivere un concetto così multiforme come l'IA potrebbe creare dei vuoti normativi che consentano a chi produce questi sistema di eludere i requisiti predisposti, con il rischio di rendere sostanzialmente inefficace l'intera regolamentazione (anche vi è da sottolineare che, la semplice corrispondenza di un sistema alla definizione di IA fornita dal legislatore europeo non necessariamente determina l'insorgere di stringenti requisiti e oneri, in quanto ciò accade quando un sistema pone dei determinati "rischi"¹⁶⁴). Lo stesso rischio di rapida obsolescenza delle normative sull'IA potrebbe nel caso in cui la definizione normativa di IA dovesse contenere al contrario tecnicismi eccessivi, incatenando la nozione di IA allo stato dell'arte al momento dell'entrata in vigore della normativa¹⁶⁵.

Per fare un esempio del tipo di conseguenze derivanti da una definizione eccessivamente ristretta, se l'IA fosse stata definita come un "*software*" anziché un "sistema" (*infra*), si sarebbero potuti escludere dall'applicabilità della normativa quei sistemi che integrano l'IA in un *hardware*, come ad esempio i dispositivi del c.d. *Internet of Things*, che al contrario costituiscono un settore cruciale delle potenziali applicazioni pratiche di questa tecnologia, presente negli odierni *smart devices*, che lentamente stanno sempre più iniziando a far parte della quotidianità¹⁶⁶.

Una circostanza così descritta, dove la regolamentazione e l'innovazione cercano di inseguirsi per non rimanere l'una succube dell'altra, costituisce un tema comune a tutto il mondo del digitale, e un dilemma che ha caratterizzato e tutt'ora contraddistingue questo settore¹⁶⁷. Questo tema è presente "in sottofondo" all'interno di questo intero elaborato: si può preliminarmente considerare, in questo momento, come secondo alcuni l'apparente contrasto tra innovazione e regolamentazione non dovrebbe essere percepito come tale: la normazione non dovrebbe infatti incidere sul ritmo, ma sulla direzione dell'innovazione, tendendo a disegnare un corretto sviluppo della società¹⁶⁸.

A causa di queste problematiche, dunque, si possono presentare due definizioni normative di IA, fornite l'una dall'Unione Europea nella proposta di

¹⁶⁴ Schepisi (2023), 60.

¹⁶⁵ Schepisi (2023), 58. Vedi anche Benbouzid et al. (2022), 21, 28.

¹⁶⁶ Floridi (2023), 5. L'*Internet of Things* (IoT) o "Internet delle cose" consiste nell'insieme di dispositivi fisici connessi alla rete ed interconnessi tra loro che popolano da ormai diverso tempo la realtà quotidiana, consentendo, fra le altre cose, la conservazione e l'analisi di dati in tempo reale.

¹⁶⁷ Bradford, (2023) 136-138, 370.

¹⁶⁸ Come riportato da Floridi, "se ci piace dove stiamo andando, possiamo andarci alla velocità che vogliamo." Floridi (2022), 86.

regolamento europeo, e l'altra dalla più recente fonte normativa proveniente dagli USA che ha come oggetto l'IA, cioè il c.d. *Executive Order on the Safe, Secure and Trustworthy Development and use of Artificial Intelligence* (di seguito denominato l'*Executive Order*¹⁶⁹). Sulla diversità di questi testi normativi si tornerà nei successivi capitoli dell'elaborato: tuttavia, lo scopo di questo confronto di definizioni in questa fase prodromica è quello di sottolineare la tendenziale convergenza tra le definizioni adottate da USA e UE che, seppure approssino in maniera completamente diversa il tema della regolamentazione dei sistemi *AI-powered* e del digitale in generale¹⁷⁰, hanno adottato (o stanno per adottare, nel caso dell'*AI Act*), definizioni molto simili del concetto normativo di IA¹⁷¹. L'ultima versione testo dell'*AI Act*, mentre i negoziati per la finalizzazione del suo contenuto sono ancora in corso al momento in cui si scrive¹⁷², definisce l'IA alla seguente maniera:

*“Un sistema progettato per operare con elementi di autonomia e che, sulla base di dati e input forniti dalla [stessa] macchina e/o dall'uomo, deduce come raggiungere un determinato insieme di obiettivi utilizzando approcci di apprendimento automatico e/o di logica e conoscenza, e produce output generati dal sistema, come predizioni di contenuto (sistemi di IA generativi), raccomandazioni o decisioni che influenzano gli ambienti con cui il sistema di IA interagisce.”*¹⁷³

In *primis*, tratto saliente di questa definizione l'approccio tecnologicamente “neutrale” nell'indicare l'IA come un “sistema” (senza ulteriori specifiche tecniche). Tuttavia, nella versione precedente dell'*AI Act*, era stata diversamente definita come un *software*¹⁷⁴. Viceversa, l'IA è qui delineata quindi come un

¹⁶⁹ Exec. Order No. 14110, Fed. Reg. 75191 (30 ottobre 2023) (d'ora in poi abbreviato in Exec. Ord. No. 14110).

¹⁷⁰ Gli Stati Uniti infatti hanno un approccio alla normazione del digitale c.d. *market-driven*, (guidato dal mercato), lasciando che regole e limitazioni alla tecnologia siano appannaggio della *tech industry*, piuttosto che dello stato, nella (pressoché) ferma convinzione che questa sia la migliore strategia per non porre un freno all'innovazione. Al contrario, l'Unione Europea ha un approccio regolamentare al digitale sostanzialmente incentrato sulla tutela dei diritti fondamentali contenuti nelle fonti normative europee di più alto rango, con un intervento nel mercato certamente più incisivo rispetto agli statunitensi. *Vedi in generale* Bradford (2023).

¹⁷¹ Floridi (2023) 3-14.

¹⁷² Per il punto sul percorso legislativo dell'*AI Act* e lo stato attuale della proposta di legge, *vedi* il paragrafo 3.1. del Capitolo 2.

¹⁷³ Art. 3(1), *AI Act* (modificato in ultimo nella terza versione del testo del regolamento, cioè quella oggetto del mandato del Consiglio, d'ora in avanti denominata “Versione III”).

¹⁷⁴ Art 3(1), *AI Act* (prima versione della Commissione Europea, d'ora in avanti denominata “Versione I”). *Vedi anche* Floridi (2023), 12.

sistema che, attraverso le computazioni di *input* e *output* di ML, è in grado di interagire con l'ambiente circostante tramite predizioni di contenuti o decisioni algoritmiche. Si è dunque scelto di adottare una definizione ampia e sostanzialmente elastica, che, tra l'altro, potrebbe favorire anche l'intervento da parte della Corte di Giustizia dell'UE, nell'ottica di precisare e chiarire di volta in volta determinati profili, e operare un adeguamento di significato di talune definizioni¹⁷⁵.

Tra i tratti distintivi della nozione appena esaminata, spicca la menzione della IA generativa, di recente introduzione nella bozza del regolamento¹⁷⁶: questo è da considerarsi come un chiaro indicatore della popolarità acquisita dall'IA generativa mentre il *drafting* dell'*AI Act* è ancora assolutamente *in itinere*, soggetto a modifiche continue che si muovono parallelamente all'evoluzione di questa tecnologia e del dibattito pubblico in merito¹⁷⁷.

Tra le criticità che si riscontrano nella definizione europea invece, vi è certamente l'espressione, finanche tautologica, per cui l'IA è un “*sistema [...] che [...] produce output generati dal sistema.*” Non si vede quali altri *output* potrebbe produrre il sistema stesso, se non quelli da esso stesso prodotti¹⁷⁸. Inoltre, rispetto alla prima versione dell'*AI Act*, la nozione di IA ha perso la previsione degli

¹⁷⁵ Schepisi (2023), 60.

¹⁷⁶ Art. 3(1), *AI Act* (Versione III).

¹⁷⁷ A questo proposito, si intende sottolineare l'ulteriore modifica subita dalla definizione prevista nell'*AI Act*, secondo quell'ultima versione della proposta di regolamento divulgata nel gennaio 2024 (oggetto della nota 411). Secondo quest'ultima versione della normativa, l'IA si definisce come “*sistema basato su una macchina progettato per operare con vari livelli di autonomia che può manifestare adattabilità dopo la messa in servizio e che, per obiettivi espliciti o impliciti, deduce, dagli input che riceve, come generare output come previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.*” Art. 3(1)(1), *AI Act*, (Versione IV). La definizione finale scelta in quest'ultima versione sembra discostarsi da quella dell'OCSE, per accogliere maggiormente la prospettiva statunitense, in ultimo, derivante dagli standard definiti dal *National Institute of Standards and Technology*. Quest'ultima definizione di IA fornita sembrerebbe potenzialmente escludere ad un primo impatto, i sistemi che non hanno un visibile meccanismo di adattabilità dopo la messa in servizio, andando dunque a ricomprendere solo quei modelli che si adattano a quanto “appreso” dopo l'immissione sul mercato. Ciò, dunque, potrebbe portare all'esclusione dall'applicabilità dell'intero *AI Act* quei sistemi di IA c.d. “*rule-based,*” cioè il cui *output* si basa principalmente su una amplissima serie di regole piuttosto che su un apprendimento continuo e crescente a seguito della messa in servizio. Un'esclusione di questo tipo potrebbe interessare quei sistemi *rule-based* complessi di IA che svolgono valutazioni automatizzate, diminuendo sensibilmente l'ampiezza dell'applicabilità della norma a sistemi potenzialmente dannosi per i diritti fondamentali (sistemi di IA come quelli che hanno determinato il “*toeslagenaffaire*” in Olanda menzionato nel paragrafo 3.2 sono *rule-based*). Vedi Tiulkanov, A. (23 gennaio 2024). *According to the leaked #AIAct text, the AI system's definition has undergone a controversial change that may potentially narrow the scope of the regulation.* LinkedIn. https://www.linkedin.com/feed/update/urn:li:activity:7155481476115873792/?updateEntityUrn=urn%3Ali%3Afs_updateV2%3A%28urn%3Ali%3Aactivity%3A7155481476115873792%2CFEED_DETAIL%2CEMPTY%2CDEFAULT%2Cfalse%29.

¹⁷⁸ Floridi (2023), 12.

obbiettivi che questa è programmata per perseguire, precedentemente presenti nel testo normativo e identificati come “obbiettivi definiti da umani” (*human-defined objectives*¹⁷⁹).

Il concetto di *human-defined objectives* è invece stato inserito e mantenuto nella più recente definizione adottata dagli USA e riportata nell'*Executive Order* emanato dal Presidente degli Stati Uniti nell'ottobre del 2023, che descrive così la medesima tecnologia:

*“Un sistema basato su una macchina che può, per un dato insieme di obiettivi definiti dall'uomo, fare previsioni, raccomandazioni o decisioni che influenzano ambienti reali o virtuali. I sistemi di intelligenza artificiale utilizzano input basati sulla macchina e sull'uomo per percepire ambienti reali e virtuali, astrarre tali percezioni in modelli attraverso l'analisi in modo automatizzato e utilizzare l'inferenza del modello per formulare opzioni di informazione o azione.”*¹⁸⁰

La nozione di IA fornita dalla disciplina statunitense risulta per certi versi molto simile a quella comunitaria, segnale di una convergenza o comunque di un consenso che coinvolge sia Bruxelles che Washington¹⁸¹. Nonostante i diversi approcci alla regolamentazione dell'UE e degli USA¹⁸², è certamente un segnale positivo di coordinamento a livello internazionale l'adozione di una simile determinazione a livello legislativo. Sarebbe da auspicare una conservazione di tale consenso non solo a livello internazionale; quindi, con il coinvolgimento di altri paesi interessati a regolamentare la tecnologia, ma anche *pro futuro* per far sì che anche i testi normativi successivi possano conservare un fondamento comune e condiviso. Tecnologie come l'IA non conoscono limitazioni territoriali applicative, motivo per il quale propendere per una definizione comune a livello globale aiuterebbe fortemente il coordinamento e, in ultimo, l'efficace regolamentazione di questi sistemi anche in mancanza di qualsiasi cornice o *standard* adottato a livello sovranazionale.

In ultimo, un'interessante sintesi delle definizioni statunitense, europea, con un influsso derivante anche da ulteriori nozioni provenienti da organizzazioni internazionali tra cui l'Organizzazione per la Cooperazione e lo Sviluppo

¹⁷⁹ Floridi (2023), 6, 10, 12.

¹⁸⁰ Exec. Ord. 14110, § 3(b).

¹⁸¹ Floridi (2023). A riprova dell'avvenuta convergenza tra USA e UE, *vedi* nota 177.

¹⁸² *Vedi* nota 170.

Economico (OCSE),¹⁸³ è stata offerta dal Luciano Floridi, professore presso l'Università di Oxford, filosofo, e fondatore del *Digital Ethics Center* presso l'università di Yale, nonché uno dei maggiori esperti italiani del settore:

*“L’Intelligenza Artificiale (IA) si riferisce a un sistema ingegnerizzato in grado, per un determinato insieme di obiettivi definiti dall'uomo, di generare output - come contenuti, previsioni, raccomandazioni o decisioni - di imparare dai dati storici, di migliorare il proprio comportamento e di influenzare persone e ambienti.”*¹⁸⁴

Questa lineare definizione sembra contenere i punti di forza presenti nei già menzionati testi normativi europeo e statunitense, come la neutralità tecnologica o l'inserimento degli “obiettivi” dell'IA definiti dall'uomo (*human-defined objective*), senza tuttavia replicarne le debolezze che si erano precedentemente considerate. È, dunque, da auspicarsi un avvicinamento ad una nozione normativa di IA il più possibile simile a questa, sempre tenendo in considerazione la necessità di convergenza a livello globale verso una definizione unica.

A questo proposito si può accogliere favorevolmente la scelta dell'OCSE di modificare la definizione generale di IA, da quella adottata nei Principi del 2018, per aggiornarla e renderla maggiormente in linea con quanto predisposto dall'*AI Act*¹⁸⁵: un avvicinamento delle definizioni adottate da varie istituzioni sembra essere la giusta strada per assicurare una maggiore armonizzazione e certezza del diritto in un contesto di rilevanza transnazionale.

3. *Le opportunità e i rischi dell'intelligenza artificiale*

È necessario ora soffermarsi sulle potenzialità e i rischi che l'avvento di una tecnologia come l'Intelligenza Artificiale può portare con sé. L'adozione di essa può determinare l'automatizzazione e l'ottimizzazione di molteplici processi all'interno di settori più svariati. Tuttavia, per la seguente analisi circa le opportunità di significativo progresso scientifico e sociale che l'IA può portare con sé, si sono selezionati tre ambiti di interesse: la medicina e la ricerca medica, il mondo del lavoro e, in ultimo, il cambiamento climatico. Questi tre campi sono stati

¹⁸³ Ma anche organizzazioni come per la creazione di standard normativi ISO/IEC, come predisposto dallo standard ISO/IEC 23894. Floridi (2023), 7. In merito al tema della fissazione di regole attraverso standardizzazione, vedi anche il paragrafo III.2 del Capitolo 3.

¹⁸⁴ *Id.*, 13.

¹⁸⁵ *OECD AI Principles* (OECD18), *OECD AI Principles* (OECD23).

selezionati per mostrare in cosa si traduce e può tradursi l'utilizzo dell'IA, se impiegata nel rispetto di regole a protezione non solo degli individui e dei diritti fondamentali a loro ascrivibili, ma anche dello stato di diritto, e della democrazia: non solo innovazione e progresso, ma anche lo sviluppo di una società più equa e giusta.

Con la consapevolezza di suddette opportunità, sarà possibile individuare i rischi inerenti a questo tipo di tecnologia, soprattutto per quanto riguarda il loro utilizzo in determinati settori di importanza cruciale che impattano direttamente sui diritti fondamentali degli individui: a titolo esemplificativo e non esaustivo figurano pubblica sicurezza, lavoro, *welfare* economico, salute. Lo scopo di una trattazione simile non deve rinvenirsi nella volontà di suscitare sentimento di rifiuto o di timore nei confronti del progresso e dell'utilizzo di una tecnologia capace di sì di innovare, ma anche di incidere sui diritti fondamentali e causare conseguenze potenzialmente perniciose. Diversamente, l'obiettivo è quello di portare avanti un ragionamento informato sull'IA e sulle sue implicazioni sia positive che negative: quest'ultime in grado di essere quantomeno arginate grazie all'utilizzo di una puntuale regolamentazione con lo sguardo rivolto ai diritti fondamentali e all'utilizzo etico di questi strumenti tecnologici.

A questo proposito, si è scelto di trattare separatamente i rischi attuali dell'IA e quelli considerati "virtuali", cioè quelli che, sebbene percepiti come i più opprimenti e di natura esistenziale, oggetto delle più recenti cronache nazionali e internazionali¹⁸⁶, sono quelli assolutamente più improbabili. Essi, tuttavia, non sono

¹⁸⁶ Si pensi in *primis* alla lettera pubblicata dal Future of Life Institute nel marzo 2023, in cui sia aziende produttrici di sistemi di IA che parte del mondo accademico hanno chiesto una "pausa" nello sviluppo dei grandi progetti di IA quali ad esempio i c.d. Large Language Models (LLMs) come ChatGPT, a causa degli altissimi rischi "esistenziali" che questi possono porre per la società. Nonostante alcuni punti sottolineati dalla lettera aperta siano condivisibili, quali la presenza di potenziali criticità con lo sviluppo di modelli di grandi dimensioni allenati secondo modalità poco trasparenti, modelli operanti come *black boxes* le cui decisioni sono impossibili da comprendere a pieno (mancanza di *explainability*), le conclusioni relative alla necessità di bloccare lo sviluppo di questi a causa di imprevedibili rischi esistenziali per l'umanità, sono da ritenersi parte di un pensiero catastrofista di matrice sensazionalistica. A riprova della diffusione di questo genere di retorica fatalista e della sua diffusione nel mondo *tech*, anche a seguito del rilascio di ChatGPT-3.5 nell'ottobre del 2022, possono essere considerate le recenti dichiarazioni del miliardario e CEO di Tesla (e altre società) Elon Musk (in ultimo, ospite all'AI Summit tenutosi nel Regno Unito nell'autunno del 2023 e culminato con la c.d. Dichiarazione di Bletchley che considera l'IA "*una delle più grandi minacce per l'umanità*," sostenendo che allo stato di cose, vi è una concreta possibilità che questa possa sfuggire al controllo umano. Vedi *Pause Giant AI Experiments: An Open Letter*, *The Future of Life Institute*, (22 marzo 2023) <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>. Vedi anche Milmo, D., & Stacey, K. (1° novembre 2023). 'It's not clear we can control it': what they said at the Bletchley Park AI summit. *The Guardian*. <https://www.theguardian.com/technology/2023/nov/01/elon-musk-calls-ai-one-of-the-biggest-threats-to-humanity-at-summit>. Vedi anche il paragrafo 3.3.

privi di rilevanza nel dibattito pubblico. Infatti, possono costituire una notevole distrazione che non consente di sfruttare a dovere le potenzialità dell'IA, anche con l'approvazione di una necessaria cornice normativa per mitigare i rischi più svariati.

Manifestare un eccessivo scetticismo nei confronti dell'IA e delle sue applicazioni per il semplice fatto che questa potrebbe porre dei rischi vitali nei confronti della sopravvivenza umana è non solo un ragionamento sostanzialmente fallace¹⁸⁷, ma anche in grado di distrarre l'opinione pubblica e deviare l'attenzione dei governi nazionali e delle organizzazioni internazionali dall'intento di regolamentare correttamente l'IA per poterne garantire uno sviluppo ed un utilizzo nel pieno rispetto dei diritti fondamentali¹⁸⁸. Ciò che realmente rappresenta una minaccia di natura esistenziale per la sopravvivenza umana (ad esempio la crisi climatica, il dilagare sempre più frequente di guerre, carestie, epidemie e ingiustizie sociali largamente diffuse a livello globale¹⁸⁹) possono essere affrontati con esiti positivi con l'utilizzo delle tecnologie più avanzate.

Per cui, l'atteggiamento catastrofista e fatalista nei confronti dell'IA non deve distrarre dalla possibilità di applicare la tecnologia in maniera utile ed efficace a questi settori di critica importanza, nell'ottica della creazione di società più equa e giusta¹⁹⁰. Questo è l'obiettivo della c.d. *AI for Social Good (AI4SG)*, teorizzata da alcuni studiosi del settore¹⁹¹: essa non si limita a seguire una cornice di principi etici, ma rappresenta l'effettiva concretizzazione di questi principi in modo tale che il suo impiego sia volto a “(i) prevenire, mitigare o risolvere i problemi che incidono negativamente sulla vita umana e/o sul benessere del mondo naturale e/o (ii) consentire sviluppi preferibili dal punto di vista sociale e/o sostenibili dal punto di vista ambientale.”¹⁹²

È dunque da auspicare un tale utilizzo dell'IA: non un sottoutilizzo, causato da eccessive paure o preoccupazioni esistenzialiste nei confronti di questa

¹⁸⁷ Come si vedrà nel paragrafo 3.3, le possibilità che l'IA sfugga al controllo umano e si abbiano scenari di “singolarità” tecnologica, sono frutto principalmente di scenari fantascientifici più che reali possibilità.

¹⁸⁸ Madung, O. (9 agosto 2023). AI hysteria is a distraction: algorithms already sow disinformation in Africa. *The Guardian*. <https://www.theguardian.com/global-development/2023/aug/09/ai-chatgpt-doomerism-threat-already-here-big-tech-algorithms-sow-disinformation>. Vedi anche Floridi (2022), 262-263.

¹⁸⁹ Floridi (2022), 86.

¹⁹⁰ Saujani, R. (5 dicembre 2023). We don't have to choose between ethical AI and innovative AI. *TIME*. <https://time.com/6342280/ai-paid-leave-social-good/>.

¹⁹¹ *Ex multis* Hager, G., Drobniš, A.W., Fang, F., Ghani, R., Greenwald, A., Lyons, T., Parkes, D.C., Schultz, J., Saria, S., Smith, S.F., & Tambe, M. (2019). Artificial Intelligence for Social Good. *ArXiv*, abs/1901.05406.

¹⁹² Floridi (2022), 222-229.

tecnologia, il quale potrebbe causare rilevanti costi-opportunità¹⁹³; né un sovra-utilizzo/cattivo utilizzo completamente deregolamentato e senza limiti dell'IA in tutti i settori, che invece può portare alle già menzionate conseguenze negative a livello sociale, collettivo o individuale¹⁹⁴.

3.1 Le opportunità: applicazioni nel settore della medicina, del lavoro, e della crisi climatica

I settori in cui l'utilizzo di sistemi *AI-powered* di molteplice natura possono portare notevoli progressi e ottimizzazioni sono numerosissimi. La caratteristica intrinseca dell'IA di essere uno strumento *general-purpose*¹⁹⁵ ne consente la sua adattabilità in vari contesti, perseguendo scopi apparentemente non necessariamente conciliabili. Per capire meglio la sua versatilità, basti paragonarla ad una delle innovazioni del secolo precedente, di una tecnologia che, al pari dell'IA, può essere utilizzata per gli scopi più disparati: Internet¹⁹⁶. Queste due innovazioni epocali hanno in comune l'alto tasso di adattabilità¹⁹⁷ che consente ne consente l'applicazione per perseguire un ampio ventaglio di possibili obiettivi. Non a caso, l'*AI Index Report 2023* dell'HAI di Stanford, ha riportato che l'IA potrebbe essere considerata il nuovo “miglior scienziato” al mondo¹⁹⁸.

Un settore in cui l'impatto della recente adozione di sistemi di IA è particolarmente fruttuoso e sta determinando dei progressi di inestimabile valore scientifico è il campo medico-sanitario. Su di esso, più di ogni altro ambito, si sono concentrati gli investimenti a livello globale nel 2022: più di 6,1 miliardi di dollari sono stati infatti dedicati allo sviluppo dell'IA nell'*healthcare*¹⁹⁹.

¹⁹³ Floridi (2022), 280-282.

¹⁹⁴ *Id.*

¹⁹⁵ Edwards (2022), 6-7.

¹⁹⁶ Nel 2005, il Prof. Johnatan Zittrain, fondatore del Berkman Klein Center for Internet & Society dell'università statunitense di Harvard, coniò il concetto delle tecnologie c.d. “generative,” cioè in grado di combinare i seguenti caratteri: capacità di sfruttamento (*capacity for leverage*, intesa come capacità di rendere meno complessi lavori difficili), adattabilità (*adaptability*, intesa come l'ampiezza degli scopi per l'utilizzo della tecnologia in esame), facilità di padronanza (*ease of mastery*, intesa come facilità per il grande pubblico di poterne sfruttare l'utilizzo senza particolari requisiti di conoscenze tecniche), e accessibilità (*accessibility*). La generatività di una certa tecnologia è infatti definita come “una funzione della capacità di una tecnologia di fare leva su una serie di compiti, dell'adattabilità a una serie di compiti diversi, della facilità di padronanza e dell'accessibilità.” Zittrain, J. (2005). The generative Internet. *Social Science Research Network*. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID921636_code177769.pdf?abstractid=847124&mirid=1&type=2.

¹⁹⁷ Intesa come *adaptability*, vedi nota precedente.

¹⁹⁸ *The Artificial Intelligence Index Report 2023, AI Index*, Top Ten Takeaways (n. 4) (2023)

¹⁹⁹ *Id.*, Chapter 4: the economy, 171, 195.

Astrattamente, è molto complesso immaginare una delimitazione dei possibili usi dell'IA in questo settore, poiché si è osservato che questa può essere integrata idealmente nei più disparati aspetti della medicina, con un uguale tasso di successo: dall'utilizzo di strumenti di IA in grado di riconoscere la voce dei pazienti e automatizzare il *note-taking* del medico curante durante una visita di controllo, fino all'utilizzo di modelli predittivi allenati su immense quantità di dati sanitari in grado di ideare soluzioni farmacologiche innovative a determinate malattie, o in grado di riconoscere l'insorgere di determinate condizioni prima ancora che queste siano rilevabili anche dall'occhio medico più esperto²⁰⁰. Per quanto riguarda questa seconda classe di applicazioni, si fa principalmente riferimento allo sviluppo di modelli di IA da utilizzare nella ricerca medico-scientifica. Come dato a favore della portata innovativa di questa tecnologia nella ricerca, sono diverse le riviste di letteratura scientifica medica che hanno recentemente aperto (o sono in procinto di aprire) delle sezioni editoriali dedicate all'intersezione tra l'IA nel campo della medicina e della ricerca²⁰¹.

Nonostante ciò, va sottolineato che l'adozione di specifiche applicazioni dell'IA non è intrinsecamente e necessariamente connessa alle più recenti innovazioni nel campo della medicina. Le radici di tali impieghi talvolta possono anche risalire a contesti e periodi precedenti, suggerendo che la storia e lo sviluppo dell'IA non sono guidati esclusivamente dalle scoperte degli ultimi anni, ma piuttosto da un continuo processo evolutivo iniziato diverse decadi fa. Ad esempio, nel campo delle neuroscienze, le reti neurali di apprendimento sono state impiegate per simulare il comportamento del cervello umano per poterne osservare i tratti salienti, fin dall'inizio degli anni '90²⁰². Già nel 1992 si parlava di due diverse applicazioni di reti neurali per osservare da più vicino il funzionamento del cervello: da un lato quelle che consentivano al computer di “*simulare le informazioni secondo un pattern tipico del cervello, o a blocchi, piuttosto che elaborare una porzione di dati alla volta*”²⁰³ e quelle che invece consistevano in un

²⁰⁰ Beam, A. L., et al. (2023). Artificial intelligence in medicine. (Editorial) *The New England Journal of Medicine*, 388(13), 1220–1221. <https://doi.org/10.1056/nejme2206291>.

²⁰¹ Beam, A. L., et al. (2023). Artificial intelligence in medicine. (Editorial) *The New England Journal of Medicine*, 388(13), 1220–1221. <https://doi.org/10.1056/nejme2206291>, con cui si è annunciata l'apertura della sezione “Artificial Intelligence in medicine” in *The New England Journal of Medicine*. O ancora, si veda la recente nascita di riviste scientifiche quali *Artificial intelligence in Medicine: Artificial Intelligence in Medicine | Journal | ScienceDirect.com by Elsevier*. (n.d.). <https://www.sciencedirect.com/journal/artificial-intelligence-in-medicine>.

²⁰² Damasio, A. & Damasio, H. (1992) Brain and Language, *Scientific American*, Vol. 267 No. 3 Special Issue: Mind and Brain, 88-109, 106.

²⁰³ Damasio & Damasio (1992), 106.

“tipo speciale di computer con migliaia di semplici dispositivi connessi l’uno all’altro in una maniera simile al cervello.”²⁰⁴ Il successo di applicazioni dell’IA di questo tipo non è potuto che aumentare esponenzialmente con l’incremento della *performance* dei computer e la capacità di gestione di *datasets* sempre più ampi²⁰⁵.

Un altro settore in cui l’IA è un acceleratore di innovazione è quello della cosiddetta *drug discovery*. La scoperta di nuovi medicinali mediante l’impiego di modelli di IA consiste nell’ultima frontiera della farmacologia. Ciò infatti può consentire di superare quegli ostacoli posti dalla chimica classica²⁰⁶, avendo essa un ruolo cruciale nel tagliare le tempistiche necessarie a raggiungere certi obiettivi, abbassare i costi computazionali e incrementare l’affidabilità e l’accuratezza di determinati processi di scoperta di nuove terapie farmacologiche e nuovi medicinali²⁰⁷.

D’altra parte, la ricerca non è l’unico settore in cui l’IA fornisce un importante supporto nel campo medico: l’integrazione di questa tecnologia in dispositivi fisici, può essere per alcuni pazienti uno strumento in grado di migliorare la qualità della vita. Motivo per cui la *Food and Drug Administration* statunitense (FDA), tra il 2012 e il 2018, ha approvato numerosi *smart devices* integrati con applicazioni di IA da utilizzare nel settore medico²⁰⁸. Ad esempio, si pensi a quei dispositivi che consentono di controllare stabilmente il decorso di alcune patologie in pazienti particolarmente vulnerabili. Strumenti come *Advisor Pro*, prodotto da *DreaMed Diabetes Ltd.*, fornisce un costante monitoraggio della glicemia dei pazienti affetti da diabete; il sistema, inviando i dati ad un *server cloud* che sfrutta l’IA, consente di calibrare da remoto la dose di insulina da somministrare al paziente²⁰⁹. Ancora, sistemi come il *Guardian Connect System* prodotto dalla statunitense *Medtronic plc.*, utilizza l’IA nel monitoraggio dei valori di glucosio del

²⁰⁴ *Id.*

²⁰⁵ Nomura, A., *et al.* Artificial Intelligence in Current Diabetes Management and Prediction. *Curr Diab Rep* 21, 61, 1 (2021). <https://doi.org/10.1007/s11892-021-01423-2>.

²⁰⁶ Gupta, R. *et al.* (2021), Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers* 25, 1315–1360, 1320-1323. <https://doi.org/10.1007/s11030-021-10217-3>

²⁰⁷ Gupta *et al.* (2021), 1323. *Vedi anche* Heaven, W. D. (5 ottobre 2023). AI is dreaming up drugs that no one has ever seen. Now we’ve got to see if they work. *MIT Technology Review*. <https://www.technologyreview.com/2023/02/15/1067904/ai-automation-drug-development/>

²⁰⁸ Nomura *et al.* (2021), 3.

²⁰⁹ *Id.*, 3. *Vedi anche* DreaMed. (31 luglio 2023). *DREAMED MAIN Home 3* - DreaMed. DreaMed - Cutting-edge Diabetes Technology for Insulin Treatment Optimization. <https://dreadmed.ai/>.

diabetico, in modi di riuscire a prevedere la prossima crisi ipoglicemica del paziente con un anticipo fino a 30 minuti prima della sua insorgenza²¹⁰.

In ultimo, è opportuno fornire un *caveat* che si ricollega direttamente con l'importanza della regolamentazione di questi sistemi che utilizzano di IA. Infatti, questo tipo di tecnologie, sebbene senza dubbio foriere di un impatto positivo nella gestione di patologie debilitanti nei confronti di coloro che ne sono affetti, possono portare con sé delle criticità, per quanto riguarda il trattamento di dati altamente sensibili²¹¹ e, conseguentemente, la loro sicurezza. Il livello di cybersicurezza presente in questi dispositivi deve essere a prova di qualsiasi vulnerabilità, fissato secondo standard che abbiano in *primis* l'obiettivo di tutelare a dovere l'interessato. Allo stato attuale dell'arte, tale livello di sicurezza non è sempre assicurato, determinando degli effetti potenzialmente dannosi nei confronti dei pazienti²¹².

Nella conversazione sul legame tra l'intelligenza artificiale e il mondo del lavoro, si ripresenta spesso un interrogativo che cattura l'attenzione dell'opinione

²¹⁰ *Id. Vedi anche GuardianTM Connect Continuous Glucose Monitor | Medtronic. (27 novembre 2023).* <https://www.medtronicdiabetes.com/products/guardian-connect-continuous-glucose-monitoring-system>.

²¹¹ Si faccia riferimento alla definizione implicita di dati "sensibili" (tra cui quelli relativi alla salute della persona, quelli biometrici o genetici) fornita all'art. 9 del GDPR, in cui si prevede che il trattamento degli stessi è assolutamente vietato se non per motivazioni ben precisate nel comma due dello stesso, o con l'esplicito consenso dell'interessato. Art. 9(1)(2) GDPR.

²¹² Nel 2019, la Food and Drug Administration (FDA) statunitense ha richiamato l'attenzione dei pazienti affetti da diabete e della comunità medica sui potenziali rischi di vulnerabilità in ambito *cybersecurity* che alcuni dispositivi medici possono comportare. Nel caso di specie, la statunitense Medtronic, ha richiamato dal mercato una serie di *devices* utilizzati per il monitoraggio dei livelli di insulina a causa della presenza di grosse falle di sicurezza che avrebbero astrattamente consentito ad un *hacker* di introdursi nel sistema interno del dispositivo e modificare i livelli di insulina somministrati al paziente. Anche se ciò non è mai accaduto, la sola presenza di una tale vulnerabilità ha non solo portato l'azienda a richiamare dal mercato l'intera gamma di prodotti, ma ha anche consentito l'incremento dell'attenzione pubblica sulla rilevanza di profili di cybersicurezza. Anche se i dispositivi ora in esame non utilizzavano sistemi di IA al loro interno, la maggiore complessità di questi, in combinazione con la quantità di dati sensibili che questi sono in grado di conservare, porta a ritenere che l'attenzione sui profili di sicurezza di questi *smart devices* deve rimanere alta. *Vedi* Klonoff, D., & Han, J. (2019). The First Recall of a Diabetes Device Because of Cybersecurity Risks. *Journal of diabetes science and technology*, 13(5), 817–820. <https://doi.org/10.1177/1932296819865655>.

pubblica²¹³, suscitando più o meno preoccupazioni²¹⁴: “L’IA ‘ruberà’ posti di lavoro?” Sebbene vi sia un fondo di verità in questa valutazione, formulata in maniera più o meno catastrofista, il World Economic Forum ha recentemente invitato l’opinione pubblica a valutare la quantità di nuovi posti di lavoro, nuove mansioni e posizioni lavorative che l’IA è in grado di generare²¹⁵. Ciò porterebbe ad una compensazione della perdita di quegli impieghi che, ad oggi, sono particolarmente minacciati dall’avvento di questa tecnologia. Se da un lato alcune occupazioni come, ad esempio, i ruoli impiegatizi amministrativi, i ruoli di segreteria o contabilità²¹⁶, tenderanno a diminuire progressivamente, surclassati dall’utilizzo di IA, (soprattutto quella generativa), la domanda di altri ruoli, invece, aumenterà esponenzialmente²¹⁷.

In *primis*, è da prevedere che i ruoli direttamente correlati al settore dell’IA e del ML siano quelli ad avere una richiesta più impellente di nuovi e qualificati esperti²¹⁸. La crescente complessità e la rapida evoluzione di questa tecnologia richiederanno un costante afflusso di professionisti altamente specializzati in grado di affrontare sfide e opportunità emergenti²¹⁹. Non solo, altre figure tecniche ad altri settori tecnico-scientifici saranno necessarie per assicurare il rispetto degli standard di sicurezza, trasparenza, robustezza necessari al mantenimento di tutta la filiera di sviluppo, impiego e manutenzione dell’IA: ad esempio, tra gli altri, professionisti di analisi di dati e *database*, matematici, statistici, così come esperti di etica delle tecnologie²²⁰.

²¹³ L’interrogativo riportato compare su numerosi articoli di testate giornalistiche nazionali ed internazionali. Vedi Ball, J. (16 agosto 2023). Will AI steal my job? Maybe – but here are some possible new opportunities. *The Guardian*. <https://www.theguardian.com/global-development/2023/aug/16/will-ai-steal-my-job-maybe-but-here-are-some-possible-new-opportunities>; vedi anche Semuels, A. (6 agosto 2020). Millions of Americans have lost jobs in the Pandemic—And robots and AI are replacing them faster than ever. *Time*. <https://time.com/5876604/machines-jobs-coronavirus/>; vedi anche Spagnuolo, E. (29 settembre 2023). Intelligenza artificiale: i lavori a rischio e quelli emergenti secondo il World Economic Forum. *Corriere Della Sera*. https://www.corriere.it/tecnologia/cards/intelligenza-artificiale-i-lavori-a-rischio-e-quelli-emergenti/il-lavoro-al-tempo-di-chatgpt_principale.shtml.

²¹⁴ Secondo l’AI Index Report dell’HAI di Stanford, la prima preoccupazione ad interessare la popolazione americana circa l’utilizzo dell’IA è proprio la potenziale perdita di posti di lavoro. *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 8: public opinion, 331 (2023)

²¹⁵ *We always hear that AI will take our jobs. But what jobs will it create?* (6 ottobre 2023). World Economic Forum. <https://www.weforum.org/agenda/2023/09/jobs-ai-will-create/>.

²¹⁶ *Id.* (Ivi, vedi figure 7).

²¹⁷ *Id.*

²¹⁸ *The Artificial Intelligence Index Report 2023, AI Index*, Top Ten Takeaways (n. 6), (2023).

²¹⁹ *We always hear that AI will take our jobs. But what jobs will it create?* (6 ottobre 2023). World Economic Forum

²²⁰ *Id.*

Questa crescita è da osservare in combinazione con l'aumento della richiesta di ruoli negli ambiti più disparati (amministrativo, legale,²²¹ finanziario, assicurativo, delle infrastrutture) che integrino l'utilizzo dell'IA nel *day-to-day* delle loro operazioni. Secondo il sopracitato AI Index Report del 2023, in tutti i settori dell'industria degli Stati Uniti per i quali sono disponibili dati (ad eccezione di agricoltura, silvicoltura, pesca e caccia), il numero di offerte di lavoro legate all'IA è aumentato in media dell'1,7% nel 2021 all'1,9% nel 2022²²².

Infine, l'IA non si limiterà ad avere un impatto sull'aumento degli impieghi, delle tipologie di mansioni specifiche, delle occupazioni, presenti nell'attuale mondo del lavoro. In aggiunta, più le aziende integreranno l'IA nei loro *workflows* e operazioni, aumentando quindi anche la domanda per le tipologie di occupazioni menzionate poco fa, più esse contribuiranno direttamente all'avanzamento della ricerca e dell'applicazione dei sistemi di IA, con la creazione di una sorta di gioco a somma sempre positiva, che velocizza il ritmo dell'innovazione. Nel 2022, lo sfruttamento da parte delle aziende di sistemi che impiegano l'IA ha interessato le seguenti aree applicative: l'automazione dei processi robotici (39%), la visione artificiale o c.d. *computer vision* (34%), la comprensione del testo in linguaggio naturale (33%) e gli agenti virtuali (33%²²³). Inoltre, nel 2022, le attività che hanno incorporato l'IA sono state, *inter alia*, l'ottimizzazione delle operazioni di servizio (24%), seguita dalla segmentazione della clientela (19%), l'analisi del servizio clienti (19%²²⁴). L'IA può dunque essere iniziata a considerare un vero e proprio *asset* aziendale²²⁵, la cui applicazione consentirà di implementare la stragrande maggioranza dei processi

²²¹ L'utilizzo dell'IA nel settore legale costituisce non solo una frontiera estremamente affascinante, ma consentirebbe anche di implementare la fornitura dei servizi legali per mirare a chiudere progressivamente il cosiddetto "Access to Justice Gap" ("A2J"). Per "A2J" si intende cioè quel divario all'accesso dei servizi di giustizia presente in molti paesi, tra cui gli USA. Proprio negli USA, il dibattito sull'impiego di sistemi integrati con IA per consentire la prestazione dei servizi legali consiste nel settore innovativo della "Legaltech," in cui scienza informatica e giurisprudenza si combinano per la creazione di *software*, servizi online che possano essere di supporto per coloro che necessitano l'erogazione di determinati servizi legali. Vedi *Technology to solve the access to justice gap* | Suffolk Legal Tech Class 2023. Quinten Steenhuis, <https://suffolklitlab.org/legal-tech-class/docs/a2j/a2j-technology>.

²²² *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 4: the economy, 171 (2023)

²²³ *Id.*, 172.

²²⁴ *Id.*

²²⁵ Così è stata denominata dalla Prof.ssa Daniela Caterino nell'ambito del suo intervento dal titolo "Deumanizzazione" degli organi societari e profili di responsabilità nella corporate governance", in occasione Stati Generali del Diritto di Internet e dell'IA, tenutisi presso l'Università LUISS Guido Carli il 15 dicembre 2023. LuissGuidoCarli. (22 dicembre 2023). *Gli Stati Generali del Diritto di Internet e della Intelligenza Artificiale (Giorno 2)* [Video]. Min. 5:19:11-5.44.54 YouTube. <https://www.youtube.com/watch?v=vEnAZFHokm8>.

produttivi, ma anche amministrativi o gestori, delle imprese. Tuttavia, considerando l'assoluta difficoltà nell'effettuare la somma algebrica dei i posti di lavoro che verranno creati e quelli che comunque saranno inevitabilmente impattati dal sempre più largo uso di dell'IA (soprattutto quella generativa) e dunque la generale complessità di formulare lo scenario che attende il mercato del lavoro, bisognerebbe responsabilizzare istituzioni e governi a prendere in considerazione misure di protezione a sostegno dei lavoratori, nell'ottica di minimizzare il più possibile gli effetti potenzialmente negativi²²⁶. D'altronde, anche l'opinione pubblica, chiamata a rispondere in relazione al quesito in analisi, ha dimostrato una forte incertezza nello stabilire quale risultato sarà più probabile, se una diminuzione o un aumento complessivo dei posti di lavoro²²⁷.

In ultimo, si consideri l'opportunità di utilizzare sistemi di IA per arginare i perniciosi effetti della crisi climatica, nell'ottica di contribuire a trovare delle modalità di gestione più efficienti delle risorse ambientali utilizzando la tecnologia, e contemporaneamente limitare i danni che il riscaldamento globale è in grado di causare. L'accelerazione delle azioni per contrastare la crisi climatica è infatti necessaria, in quanto, la probabilità che l'obiettivo di mantenere l'aumento delle temperature globali al di sotto di 1,5° Celsius, come stabilito dall'Accordo di Parigi²²⁸, sia rispettato, è sostanzialmente nulla²²⁹. A questo proposito, l'IA si pone come strumento cruciale per poter al meglio strutturare azioni dirette che possano coadiuvare lo sforzo per la sostenibilità delle attività umane, in numerosi campi²³⁰.

²²⁶ Sono state avanzate numerose ipotesi di misure a sostegno dei lavoratori che possano mitigare le conseguenze potenzialmente negative del sempre maggiore utilizzo di IA il sostegno al *re-skilling* (cioè l'aggiornamento di nuove competenze) di coloro la cui professione è stata integralmente automatizzata, o l'imposizione di una c.d. "robot tax" per le aziende che automatizzano integralmente determinati processi tagliando l'organico; alcuni, ritengono che qualora la perdita di posti di lavoro a causa dell'automazione diventi un problema da affrontare su larghissima scala, una misura potenzialmente utile possa essere l'erogazione di sostegni come il Reddito di Base Universale, o Universal Basal Income (UBI). *Vedi in generale* Schaake, M. (9 gennaio 2023) It's already time to think about an AI tax. *Financial Times*. <https://www.ft.com/content/242c8f5a-43af-43d5-875f-261a0841045a>; *vedi anche* Przegalińska, A., & Wright, R. E. (2021). AI: UBI Income Portfolio Adjustment to Technological Transformation. *Frontiers in Human Dynamics*, 3. <https://doi.org/10.3389/fhumd.2021.725516>.

²²⁷ Do Americans think artificial intelligence will lead to there being more jobs for people or fewer jobs? [Grafico]. In *Statista*. (30 dicembre 2022) <https://www.statista.com/statistics/1357711/opinions-on-artificial-intelligence-impact-on-jobs-by-age-us/>.

²²⁸ Accordo di Parigi, così come ratificato dall'UE; EurLex GU L 282 del 19 ottobre 2016, 4–18.

²²⁹ *Accelerating Climate Action with AI*, Boston Consulting Group, 5 (novembre 2023). <https://web-assets.bcg.com/72/cf/b609ac3d4ac6829bae6fa88b8329/bcg-accelerating-climate-action-with-ai-nov-2023-rev.pdf>.

²³⁰ Floridi (2022), 297-298.

Secondo un recente report redatto dal *Boston Consulting Group* a novembre 2023, sono emerse le potenziali capacità di questo strumento tecnologico di essere impiegato in tre settori cruciali per la sostenibilità ambientale. In *primis* nella c.d. “*mitigation*,” tra cui figura, la riduzione delle emissioni²³¹; in secondo luogo, l’IA può essere sfruttata per coadiuvare i cittadini, i Paesi, le città, le imprese in attività di “*adaptation and resilience*,” per imparare come rispondere in maniera efficace agli effetti immediatamente dannosi della crisi climatica²³²; ancora, essa può essere attivamente impiegata per consentire l’azione per il clima attraverso miglioramenti nella modellistica climatica, nell’ecologia climatica e nell’educazione al clima, nonché accelerare le innovazioni rivoluzionarie che apriranno nuovi orizzonti alla tutela dell’ambiente²³³. Infatti, il clima è un sistema complesso, interconnesso e assolutamente multiparametrico, che ben si presta ad essere analizzato secondo le logiche dell’IA, soprattutto con i modelli più complessi che consentono di prendere in considerazione un’ampia gamma di variabili²³⁴.

Ancora, secondo un’analisi del filosofo Floridi, l’IA pone due opportunità nel settore della sostenibilità ecologica per far fronte ad una delle sfide esistenziali che mette in discussione la stessa sopravvivenza dell’uomo nell’antropocene²³⁵. Da un lato, l’IA può essere, come è stato più volte ripetuto, impiegata nella ricerca scientifica, in quanto può contribuire a migliorare l’attuale comprensione di fenomeni meteorologici complessi, variabili, e tendenzialmente sempre più estremi e imprevedibili a causa proprio del riscaldamento globale²³⁶: a questo proposito, sono già stati sviluppati algoritmi di IA in grado di mappare le variazioni di temperature a livello globale, di predire onde di instabilità tropicale, precipitazioni, e altri fenomeni²³⁷.

La seconda opportunità consiste invece in un risvolto prettamente pragmatico nello strutturare soluzioni più ecologiche, come il miglioramento e l’ottimizzazione della generazione e dell’uso dell’energia, utilizzando proprio la tecnologia. Ad oggi esistono già dei modelli di IA programmati per essere impiegati per questi scopi. Ad esempio, il laboratorio di IA di casa Google, DeepMind, ha recentemente sviluppato un sistema per programmare un’ottimale gestione della

²³¹ *Id.*, 10.

²³² *Id.*

²³³ *Accelerating Climate Action with AI*, Boston Consulting Group, 10 (novembre 2023).

²³⁴ *Id.*, 9.

²³⁵ Floridi, 297-300.

²³⁶ Floridi (2022), 298.

²³⁷ *Id.*

produzione energetica tramite fusione nucleare (una significativa alternativa ai combustibili fossili), in un reattore denominato *tokamak*; produzione che ad oggi non risulta ancora realizzabile a causa della mancanza di dati sperimentali, che tuttavia l'IA è stata in grado di fornire tramite i suoi calcoli, consentendo di risparmiare tempo e risorse nello sviluppo di uno studio per ottenere i suddetti dati²³⁸.

Tuttavia, l'utilizzo di modelli basati sull'IA nell'ecologia solleva una considerazione particolare, come evidenziato nel AI Index Report dell'Università di Stanford. Infatti, l'IA è in grado di apportare contemporaneamente significativi vantaggi e svantaggi all'ambiente²³⁹. Le emissioni derivanti dal *training* algoritmico, dal mantenimento di computer, o di *server*, dei relativi sistemi di raffreddamento, nonché l'estrazione dei materiali utilizzati nella costruzione degli *hardware* di vario genere che supportano i modelli di IA²⁴⁰, hanno collettivamente un impatto non indifferente sull'ambiente, sia a livello di emissioni prodotte, che di complessivi effetti anche dannosi sugli ecosistemi²⁴¹. Nello specifico, maggiore l'ampiezza dei modelli di IA da sviluppare, maggiore sarà l'incidenza delle emissioni prodotte: di conseguenza, l'esponenziale crescita di grandezza dei modelli di IA dell'ultimo decennio (con la sempre maggiore popolarità delle tecniche di *deep learning* e la creazione di grandi sistemi come i *Large Language Models*²⁴²) ha contribuito ad aumentare a dismisura il potenziale impatto sull'ecosistema²⁴³.

A ragione, l'inquinamento direttamente derivante dallo sviluppo e dall'utilizzo dell'IA e la sua relativa sostenibilità costituisce un settore poco esplorato, principalmente a causa della mancanza di trasparenza nelle informazioni fornite dagli sviluppatori²⁴⁴; motivo per il quale è complesso definire, al netto di quanto appena considerato, quale sia il *reale* impatto dell'IA sugli ecosistemi del

²³⁸ *The Artificial Intelligence Index Report 2023*, Chapter 2: technical performance, 74 (2023).

²³⁹ *The Artificial Intelligence Index Report 2023*, *AI Index*, Top 10 Takeaways (n. 3) (2023).

²⁴⁰ *Costi ambientali dei dispositivi di IA - Privacy Network*. (24 ottobre 2022). Privacy Network. <https://privacy-network.it/costi-ambientali-dei-dispositivi-di-ia/>.

²⁴¹ Floridi (2022), 302.

²⁴² Per una definizione dei Large Language Models (LLMs), vedi nota 92.

²⁴³ *Id.*, 308.

²⁴⁴ Floridi fornisce l'esempio del *training* dell'LLM di OpenAI, GPT-3.5. Per calcolare infatti le emissioni prodotte da una sessione di allenamento di questo modello di IA generativa, che è in grado di riconoscere il linguaggio naturale e rispondere producendo dei testi simili a quelli umani, degli studiosi hanno dovuto stimare il dato delle emissioni a partire da calcoli probabilistici, non potendo ricorrere all'utilizzo di dati forniti dalla casa madre, in quanto questi non sono stati condivisi con il pubblico. A partire da ciò hanno stimato che una singola sessione di addestramento produrrebbe 223.920 kg di CO₂, cioè quanto producono 49 vetture in un anno negli Stati Uniti. *Id.*, 302, 308-310.

pianeta. Tuttavia, secondo alcuni, sfruttare le opportunità offerte dall'implementazione di questa tecnologia per la tutela del clima è sia fattibile che auspicabile, sebbene vi sia la possibilità di dover sostenere un rilevante sacrificio in termini di prevedibile incremento dell'impronta ecologica²⁴⁵. Anche in questo campo, il punto di equilibrio si trova nel delineare una corretta *governance* “reattiva ed efficace”²⁴⁶ degli strumenti di IA, che tenga in considerazione il potenziale *trade-off* che può verificarsi tra effettiva sostenibilità ambientale ed innovazione e derivante dall'incremento delle potenzialità dei sistemi, come ad esempio tramite il miglioramento di parametri algoritmici come l'accuratezza²⁴⁷.

3.2 Le categorie dei rischi “attuali” dell'IA: gli elementi critici intrinseci e quegli elementi derivanti da precise applicazioni dei sistemi

Secondo l'AI Index Report più volte menzionato nei paragrafi precedenti, con l'incremento dell'utilizzo di sistemi *AI-powered*, stanno rapidamente aumentando anche gli “incidenti” e le polemiche pubbliche che un utilizzo *contra etica* dell'IA può portare con sé (prima ancora che *contra legem*, a causa della generale assenza di una cornice normativa che regoli questa tecnologia nello specifico²⁴⁸). Nel *database* indipendente e aperto al pubblico predisposto dall'*AI, Algorithmic, and Automation Incidents and Controversies* (AIAAIC²⁴⁹), le controversie relative a questi sistemi sono aumentate di 26 volte nel periodo dal 2012 al 2021. L'aumento degli incidenti segnalati indica non solo una maggiore integrazione dell'IA nel mondo reale e una crescente consapevolezza delle possibili utilizzazioni eticamente scorrette, ma anche un miglioramento nella tracciabilità dei suddetti incidenti, suggerendo che quelli più risalenti potrebbero essere passati inosservati o probabilmente sotto-segnalati²⁵⁰.

L'esame dei rischi attuali, in contrapposizione con quelli virtuali che saranno accennati al paragrafo 3.3, non ha un mero scopo espositivo, bensì ne serve un altro ben preciso: quello di indicare inequivocabilmente l'ormai impellente

²⁴⁵ *Id.*, 302.

²⁴⁶ *Id.*, 302-303.

²⁴⁷ *Id.*, 311.

²⁴⁸ *The Artificial Intelligence Index Report 2023, AI Index, Top 10 Takeaways* (n. 5).

²⁴⁹ L'*AI, Algorithmic, and Automation Incidents and Controversies* (AIAAIC) è un progetto che ha costituito un *database* pubblico, aperto e indipendente, inizialmente costituito come *repository* privata, per comprendere meglio i rischi reputazionali dell'IA; successivamente, con lo sviluppo di questa tecnologia, il progetto si è evoluto in un'iniziativa più completa che segue le questioni etiche associate all'AI. *Id.*, Chapter 3: Technical AI Ethics, 133 (2023).

²⁵⁰ *Id.*

necessità di fornire una cornice regolatoria di queste tecnologie, che possa consentire contemporaneamente la mitigazione dei rischi ad essa associati, e la possibilità di un impiego in totale sicurezza, nell’ottica di sfruttare correttamente le sue potenzialità. Non solo: osservando la diversa tipologia delle problematiche prettamente pratiche che i sistemi di IA possono causare lungo la loro intera *supply chain*, meglio si comprende la scelta di alcuni regolatori internazionali (in *primis* l’Unione Europea), di articolare la normativa sui sistemi di IA sulla base di requisiti da rispettare a seconda di una classificazione dei sistemi in base al rischio. Come infatti si vedrà nel paragrafo 3.3.2 del Capitolo 2, la proposta di regolamentazione europea si fonda su un approccio *risk-based*: i vari requisiti di certificazione, trasparenza, valutazioni d’impatto preventive all’impiego di un determinato sistema previste nell’*AI Act* mutano a seconda delle potenziali criticità che questi possono generare nei confronti della collettività.

All’interno delle criticità che i sistemi di IA possono recare, vi è da sottolineare una sorta di bipartizione ideale, come riportato dal titolo del presente paragrafo. Da un lato, infatti, vi sono quei caratteri congeniti a questa tecnologia, come l’effetto “scatola nera”²⁵¹, o la difficoltà nel controllare dettagliatamente tutti i vari dataset utilizzati nei *training* algoritmici, che aumentano la possibilità di pregiudizi interni nel set di dati storici impiegati²⁵²; mentre dall’altro vi sono rischi applicativi dell’IA in una serie di settori strategici. Si è scelto di analizzare queste tipologie di pericoli dell’IA in maniera separata.

I primi rischi, cioè quelli intrinseci dell’IA, si possono combinare in maniera tale da incidere gravemente su coloro che sono i destinatari della decisione algoritmica, portando talvolta a dei risultati sostanzialmente discriminatori a livello sociale ed economico. Un esempio di come queste conseguenze dell’uso dell’IA possano emergere, si è verificato diversi anni fa nei Paesi Bassi. Si tratta del caso conosciuto come il “*toeslagenaffaire*.”²⁵³ Tra il 2013 e il 2019, l’autorità fiscale olandese ha impiegato un algoritmo di IA multiparametrico con lo scopo individuare preventivamente quei nuclei familiari che avessero maggiori

²⁵¹ Vedi il sottoparagrafo 2.2.1, in parte dedicato *black box effect*, intimamente connaturato nelle tecniche di apprendimento automatico di ML.

²⁵² Vedi il sottoparagrafo 2.2.2, in parte dedicato ai dati dotati di pregiudizi storici intrinseci che si propagano e ripropongono sistematicamente all’interno dei risultati e delle decisioni algoritmiche di questi sistemi.

²⁵³ Traducibile dall’olandese come il “caso dei *benefits*.”

probabilità di richiedere i sussidi per l'assistenza all'infanzia senza averne diritto²⁵⁴. L'algoritmo predittivo alimentato dall'IA era stato allenato su un set di dati intrinsecamente pieno di pregiudizi economici e sociali: il possesso di una doppia nazionalità, così come un basso reddito, erano considerati parametri in grado di alzare il rischio di potenziale frode nella richiesta dei sussidi. In aggiunta, la natura prettamente auto-apprendente, insita in questo genere di sistemi, non ha fatto altro che confermare e, in ultimo, istituzionalizzare il *bias* già presente nei confronti delle minoranze e dei nuclei familiari di per sé in difficoltà²⁵⁵. L'autorità olandese ha dunque sanzionato ingiustamente decine di migliaia di famiglie, spingendo alcune di queste a condizioni di estrema indigenza per ottemperare alle sanzioni. Il caso in analisi è dunque indicativo di come possono combinarsi, a danni della collettività, quelle caratteristiche congenite dell'IA quali l'auto-apprendimento, la presenza di pregiudizi nei dati di allenamento, e l'amplificazione su larga scala degli stessi pregiudizi a causa dell'automatizzazione dei processi. Inoltre, questo scandalo è servito da caso emblematico per porre un monito a tutti i paesi europei impegnati nello sviluppo di IA da implementare nel settore della pubblica amministrazione, a tener conto dei potenziali rischi senza un adeguato sistema di *checks and balances* a tutela della collettività contro la discriminazione²⁵⁶.

Un altro aspetto connaturato alla natura dell'IA e potenzialmente foriero di conseguenze negative per la collettività consiste nella difficile conciliazione tra *privacy* e *training* algoritmico. In *primis*, la relazione tra questa tecnologia e la tutela della *privacy* viene sempre più percepita come un vero e proprio *trade-off*: alcuni sostengono che i sistemi di IA, in quanto tecnologie che operano uno sfruttamento intensivo di dati, hanno minato irrimediabilmente la capacità

²⁵⁴ Heikkilä, M. (13 aprile 2022). Dutch scandal serves as a warning for Europe over risks of using algorithms. *POLITICO*. <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>. Vedi anche Huyskes, D. (2022). Resistere all'inevitabile, - *Luiss University Press – le meraviglie del possibile*, <https://luissuniversitypress.it/resistere-allinevitabile/>.

²⁵⁵ *Id.*

²⁵⁶ Per quanto riguarda l'Italia, come d'altronde il resto dei paesi europei, l'utilizzo di sistemi automatizzati che impieghino l'IA nel settore della pubblica amministrazione è in rapida crescita. Per questo motivo sono nate iniziative come l'Osservatorio Amministrazione Automatizzata (OAA) di Privacy Network, associazione no profit che si occupa di *privacy* e tutela dei diritti digitali, che ha l'obiettivo di fornire delle "mappature" dei sistemi di IA che vengono utilizzati nel settore pubblico. Vedi Osservatorio Amministrazione automatizzata. *Privacy Network*. <https://privacy-network.it/osservatorio/>.

individuale di mantenere un controllo sui dati personali, facendo venir meno la sovranità individuale sugli stessi²⁵⁷.

A ben vedere, alcuni principi cardine della normativa europea in materia di protezione dei dati personali sono in aperto contrasto con elementi essenziali che caratterizzano l'IA: ad esempio, principi come quello di *data minimisation*²⁵⁸, sono di non facile conciliazione con le basi di funzionamento dell'apprendimento automatico, per il quale maggiori sono le quantità di dati forniti, maggiore sarà la qualità dell'*output* del sistema, e dunque, la sua accuratezza²⁵⁹. Ancora, risulta complesso assicurare correttamente la tutela della privacy di coloro i cui dati personali sono confluiti negli immensi *dataset* impiegati per l'allenamento dei sistemi. Allo stesso modo, non è semplice assicurare il rispetto dei diritti degli interessati al trattamento, come il diritto alla cancellazione o alla rettifica di dati personali non corrispondenti a realtà²⁶⁰. A causa di una generale e diffusa riluttanza delle case sviluppatrici a divulgare le informazioni sul *training* algoritmico, così come la mancanza di chiari requisiti normativi sulla trasparenza, non vi è quella completa informazione necessaria nei confronti del pubblico per quanto concerne il trattamento dei dati²⁶¹. Proprio sulla base di queste motivazioni relative al complesso bilanciamento tra privacy e allenamento dei modelli di IA, specialmente quelli che utilizzano maggiori quantità di dati come i *Large Language Models*²⁶² (LLMs²⁶³), l'Autorità Garante Italiana, con il già menzionato prov. n. 112 del 2023, ha temporaneamente inibito il colosso statunitense OpenAI dal trattare i dati (e quindi sostanzialmente offrire la piattaforma di IA conversazionale ChatGPT) in Italia²⁶⁴.

²⁵⁷ Mantelero, A. (2022). Beyond data: Human Rights, Ethical and Social Impact Assessment in AI. *Springer Nature*, Chapter 1: Beyond data, 2.

²⁵⁸ Il principio di minimizzazione dei dati o *data minimisation* prevede che coloro che svolgono un qualsiasi trattamento sui i dati personali (i c.d. "titolari del trattamento") si limitino a quanto strettamente necessario rispetto alle finalità per le quali gli stessi dati sono trattati Art. 5(1)(c) del GDPR.

²⁵⁹ Intervento del Prof. Filiberto Brozzetti "*Il dubbio AI-mletico: Accuratezza o minimizzazione?*", in occasione Stati Generali del Diritto di Internet e dell'IA, tenutisi presso l'Università LUISS Guido Carli il 15 dicembre 2023. LuissGuidoCarli. (22 dicembre 2023). *Gli Stati Generali del Diritto di Internet e della Intelligenza Artificiale (Giorno 2)* [Video]. YouTube., min. 8:32:39- 8:46:48. <https://www.youtube.com/watch?v=vEnAZFHokm8>. Vedi anche Mészáros et. al (2023), 16-17, 24.

²⁶⁰ Artt. 16-17 GDPR.

²⁶¹ Mészáros et. al (2023), 13-14.

²⁶² Per una definizione dei Large Language Models (LLMs), vedi nota 92.

²⁶³ Vedi Mészáros et. al (2023), 3.

²⁶⁴ Provvedimento n. 112 del 30 marzo 2023, Autorità Garante per la protezione dei dati personali. Già citato nella nota 140.

L'assenza di chiari requisiti di legge per minimizzare le frizioni tra privacy e sviluppo dell'IA ha consentito, nel tempo, il proliferare di incidenti relativi al corretto trattamento delle informazioni personali. Uno di questi ha avuto luogo nel 2016, nel Regno Unito. L'azienda DeepMind Technologies (ora Google DeepMind) aveva stipulato un accordo con il Royal Free London NHS Foundation Trust per accedere ai dati di tre ospedali londinesi, che trattavano oltre 1,6 milioni di pazienti all'anno²⁶⁵. I dati sanitari non erano stati adeguatamente de-identificati, e i pazienti non erano a conoscenza del fatto che le proprie informazioni mediche venivano utilizzate anche da terzi, tra cui anche DeepMind²⁶⁶. Sebbene inizialmente dichiarato che i dati sanitari sarebbero stati elaborati per sviluppare una app priva di sistemi di IA, un memo interno ha suggerito la presenza di un'intesa riservata tra le due entità dalla più ampia portata, che implicava la creazione di algoritmi di grandi dimensioni da allenare anche con i dati sanitari in questione²⁶⁷. Tutto ciò, a completa insaputa dei pazienti a cui le informazioni sanitarie erano direttamente riconducibili.

Le tensioni tra tutela della privacy e sviluppo dei modelli di IA non esauriscono l'ampia gamma di rischi intrinseci che l'intera catena produttiva e di dispiegamento di questa tecnologia è in grado di causare su diversi piani (individuale, collettivo, e sociale). La tabella sottostante rappresenta una mappatura complessiva dei potenziali rischi lungo tutta la *supply chain* dell'IA. Infatti, sebbene quelli analizzati rappresentino le principali preoccupazioni che il funzionamento intrinseco di questi sistemi può portare con sé, vi sono una serie di altri elementi critici che possono contribuire ad aumentare il rischio delle implicazioni negative di questi sistemi, riportati analiticamente nella tabella 1 nella pagina successiva.

²⁶⁵ Hodson, H. (6 maggio 2016). Revealed: Google AI has access to huge haul of NHS patient data. *New Scientist*. <https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/>

²⁶⁶ *Id.*

²⁶⁷ Lomas, N. NHS memo details Google/DeepMind's five year plan to bring AI to healthcare, *Techcrunch* (8 giugno 2016) <https://techcrunch.com/2016/06/08/nhs-memo-details-googledeepminds-five-year-plan-to-bring-ai-to-healthcare/>.

	Resource extraction	Resource processing	Deployment
Individual	<ul style="list-style-type: none"> • Exploitative labour • Data theft 	<ul style="list-style-type: none"> • Exploitative labour • Non-consensual data processing 	<ul style="list-style-type: none"> • Behavioural manipulation • Physical harm • Unfair surveillance
Collective	<ul style="list-style-type: none"> • Exploitation of marginalized or indigenous groups • Group privacy harms from data collection • Exploitation of community resources 	<ul style="list-style-type: none"> • Exploitation of marginalized or indigenous groups • Group privacy harms from data processing 	<ul style="list-style-type: none"> • Unfair surveillance of groups • Exploitation of gig workers • Algorithmic group discrimination
Societal	<ul style="list-style-type: none"> • Institutionalised labour exploitation • Depletion of natural resources • Biodiversity loss • Injury to animals • Extraction emissions 	<ul style="list-style-type: none"> • Institutionalised labour exploitation • Contamination of natural resources • Model training emissions • Chemical waste • Water use 	<ul style="list-style-type: none"> • Institutionalised labour exploitation • Widening structural inequalities • Model use emissions • Increasing digital divide • Eroding trust in social/political institutions

Tabella 1: Mapping intersecting dimensions of harm across the AI supply chain.” Fonte: □ Thomas, C., Roberts, H., Mökander, J., Tsamados, A., Taddeo, M., Floridi, L., (2023). The case for a broader approach to AI assurance: addressing “hidden” harms in the development of Artificial Intelligence, 4. Disponibile su SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4660737.

Tutti questi fattori incidono in maniera assolutamente trasversale in molteplici contesti; necessitano tutti di essere tenuti da conto nel momento in cui si ritiene necessario lo sviluppo di regole e limitazioni per l’IA attraverso la creazione di legislazione ad essa rivolta.

Avendo ora esaminato i rischi derivanti da quegli elementi “intrinseci” dell’IA, è necessario soffermarsi, seppur brevemente, su quei pericoli insiti in precise applicazioni di questa tecnologia, in settori ad alto rischio. Alcune ipotesi, a titolo meramente esemplificativo, possono essere l’utilizzo dell’IA in settori di pubblica rilevanza come il *welfare*, come nel già menzionato caso olandese o la sanità pubblica. Tuttavia, altri impieghi sono recentemente emersi, che di per sé possono portare a delle conseguenze nocive strettamente legate al modo in cui l’IA può essere impiegata. Si fa riferimento all’utilizzo di questi sistemi nell’ambito della criminalità, o l’impiego per scopi militari o di pubblica sicurezza.

In *primis*, così come sono emerse le potenzialità di sviluppo economico che l’IA è in grado di portare con sé, così sono emerse le “cattive pratiche” che possono essere portate a termine utilizzando questi sistemi “per il male sociale.”²⁶⁸ Sebbene nella prima letteratura emergente si è data una particolare attenzione agli usi civili dell’IA e la possibilità di configurare ipotesi di illeciti civili (e conseguente responsabilità per i danni causati), non va dimenticata la portata distruttiva che un

²⁶⁸ Floridi (2022), Capitolo 8. Cattive pratiche: l’uso improprio dell’IA per il male sociale, 177-220.

uso dell'IA per scopi criminali può arrecare, anche su larga scala²⁶⁹. Floridi ha infatti individuato una serie di aree di incidenza in cui i sistemi *AI-powered* possono avere rilevanza penale. Tra queste figurano il settore finanziario e dei mercati²⁷⁰, i reati contro la persona²⁷¹, la produzione e la distribuzione di sostanze stupefacenti²⁷², ma anche frodi e contraffazione²⁷³. Anche qui, l'incidenza in questa molteplicità di ambiti di algoritmi “malevoli” dipende sempre dal fatto che l'IA è contemporaneamente uno strumento *general-purpose* e *multi-purpose*²⁷⁴. Tuttavia, non è possibile ricondurre alcuna tipologia di responsabilità, ancorchè penale, nei confronti dei sistemi di IA impiegati in questi settori. Soprattutto nei reati punibili solo con l'insorgere di un certo elemento psicologico quale il dolo, è impossibile configurare in capo ad un sistema automatizzato di IA una qualsivoglia intenzionalità alla quale ricollegare l'ipotesi di responsabilità, in quanto il sistema in sé “*non ha la capacità di attribuire un significato alle proprie azioni [...] In quanto sistemi non pensanti, privi di intenzioni e semplice riserva di capacità di agire, [l'IA] non può assumersi la responsabilità delle proprie azioni.*”²⁷⁵

Al di là delle applicazioni strettamente criminose dell'IA, vanno considerati quegli utilizzi nell'ambito militare e di polizia, i quali, sebbene servano uno scopo di pubblica sicurezza e difesa, devono essere prodotti e impiegati nel rispetto dei diritti umani, e dei diritti fondamentali sanciti dalle fonti primarie dell'ordinamento nazionale ed europeo. L'integrazione in ambito militare dell'IA

²⁶⁹ *Id.*, 179.

²⁷⁰ Ad esempio, alcuni sistemi automatici sono in grado di influenzare le scelte degli agenti “umani” a mezzo di sistemi di *trading* automatizzato che hanno il solo scopo di manipolare il mercato. *Id.*, 189-191.

²⁷¹ Un esempio di reati contro la persona commessi da parte dell'IA sono le ipotesi di molestie, ma anche di tortura. Le prime sono state commesse da agenti autonomi quali bot conversazionali che hanno sviluppato delle modalità di interazione aggressive e lesive dei diritti degli *end-users* sia indipendentemente dall'ambiente in cui venivano dispiegati, sia proprio a causa di questo: è il caso di Tay, bot conversazionale di IA progettato da Microsoft e implementato su Twitter nel 2016, che ha appreso e replicato comportamenti estremisti, razzisti ed antisemiti, proprio a partire dalla propria interazione con gli utenti della piattaforma social appartenenti alle frange dell'ultradestra statunitense. *Id.*, 195. *Vedi anche* Hunt, E. (24 marzo 2016). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. *The Guardian*. <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>.

²⁷² Il commercio di sostanze stupefacenti è un altro settore, ancorchè illecito, in cui l'IA ha potuto dimostrare le sue potenzialità di automazione ed ottimizzazione dei processi. Ad esempio, il commercio della droga ha registrato miglioramenti significativi grazie all'impiego di veicoli autonomi pilotati esclusivamente dall'IA (droni o sistemi di navigazione). Tuttavia, i sistemi di IA sono sfruttati in questo campo anche per aumentare il bacino di utenza del commercio illecito di droga, grazie all'utilizzo di pubblicità indirizzata agli acquirenti e bot conversazionali che interagiscono direttamente con la potenziale clientela. Floridi (2022), 193-194.

²⁷³ Floridi (2022), 189-205.

²⁷⁴ Edwards (2022), 2, 6-7.

²⁷⁵ Floridi (2022), 197-198

consente la sua implementazione secondo una molteplicità di scopi, tra cui l'automatizzazione dei processi e una maggiore indipendenza nell'operare vari tipi di armamenti (tra cui veicoli autonomi, droni, o i c.d. *lethal autonomous weapons systems*²⁷⁶, LAWS²⁷⁷). Sebbene queste applicazioni possano portare ad un'ottimizzazione dei processi militari e di difesa, non vanno dimenticati gli effetti assolutamente devastanti che essi sono in grado di provocare in forza del loro potenziale dispiegamento su larga scala. Alcuni sostengono che l'utilizzo dei LAWS dovrebbero allertare la comunità internazionale, a causa della possibilità, per questi sistemi automatizzati o completamente autonomi, di commettere azioni in grado di infrangere i diritti umani²⁷⁸: l'assenza di una cornice normativa internazionale che possa fissare i parametri di utilizzo di queste tecnologie è un dato preoccupante, così come lo sforzo, per ora vano, di alcuni paesi di cercare un terreno comune di disciplina²⁷⁹.

Un esempio quanto meno attuale, strettamente riconnesso con le potenzialità distruttive dell'IA in ambito militare è costituito da quei sistemi come l'"*haBzora*", anche conosciuto con il nome di "*Gospel*," progettato e utilizzato dallo Stato di Israele. Questo sistema ha lo scopo di individuare, a mezzo di inferenze probabilistiche basate su algoritmi di ML alimentati da grandi quantità di dati e metadati, gli obiettivi strategici da colpire con i bombardamenti sul territorio di Gaza iniziati a seguito degli attacchi sferrati da Hamas il 7 ottobre 2023²⁸⁰. L'obiettivo dell'algoritmo sembrerebbe quello di limitare il più possibile le vittime civili, prediligendo una modalità di operatività chirurgica nei confronti dei membri

²⁷⁶ Gli *Autonomous Weapons Systems* (AWS) sono strumenti così definiti: "agenti artificiali che, come minimo, sono in grado di modificare i propri stati interni per raggiungere un determinato obiettivo, o una serie di obiettivi, all'interno del proprio ambiente operativo dinamico e senza l'intervento diretto di un altro agente, che possono anche essere dotati di alcune capacità di modificare le proprie regole di transizione senza l'intervento di un altro agente, e che vengono impiegati con lo scopo di esercitare una forza cinetica contro un'entità fisica (sia essa un oggetto o un essere umano) e a tal fine sono in grado di identificare, selezionare o attaccare il bersaglio senza l'intervento di un altro agente." Gli LAWS, inoltre, costituiscono un sottoinsieme dei AWS, dotati per l'appunto di capacità di eliminare *target* umani. Taddeo, M., Blanchard, A. A Comparative Analysis of the Definitions of Autonomous Weapons Systems. *Sci Eng Ethics* 28, 15 (2022). <https://doi.org/10.1007/s11948-022-00392-3>.

²⁷⁷ Lingevicius, J. Military artificial intelligence as power: consideration for European Union actorness. *Ethics Inf Technol* 25, 19 (2023). <https://doi.org/10.1007/s10676-023-09684-z>

²⁷⁸ Mantelero (2022) Foreword viii.

²⁷⁹ *Id.*

²⁸⁰ Le informazioni sul sistema di IA impiegato dallo Stato di Israele sono frutto del *reportage* redatto dalle testate israelo-palesinesi +972 Magazine e Local Call. Reiff, B. (3 dicembre 2023). 'A mass assassination factory': Inside Israel's calculated bombing of Gaza. +972 Magazine. <https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/>.

strategici di Hamas²⁸¹. Tuttavia, con la crescente richiesta di obiettivi da parte da parte dello Stato di Israele e l'intensificarsi dell'offensiva sul territorio di Gaza, il sistema ha permesso l'individuazione di più di cento obiettivi strategici al giorno, aumentando la portata e la frequenza degli attacchi, a scapito della loro accuratezza, della minimizzazione delle vittime civili e della possibilità di sottoporre ad una revisione umana i risultati algoritmici. Le conseguenze sono state assolutamente devastanti per la popolazione civile palestinese, che conta più di 20.000 di vittime civili in soli 80 giorni di invasione di Gaza²⁸².

Inoltre, non è da dimenticare uno degli aspetti più controversi dell'utilizzo di tecnologie *AI-powered* in ambito bellico, specialmente per quanto riguarda l'utilizzo dei LAWS. Essi, infatti, possono causare la creazione di un divario sempre più profondo tra le attività militari che l'IA pone in essere e le implicazioni etiche morali delle stesse. Se già la frammentazione gerarchica e burocratica dei processi bellico-militari funge da strumento notoriamente efficace per creare una dissonanza cognitiva tra il mero svolgimento di compiti e la causazione di attività moralmente insostenibili²⁸³, l'automazione tramite l'IA di questi processi non fa che aumentare esponenzialmente questo divario. Gli aspetti etici dello svolgimento di determinate attività scompaiono in quanto rifiutati dalla stessa componente umana, che può delegare la commissione di questi atti ad enti strutturalmente incapaci di valutarne le implicazioni morali.

A ben vedere, l'uso dell'IA in ambito militare costituisce una delle prime vistose eccezioni di applicabilità della proposta di regolamento europeo dell'*AI Act*²⁸⁴. Questa scelta, oggetto di un compromesso politico tra i vari Stati membri dell'UE nell'ambito della discussione della proposta di regolamento, ha sostanzialmente ricavato un'area protetta, in cui i requisiti previsti per i sistemi "ad alto rischio" non sono applicabili, controbilanciando la tutela dei diritti

²⁸¹ Sebbene questi siano obiettivi dichiarati dal governo israeliano, non vi sono particolari prove a favore di un funzionamento del sistema di tale sorta; è piuttosto evidente un sovra-utilizzo del sistema per massimizzare l'incidenza dei danni sulla popolazione e sul territorio. Fenton-Harvey, J. (13 dicembre 2023). The Gospel: Israel's controversial AI used in the Gaza war. *The New Arab*. <https://www.newarab.com/analysis/gospel-israels-controversial-ai-used-gaza-war>

²⁸² I dati sono aggiornati al giorno 25 dicembre 2023. *Vedi*

²⁸³ Bauman, Z. (2013). *Modernity and the Holocaust*. John Wiley & Sons, Chapter 4. Uniqueness and Normality of Holocaust, 93-106.

²⁸⁴ Art. 2(3), *AI Act* (Versione III)

fondamentali a cui sono informati i principi di certificazione, controllo e trasparenza dell'*AI Act* con la necessità di sviluppo in campo militare²⁸⁵.

Uno dei temi attualmente più discussi nell'ambito della valutazione dei rischi applicativi dell'IA, soprattutto in ambito europeo, è anche quello relativo all'adozione di tecnologie nella pubblica sicurezza che possano incidere fortemente sui diritti fondamentali. Proprio seguendo quell'approccio basato sul rischio utilizzato dall'UE nell'*AI Act*²⁸⁶, l'utilizzo di IA in questo settore può porre pericoli considerati inaccettabili all'interno della normativa. Ad esempio, inaccettabile è il rischio posto dai sistemi di IA in grado di consentire, a mezzo di tecnologie di riconoscimento facciale, l'identificazione biometrica di soggetti in tempo reale salvo alcune eccezioni previste specificamente dalla legge²⁸⁷. Consentire un'identificazione basata su dati biometrici (ancorchè considerati dati sensibili dalla normativa europea, ai sensi dell'art. 9 GDPR²⁸⁸) in tempo reale con una sostanziale sorveglianza biometrica indiscriminata, non risponde ad un bilanciamento tra il pubblico interesse alla sicurezza e alla prevenzione dei reati, con il diritto alla privacy e alla non discriminazione (dal momento che come accade spesso queste tecnologie risultano essere costellate di pregiudizi interni che ne influenzano il funzionamento²⁸⁹).

²⁸⁵ Tuttavia, alcune organizzazioni internazionali ed europee come European Digital Rights (EDRi), AccessNow ed Algorithm Watch, durante le fasi iniziali del processo di *drafting* dell'*AI Act*, hanno esortato il legislatore a formulare diversamente l'eccezione di applicabilità ai modelli di IA utilizzati per "scopi militari," in quanto la dicitura presente nell'attuale bozza è troppo ampia e facilmente soggetta a interpretazioni a seconda della casistica. Secondo le organizzazioni che hanno presentato un emendamento all'art. 2(3) dell'*AI Act*, la caratteristica *multi-purpose* di alcuni sistemi di IA potrebbe portare a situazioni per cui modelli utilizzati per scopi militari vengano adottati da governi ed organizzazioni per scopi anche diversi da quelli militari, che potrebbero anche essere considerati "ad alto rischio," ma che dunque sfuggirebbero ai requisiti previsti dalla normativa proprio in forza dello scopo iniziale per i quali son stati impiegati. Questo determina un *loophole* non indifferente nella normativa, dando motivo di preoccupazione. Vedi European Digital Rights, Access Now, Algorithm Watch, Panoptykon Organization, Article 19, Electronic Frontier Finland, & Electronic Privacy Information Center. (2022). *Scope of the EU Artificial Intelligence Act (AIA): Military Purposes and National Security*. European Center for Not-for-profit Law. https://ecnl.org/sites/default/files/2022-03/ECNL%20Papers%20on%20scope%20of%20AIA%20ECNL_FINAL.pdf.

²⁸⁶ Per un'analisi sull'approccio *risk-based* adottato dall'UE nella stesura dell'*AI Act*, vedi il sottoparagrafo 3.3.2 del Capitolo 2.

²⁸⁷ Vedi il sottoparagrafo 3.3.2 del Capitolo 2 sulle tecnologie di IA considerate "inaccettabili" a causa del rischio posto nei confronti della società.

²⁸⁸ Art. 9(1) GDPR.

²⁸⁹ Uno studio effettuato nel 2019 dall'agenzia statunitense NIST (*National Institute of Standards and Technology*) ha verificato che la maggior parte di sistemi di riconoscimento facciale esistenti nel mercato statunitense hanno ottenuto prestazioni peggiori su volti non bianchi, generando dei risultati discriminatori a causa del maggior numero di falsi positivi e falsi negativi riscontrati nell'analisi di volti appartenenti al gruppo BIPOC (*black, indigenous and people of color*). Vedi Grother, P., Ngan, M., Hanaoka, K., (2019) *NISTIR 8280: Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, <https://doi.org/10.6028/NIST.IR.8280>.

In conclusione, è d'uopo una postilla finale. La precedente elencazione, ancorchè dettagliata, dei potenziali rischi intrinseci e di applicazione dell'IA coesiste perfettamente con le opportunità di sviluppo e di progresso che essa è in grado di portare con sé. Opportunità e rischi non devono percepirsi come due aspetti opposti e inconciliabili dell'IA: una visione simile porterebbe inevitabilmente a sostenere visioni riduzioniste verso un approccio di eccessiva cautela o, dall'altro lato, ad un sovra utilizzo senza criterio e protezione della collettività²⁹⁰. Per una corretta implementazione dell'IA nella società digitale odierna, è da auspicarsi un approccio intermedio tra i due estremi, correttamente informato sia sulle opportunità che dei potenziali rischi, per poter sfruttare al meglio una tecnologia astrattamente e concretamente in grado di servire l'intera società²⁹¹.

3.3 Un cenno ai rischi virtuali: la singolarità tecnologica²⁹²

Si è scelto di riportare un breve cenno ad uno di quei rischi denominati come “virtuali,” in diretta contrapposizione con quelli considerati “attuali” del paragrafo precedente. Questa trattazione, ancorchè breve, si ritiene necessaria anche solo per sottolineare un aspetto: le ipotesi dei rischi che l'IA pone nei confronti della sopravvivenza della società umana, seppure intriganti, non corrispondono allo stato attuale della realtà tecnologica, né sono tantomeno da considerare in cima a quella vasta gamma di problematiche attuali e reali (precedentemente analizzate) che i sistemi di IA possono portare con sé.

Il tema della singolarità, nonostante la sua scarsa possibilità di verifica, viene spesso rilanciato da individui estremamente influenti, talvolta potenti imprenditori che godono di fama globale²⁹³. A questo proposito, hanno suscitato particolare clamore le note dichiarazioni sui rischi esistenziali dell'IA rese da Elon Musk, CEO di Tesla, Space X e imprenditore da alcuni considerato un assoluto visionario del progresso tecnologico. In un'intervista seguita all'AI

²⁹⁰ *Id.*, 280-282.

²⁹¹ *Id.*, Capitolo 9. Buone pratiche: l'uso dell'IA per il bene sociale, 221-258.

²⁹² Il concetto di singolarità tecnologica consiste in un evento ipotetico nel futuro in cui il progresso tecnologico accelera oltre la comprensione umana, portando a cambiamenti radicali e imprevedibili, che coinvolge l'avvento di un'intelligenza superiore (artificiale) e portare a progressi tecnologici rapidi. La possibilità e l'effetto di una singolarità sono argomenti dibattuti, ma non nuovi: essi ricorrevano già nel 1950 con i dibattiti di Nobeert Wiener sulla prima cibernetica, ritornando in auge sia a metà degli anni '60, durante una delle numerose “estati” dell'IA, per poi tornare ad affascinare il pubblico nella fine degli anni '80, fino ad oggi. Benbouzid *et al.* (2022), 9-10.

²⁹³ *Id.*, 9.

Summit tenutosi a Bletchley nell'autunno 2023²⁹⁴, l'imprenditore ha indicato l'IA come “una delle più grandi minacce per l'umanità”²⁹⁵, nonché “qualcosa che sarà molto più intelligente di noi [...], [che] non è chiaro se possiamo controllare.”²⁹⁶ Questa visione ricomprende un'ideologia, sposata anche da altri imprenditori del settore ed esperti, che vede vicino l'arrivo di qualche forma ultra intelligente di IA in grado di superare le capacità intellettive umane e il loro controllo. Questa potenziale forma di sviluppo di questa tecnologia riappare ciclicamente nel dibattito sull'IA, talvolta come IA “forte,”²⁹⁷ talvolta come IAG, o *AGI* (Intelligenza Artificiale Generale, *Artificial General Intelligence* ²⁹⁸). Interrogarsi principalmente sulla possibilità dell'IA di mettere in pericolo la sopravvivenza umana stessa, sfuggendo al controllo umano, tralasciando le preoccupazioni per le reali implicazioni dell'IA per quanto riguarda la sua capacità di aumentare la discriminazione, minare la privacy individuale e di gruppo, o influire negativamente su sostenibilità economica, ambientale e sociale (*inter alia*), è potenzialmente fuorviante, soprattutto nel momento in cui sorge la necessità di discutere se, e in che misura, configurare una disciplina di regolamentazione per questa tecnologia. Esponenti accademici come Fei-Fei Li, co-direttrice dello *Stanford Human-centric AI Institute* nonché tra i maggiori esperti mondiali di sistemi di IA, sostengono che queste preoccupazioni esistenzialiste dovrebbero essere sì oggetto di discussioni intellettuali, ma non monopolizzare la conversazione circa i rischi dell'IA. Floridi, dal canto suo, avanza un ulteriore aspetto: limitare l'attenzione sui rischi esistenziali non solo è dannoso ma anche “irresponsabile” in quanto costituisce una forte distrazione sulle reali sfide poste dall'IA²⁹⁹.

Ad ogni modo, salta all'occhio come, coloro che sono intenti ad avvertire l'opinione pubblica per i potenziali rischi devastanti nei confronti della stessa sopravvivenza dell'umanità come la conosciamo a causa della larga implementazione di sistemi di IA sono gli stessi che investono miliardi nello

²⁹⁴ Vedi nota 186.

²⁹⁵ Heikkilä, M. (14 novembre 2023). AI is at an inflection point, Fei-Fei Li says. *MIT Technology Review*. <https://www.technologyreview.com/2023/11/14/1083352/ai-is-at-an-inflection-point-fei-fei-li-says/>.

²⁹⁶ *Id.*

²⁹⁷ Per la distinzione tra IA “forte” e “debole,” vedi il paragrafo 2.1.

²⁹⁸ Floridi (2022), 263.

²⁹⁹ *Id.*, 262.

sviluppo di questo settore³⁰⁰. Lo stesso Musk investe nello sviluppo dell'IA su molteplici fronti: con la fondazione della nuova start-up xAI³⁰¹, nonché in altri progetti come quelli portati avanti da Neuralink, società che si occupa dello sviluppo di neuro-tecnologie ed IA³⁰². Per questo motivo, alcuni sostengono che la preoccupazione sui rischi esistenziali dell'IA è *“tipica del mondo ricco, che probabilmente impensierirà le persone nelle società opulente, che sembrano dimenticare i mali reali che [già] opprimono l'umanità e il nostro pianeta.”*³⁰³

³⁰⁰ Capone, E. (30 maggio 2023). Lo strano appello contro le IA: “Rischiando l'estinzione”. Ma le aziende continuano a svilupparle. *La Repubblica*. https://www.repubblica.it/tecnologia/2023/05/30/news/ia_rischio_estinzione_appello_altman_hint_on_musk_regole-402536379/.

³⁰¹ Milmo, D. (21 agosto 2023). Elon Musk launches AI startup and warns of a 'Terminator future.' *The Guardian*. <https://www.theguardian.com/technology/2023/jul/13/elon-musk-launches-xai-startup-pro-humanity-terminator-future>

³⁰² Waltz, E. (24 giugno 2021). Elon Musk announces Neuralink Advance toward syncing our Brains with AI. *IEEE Spectrum*. <https://spectrum.ieee.org/elon-musk-neuralink-advance-brains-ai>

³⁰³ Floridi (2022), 269.

Capitolo 2: la regolamentazione dell'IA nell'Unione Europea

1. Premesse: la corsa alla regolamentazione dell'IA e il modello regolamentare europeo nel campo dell'economia digitale

Il presente capitolo si pone l'obiettivo di procedere ad una analisi dell'approccio adottato dall'Unione Europea nel fornire una specifica regolamentazione di quei sistemi che utilizzano l'Intelligenza Artificiale così come correttamente individuata nelle varie definizioni fornite in quello precedente. Dopo aver precisato non solo gli elementi fondanti, ma anche le implicazioni che definiscono e contraddistinguono questi sistemi, che li caratterizzano, comportandone il loro utilizzo più o meno rischioso nell'odierna società digitale, è necessario passare in rassegna dei tipi di regolamentazione al momento esistenti nei vari ordinamenti, volti a governare questa tecnologia peculiare la cui necessità di regolamentazione è solo recentemente emersa. Tutto ciò, nell'ottica di presentare, al termine dell'elaborato, un quadro dell'attuale *state of the art* in tema di disciplina dell'IA, per poi indicare alcuni aspetti normativi (e non) che, se integrati, possono consentire lo sviluppo e l'utilizzo di IA più possibile in linea con i diritti umani e fondamentali, l'etica, e in ultimo foriera di quel benessere sociale che l'innovazione è talvolta in grado di generare, se indirizzata sul giusto binario³⁰⁴.

A partire da quelle opportunità di progresso in svariati settori che l'IA ha dimostrato di poter comportare³⁰⁵, ciò ha determinato il progressivo aumento di impiego di tecnologie basate su algoritmi che lentamente entrano a far parte della quotidianità di un largo numero di individui, accentuando una chiara corsa allo sviluppo di sistemi di IA (c.d. “*Race to AI*”³⁰⁶ o anche “*AI arms race*”) a cui hanno preso parte i maggiori paesi a livello globale, nonché le aziende che dominano questo settore. Questa “corsa” è giustificata dalle innegabili opportunità non solo di progresso, ma anche di profitto, che l'impiego di sistemi *AI-powered* può portare con sé nei campi più svariati. Molti paesi hanno cercato di sviluppare questi sistemi rapidamente, alcuni con più successo di altri, ma tutti prefigurandosi l'obiettivo di

³⁰⁴ Si pensi a quella nozione di *AI For Social Good* (AI4SG) presentata nel capitolo precedente. Vedi p. 37. Vedi capitolo 3 parte III per questi spunti sugli strumenti da integrare nella regolamentazione dell'IA.

³⁰⁵ Vedi Capitolo 1 paragrafo 3.1.

³⁰⁶ Smuha, N. A. (2021b). From a ‘race to AI’ to a ‘race to AI regulation’: regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), 57–84, <https://doi.org/10.1080/17579961.2021.1898300>.

diventare *leader* nell'avanzamento tecnologico e nel progresso di questa tecnologia, che fosse specificamente nello sviluppo di sistemi di IA, ovvero la creazione di start-up a ciò dedicate, ovvero nell'avanzamento del mondo accademico e del capitale umano nell'IA come disciplina, ovvero, ancora nella progettazione di *software* e *hardware* da impiegare per questi scopi³⁰⁷.

La pervasività sempre maggiore di questa corsa allo sviluppo di sistemi *AI-powered* è giustificata non solo dalle opportunità di produrre dei benefici economici e persino sociali, ma anche dai rilevanti costi-opportunità che la mancata adozione di questa tecnologia può causare³⁰⁸. A questo proposito, il gruppo di esperti sull'IA costituito dalla Commissione Europea (il c.d. *High-Level Expert Group on AI, AI HLEG*³⁰⁹) sottolineò già nel 2019 che l'inefficace sfruttamento economico dell'IA a livello europeo poteva generare una perdita complessiva di più di €400 miliardi in valore aggiunto al GDP dell'Unione Europea entro il 2030³¹⁰. Questo spiega la generale riluttanza, non solo a livello comunitario bensì appartenente a tutto il contesto mondiale, nell'approcciare questo settore dell'innovazione tecnologica con eccessiva cautela. Essendo l'IA una delle forze trasformative di importanza strategica più significative del nostro tempo, nessuno, a livello globale, ha voluto lasciarsi sfuggire l'occasione di saltare sul c.d. "*treno dell'IA*."³¹¹

Tuttavia, in tempi più recenti, questa "*Race to AI*" ha innescato una parallela "corsa alla regolamentazione" (c.d. "*Race to AI regulation*")³¹² di questa tecnologia³¹³. Come descritto da alcuni accademici "*sembra che stia emergendo un nuovo campo di gioco per la competizione normativa globale che, nel migliore dei*

³⁰⁷ Il rapido sviluppo dell'IA, sulla spinta propulsiva dell'innovazione di molti paesi, è una realtà abbastanza consolidata nel panorama tecnologico attuale. Si pensi che già tra il 2017 e il 2019, Cina, Stati Uniti e Unione Europea (tra gli altri) si erano tutti attivati, secondo i rispettivi approcci alla produzione e alle diverse capacità di sfruttamento dell'IA, ad immaginare delle strategie per mantenere una posizione di rilevanza nel settore di sviluppo di questa tecnologia (la Cina con l'emanazione del suo programma di sviluppo denominato "*Next Generation AI Development Plan*" nel 2017; gli USA con l'*Executive Order No. 13859* del 2019, a firma dell'allora Presidente degli Stati Uniti Donald Trump, che ha predisposto l' "*American AI Initiative*;" e l'UE con il *drafting* della c.d. "*European AI Strategy*" nel 2018). Vedi Smuha (2021b), 3.

³⁰⁸ Floridi (2022), 280-282. Vedi anche Smuha (2021b), 3.

³⁰⁹ Per la definizione dell'*High-Level Expert Group* costituito dalla Commissione Europea nel 2018 e il riassunto dei suoi lavori, vedi paragrafo 2.

³¹⁰ Smuha (2021b), 4. Vedi anche AI HLEG, Policy Recommendations for Trustworthy AI (giugno 2019) <https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.

³¹¹ Commissione Europea, 'Artificial Intelligence for Europe' (Comunicazione) COM (2018) 237 final.

³¹² Smuha (2021b), 4.

³¹³ *Id.*

casi, spinge i governi – nell’incertezza dell’impatto della tecnologia, dell’impatto dell’intervento normativo e del costo del mancato intervento - a trovare il bilanciamento più appropriato tra protezione e innovazione.”³¹⁴ Consci delle criticità su molteplici fronti che l’IA può causare, svariati ordinamenti hanno preso parte a questa accelerazione normativa, ognuno applicando a questa spinta la propria cultura legislativa orientata verso alcuni strumenti piuttosto che altri. Tuttavia, alcuni temono che la convergenza di queste due tendenze (*Race to AI* da un lato e *Race to AI regulation* dall’altro) possa scatenare una concorrenza normativa “al ribasso” per quanto riguarda la protezione contro i potenziali rischi dell’IA³¹⁵. Per questo, è auspicabile che la regolamentazione tenga conto sì delle opportunità, e dei relativi costi-opportunità, ma anche delle necessità strutturali di limitare i fenomeni rischiosi e dannosi che l’IA porta con sé, cercando di tutelare a tutto tondo non solo gli individui che vengono in contatto con i sistemi, ma l’intera società civile e, in ultimo, l’assetto democratico e lo stato di diritto³¹⁶.

D’altronde, quest’ultimi costituiscono le aree di rischio che hanno ricevuto e continuano a ricevere scarsa considerazione da parte dell’opinione pubblica, soprattutto per quanto riguarda l’intera *supply chain* dell’IA, mentre si presta maggiore attenzione a quei rischi *ex post* che influiscono direttamente sulla sfera personale dei singoli individui che interagiscono con questi sistemi³¹⁷.

L’obiettivo del presente capitolo è dunque quello di analizzare l’approccio tenuto dall’Unione Europea in questo settore: in *primis* ripercorrere quelle tappe che hanno portato alla scelta di ricorrere alla legislazione e gli obiettivi a cui essa mira, e poi analizzare il contenuto di quella che, attualmente, è la prima ipotesi di regolamentazione onnicomprensiva, basata su diversi standard di rischio e corredata di strumenti di *enforcement* quali sanzioni³¹⁸, regolamentazione non solo coinvolge l’IA nel suo complesso, ma che fornisce un quadro di requisiti graduati a seconda del tipo di sistema: la Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull’Intelligenza Artificiale, cioè il c.d. *AI Act*.

³¹⁴ *Id.*, 5.

³¹⁵ *Id.*

³¹⁶ *Id.*, 22.

³¹⁷ Tsamados, A., Aggarwal, N., Cowls, J., Morley J., Roberts H., Taddeo, M., Floridi, L. (2021). The Ethics of Algorithms: Key Problems and Solutions. In: Floridi, L. (eds) *Ethics, Governance, and Policies in Artificial Intelligence. Philosophical Studies Series*, vol 144. Springer, Cham. https://doi.org/10.1007/978-3-030-81907-1_8. Vedi anche Thomas et. al (2023), 4.

³¹⁸ Art. 71, *AI Act* (Versione I).

Alcune considerazioni preliminari sull'approccio dell'UE sono necessarie per comprendere come si è evoluto quel percorso che ha portato, ad oggi, alla scelta dell'Unione di fornire una regolamentazione dell'IA così strutturata. Queste considerazioni muovono a partire dal generale approccio che l'Unione Europea ha assunto nel mondo dell'industria del digitale nello scorso decennio.

Secondo le teorie di Anu Bradford, Prof.essa della Columbia Law School e direttrice dell'*European Legal Studies Center* della stessa università newyorkese, celebre per aver coniato la nozione del c.d. "effetto Bruxelles"³¹⁹ (*Brussels Effect*), il modello di approccio al mondo del digitale dell'UE differisce dagli altri due (principali) approcci portati avanti dalle due potenze globali nel settore della tecnologia, gli USA e la Cina. I primi, nello sviluppare gran parte dell'economia digitale del mondo contemporaneo, hanno adottato un approccio normativo denominato da Bradford "market-driven", cioè guidato dalle logiche di mercato e ispirato alla protezione di valori fondamentali quali la libertà di parola, la libertà di Internet e il mantenimento di incentivi all'innovazione, modellati sulla scorta di una dottrina di tangibile tecno-ottimismo, che ben si accorda con la fede intransigente nei mercati rispetto alla disciplina governativa, che è dunque limitata il più possibile, dando spazio alle dinamiche di mercato³²⁰.

La Cina, al contrario, ha abbracciato un approccio denominato "state-driven," incentrato sullo sfruttamento dell'industria tecnologica come punto di forza per consolidare il controllo del governo sull'economia e, allo stesso tempo, la

³¹⁹ In breve, per "Effetto Bruxelles" si intende quell'adeguamento a livello globale ai principi europei come risposta alla regolamentazione di un determinato settore emanata dall'Unione. Esso si può verificare qualora l'Unione adotti una regolamentazione destinata a disciplinare un certo settore economico le cui aziende traggono maggiore vantaggio dal seguire un unico standard globale, piuttosto che approfittare dei mercati extra-UE più lassisti, in quanto meno regolamentati. Il tipico esempio di Effetto Bruxelles è quello che si è manifestato a seguito dell'emanazione del Regolamento generale sulla protezione dei dati del 2016 (GDPR). Dall'entrata in vigore del regolamento nel 2018, più di 100 Paesi hanno adottato norme sulla privacy modellate sulla scorta dei principi contenuti nel GDPR, quasi secondo un principio di c.d. *mirror legislation*. Questo meccanismo di adeguamento, tuttavia, non era per nulla scontato, a causa della diversità di questi paesi dall'Unione Europea e dalla sua traduzione legislativa. Poiché nel settore della protezione dei dati era dunque più conveniente per le aziende che operano a livello internazionale conformarsi a uno standard unico, altri Paesi sono stati ispirati a adottare politiche sulla privacy con caratteristiche simili a quelle europee. Il settore della protezione dei dati non è tuttavia l'unico in cui è stato possibile osservare un effetto Bruxelles, dato che questo ha avuto logo anche per regolamentazioni di tipo molto diverso dal GDPR, come, *ex multis*, il Regolamento (CE) n. 1907/2006 del Parlamento europeo e del Consiglio, del 18 dicembre 2006, concernente la registrazione, la valutazione, l'autorizzazione e la restrizione delle sostanze chimiche (REACH, acronimo di *Registration, Evaluation, Authorisation and Restriction of Chemicals*, d'ora in poi denominato "Regolamento REACH"). Vedi in generale Bradford, A. (2019b). *The Brussels effect: How the European Union Rules the World*, In Oxford University Press eBooks.

³²⁰ Bradford (2023), *Chapter 1: the American Market-driven Regulatory Model*, 34-68.

società civile cinese (attraverso l'utilizzo della tecnologia come strumento di propaganda e sorveglianza, secondo principi dell'autoritarismo digitale³²¹); ciò a scapito della protezione delle libertà individuali, in quanto la *governance* del digitale, in Cina, è fortemente ricollegata alla sovranità dello stato autoritario³²².

Diversamente da entrambi i paesi, l'Unione Europea, nonostante non abbia un ruolo di spicco nel panorama dell'innovazione tecnologica rispetto agli USA e alla Cina³²³, ha scelto di assumere un approccio alla regolamentazione del settore differente dai suoi *competitors*, denominato da Bradford il “*rights-driven model*”³²⁴. Nella propria riluttanza ad allinearsi ad uno degli schemi basati rispettivamente sul mercato (USA) o sul controllo del governo autoritario (Cina), l'UE “*ritiene che i governi abbiano un ruolo centrale sia nel guidare l'economia digitale sia nell'utilizzare l'intervento normativo per sostenere i diritti fondamentali degli individui, preservare la struttura democratica della società e garantire un'equa distribuzione dei benefici nell'economia digitale.*”³²⁵

Le motivazioni alla base dello sviluppo di un approccio simile alla regolamentazione della tecnologia sono molteplici, riscontrabili sia nella tradizione legislativa dei paesi europei, che in una generale esigenza di armonizzazione del mercato interno. Tuttavia, una ragione centrale è la convinzione tutta europea del fatto che i mercati, lasciati a sé stessi, non producano necessariamente risultati ottimali e che sia necessario l'intervento del governo per preservare e rafforzare questi diritti sopracitati a tutela della comunità europea³²⁶. Secondo questa visione, lo stato o in questo caso, l'entità sovranazionale che legifera, ha un ruolo centrale nell'assicurare sì il rispetto della concorrenza nei mercati, ma anche (e soprattutto, secondo Bradford) quello dei diritti fondamentali, i quali sono spesso compromessi dagli stessi prodotti e servizi del settore digitale³²⁷.

³²¹ *Id.*, Chapter 2: *the Chinese State-driven Regulatory Model*, 69-104.

³²² *Id.*, 70.

³²³ Se si prendessero come punto di riferimento le aziende *leader* del digitale, queste sono nate principalmente negli USA o in Cina, salvo pochissime eccezioni europee (tra cui figura, ad esempio, la piattaforma svedese di *streaming* musicale ampiamente diffusa a livello globale, Spotify). La lista “*The World's Largest Technology Companies*” redatta annualmente dalla rivista *Forbes* figurava due aziende europee nelle prime 20, durante il 2021; la stessa lista redatta nel 2023, non indica alcuna società *tech* sviluppata da un paese dell'Unione. *Vedi anche* Bradford (2023), 136.

³²⁴ Bradford (2023), Chapter 3: *the European Rights-driven Regulatory Model*, 105-145.

³²⁵ *Id.*, 105.

³²⁶ *Id.*, 107. *Vedi anche* Nemitz, P., (2018) *Constitutional Democracy and Technology in the Age of Artificial Intelligence*, *Phil. Trans. R. Sc. A.* 376, 1.

³²⁷ Si pensi all'avvento delle piattaforme *social* durante gli scorsi anni '10 del ventunesimo secolo. Nonostante il loro *business model* si basasse (e tutt'oggi si basi ancora, in gran parte) sullo sfruttamento dei dati personali degli utenti a fronte del servizio di *social media* fornito, l'Unione

Per questo motivo, utilizzare la regolamentazione del digitale per salvaguardare il rispetto di quei diritti stabiliti nella Carta Europea dei diritti fondamentali³²⁸ è un tratto assolutamente distintivo dell'ordinamento comunitario³²⁹. A supporto di questa scelta, vi sono le numerose normative che l'UE ha emanato nel settore digitale volte a tenere fede a questo scopo: *in primis* il GDPR, ma anche il c.d. “*Data Package*”³³⁰ precedentemente menzionato, o anche la coppia di regolamenti *Digital Service Act*³³¹ e *Digital Market Act*³³², per citarne alcuni. Tra i progetti regolamentari ancora *in itinere* dell'Unione nel settore del digitale, figura senza dubbio anche il progetto di regolamento sull'IA, l'*AI Act*, il quale, sebbene non ancora approvato³³³, sembra combinare sia lo scopo di armonizzazione del mercato interno per rafforzare la posizione economica comunitaria a livello globale nel campo dello sviluppo e dell'impiego di IA, ma anche lo scopo di tutela dei diritti di coloro che risiedono nel territorio comunitario³³⁴, nonché dello stato di diritto e della democrazia. Almeno, questo è quanto l'UE dichiara di voler perseguire nel suo disegno legislativo ideale per

Europea è quantomeno intervenuta a legiferare per tutelare maggiormente i diritti degli utenti che risiedono nell'UE, conferendo loro diritti e modalità per farli valere a mezzo del Regolamento generale sulla protezione dei dati (GDPR, che sostituì l'allora già obsoleta “*Data Protection Directive*”, cioè la direttiva 95/46/EC). Sebbene si possa discutere dell'efficacia di una normativa come il GDPR a fornire una tutela effettiva dei dati personali degli utenti, il cui sfruttamento non è certo terminato dopo l'emanazione del GDPR, è innegabile che l'UE abbia mandato un segnale forte con l'approvazione del regolamento, spostando l'attenzione della *digital industry* su una categoria di soggetti che, fino a quel momento, non erano stati destinatari di un *empowerment* a livello di diritti digitali: i c.d. *data subjects*, cioè i singoli individui, denominati “interessati al trattamento dei dati personali.” *Vedi in generale* Zuboff (2019).

³²⁸ Carta dei diritti fondamentali dell'Unione Europea, 7 giugno 2016, 2012 O.J. (C 326/02) (anche conosciuta come “Carta di Nizza”).

³²⁹ Bradford (2023), 110-118.

³³⁰ Contente il *Data Governance Act* e la proposta di regolamento *Data Act*. Smuha (2023), 3. *Vedi anche* nota 116.

³³¹ Il *Digital Service Act* (DSA) è un regolamento dell'UE che mira a moderare le piattaforme online, ad affrontare i contenuti illegali generati dagli utenti, la pubblicità trasparente e la disinformazione. Il regolamento modifica l'obsoleta direttiva c.d. *e-commerce* del 2000 (direttiva 2000/31/CE), introducendo nuove regole per la responsabilità delle piattaforme e armonizzando le leggi in materia tra i 27 Stati membri. *Vedi in generale* Regolamento 2022/2065 relativo al mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (*Digital Service Act*) 2022 GU (L 277).

³³² Il *Digital Market Act* (DMA) è un regolamento dell'UE che mira a garantire un ambiente più competitivo nel mercato digitale dell'Unione. Esso crea otto categorie di servizi di piattaforma fondamentali (motori di ricerca online, servizi di intermediazione online, piattaforme di condivisione video, piattaforme di comunicazione, sistemi operativi, servizi *cloud* e servizi pubblicitari) e impone obblighi più severi alle principali aziende che vi rientrano. *Vedi in generale* Regolamento 2022/1925 sui mercati contendibili ed equi nel settore digitale (*Digital Market Act*), GU (L 265).

³³³ Per lo approfondire lo status dei lavori relativo all'approvazione dell'*AI Act*, *vedi* il paragrafo 3.1.

³³⁴ *Vedi* Smuha (2023), 8.

l'IA³³⁵. Tuttavia, alcuni hanno argomentato che la pretesa di proteggere i diritti fondamentali dei residenti europei mal si concilia con il fatto che il regolamento è strutturato, in realtà, come una normativa sulla sicurezza e la conformità dei prodotti (in questo caso i sistemi *AI-powered*) agli standard prestabiliti³³⁶.

Un'ulteriore considerazione è d'uopo per meglio inquadrare la peculiarità del modello regolamentare di matrice europea, sempre in relazione a quell'obiettivo di perseguire, a mezzo della regolamentazione dell'IA, la protezione di diritti fondamentali cristallizzati nella Carta di Nizza. A fronte degli evidenti rischi che l'IA può porre, molte case produttrici di servizi che impiegano IA hanno creato i propri codici etici, strumenti *market-driven* di autoregolamentazione interna, nell'ottica di condurre, in ultimo, a delle migliori *corporate practices*³³⁷. Tuttavia, la scelta dell'Unione di presentare il disegno di legge dell'*AI Act* per definire la cornice di operatività nella legalità dell'IA ha inequivocabilmente dimostrato agli altri attori del mercato che i codici etici emanati dai privati, da soli, sono insufficienti³³⁸: per assicurare una tutela a tutto tondo dei diritti fondamentali sono necessari impegni legali vincolanti, che derivino direttamente da un processo legislativo democratico³³⁹.

Oltre l'approccio europeo *rights-driven* alla regolamentazione del digitale in generale, vi è un altro obiettivo perseguito dall'Unione nel predisporre, soprattutto negli ultimi anni, diversi testi normativi mirati alla creazione di regole per questa particolare *industry*: cioè la necessità di affermare la sovranità digitale europea³⁴⁰. La crescente proliferazione di disegni legislativi dell'UE in questa materia non è casuale, ma risponde alla ben precisa prerogativa di affermare la propria autonomia sia verso l'esterno, in un settore dominato da potenze industriali non europee come la Cina e gli Stati Uniti, ma anche verso gli stessi ordinamenti

³³⁵ Vedi Relazione di accompagnamento, *AI Act* (Versione I), 1, in cui viene disposto che “l'interesse dell'Unione è quello di preservare la leadership tecnologica dell'UE e assicurare che i cittadini europei possano beneficiare di nuove tecnologie sviluppate e operanti in conformità ai valori, ai diritti fondamentali e ai principi dell'Unione.”

³³⁶ Questo profilo critico, insieme ad altri ascrivibili all'*AI Act*, saranno analizzati nel paragrafo 3.4 del presente capitolo.

³³⁷ Per esempio, come riportato da Bradford, numerose *Big tech* (tra cui Google, Meta, Microsoft etc.) hanno prodotto, in tempi non sospetti, codici etici contenenti le *corporate best practices* in materia. Bradford (2023), 115, 441, nota 51.

³³⁸ *Id.*, 115.

³³⁹ *Id.*

³⁴⁰ Smuha (2023).

interni dei singoli Stati membri³⁴¹. È l'UE che, per evitare spaccature nel mercato interno dell'Unione conseguente anche ad una potenziale frammentazione dal punto di vista normativo *ex ante*, detta la linea di regolamentazione per il digitale in tutta Europa: questa è la c.d. *normative digital sovereignty*, consistente nella capacità dell'UE di controllare quale impostazione di disciplina operi in questo delicato settore³⁴². Specificamente nel settore dell'IA e della sua regolamentazione, l'UE ha voluto scongiurare il più possibile la possibilità di frammentazioni del mercato interno che potevano essere causati da Stati membri i quali, coinvolti anch'essi nella *Race to AI regulation*, potessero emanare disposizioni mirate a regolare questo genere di tecnologie, causando in ultimo dei discostamenti *ex ante* o *ex post* rispetto la linea direttiva dell'Unione³⁴³.

Dunque, per tutti i motivi appena esposti, l'Unione Europea ha cercato di prendere parte il prima possibile alla corsa globale per la creazione di una cornice regolamentare per l'IA. La Proposta di Regolamento dell'*AI Act* della primavera del 2021 (la Versione I presentata dalla Commissione europea) è frutto di un lavoro intenso dell'Unione iniziato diversi anni prima. Questo percorso di ricerca e di studio ha avuto un ruolo centrale nel modellare l'approccio europeo all'IA e la forma, la *ratio*, la struttura e gli obiettivi del summenzionato *AI Act*.

2. *La strada per la Proposta di Regolamento (2017-2021); la scelta dello strumento legislativo da adottare e la base giuridica idonea*

Il punto di partenza dell'interesse dell'Unione Europea nei confronti dei sistemi di IA risale al 2017. Infatti, già in una delle sue riunioni di *routine* di quell'anno, il Consiglio Europeo espresse la necessità che si diffondesse “*un senso di urgenza per affrontare le tendenze emergenti come l'intelligenza artificiale.*”³⁴⁴

Successivamente, nell'aprile del 2018, la Commissione Europea ha definito una strategia per l'IA³⁴⁵ che affrontasse i suoi aspetti socioeconomici,

³⁴¹ Il concetto della sovranità del sistema europeo rispetto ai singoli ordinamenti nazionali (*inter alia*) è stato più volte ribadito dalla Corte di Giustizia dell'Unione Europea (CGUE) fin dalla nascita della CE, con la celebre sentenza Costa contro Enel del 1964. Vedi C 6/64, Sentenza della Corte di Giustizia dell'Unione Europea del 15 luglio 1964, *Flaminio Costa v. E.N.E.L.* ECLI:EU:C:1964:66.

³⁴² Smuha (2023), 8-9.

³⁴³ Schepisi (2023), 56. Vedi anche nota 365 sulle iniziative normative danesi e maltesi.

³⁴⁴ Consiglio Europeo, *Riunione del Consiglio Europeo* (19 ottobre 2017) – *Conclusioni EUCO* 14/17, 2017, 8.

³⁴⁵ Commissione Europea, ‘Artificial Intelligence for Europe’ (Comunicazione) COM (2018) 237 final.

parallelamente ad un aumento di investimenti nella ricerca, nell'innovazione, e nelle capacità di sviluppare IA di tutta l'Unione. La strategia, intitolata “*Intelligenza artificiale per l'Europa*,”³⁴⁶ per la prima volta, affermava un imperativo fondamentale da tener presente come monito per gli anni a venire: “[è necessario] mettere la potenza dell'IA al servizio del progresso umano”, mentre “i valori e i diritti fondamentali dell'Unione sono posti in primo piano”³⁴⁷.

Circa nello stesso periodo, 23 Stati membri dell'UE (più la Norvegia e il Regno Unito, da allora già annoverato come ex-membro dell'Unione a seguito della *Brexit*) hanno firmato una “*Dichiarazione di cooperazione sull'intelligenza artificiale (AI)*.”³⁴⁸ La Dichiarazione mirava a riunire i firmatari verso una strategia europea congiunta di ricerca, sviluppo e investimento (R&S&I) per i sistemi di IA. Il documento aveva il merito di essere il primo accordo di collaborazione tra gli Stati membri che avesse l'obiettivo di instaurare una cooperazione internazionale in questo settore.³⁴⁹ Tuttavia, la dichiarazione non dimostrava alcuna aspirazione dei paesi firmatari a disciplinare tali tecnologie con l'utilizzo di *hard law*³⁵⁰.

Seguendo questa prospettiva, il passo successivo compiuto dall'Unione Europea è stato quello di istituire il c.d. *High-Level Expert Group on Artificial Intelligence*³⁵¹ (Gruppo di esperti di alto livello sull'IA, di seguito abbreviato in “HLEGAI”). Il gruppo si è occupato, *inter alia*, di redigere le celebri “*Ethics Guidelines For Trustworthy AI*”³⁵² (Linee guida di etica per un'IA affidabile). Esse sono uno strumento di *soft law*³⁵³, redatte allo scopo di indicare il percorso da seguire per lo sviluppo di un'IA non solo lecita, ma anche etica e solida. Le stesse, inoltre, sono state il risultato di mesi di lavoro dell'HLEGAI, costituito da 52

³⁴⁶ *Id.*

³⁴⁷ *Id.*

³⁴⁸ *Declaration of Cooperation on Artificial Intelligence*, 4 (10 aprile 2018), <https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence#:~:text=Declaration%20signed%20at%20Digital%20Day%20on%2010th%20April%202018.&text=This%20Declaration%20builds%20on%20the.of%20a%20Digital%20Single%20Market>.

³⁴⁹ Moltzatu, A. (22 dicembre 2019) *EU Strategy on Artificial Intelligence In 2018: A Closer Look at the Strategy by the European Union*, Medium. <https://medium.com/towards-data-science/the-eu-strategy-on-artificial-intelligence-ab51ce5ce5bd>.

³⁵⁰ *Id.* La c.d. *hard law*, ossia l'insieme di strumenti di diritto tradizionale aventi forza di legge, viene solitamente definita in contrapposizione alla *soft law*. Quest'ultima consiste in strumenti *quasi*-legali senza forza giuridica sostanziale, come dichiarazioni, libri bianchi o linee guida redatte da enti pubblici, istituzioni, ma anche da privati. Vedi Druzin, B.H. (2016) *Why Does Soft Law Have Power Anyway*, 7 *Asian JIL*, 361-362. Vedi anche Floridi (2022), 130.

³⁵¹ *High-level expert group on artificial intelligence, Shaping Europe's Digital Future*. <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>

³⁵² *Ethics guidelines for trustworthy AI*. (8 aprile 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

³⁵³ Per la definizione di *soft law*, vedi nota 350.

esperti, di cui la maggioranza appartenenti all'*industry*³⁵⁴ (la cui componente accademica per l'Italia era rappresentata dal Prof. Luciano Floridi³⁵⁵). Le *guidelines* sono state prima redatte autonomamente dall'HLEGA; successivamente si è aperta una fase di consultazione pubblica aperta a cui hanno potuto contribuire oltre 350 organizzazioni, le quali hanno testato le linee guida nella pratica e hanno fornito più di 500 commenti³⁵⁶. Questo processo di coinvolgimento di esterni ha consentito un proficuo dialogo trasversale e assolutamente *multistakeholder* tra i partner industriali e il gruppo di lavoro creato dalle istituzioni, in un'utile collaborazione per la creazione di *best practices* per avere sistemi di IA migliori e che ispirassero maggiore fiducia nel loro utilizzo.

Lo sforzo europeo di avviare un discorso politico sulla regolamentazione dell'IA ha portato alla pubblicazione di un documento che ha definitivamente aperto la strada all'attuale proposta di legge dell'*AI Act*: il “*Libro Bianco del 2020 sull'intelligenza artificiale - Un approccio europeo all'eccellenza e alla fiducia.*”³⁵⁷ Il Libro Bianco emanato dalla Commissione UE ha, tra l'altro, portato avanti questo approccio inclusivo che ha connotato anche la redazione delle *Ethics Guidelines* precedentemente menzionate, coinvolgendo oltre 1250 “portatori di interessi” pubblici e privati, che hanno avuto l'opportunità di presentare le proprie osservazioni in materia, così come era accaduto per il *drafting* finale delle linee guida del 2019³⁵⁸.

Il Libro Bianco, inoltre, discute ampiamente le basi di un quadro normativo, ma, in *primis* descrive la posizione dell'UE nel mercato dell'IA nel periodo che precede il 2020³⁵⁹. Anche se l'UE non può essere considerata tradizionalmente una potenza produttrice di innovazione tecnologica nel campo dell'IA³⁶⁰ rispetto ad altri paesi come gli Stati Uniti, nel Libro Bianco viene presentata un'UE *leader* della produzione di IA utilizzata nel settore manifatturiero, così come quella impiegata nello sviluppo di applicazioni *software* per imprese e organizzazioni. Tuttavia, gli altri paesi come gli Stati Uniti ottengono risultati

³⁵⁴ Mantelero (2022), 47.

³⁵⁵ Floridi, L. (2021) The End of an Era: from Self-Regulation to Hard Law for the Digital Industry, *Philos. Technol.* 619-622, 621.

³⁵⁶ Relazione di accompagnamento, *AI Act* (Versione I), 9-10.

³⁵⁷ *Libro Bianco della Commissione Europea sull'Intelligenza Artificiale – Un approccio europeo all'eccellenza e alla fiducia*, 11 COM (2020) 65 final, 19 febbraio 2020 (d'ora in avanti “Libro Bianco sull'IA”).

³⁵⁸ Relazione di accompagnamento, *AI Act* (Versione I), 10. Vedi anche Libro Bianco sull'IA, 28-29.

³⁵⁹ Libro Bianco sull'IA, 3-4.

³⁶⁰ Vedi nota 323.

migliori nella creazione di piattaforme online e applicazioni per i consumatori, elementi che rappresentano una fonte consistente di dati aggregati da impiegare nell'allenamento dell'IA. Il Libro Bianco non fa mistero che l'UE, a differenza degli altri paesi *leader* del settore, non dispone di investimenti pubblici e privati sufficienti per lo sviluppo di tali tecnologie: ad esempio, già nel 2016, gli Stati Uniti avevano stanziato più di 12,1 miliardi di dollari per lo sviluppo dell'IA, mentre l'UE ne aveva allocati solo 3,2³⁶¹.

In secondo luogo, il Libro Bianco espone l'intenzione dell'Unione di cogliere l'opportunità di trarre profitto dalla “prossima ondata di dati”³⁶² (“*data wave*”). Dal momento che la quantità complessiva di dati presenti nell'Infosfera aumenterà in modo esponenziale³⁶³, è necessario sfruttare correttamente questo incremento, a favore dello sviluppo di una tecnologia “affamata” di dati come l'IA³⁶⁴.

Infine, un'altra ragione per definire un quadro normativo a livello europeo che il Libro Bianco indica si rinviene, come anticipato nel paragrafo precedente, nella necessità di armonizzare le leggi tra gli Stati membri. Infatti, già nel periodo antecedente alla redazione del Libro Bianco, gli Stati membri di Malta e Danimarca avevano già approvato (o erano in procinto di approvare) disegni di legge a livello nazionale specificamente indicati a disciplinare alcuni aspetti l'IA³⁶⁵. D'altra parte, anche la Germania nello stesso periodo aveva dichiarato di volersi dotare al più presto di un insieme di norme, basate sul rischio, che regolassero i sistemi di IA³⁶⁶. Per evitare di soffocare lo sviluppo dell'IA e per scongiurare una potenziale frammentazione del mercato interno, le istituzioni europee già nel 2020 avevano avvertito la necessità di armonizzare la disciplina per l'IA con un regolamento.

³⁶¹ Libro Bianco sull'IA, 4.

³⁶² *Id.*, 4-5.

³⁶³ Floridi (2022), 26. *Vedi anche* il sottoparagrafo 2.2.2., Capitolo 1.

³⁶⁴ Libro Bianco sull'IA, 4.

³⁶⁵ Lo Stato membro di Malta, già nel 2018, aveva introdotto un sistema di certificazione, ancorché volontaria, per i sistemi di IA. La Danimarca, invece, avviava in quegli anni un disegno di legge che fornisse un prototipo per la creazione di un “marchio” per l'IA che avesse un elevato grado di rispetto degli standard etici. Per la legislazione maltese, *vedi in generale Innovative Technology Arrangements and Services Act, Legizlazzjoni Malta* (1° novembre 2018), <https://legislation.mt/eli/cap/592/eng/pdf>. Per la legislazione danese *vedi in generale New seal for IT-security and responsible data use is in its way, Ministry of industry, business and financial affairs* (31 ottobre 2019), <https://eng.em.dk/news/2019/oktober/new-seal-for-it-security-and-responsible-data-use-is-in-its-way/>.

³⁶⁶ *Opinion of the data Ethics Commission (Data Ethik Kommission) – Executive Summary*, https://assets.contentstack.io/v3/assets/blt3de4d56151f717f2/blt300ce23c9789e0f3/5e5cfe13fa08326331360f93/191023_DEK_Kurzfassung_en_bf.pdf.

L'insieme di queste istanze ha portato dunque alla proposta di legge sull'IA del 2021, oggetto dell'analisi del prossimo paragrafo.

In conclusione, è da notare quindi il cambiamento di approccio al tema dell'IA che l'Unione ha prospettato, dalla pubblicazione delle *Ethics Guidelines* nel 2019 in poi. Quella strategia incentrata sull'adozione dei soli mezzi di *soft law* per implementare un utilizzo "etico" della tecnologia integralmente considerata è stata abbandonata³⁶⁷ in favore della creazione di una struttura di *governance risk-based* più articolata e soprattutto che scaturisce da fonti del diritto aventi forza di legge.

Il sostanziale abbandono del ricorso ai soli strumenti di *soft law* come le Linee Guida prodotte dall'HLEGAI, non si è limitato solo al settore istituzionale. Secondo alcuni, la *soft law* nel settore dell'economia digitale integralmente considerata, sebbene sia stata largamente impiegata negli ultimi decenni, con un (ancorchè eccessivo) proliferare di codici etici e linee guida provenienti da una grande molteplicità di entità diverse³⁶⁸, ha oramai dimostrato i suoi limiti³⁶⁹. Soprattutto per quanto riguarda il tema dell'IA, in cui si è osservato questo eccessivo proliferare di principi e linee guida in campo etico³⁷⁰, si è potuto parlare di due fenomeni che, in ultimo, hanno diminuito di gran lunga il valore, nonché la genuinità di queste iniziative di *self-regulation*: da un lato il c.d. "*shopping etico*,"³⁷¹ e dall'altro, il c.d. "*bluwashing*."³⁷² L'etica, in questo campo, è stata

³⁶⁷ Mantelero (2022), 166-167. Vedi anche Floridi (2021), 621-622.

³⁶⁸ Floridi (2022), 93.

³⁶⁹ Durante l'avvento di Internet e delle piattaforme social, è infatti stato fatto largo utilizzo di strumenti di auto-regolamentazione da parte delle aziende per imprimere un approccio più "etico" alla tecnologia. L'idea che fossero le stesse aziende a dotarsi di linee guida da seguire sembrava un utile *escamotage* per far sì che queste adottassero un approccio non solo rispettoso della legalità, ma anche dell'etica stessa. Essendo l'autoregolamentazione più snella, più semplice da implementare, da rispettare e finanche da aggiornare con il cambiamento nel tempo di questi strumenti, sembrò, all'epoca, lo strumento più tagliato per essere utilizzato nel settore del digitale. Tuttavia, negli ultimi tempi, soprattutto con lo sviluppo della tecnologia e degli strumenti di IA, si è compresa la sostanziale inefficacia di un utilizzo della sola autoregolamentazione a mezzo di strumenti di *soft law*. Ciò ha dunque determinato quell'inversione di rotta che ha portato a testi legislativi come il GDPR, il DSA/DMA, precedentemente menzionati. Essi sono infatti il frutto della convinzione che l'era della sola *self-regulation* per le aziende del settore digitale sia sostanzialmente terminata. Vedi Floridi (2021).

³⁷⁰ Floridi (2022), 92-94.

³⁷¹ Per shopping etico si intende "il malcostume di scegliere, adattare o rivedere "mescolare e abbinare" principi etici, linee guida, codici, quadri di riferimento o altri standard simili (specialmente ma non solo nell'etica dell'IA) estraendoli da una varietà di offerte disponibili, per conferire una patina nuova a dei comportamenti preesistenti (scelte, processi, strategie ecc.) e in tal modo giustificarli a posteriori, invece di implementare o affinare nuovi comportamenti confrontandoli con standard etici pubblici." Floridi (2022), 110. Vedi anche Floridi (2021), 621-622.

³⁷² Nell'etica ambientale, il *greenwashing* è una forma di sostanziale disinformazione, costituito da quel comportamento rappresentato da un'entità che intende apparire al pubblico più sostenibile e rispettoso dell'ambiente di quanto effettivamente non sia. Nel caso dell'etica, Floridi ha coniato il

largamente utilizzata come strategia per ritardare quanto più possibile gli interventi normativi, “*come se fosse un’alternativa al diritto in nome dell’innovazione tecnologica e del suo impatto positivo sulla crescita economica.*”³⁷³

Per cui, il futuro della regolamentazione dell’IA risiede nell’impiego dell’*hard law*. Ciò non significa, tuttavia, che la *soft law* non possa comunque essere una risorsa per entità sia pubbliche che private, contestualmente affiancata ad altre iniziative che tendano a implementare comportamenti e pratiche in linea con l’etica, ma avendo sempre una solida cornice legislativa di *AI governance*³⁷⁴ di sottofondo.

A corroborare la necessità di uno strumento legislativo orizzontale di *hard law*, la Commissione ha condotto una valutazione d’impatto in relazione alla Proposta di regolamento stessa, prima della sua presentazione, conclusasi a marzo 2021 (un mese prima della presentazione dell’*AI Act*)³⁷⁵. Nella valutazione di impatto, venivano presentate quattro opzioni strategiche di regolamentazione, dotate di diversi gradi di intervento ed incisività normativa: da uno strumento legislativo dell’UE che istituiva un sistema di etichettatura volontario dell’IA per le aziende (l’ “opzione 1”), a uno strumento legislativo orizzontale dell’UE che stabilisse dei requisiti obbligatori per tutti i produttori di sistemi di IA, indipendentemente dal rischio che questi sono in grado di porre nei confronti degli individui (l’ “opzione 4”)³⁷⁶. L’opzione scelta dalla Commissione, a fronte della valutazione d’impatto, è stata quella “3+”, cioè quella che prevedesse l’adozione di uno strumento legislativo orizzontale basato su un approccio *risk-based* con dei requisiti più stringenti per le tecnologie considerate ad alto rischio, e la contestuale previsione di codici di condotta volontari per i sistemi meno problematici³⁷⁷. La scelta di questa opzione strategica rispetto alle altre è stata dunque il risultato di un bilanciamento tra vari interessi in gioco, a partire dai rischi di violazione dei diritti fondamentali, cercando di mantenere sotto una determinata soglia i costi di *compliance*³⁷⁸. Questa scelta risulta quella maggiormente di favore

neologismo *bluwashing* per indicare “*il malcostume di fare affermazioni infondate o fuorvianti al riguardo (o di attuare misure superficiali a favore) dei valori etici e dei benefici dei processi, prodotti, servizi, o altre soluzioni digitali al fine di apparire più etici dal punto di vista digitale di quanto non sia effettivamente.*” Floridi (2022), 112.

³⁷³ *Id.*, 114.

³⁷⁴ *Id.*, 127-130.

³⁷⁵ Relazione di accompagnamento, *AI Act* (Versione I), 10-11.

³⁷⁶ Relazione di accompagnamento, *AI Act* (Versione I), 10.

³⁷⁷ *Id.*

³⁷⁸ *Id.*, 11. Vedi anche Benbouzid *et al.*, (2022), 25.

per l'*industry*, che avrebbe dunque sofferto di un approccio eccessivamente rigido di previsione di standard identici per qualsivoglia sistema di IA³⁷⁹.

Una volta stabilito che l'impiego di uno strumento normativo orizzontale che armonizzasse e a livello europeo della disciplina dell'IA fosse lo strumento più adeguato agli obiettivi perseguiti, l'Unione si è trovata ad affrontare altre questioni preliminari: da un lato, la corretta scelta di quale strumento normativo orizzontale da impiegare, e in secondo luogo la base giuridica su cui lo stesso si dovesse fondare³⁸⁰.

Per quanto riguarda la prima scelta, cioè l'adozione di un regolamento piuttosto che altri strumenti di diritto derivato dell'UE, quale, ad esempio, una (o una serie) di direttive, la Commissione ha scelto di prediligere uno strumento che assicurasse l'applicazione non solo uniforme all'interno dei vari Stati membri (evitando quindi frammentazioni su temi fondamentali quali la definizione di IA o l'individuazione delle pratiche dannose), ma anche diretta, ai sensi di quanto disposto dall'art. 288 del Trattato sul Funzionamento dell'Unione Europea³⁸¹. Il regolamento così strutturato è mirato a sviluppare un mercato unico per i sistemi di IA, garantendo contemporaneamente la tutela dei diritti fondamentali ai residenti dell'UE e la certezza del diritto per operatori e consumatori³⁸². Questa scelta risulta ad ogni modo in conformità con la tendenza europea di legiferare nel settore del digitale tramite l'adozione di diritto derivato quanto più omnicomprensivo possibile³⁸³.

L'*AI Act*, tuttavia, mantiene un'ampia discrezionalità in una grande varietà di temi la cui definizione è, di fatto, sostanzialmente rimessa agli Stati Membri, finendo altresì per somigliare ad una direttiva³⁸⁴. Ad esempio, l'*AI Act*, nel disporre

³⁷⁹ Mantelero (2022), 141-142.

³⁸⁰ Infatti, la proposta di regolamento sull'IA, come ogni altra fonte di diritto derivato dell'UE, necessita la corretta individuazione della "base giuridica", intesa come la disposizione dei Trattati dell'UE che attribuiscono alle istituzioni il potere di adottarli. Il passaggio di individuazione della corretta base giuridica costituisce non solo un elemento di legittimazione ancorchè costituzionale, ma anche un importante passaggio che chiarisce elementi sostanziali (come possono essere gli obiettivi perseguiti dalla norma stessa) quando formali (come, ad esempio, la corretta procedura di formazione del consenso da utilizzare per l'adozione di un certo atto). Strozzi, G., & Mastroianni, R., (2019), *Diritto dell'Unione Europea: parte Istituzionale*, ottava edizione, G. Giappichelli Editore, 282. In merito alla rilevanza "costituzionale" della base giuridica adottata per un certo atto di diritto derivato il Parere della CGUE sul protocollo di Cartagena, *vedi anche* CGUE, Parere n. 2/00 del 6 dicembre 2001.

³⁸¹ Versione consolidata del Trattato sul funzionamento dell'Unione europea, art. 288, 9 maggio 2008. 288, 9 maggio 2008, 2008 G.U. (C 115) (d'ora in avanti abbreviato come "TFUE").

³⁸² Relazione di accompagnamento, *AI Act* (Versione I), 7-8.

³⁸³ *Vedi in generale* il paragrafo 1.

³⁸⁴ Schepisi (2023), 71-73.

un modello di *AI governance* europeo³⁸⁵, prevede l'istituzione o la designazione di una autorità nazionale competente nell'ambito dell'IA³⁸⁶. Tuttavia, viene lasciato agli Stati membri un potere discrezionale nel decidere se designare un'autorità di nuova costituzione, ovvero attribuire le nuove competenze necessarie ad una autorità preesistente che sia dotata dei requisiti previsti dallo stesso regolamento³⁸⁷. L'*AI Act*, a questo proposito, si limita ad indicare che l'autorità di vigilanza, scelta dai singoli Stati Membri, dovrà essere dotata di requisiti necessari di indipendenza, autonomia e professionalità, sottolineando l'importanza che vi sia una vigilanza attenta in questo settore³⁸⁸, ma lasciando dunque gli Stati membri la libertà di definizione delle modalità di costituzione e funzionamento dell'autorità³⁸⁹. Questo è solo uno dei possibili esempi di quei tratti del regolamento che lascia una discrezionalità simile ad una direttiva agli Stati membri.

Per quanto riguarda invece la base giuridica, diverse ipotesi sono state vagliate dalla Commissione prima di adottare quella relativa all'art. 114 TFUE³⁹⁰. Ad esempio, è stata valutata la possibilità di utilizzare a questo scopo l'art. 19 del TFUE³⁹¹. L'articolo in questione prevede che il Consiglio Europeo, deliberando secondo una procedura legislativa speciale, possa adottare i provvedimenti necessari per combattere le discriminazioni a livello comunitario³⁹². Ed è proprio la procedura legislativa speciale prevista da questo articolo che non consente l'adozione di un regolamento come l'*AI Act*, i cui propositi si spingono considerevolmente oltre la semplice creazione di un IA "non discriminatoria." Un'altra ipotesi che è stata vagliata inizialmente, sempre in linea con il principio di non discriminazione, è stata la possibilità di utilizzare l'art. 2 del Trattato sull'Unione Europea (TUE)³⁹³ come base giuridica del regolamento³⁹⁴. In questo caso, l'art. 2 TUE, che espone il principio generale di non discriminazione vigente

³⁸⁵ Titolo IV, *AI Act*.

³⁸⁶ Vedi Art. 59, *AI Act* (Versione III). Vedi anche considerando 77, *AI Act* (Versione III). Vedi anche Schepisi (2023), 71. Vedi anche il paragrafo 3.3.4.

³⁸⁷ Schepisi (2023), 72-73.

³⁸⁸ Vedi Art. 59, *AI Act* (Versione III)

³⁸⁹ Schepisi (2023), 72-73.

³⁹⁰ L'art. 114 del TFEU prevede che il Parlamento europeo e il Consiglio possano adottare regolamenti relativi al funzionamento del mercato interno. Art. 114 TFUE.

³⁹¹ Art. 19 TFUE. Vedi Schepisi (2023), 71.

³⁹² Art. 19 TFUE.

³⁹³ Versione consolidata del Trattato sull'Unione europea, art. 2, GU C 202 del 7.6.2016 (d'ora in avanti abbreviato come "TUE").

³⁹⁴ Schepisi (2023), 71.

nell'intero ordinamento europeo, non consente tuttavia l'adozione di fonti di diritto derivato da parte delle istituzioni³⁹⁵.

Tuttavia, dal momento che l'obiettivo centrale dell' *AI Act* sembra essere la riduzione al minimo delle possibilità di frammentazione del mercato unico digitale all'UE, connesso alla necessità di certezza del diritto in questo settore, è stato scelto a questo proposito l'art. 114 TUE, che prevede l'adozione di misure destinate ad assicurare il funzionamento di quest'ultimo³⁹⁶: esso è sembrato l'unico fondamento normativo che legittimasse l'intera proposta di regolamento, sottolineando con forza la natura protettiva e di tutela del mercato interno dell'Unione, che rimane l'obiettivo principale perseguito³⁹⁷, come verrà analizzato nel prossimo paragrafo.

Peraltro, la base giuridica è stata anche "arricchita" di un altro riferimento ad un altro articolo, l'art. 16 del TFUE³⁹⁸, che consente alle istituzioni di adottare atti che predispongono tutele per i dati personali dei residenti dell'Unione. L'art. 16 TFUE funge da fondamento normativo a quelle previsioni dell'*AI Act* che individuano delle restrizioni sull'utilizzo di sistemi di IA per l'identificazione biometrica remota in "in tempo reale" in spazi pubblici ai fini di attività di sicurezza³⁹⁹.

3. *L'AI Act (2021)*

3.1 *Il percorso normativo dell'AI Act*

La Proposta di Regolamento Europeo è stata, come più volte rimarcato, presentata dalla Commissione Europea il 21 aprile del 2021⁴⁰⁰. La Commissione, secondo le attribuzioni previste dall'art. 17 del TUE⁴⁰¹ così come modificato in ultimo dal Trattato di Lisbona, è infatti titolare del potere di iniziativa legislativa, dal momento che essa sola provvedere a promuovere l'interesse generale dell'Unione e, nel farlo, gode di totale indipendenza dalle altre istituzioni nonché

³⁹⁵ *Id.*

³⁹⁶ Relazione di accompagnamento, *AI Act* (Versione I), 6-7.

³⁹⁷ *Id.*, 3.

³⁹⁸ Art. 16 TFUE.

³⁹⁹ Relazione di accompagnamento, *AI Act* (Versione I), 6-7. *Vedi* il paragrafo 3.3.2 per maggiori specifiche sui sistemi di identificazione biometrica in tempo reale.

⁴⁰⁰ *Timeline - Artificial intelligence*. (n.d.). European Council - Council of the European Union. <https://www.consilium.europa.eu/en/policies/artificial-intelligence/timeline-artificial-intelligence/>

⁴⁰¹ Art. 17 TUE.

dagli Stati membri in forza dell'art. 17 del TUE terzo comma⁴⁰². L'iniziativa della Commissione ha pertanto dato il via alla procedura legislativa ordinaria con la quale l'atto deve essere adottato, seguendo quanto previsto dall'art. 294 TFUE⁴⁰³, procedura che vede Parlamento Europeo e Consiglio interagire tra loro come co-legislatori.

A causa della portata del regolamento stesso (nella sua prima versione costituito da 89 considerando, 85 articoli, nonché da 7 allegati dal contenuto di natura sostanzialmente tecnica ma non per questo meno cruciali per la definizione dell'applicazione della normativa⁴⁰⁴), sia il Parlamento che il Consiglio hanno necessitato di un lungo periodo di studio della normativa prima di poter adottare le loro rispettive posizioni, ai sensi dell'art. 294 del TFUE, commi 3-6 (prima lettura).

Dopo la finalizzazione delle posizioni di Consiglio e Parlamento, che ha portato la successiva creazione di tre versioni del testo dell'*AI Act* (riportate per lo scopo di quest'analisi come, Versione I della Commissione, Versione II del Parlamento e Versione III del Consiglio), si è aperta la fase interlocutoria tra le istituzioni europee dei c.d. "*triloghi*"⁴⁰⁵, cioè quei negoziati interistituzionali che mirano alla ricerca di un accordo prettamente politico su alcune previsioni di rilievo dell'*AI Act* (dall'ampiezza della definizione di IA, ai tipi di sistemi considerati "ad alto rischio", o ancora ai requisiti necessari da individuare per regolamentare i modelli più complessi di IA, come i LLMs⁴⁰⁶). A seguito di intensi mesi di negoziazioni e del susseguirsi di numerosi triloghi durante tutto il 2023⁴⁰⁷, le istituzioni hanno raggiunto un accordo politico nel dicembre 2023, dopo una negoziazione record di 36 ore⁴⁰⁸, mettendo sostanzialmente un punto su una serie

⁴⁰² Strozzi & Mastroianni (2019), 187-189.

⁴⁰³ Art. 294 TFUE.

⁴⁰⁴ Si pensi ad esempio alla portata dell'Allegato III, in cui la Commissione ha previsto di inserire l'elenco dettagliato di tutte quelle implicazioni dell'IA che consentono di classificare un sistema "ad alto rischio." Nonostante la classificazione basata sul rischio sia prevista nello stesso testo dell'*AI Act* (Titolo III), è stato predisposto dalla Commissione un meccanismo per mantenere il regolamento aggiornato rispetto agli usi e alle applicazioni emergenti dell'IA, che consentisse di ampliare alla stessa l'elenco dei requisiti previsti nell'Allegato III per ricomprendere nella nozione di sistema "ad alto rischio" utilizzati in settori predefiniti, quelli che tuttavia non rientrerebbero inizialmente nell'applicabilità dell'*AI Act*. Vedi Relazione di accompagnamento, *AI Act* (Versione I), 14-15.

⁴⁰⁵ Come anticipato nel paragrafo 2.3 del Capitolo 1.

⁴⁰⁶ Per una definizione dei Large Language Models (LLMs), vedi nota 92.

⁴⁰⁷ Bertuzzi, L., (14 dicembre 2023b). European Union squares the circle on the world's first AI rulebook. [www.euractiv.com](https://www.euractiv.com/section/artificial-intelligence/news/european-union-squares-the-circle-on-the-worlds-first-ai-rulebook/). <https://www.euractiv.com/section/artificial-intelligence/news/european-union-squares-the-circle-on-the-worlds-first-ai-rulebook/>.

⁴⁰⁸ *Id.*

di questioni centrali che caratterizzeranno nel futuro l'*AI Act*⁴⁰⁹. Con l'accordo politico dello scorso dicembre, è stato dato inizio ad un procedimento di messa a punto della formulazione letterale delle singole previsioni e dei considerando, che precederà la definitiva adozione del testo finale da parte delle istituzioni, auspicabilmente durante la prima metà del 2024⁴¹⁰.

Nonostante lo stadio assolutamente avanzato della procedura legislativa e il fatto che il contenuto dell'*AI Act* sia oggetto di un accordo prettamente politico di fondo, i reali risultati dei precedenti mesi di intense negoziazioni saranno solo evidenti una volta che il testo finale dell'*AI Act* da sottoporre al vaglio finale delle istituzioni per la sua adozione sarà pubblicato. Pertanto, per lo scopo di questa analisi ci si è limitati ad analizzare la Versione I dell'*AI Act*, inserendo nei paragrafi successivi di descrizione meramente contenutistica della normativa, i tratti salienti che sono stati introdotti o modificati con le successive versioni II e III, cercando comunque di mantenere una visione quanto più unitaria possibile dell'iter legislativo proposto di regolamento, ancora in corso⁴¹¹. In ultimo, nel gennaio 2024 è stata resa nota una Versione IV dell'*AI Act*, oggetto di una divulgazione non ufficiale⁴¹² da parte della stampa, che verrà utilizzata per osservare *in itinere* il processo di messa a punto della formulazione letterale delle singole disposizioni, che può aver inciso in maniera determinante (talvolta negativamente, come si vedrà) anche su alcuni aspetti considerati “stabilizzati” dall'accordo politico raggiunto nel dicembre 2023.

Per quanto riguarda la sua effettiva entrata in vigore, l'*AI Act*, prevedrà un periodo di intermedio di *vacatio* di circa due anni per consentire agli sviluppatori, ai produttori e all'*industry* dell'IA in generale di predisporre le adeguate cautele per assicurare la *compliance* con il regolamento⁴¹³; allo stesso tempo, gli Stati membri

⁴⁰⁹ *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI* | News | European Parliament. (9 dicembre 2023). <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>. Vedi anche *Press release - Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*. (9 dicembre 2023). European Council - Council of the European Union. (9 dicembre 2023) <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.

⁴¹⁰ *Id.*

⁴¹¹ Vedi Bertuzzi, L. (22 gennaio 2024). *LEAK: Given the massive public attention to the #AIAct, I've taken the rather unprecedented decision to publish the final text. The agreed text is on the right-hand side for those unfamiliar with four-column documents. Enjoy the reading! Some context: 1/6*. Twitter. <https://twitter.com/BertuzLuca/status/1749326217612820558>

⁴¹² Vedi nota precedente.

⁴¹³ *Id.*

avranno il tempo di formulare i propri necessari adeguamenti in quelle materie in cui l'UE ha lasciato un potere discrezionale nazionale più o meno ampio⁴¹⁴. Tuttavia, dall'ultimo accordo politico raggiunto a dicembre 2023, alcune disposizioni del regolamento potrebbero iniziare a produrre i loro effetti ben prima del termine di due anni stabilito, come ad esempio il divieto di utilizzo nell'Unione di quei sistemi di IA che pongono un rischio c.d. "inaccettabile."⁴¹⁵

Per quanto concerne la prospettiva di anticipare in maniera sostanziale gli impatti dell'*AI Act*, soprattutto per quanto riguarda la conformità dei sviluppatori di sistemi di IA, le istituzioni dell'Unione Europea hanno introdotto l'iniziativa nota come *AI Pact*: esso consiste in un accordo ad adesione assolutamente volontaria che consente alle aziende che vi aderiscono di stimolare una c.d. *voluntary anticipated compliance* (conformità anticipata e volontaria) ai requisiti previsti dall'*AI Act* prima che questi entrino in vigore, trascorsi due anni dal voto finale sul testo della proposta di regolamento⁴¹⁶.

3.2 *Gli obiettivi perseguiti dalla proposta di regolamento: quelli espliciti, il "first mover advantage," e la ricerca di un "effetto Bruxelles"*

A fronte di quanto previsto dal Libro Bianco pubblicato nel 2020, la cui importanza per il processo che ha portato all'*AI Act* è altresì evidente, la proposta di regolamento si propone di realizzare un "*ecosistema di fiducia*"⁴¹⁷ attraverso l'introduzione di un quadro giuridico volto ad assicurare l'affidabilità dei sistemi di IA⁴¹⁸. L'Unione Europea ha cercato di promuovere questa fiducia non solo tra gli individui, cioè gli utenti singoli che adottano soluzioni basate sull'IA, ma anche tra le imprese che si impegnano nello sviluppo di tali sistemi⁴¹⁹. La proposta, scaturita dalle richieste delle altre istituzioni, è in linea sia con quanto avanzato dal Consiglio Europeo⁴²⁰, cioè la richiesta di rendere l'Unione *leader* nello sviluppo di un'IA

⁴¹⁴ Come, ad esempio, la determinazione delle autorità nazionali preposte alla vigilanza sul rispetto dell'*AI Act*. Vedi *AI Act* (Versione III). Vedi anche considerando 77, *AI Act* (Versione III). Vedi anche Schepisi (2023), 71.

⁴¹⁵ Mukherjee, S., Coulter, M., & Chee, F. Y. (14 dicembre 2023) Explainer: What's next for the EU AI Act? *Reuters*. <https://www.reuters.com/technology/whats-next-eu-ai-act-2023-12-14/>.

⁴¹⁶ Coraggio, G. (9 Gennaio 2024). *The Approval of the EU AI Act and What to expect Next - with Mr. Brando Benifei*. Min. 18.26 – 20.00

⁴¹⁷ Si guardi il titolo del Libro Bianco "un approccio all'eccellenza e alla fiducia". Vedi in generale, Libro Bianco sull'IA.

⁴¹⁸ Relazione di accompagnamento, *AI Act* (Versione I), 1.

⁴¹⁹ *Id.*

⁴²⁰ Vedi Consiglio Europeo, Riunione Straordinaria del Consiglio Europeo (1° e 2 ottobre 2020) – Conclusioni, EUCO 13/20, 2020, 7. Vedi anche Relazione di accompagnamento, *AI Act* (Versione I), 2.

sicura, affidabile ed etica, ma da quanto richiesto dal Parlamento Europeo, cioè che vi siano previste garanzie per la tutela dei principi etici.⁴²¹

Questa scelta si riflette direttamente sugli obiettivi della normativa, richiamati nella Relazione di accompagnamento all'*AI Act*, che precede la prima versione proposta dalla Commissione dell'Unione Europea nell'aprile del 2021⁴²². Gli obiettivi "specifici" della regolamentazione così come presentati dalla Commissione sono i seguenti: in *primis*, assicurare che i sistemi immessi sul mercato dell'UE siano sicuri e rispettino i diritti fondamentali e i valori fondanti dell'Unione⁴²³; in secondo luogo, assicurare la certezza del diritto per stimolare investimenti ed innovazione; in terzo luogo, migliorare la *governance* e l'applicazione della normativa già esistente in materia di diritti fondamentali e requisiti di sicurezza applicabili ai sistemi di IA⁴²⁴; e, in ultimo ma non certamente di importanza data la centralità di questo obiettivo rispetto la base giuridica del regolamento⁴²⁵, facilitare lo sviluppo di un mercato unico per applicazioni di IA lecite, sicure ed affidabili, scongiurando il più possibile la frammentazione del mercato comunitario.

Sebbene questi siano gli scopi esplicitamente riportati dalla Relazione di accompagnamento all'*AI Act*, è possibile fornire alcune precisazioni sugli obiettivi scelti dall'UE nella scelta di regolamentare l'IA, soprattutto per chiarificare quegli intenti che la Commissione non ha esplicitato nella Relazione preliminare alla proposta di regolamento, ma che tuttavia emergono da un siffatto approccio a questo tipo di tecnologie. Questi scopi sono da considerarsi in parallelo a quelle considerazioni più generali relative all'approccio tutto europeo alla regolamentazione del digitale complessivamente considerato (cioè la necessità di assumere un modello *rights-driven*, insieme a quella di raggiungere una sorta di sovranità normativa digitale⁴²⁶).

⁴²¹ Risoluzione del Parlamento Europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate (2020/2021(INL)). *Vedi anche* Relazione di accompagnamento, *AI Act* (Versione I), 2.

⁴²² Relazione di accompagnamento, *AI Act* (Versione I), 3.

⁴²³ *Id.*

⁴²⁴ *Id.*

⁴²⁵ Cioè l'art. 114 TFUE. *Vedi* paragrafo precedente.

⁴²⁶ *Vedi* paragrafo 1.

L'Unione Europea, infatti, ha puntato a sfruttare il c.d. *first mover advantage*⁴²⁷ (traducibile come “il vantaggio del pioniere”) trasposto nel campo della regolamentazione anziché in ambito unicamente macroeconomico. Infatti, il paese o l'entità sovranazionale che per prima adotti una normativa sull'IA, sarebbe in grado di beneficiare del vantaggio ad aver operato per prima rispetto agli altri. A questo proposito, la Prof.essa Smuha⁴²⁸ propone un esempio efficace per spiegare lo sfruttamento di un siffatto effetto normativo ed economico. Se un dato “Paese A” introduce nuove normative per i sistemi di IA, le imprese locali si adegueranno per necessità a questi requisiti. Ciò potrebbe influenzare anche le imprese straniere basate in un dato “Paese B” che desiderano operare nel mercato del Paese A, portandole a conformarsi alle stesse regole di quest'ultimo. Tuttavia, le imprese del Paese B affronteranno dei costi per garantire la *compliance* ai requisiti che il Paese A impone, spingendole, in ultimo, a fare pressioni sul proprio governo, quello del Paese B, per chiedere l'adozione di requisiti simili. In definitiva, entrambi i paesi potrebbero finire per adottare regole sostanzialmente simili: tuttavia, il Paese B sarà un seguace anziché un innovatore nel settore, e le sue imprese dovranno comunque recuperare lo scarto temporale a livello di conformità alle leggi rispetto a quelle del Paese A, che avranno già assorbito i costi rilevanti.⁴²⁹

È dunque evidente che l'UE abbia preso in considerazione proprio questo vantaggio di stabilire “per primi” un precedente nell'ambito della legislazione omnicomprensiva e basata sul rischio dell'IA, approfittando dunque dei benefici di un siffatto effetto di pressione sostanziale sugli altri ordinamenti⁴³⁰.

Tuttavia, il legislatore europeo ha altresì voluto sfruttare un altro dato, non tanto di metodo come può essere considerato il *first mover advantage*, ma piuttosto di merito. Infatti, la predisposizione di regole chiare che mirino il più possibile ad

⁴²⁷ Il concetto puramente economico di “vantaggio del pioniere” sta nell'indicare quel vantaggio competitivo che le aziende possono guadagnare nel mercato quando risultano essere i primi ad ottenere il controllo delle risorse o a lanciare un nuovo prodotto o servizio. Smuha (2021b), 17-18. *Vedi anche* Benbouzid *et al.* (2022), 25. *Vedi anche* Bradford (2023), 349.

⁴²⁸ La Prof.essa Nathalie Smuha è studiosa di diritto e filosofia presso l'Università KU Leuven, esperta nelle implicazioni legali dell'IA e tecnologie digitali, ha coordinato l'High-Level Expert Group (HLEGAI) sull'IA presso la Commissione Europea. Con esperienze accademiche e professionali a livello internazionale, la sua ricerca si concentra sull'impatto dell'IA su diritti umani, democrazia e stato di diritto. Nathalie Smuha | legal scholar and philosopher. (n.d.). *Nathaliesmuha*. <https://www.nathaliesmuha.com/>

⁴²⁹ Smuha (2021b), 18.

⁴³⁰ A riprova di ciò, riverberano i titoli di giornale trionfanti sulla “Prima regolamentazione omnicomprensiva dell'IA,” a seguito del raggiungimento del più recente accordo prettamente “politico” sulla definizione della linea normativa da adottare nell'ambito di alcuni temi più controversi, all'inizio dicembre 2023. *Vedi* paragrafo 3.1.

un utilizzo di questi sistemi in linea con i diritti fondamentali funge come fattore di incremento della fiducia dei consumatori, sia a livello B2B che B2C⁴³¹. La fiducia all'interno di un dato settore del mercato, di per sé, ha un valore sul piano economico non indifferente: secondo uno studio del *Capgemini Research Institute*, i consumatori hanno un'attitudine più positiva nei confronti di aziende che implementano un uso dichiaratamente "etico" dell'IA, rispetto alle altre⁴³². Per questo motivo, l'UE ha sviluppato una cornice normativa incentrata sul rispetto dei diritti fondamentali, anche per sfruttare coscientemente un vantaggio competitivo che lo sviluppo dell'IA in linea con l'etica può portare a livello economico⁴³³. Un approccio di questo tipo ha dunque rispettato quell'intento dell'UE di creare quell'ecosistema di fiducia richiamato dal Libro Bianco del 2020⁴³⁴. La stessa "fiducia" a livello collettivo sia a livello dei singoli utenti, ma anche a livello di istituzioni ed aziende private, è strettamente riconnessa alla necessaria salvaguardia dei diritti fondamentali stabiliti dal diritto dei Trattati, annoverata, come anticipato, negli obiettivi specifici della proposta di legge⁴³⁵.

Ciò, tuttavia, tenendo conto un dato: sebbene nella valutazione di impatto effettuata *ex ante* dalla Commissione per la scelta dello strumento normativo più adeguato da adottare vi fosse anche l'opzione di uno strumento normativo orizzontale che non avesse un approccio modulato sul rischio, e che tutelasse a tutto tondo gli individui tramite la previsione di requisiti obbligatori per tutti i sistemi⁴³⁶, la scelta di un'opzione simile è risultata eccessivamente stringente nonostante la maggiore protezione che avrebbe potuto assicurare. Infatti, secondo la valutazione di impatto, uno strumento orizzontale simile non avrebbe portato ad un equo bilanciamento tra gli interessi di tutela e la necessità di non costituire un ostacolo all'innovazione, a causa di un'eccessiva iper-regolamentazione.

Questa valutazione di convenienza tra le opzioni strategiche a disposizione della Commissione non ha tuttavia relegato a secondo posto l'intento di salvaguardia dei diritti fondamentali, tutt'altro. Dall'*AI Act*, al contrario, emerge che la regolamentazione avviene "attraverso il prisma dei diritti fondamentali e non

⁴³¹ Smuha (2021b), 16.

⁴³² Capgemini Research Institute, Why addressing ethical questions in AI will benefit organisations (2019) https://www.capgemini.com/wp-content/uploads/2021/02/AI-in-Ethics_Web-1.pdf.

⁴³³ Smuha (2021b),

⁴³⁴ *Vedi in generale*, Libro Bianco sull'IA.

⁴³⁵ Relazione di accompagnamento, *AI Act* (Versione I), 3.

⁴³⁶ *Id.*, 10. *Vedi anche* paragrafo 3.1.

*nel semplice rispetto degli stessi.*⁴³⁷ Infatti, l’approccio basato sul rischio adottato dall’*AI Act* impone divieti o obblighi tecnico-procedimentali ai vari soggetti, modulati con livelli di tutela differenziati in base ai settori e agli standard considerati. Dunque, una c.d. “*procedimentalizzazione dei rischi*”⁴³⁸ posti dai sistemi di IA consente l’incorporazione *by design* della salvaguardia dei diritti all’interno dei sistemi stessi⁴³⁹.

Nell’imprimere questo intento di salvaguardia all’*AI Act*, il legislatore europeo ha dovuto effettuare una valutazione *ex ante* di quali diritti sono irrinunciabili e, pertanto, non possono cedere di fronte alla invasività di talune applicazioni dell’IA⁴⁴⁰. Per esempio, questo bilanciamento *ex ante* tra diritti si è reso necessario per quanto riguarda l’utilizzo di sistemi di riconoscimento biometrico in tempo reale da dispiegare in luoghi pubblici, utilizzo che non è considerato consentito dalla normativa (salvo eccezioni, come si vedrà in seguito⁴⁴¹), dal momento che può condurre a violazioni del diritto alla privacy e alla protezione dei dati personali⁴⁴², nonché ad una possibile deriva verso alla sorveglianza di massa.

In ultimo, tra i vari scopi perseguiti con l’*AI Act*, occorre menzionare la ricerca, da parte della Commissione, di provocare nuovamente un c.d. “*effetto Bruxelles*.”⁴⁴³ La capacità dell’Unione Europea di operare come *rule-setter* a livello globale per quanto riguarda la regolamentazione del digitale è emersa già con quelle normative come il GDPR, che, a suo tempo, è stata in grado di creare una sorta di “*gold standard*” a livello globale⁴⁴⁴. Avendo quindi preso in considerazione la capacità europea di influenzare a partire dalle proprie iniziative normative, sia i mercati che i governi di paesi extra-UE, è comprensibile che l’Unione voglia rafforzare la propria posizione di *leadership* proprio rivestendo il ruolo di regolatore a livello internazionale, sempre sfruttando quel vantaggio del *first mover* che è stato appena presentato. La concreta possibilità del verificarsi di un effetto simile verrà

⁴³⁷ Schepisi (2023), 61.

⁴³⁸ *Id.*

⁴³⁹ *Id.*

⁴⁴⁰ *Id.*, 63.

⁴⁴¹ Per tutto ciò che riguarda l’utilizzo dei sistemi di identificazione biometrica in tempo reale in luoghi a cui ha accesso il pubblico, *vedi* il paragrafo 3.3.2

⁴⁴² Considerati diritti fondamentali secondo gli artt. 7-8 della Carta di Nizza. Artt. 7-8 Carta dei diritti fondamentali dell’Unione Europea.

⁴⁴³ Per la definizione di effetto Bruxelles, *vedi* nota 319.

⁴⁴⁴ *Id.*

affrontata nel paragrafo 3.5, di conclusione rispetto alla descrizione e all'analisi dell'*AI Act*.

3.3 *Il contenuto della proposta di regolamento*

Una premessa è necessaria prima di intraprendere l'analisi del contenuto effettivo dell'*AI Act*. Come premesso nel paragrafo precedente, sebbene il contenuto generale della normativa sia stato oggetto di un lungo percorso legislativo, le singole disposizioni e la loro precisa formulazione sono al momento oggetto di un processo di perfezionamento tecnico che porterà a quella versione finale del testo, oggetto del voto da parte dei colegislatori europei, auspicabilmente entro la prima metà del 2024. Intanto, è possibile analizzare la prima versione dell'*AI Act* (preparato della Commissione Europea nel 2021), prendendo in considerazione gli accordi politici e gli aggiornamenti successivi di valore più significativo che ne hanno modificato il contenuto, cercando di ricostruire la forma che il regolamento avrà una volta presentato nella sua versione finale di fronte alle istituzioni europee per il voto conclusivo.

3.3.1 *L'ambito di applicazione e definizioni rilevanti*

Innanzitutto, il Titolo I dell'*AI Act* definisce l'oggetto del regolamento e l'ambito di applicazione, nonché presenta una serie di definizioni di concetti utili in tutto il testo normativo (tra cui quello di Intelligenza Artificiale, come presentata nel Capitolo 1, paragrafo 2.3)⁴⁴⁵ L'oggetto generale del regolamento, previsto all'art. 1 è individuato nel predisporre regole armonizzate per mettere sul mercato dell'UE i sistemi di IA, proibendo certe pratiche di utilizzo dell'IA e predisponendo una serie di requisiti necessari per i sistemi considerati "ad alto rischio," nonché regole per un monitoraggio del mercato e di *governance* generale dell'IA⁴⁴⁶.

L'ambito territoriale di applicazione della proposta di regolamento presenta una peculiarità: infatti, l'*AI Act* si applica "*ai fornitori che immettono sul mercato o mettono in servizio sistemi di IA nell'Unione, indipendentemente dal fatto che siano stabiliti nell'Unione o in un paese terzo,*"⁴⁴⁷ nonché ai fornitori situati sul territorio degli Stati Membri⁴⁴⁸. La necessità di prevedere un siffatto

⁴⁴⁵ Art. 2(1), *AI Act* (Versione I). Vedi anche Relazione di accompagnamento, *AI Act* (Versione I), 13-14.

⁴⁴⁶ Art. 1, *AI Act* (Versione I).

⁴⁴⁷ Art. 2, *AI Act* (Versione I).

⁴⁴⁸ *Id.*

ambito di applicazione territoriale risponde a quanto veniva precedentemente stabilito nel Libro Bianco del 2020⁴⁴⁹, indicando la continuità ideale di questi diversi documenti. L'art. 2 dell'*AI Act* predispone un sostanziale effetto extraterritoriale che ricorda l'art. 3 del GDPR, anch'esso dotato del medesimo scopo di rendere il proprio contenuto applicabile non solo ai fornitori di sistemi di IA dell'Unione, ma anche quelli stabiliti al di fuori di essa⁴⁵⁰. Questo effetto in grado di esorbitare dai confini comunitari a preoccupare gli sviluppatori delle grandi case di produzione dell'IA, soprattutto dall'altro lato dell'Atlantico⁴⁵¹.

All'interno del Titolo I dell'*AI Act* figurano, all'art. 3, le definizioni ricorrenti nell'intera regolamentazione, tra cui figura al punto 1), la definizione dell'IA. Così come anticipato nel paragrafo 2.3. del Capitolo 1, la nozione di IA coniata dall'*AI Act* è stata a lungo oggetto del dibattito politico tra i vari Stati membri, in quanto essa è in grado di rendere più o meno ampio l'ambito di applicazione della normativa⁴⁵², nonché garantirne la sopravvivenza e la non immediata obsolescenza sul breve periodo.

Si noti che la definizione di IA prevista nell'*AI Act* (Versione I) è connotata da una neutralità tecnologica, in quanto è sostanzialmente slegata dall'esplicitazione di alcun tecnicismo che ne potrebbe vincolare o restringere l'ambito di applicazione, mettendo irrimediabilmente in discussione la capacità della normativa di resistere al progredire dell'avanzamento tecnologico⁴⁵³. Ed è per incrementare il tasso di neutralità tecnologica che, con la Versione III del Consiglio, la definizione di IA è cambiata, da un "*software*" ad un "*sistema*."⁴⁵⁴ A ragione, gli elementi costitutivi dell'IA del passato non sono gli stessi di quelli del presente, né saranno gli stessi nel futuro: eppure, la normativa deve essere pronta ad affrontare i cambiamenti di alcuna sorta⁴⁵⁵. Anzi, proprio in forza della velocità con cui questi cambiamenti tecnologici possono verificarsi, è importante mantenere una flessibilità nozionistica di questi concetti.

⁴⁴⁹ Vedi Libro Bianco sull'IA, 25.

⁴⁵⁰ Art. 3, GDPR.

⁴⁵¹ Coraggio, (2024) min. 11.25-13.30.

⁴⁵² Vedi il paragrafo 2.3. del Capitolo 1, in cui è menzionato il dissidio interno tra conservatori e progressisti che si era creato sulla definizione di IA.

⁴⁵³ Scopelliti, M.A., *È ancora possibile la neutralità tecnologica della normativa?*, 216 in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.

⁴⁵⁴ Vedi il paragrafo 2.3. Capitolo 1. Vedi anche Art. 3(1), *AI Act* (Versione III). Art. 3(1), *AI Act* (Versione III). Vedi anche Floridi (2023).

⁴⁵⁵ Scopelliti (2023), 214.

Ad esempio, si pensi alla differenza, solo recentemente emersa e riconosciuta anche dal Consiglio di Stato italiano, tra algoritmi “tradizionali,” intesi come una sequenza di istruzioni da seguire meccanicamente, e sistemi che impiegano l’IA nelle loro computazioni⁴⁵⁶. Questa sottile distinzione è l’esempio lampante di quanto possano cambiare i paradigmi definitivi anche all’interno di quelle categorie che spesso vengono percepite come sovrapposte, in tempi ancorché brevi. Ogni sistema di IA in quanto tale è “*uguale solo a sé stesso*;⁴⁵⁷ la varietà composita di questi sistemi rafforza e giustifica, dunque, la scelta di una nozione tecnologicamente il più neutrale possibile. La Commissione, inoltre, nell’indicare una definizione univoca all’art. 3(1) dell’*AI Act*, ha provveduto ad un meccanismo integrativo tramite l’Allegato I alla proposta di regolamento, che contiene un elenco dettagliato di tecniche per lo sviluppo di sistemi di IA può essere modificato in modo da mantenere la normativa aggiornata rispetto agli sviluppi tecnologici⁴⁵⁸.

3.3.2 *L’approccio risk-based: pratiche vietate, sistemi ad alto rischio e a rischio minimo*

Come è stato più volte ribadito, la proposta di regolamento dell’Unione Europea assume un approccio di classificazione dei sistemi di IA basato sul rischio che gli stessi pongono nei confronti dei diritti fondamentali degli individui⁴⁵⁹. La Commissione ha optato per un approccio basato sul rischio, strutturando un sistema a piramide con quattro livelli di rischio: “inaccettabile”, “alto rischio”, e “basso rischio”, i cui requisiti sono riportati nel titolo II, III e IV dell’*AI Act*⁴⁶⁰. A queste categorie, è stata affiancata una classe intermedia di sistemi che sono soggetti ad una serie di obblighi di trasparenza, come si vedrà in seguito.

⁴⁵⁶ Il Consiglio di Stato ha così definito la differenza tra algoritmi e IA: “*un algoritmo è una sequenza finita di istruzioni, ben definite e non ambigue, così da poter essere eseguite meccanicamente e tali da produrre un determinato risultato. Tale nozione quando è applicata a sistemi tecnologici, è ineludibilmente collegata al concetto di automazione ossia a sistemi di azione e controllo idonei a ridurre l’intervento umano. Cosa diversa, invece, è l’intelligenza artificiale. In questo caso l’algoritmo contempla meccanismi di machine learning e crea un sistema che non si limita solo ad applicare le regole software e i parametri preimpostati (come fa invece l’algoritmo “tradizionale”) ma, al contrario, elabora costantemente nuovi criteri di inferenza tra dati e assume decisioni efficienti sulla base di tali elaborazioni, secondo un processo di apprendimento automatico.*” Sentenza 25 novembre 2021, n. 7891 del Consiglio di Stato, sezione III.

⁴⁵⁷ *Id.*

⁴⁵⁸ Relazione di accompagnamento, *AI Act* (Versione I), 13-14

⁴⁵⁹ La Versione IV dell’*AI Act*, riproducendo la II (cioè quella che rappresentava la posizione interlocutoria del Parlamento Europeo), ha inoltre previsto una definizione di “rischio” in via generale, identificato nella “*combinazione della probabilità del verificarsi di un evento dannoso e la gravità del danno.*” Articolo 3(1) (1a), *AI Act* (Versione IV).

⁴⁶⁰ *Id.*, 14-16.

Così facendo, la Commissione, ha dato seguito alla scelta di adottare uno strumento legislativo graduato a seconda dei rischi posti dai vari sistemi, invece di prevedere un'unica serie, non graduata, di requisiti obbligatori per tutti i sistemi di IA⁴⁶¹. A fronte di questa scelta di adottare un *risk-based approach*, si è resa necessaria la definizione “a monte” di quali diritti fondamentali siano considerati inviolabili e che necessitano, di conseguenza, di una protezione specifica da parte della legge⁴⁶². L'anticipata individuazione dei diritti ritenuti inviolabili rispetto quelli percepiti come più flessibili e dunque cedevoli, si configura come un'operazione essenziale per la valutazione dei sistemi di IA⁴⁶³: valutazione per di più cruciale quando l'introduzione di tali sistemi sul mercato comporta un rischio giudicato inaccettabile per gli stessi diritti fondamentali, sia a livello individuale, sia a livello collettivo per quanto riguarda il mantenimento dello stato di diritto e della democrazia⁴⁶⁴. Non solo: la ricerca *ex ante* dei diritti fondamentali da proteggere si combina con un ulteriore bilanciamento, tra gli stessi diritti fondamentali, e non tra diritti fondamentali ed esigenze di mercato⁴⁶⁵. Di questi vari bilanciamenti si avrà maggiore contezza a seguito della trattazione dei sistemi considerati vietati e ad alto rischio, appena di seguito.

L'art. 5 dell'*AI Act* identifica quei sistemi considerati vietati⁴⁶⁶. Nonostante il preciso *wording* dell'articolo non sia ancora consolidato (così come d'altronde l'intero contenuto della normativa), i sistemi vietati individuati dalla norma sono quelli che comportano una “*manipolazione delle persone attraverso tecniche subliminali, senza che tali persone ne siano consapevoli, oppure di sfruttamento delle vulnerabilità di specifici gruppi vulnerabili, quali i minori o le persone con disabilità, al fine di distorcerne materialmente il comportamento in maniera tale da provocare loro o a un'altra persona un danno psicologico o fisico.*”⁴⁶⁷ Inoltre, sono vietati i sistemi in grado di istituire meccanismi di “punteggio sociale” (*social scoring systems*), in grado di determinare l'affidabilità degli individui in base al loro comportamento sociale o in base a specifici tratti della

⁴⁶¹ Vedi paragrafo 2.

⁴⁶² Schepisi (2023), 62-63.

⁴⁶³ *Id.*, 63, 66.

⁴⁶⁴ *Id.*

⁴⁶⁵ *Id.*, 67. Rispetto questa particolare critica relativa al bilanciamento di esigenze di mercato e diritti fondamentali, si veda quando riportato dalla ONG AccessNow, in relazione ai loro commenti sulla disciplina dell'*AI Act*. Vedi il paragrafo 3.5.

⁴⁶⁶ Art. 5, *AI Act* (Versione I).

⁴⁶⁷ Relazione di accompagnamento, *AI Act*, (Versione I), 14.

loro personalità conosciuti o addirittura previsti dal sistema stesso⁴⁶⁸. Con l'ultimo accordo di dicembre 2023 sul testo finale della proposta di regolamento, sono state aggiunte altre, limitate, ipotesi di sistemi completamente vietati nel mercato UE⁴⁶⁹: da quelli che consentono il riconoscimento delle emozioni sul luogo di lavoro e nelle istituzioni scolastiche, eccetto per motivi medici o di sicurezza (ad esempio, monitorare i livelli di stanchezza di un pilota); a sistemi che consentono la raccolta non mirata di dati biometrici da Internet o da sistemi di videosorveglianza per creare o ampliare *databases* contenenti dati biometrici⁴⁷⁰; in ultimo, sistemi di polizia predittiva individuale⁴⁷¹. Infine, rientra nel novero delle pratiche vietate, in forza dell'art. 5, l'uso dei sistemi di identificazione biometrica remota in tempo reale⁴⁷² da parte di autorità pubbliche in luoghi aperti al pubblico, attività che coinvolge complessi sistemi di IA per il riconoscimento biometrico con un contestuale

⁴⁶⁸ Art. 5, *AI Act*. Vedi anche considerando 17, *AI Act*, (Versione I).

⁴⁶⁹ Press corner. (n.d.). European Commission - European Commission. https://ec.europa.eu/commission/presscorner/detail/en/QANDA_21_1683

⁴⁷⁰ Ciò avviene principalmente tramite l'utilizzo di una tecnica denominata *web scraping*. Esso consiste in una procedura tecnica mediante la quale vengono acquisiti dati da pagine web utilizzando *software* specializzati, in grado di estrarre informazioni e metadati da sorgenti online. Questa pratica può avvenire attraverso diverse modalità, come l'estrazione da template di pagine web, il prelievo diretto dagli API dei siti o l'utilizzo di piattaforme automatizzate che dispiegano una serie di *bot* per raccogliere ampie quantità di dati da diverse fonti online. Un esempio comune di *web scraping* è l'operato dei motori di ricerca che indicizzano i contenuti online per fornire risultati di ricerca. Questo tipo di pratica, perfettamente lecita, diventa illecita nel momento in cui ad essere oggetto del "raschiamento del web" sono dati biometrici di individui, mirati a costituire *databases* a mo' di quello redatto dalla nota società americana Clearview AI, utilizzato per scopi di polizia in maniera assolutamente non conforme alle normative europee sulla tutela della privacy. Sibilla, F., Bruschi, B.S., (13 dicembre 2023). *Il web scraping nell'Intelligenza Artificiale: il caso Clearview AI e l'indagine del Garante*. Rivista Diritto Di Internet – Osservatorio Dati & Imprese <https://dirittodiInternet.it/il-web-scraping-nellintelligenza-artificiale-il-caso-clearview-ai-e-lindagine-del-garante/>.

⁴⁷¹ La nozione di "polizia predittiva" si riferisce generalmente a qualsiasi sistema che analizza dati disponibili per prevedere dove un crimine potrebbe verificarsi in una data finestra temporale (basato sul luogo) o chi potrebbe essere coinvolto in un crimine come vittima o anche come autore. Si tratta di una delle ultime iterazioni delle tecniche di analisi criminale basate sui dati, di cui le forze dell'ordine si possono avvalere per il controllo e la previsione dei crimini. I sistemi di IA di polizia predittiva si basano su dati storici normalmente acquisiti dalle forze dell'ordine, dati tutt'altro che imparziali e spesso intrinsecamente caratterizzati da pregiudizi politici, sociali, economici ecc. Vedi Richardson, R., Schultz, J., & Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. 94 *NYU Law Review*, 198-199. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423#.

⁴⁷² L'art. 3(36) dell'*AI Act* provvede a perimetrare la nozione di sistema di identificazione biometrica remota in generale. Esso consiste in un sistema di IA "finalizzato all'identificazione a distanza di persone fisiche mediante il confronto dei dati biometrici di una persona con i dati biometrici contenuti in una banca dati di riferimento, e senza che l'utente del sistema di IA sappia in anticipo se la persona sarà presente e può essere identificata." L'art. 3(37) inoltre, permette la distinzione tra i sistemi di identificazione biometrica *real-time* (attivamente vietati dalla normativa) e quelli che invece operano a posteriori (considerati solamente dei sistemi ad alto rischio): con i sistemi di identificazione in tempo reale l'operazione di *cross-examination* tra i dati biometrici di un individuo e quelli contenuti nella banca dati di riferimento avviene "senza ritardi significativi." Artt. 3(36)-(37), *AI Act* (Versione I).

raffronto tra *database*⁴⁷³. Tuttavia, questa pratica non è completamente vietata, ma limitato entro confini predefiniti dalla legge, in quanto può essere comunque legalmente effettuata nelle (ancorchè ampie) eccezioni stabilite nello stesso art. 5⁴⁷⁴.

La definizione delle eccezioni al divieto generale di utilizzare sistemi di identificazione biometrica remota *real time* consiste in uno dei temi più dibattuti durante l'intero percorso di genesi e modifica dell'*AI Act*: questo perché l'utilizzo di una simile tecnologia può portare a delle conseguenze negative immediate sulla vita delle persone, oggetto di una pressoché costante sorveglianza, nonché provocare risultati discriminatori basati su *bias* a causa di caratteristiche intrinseche di funzionamento dei sistemi⁴⁷⁵.

Ad esempio, nella versione adottata Parlamento Europeo dell'*AI Act* le stesse ipotesi eccezionali in cui questa pratica era consentita erano state integralmente rimosse⁴⁷⁶, favorendo un divieto a tutto tondo dell'utilizzo di questa tecnologia e consentendo un'ampia tutela del diritto fondamentale alla vita privata e della protezione dati personali: il Parlamento, nell'ottica del bilanciamento tra diritti fondamentali, aveva considerato primario il diritto alla privacy rispetto alla pubblica sicurezza, che poteva essere rafforzata tramite il dispiegamento di una sostanziale sorveglianza biometrica continua in tempo reale su larga scala⁴⁷⁷.

Al momento in cui si scrive non si conosce ancora con certezza l'ampiezza delle eccezioni previste al divieto generale di utilizzo di sistemi di sorveglianza biometrica in tempo reale⁴⁷⁸. Ciò che è certo, tuttavia, è lo sforzo che varie

⁴⁷³ *Id.*

⁴⁷⁴ *Id.*

⁴⁷⁵ *Vedi* considerando 18, *AI Act*, (Versione IV).

⁴⁷⁶ Art. 5, *AI Act*, (Versione II).

⁴⁷⁷ Schepisi (2023), 67-68. *Vedi anche* considerando 18, *AI Act*, (Versione II).

⁴⁷⁸ Nella prima versione dell'*AI Act*, l'identificazione biometrica remota in tempo reale negli spazi accessibili al pubblico è generalmente vietata, ma con le seguenti eccezioni di utilizzo: i) *la ricerca mirata di potenziali vittime specifiche di reato, compresi i minori scomparsi*; ii) *la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o di un attacco terroristico*; iii) *il rilevamento, la localizzazione, l'identificazione o l'azione penale nei confronti di un autore o un sospettato di reato di cui all'art. 2 paragrafo 2 della decisione quadro 2002/584/GAI del Consiglio, punibile nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno tre anni, come stabilito dalla legge di tale stato membro.*" Art. 5(1)(d), *AI Act*, (Versione I).

Le suddette eccezioni, tuttavia, devono tenere conto dei seguenti elementi: "*la natura della situazione che dà luogo al possibile uso, in particolare la gravità, la probabilità e l'entità del danno causato dal mancato uso del sistema*" Art. 5(2)(a), (b), *AI Act*, (Versione I).

Inoltre, un suddetto utilizzo secondo le eccezioni previste deve essere necessariamente autorizzato in via preventiva rilasciata da un'autorità giudiziaria, eccetto vi sia una situazione di urgenza debitamente giustificata. Art. 5(3), *AI Act*, (Versione I). In questo caso, gli Stati membri hanno a

associazioni della società civile e impegnate nella protezione dei diritti digitali ed esponenti del mondo accademico hanno portato avanti, per rammentare alle istituzioni europee l'importanza di restringere quando più possibile le suddette eccezioni all'utilizzo di sistemi di *facial recognition*, che di fatto consentono pratiche di sorveglianza biometrica⁴⁷⁹, finanche discriminatorie⁴⁸⁰, e permettendo, di fatto, la compressione dei diritti fondamentali degli individui⁴⁸¹. Del risultato di questo sforzo collettivo non si ha evidenza; tuttavia, c'è una scarsa fiducia nell'effettivo mantenimento delle eccezioni all'utilizzo di una simile tecnologia invasiva così come erano state ideate dalla Commissione Europea nella prima versione dell'*AI Act*⁴⁸². A sostegno di questa sfiducia generale, vi è quanto emerso nella più recente versione (IV). In quest'ultima bozza, le eccezioni di utilizzo dei sistemi di riconoscimento facciale in tempo reale in spazi pubblici (insieme alle previsioni sulla più ampia disciplina del riconoscimento biometrico in luoghi accessibili al pubblico non in tempo reale bensì *ex post*⁴⁸³) sono addirittura ampliate, creando una serie di vere e proprie scappatoie normative (*loopholes*) le quali, anziché assicurare un perimetro ristretto di applicabilità e di utilizzo delle tecnologie di *facial recognition*, ne rendono ancora più ampio il ricorso⁴⁸⁴, minando

disposizione una certa discrezionalità di decidere come implementare l'utilizzo di questi sistemi nell'ambito della cornice di divieti delineata dall'art. 5, essendo questo uno dei temi per i quali l'Unione Europea prevede per l'appunto una cornice normativa da riempire con le definizioni nazionali, come accadrebbe all'interno di una direttiva. Art.5(4), *AI Act*, (Versione I).

⁴⁷⁹ La redazione della lettera, fermata da più di 70 organizzazioni della società civile e da 34 esperti del settore, è stata coordinata da European Digital Rights (EDRi). (8 dicembre 2023). *Open letter to EU AI Act negotiators: Do not trade away our rights!* https://edri.org/wp-content/uploads/2023/12/Open-letter-to-EU-AI-Act-negotiators_Do-not-trade-away-our-rights-8-December-2023.pdf

⁴⁸⁰ Vedi considerando 18, *AI Act*, (Versione IV).

⁴⁸¹ *Id.*

⁴⁸² Tra coloro che esprimono maggiore pessimismo riguardo al mantenimento del divieto generale dell'uso del riconoscimento facciale in tempo reale negli spazi accessibili al pubblico, spicca il parlamentare Patrick Breyer, il quale ha recentemente diffuso un comunicato in cui mette in evidenza la crescente probabilità che l'interdizione complessiva di questo sistema (salvo le eccezioni precedentemente indicate) possa subire un notevole indebolimento nel testo definitivo dell'*AI Act*. *Vedi AI Act threatens to make facial surveillance commonplace in Europe*. (16 gennaio 2024). Patrick Breyer. <https://www.patrick-breyer.de/en/ai-act-threatens-to-make-facial-surveillance-commonplace-in-europe/>. Vedi anche Volpicelli, G. (16 gennaio 2024). *EU set to allow draconian use of facial recognition tech, say lawmakers. POLITICO*. <https://www.politico.eu/article/eu-ai-facial-recognition-tech-act-late-tweaks-attack-civil-rights-key-lawmaker-hahn-warns/>.

⁴⁸³ *Id.*

⁴⁸⁴ Le eccezioni contenute nell'art. 5, secondo quanto previsto dalla IV bozza dell'*AI Act*, sono significativamente più ampie (confrontandole con quelle riportate nella nota 478, le cui modifiche sono sottolineate): *i) la ricerca mirata di potenziali vittime specifiche di sequestro, tratta di esseri umani e sfruttamento sessuale di esseri umani, nonché ricerca di persone scomparse; ii) la prevenzione di una minaccia specifica, sostanziale e imminente per la vita o l'incolumità fisica delle persone fisiche o una minaccia genuina e attuale o una minaccia genuina e prevedibile di un attacco terroristico; iii) la localizzazione o l'identificazione di un sospettato di aver commesso un reato,*

fortemente i diritti fondamentali e, in ultimo, legittimando misure di sorveglianza biometrica⁴⁸⁵.

Le tecnologie identificate come “ad alto rischio” sono invece consentite sul mercato europeo, se rispettano dei precisi requisiti obbligatori e sono prodotti in conformità a questi, precedentemente la loro immissione in commercio o il loro utilizzo nei confronti della collettività⁴⁸⁶. La determinazione “ad alto rischio” corrisponde alle potenzialità del sistema di IA in analisi di danneggiare la salute, la sicurezza o i diritti fondamentali delle persone fisiche, non solo per quanto riguarda gli scopi che questo può perseguire, ma anche le modalità con cui lo stesso opera⁴⁸⁷. Per cui, la classificazione di un sistema “ad alto rischio” può dipendere da molteplici fattori, non solo la mera funzione svolta da un sistema⁴⁸⁸. Per questo motivo, la Commissione Europea ha scelto un approccio di identificazione di questi sistemi tramite un elenco (aperto e facilmente ampliabile da parte dell’istituzione⁴⁸⁹) di sistemi ad alto rischio contenuti nell’allegato III dell’*AI Act*⁴⁹⁰. Inoltre, sono annoverati nell’elenco dei sistemi ad alto rischio anche quei sistemi di IA che sono componenti di sicurezza di prodotti più complessi, soggetti a legislazione settoriale dell’Unione⁴⁹¹. Essi saranno sempre considerati ad alto rischio quando sottoposti a valutazione di conformità da parte di terzi ai sensi di tale legislazione settoriale.

I requisiti che i sistemi “*high-risk*” devono necessariamente seguire attengono a diversi profili di criticità⁴⁹². Tra questi vi sono, la *governance* dei dati, la documentazione tecnica e la conservazione delle attività attraverso dei *log*, la trasparenza e l’informazione a disposizione degli utenti, nonché la sorveglianza da parte di esseri umani, la robustezza, la sicurezza e l’accuratezza generale di questi sistemi⁴⁹³. La normativa europea, dunque, mira a fornire dei parametri di

con lo scopo di condurre un’indagine, di esercitare l’azione penale o eseguire una pena per reati, indicati nell’Allegato IIa, e punibili nello Stato membro interessato con una pena o una misura di sicurezza privativa della libertà della durata massima di almeno quattro anni.” Art. 5(1)(d), *AI Act*, (Versione IV).

⁴⁸⁵ Vedi Volpicelli (2024).

⁴⁸⁶ Relazione di accompagnamento, *AI Act*, (Versione I), 14.

⁴⁸⁷ *Id.*

⁴⁸⁸ *Id.*

⁴⁸⁹ Art. 7, *AI Act*, (Versione I).

⁴⁹⁰ Relazione di accompagnamento, *AI Act*, (Versione I) 14-15.

⁴⁹¹ Art. 6(1)(a)(b), *AI Act*, (Versione I).

⁴⁹² Relazione di accompagnamento, *AI Act*, (Versione I), 15. Vedi anche Capo II, *AI Act*, (Versione I).

⁴⁹³ *Id.*

riferimento obbligatori per i sistemi ad alto rischio, i quali, senza le accortezze previste nei settori di criticità appena menzionati, risultano porre un pericolo eccessivo e non correttamente controbilanciato da una serie di tutele. Ad esempio, tra i vari requisiti previsti per quanto riguarda la *governance* dei dati utilizzati nei *training* algoritmici⁴⁹⁴, vi è l'indicazione secondo la quale i set di dati di addestramento devono essere pertinenti, rappresentativi, esenti da errori e completi⁴⁹⁵. Questo requisito insieme alle altre specifiche indicazioni relative ai dati di addestramento, se adeguatamente rispettato, potrebbe portare ad una riduzione degli aspetti unicamente inerenti alla qualità dei dati che possono portare ad un *output* dell'IA in grado di danneggiare gli individui e la società.

Ancora, un altro dato interessante è rappresentato dal requisito di “sorveglianza umana” dei sistemi ad alto rischio. L'art. 14 dell'*AI Act*, infatti, descrive che questa attività di supervisione da parte di essere umani è mirata a prevenire o a ridurre al minimo i rischi per la salute, la sicurezza e i diritti fondamentali, ed essa è assicurata mediante modalità integrate nel sistema di IA fin dalla sua immissione sul mercato, oppure tramite modalità che possono essere attivate dagli stessi utenti finali⁴⁹⁶. In ultimo, i sistemi ad alto rischio sono oggetto sia di una valutazione dei rischi *ex ante*, ma anche di un monitoraggio continuo durante il corso di utilizzo e dispiegamento sul mercato⁴⁹⁷.

Tra le valutazioni necessariamente svolte *ex ante*, è stata recentemente inserita una procedura denominata “*fundamental rights impact assessment*,” (FRIA) cioè valutazione di impatto sui diritti fondamentali dei sistemi ad alto rischio⁴⁹⁸. L'analisi sulle conseguenze eventualmente lesive sui diritti inviolabili prevede che gli sviluppatori di sistemi “*high-risk*” esplicitino una serie di elementi, tra cui: una chiara delimitazione dello scopo perseguito dall'IA, nonché il suo ambito di applicazione territoriale e temporale; le categorie di persone fisiche che possono essere probabilmente influenzati da questi sistemi; il prevedibile impatto sui diritti fondamentali se il sistema ad alto rischio fosse effettivamente implementato, nonché gli eventuali rischi aggiuntivi con cui gruppi di individui vulnerabili e marginalizzati possono doversi confrontare⁴⁹⁹. Secondo la valutazione

⁴⁹⁴ Per un'accurata analisi del *training* algoritmico, vedi il sottoparagrafo 2.2.1 del Capitolo 1.

⁴⁹⁵ Art. 10(3), *AI Act*, (Versione I).

⁴⁹⁶ Art. 14(2), art. 14(3)(a)(b), *AI Act*, (Versione I).

⁴⁹⁷ Vedi il Titolo VIII dell'*AI Act*, denominato “monitoraggio successivo all'immissione sul mercato, condivisione delle informazioni, vigilanza del mercato). Artt. 61 e ss., *AI Act*, (Versione I).

⁴⁹⁸ Art. 29(a), *AI Act*, (Versione II).

⁴⁹⁹ *Id.*

di impatto così strutturata, se i pericoli non possono essere in alcun modo eliminati o comunque mitigati da parte degli sviluppatori, questi devono astenersi dal mettere sul mercato il determinato sistema⁵⁰⁰.

Questo processo di valutazione di impatto, implementato con la versione dell'*AI Act* adottata dal Parlamento Europeo, sembra essere confluita nell'accordo politico di dicembre 2023⁵⁰¹. Pertanto, idealmente, dovrebbe essere integrato nel testo definitivo della normativa. Tuttavia, emerge dalla IV versione della proposta di regolamento, che la FRIA può anche essere svolta sulla base di questionari precedentemente redatti da parte di un organo specialmente adibito a livello europeo (l'*AI Office*)⁵⁰². È tuttavia da osservare che, nonostante l'incentivo a portare a termine le valutazioni di impatto, facilitando la *compliance* da parte delle imprese, la FRIA così strutturata rischia di essere svuotata di significato e ridotta unicamente a mero adempimento formale⁵⁰³.

Per quanto riguarda i sistemi di IA considerati di “rischio minimo”, l'*AI Act* indica l'adozione unilaterale da parte degli sviluppatori (o delle organizzazioni che li rappresentano) di codici di condotta che possano promuovere l'applicazione volontaria dei requisiti relativi, ad alcuni ambiti di interesse quali (a titolo meramente esemplificativo e comunque non esaustivo) la sostenibilità ambientale, l'accessibilità per persone con disabilità, o la partecipazione di gruppi di minoranze al processo di creazione e utilizzo dei modelli di IA⁵⁰⁴. Questa previsione sulla possibilità di sviluppare codici di condotta volontari è la chiave per comprendere l'intenzione di perseguire un modello sostanziale co-regolamentazione da parte della Commissione Europea⁵⁰⁵. Tale approccio combina strumenti di *hard law* per i sistemi ad alto rischio con quelli di *soft law* per i sistemi meno problematici, evidenziando una prospettiva equilibrata nella definizione delle regole e delle pratiche da seguire a seconda dei rischi posti dai vari modelli⁵⁰⁶.

⁵⁰⁰ *Id.*

⁵⁰¹ *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI | News | European Parliament.* (9 dicembre 2023). <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>.

⁵⁰² L'*AI Office* è una delle entità costituite dall'*AI Act* nel modello comunitario di *governance* delineato nel Titolo IV. *Vedi* Art. 55(b), *AI Act*, (Versione IV).

⁵⁰³ Mantelero (2023), 37, 192.

⁵⁰⁴ Art. 69, *AI Act*, (Versione I).

⁵⁰⁵ Mantelero (2022), 142.

⁵⁰⁶ *Id.*

3.3.3. *Gli obblighi di trasparenza per sistemi intermedi e regolamentazione dell'IA generativa*

Sussiste una categoria intermedia di sistemi a sé stante che, a motivo delle proprie caratteristiche peculiari, è vincolata da obblighi di trasparenza, dal momento che il loro utilizzo potrebbe causare di per sé dei rischi intrinseci di manipolazione a livello individuale e collettivo⁵⁰⁷. Gli obblighi di trasparenza previsti dal Titolo IV dell'*AI Act* lasciano tuttavia impregiudicati tutti i requisiti previsti nel Titolo III per i sistemi ad alto rischio⁵⁰⁸. Questi obblighi relativi alla trasparenza e all'informazione si applicano a quei sistemi che in *primis* interagiscono con gli esseri umani (in quanto utenti finali o intermedi degli stessi sistemi), che sono utilizzati per rilevare le emozioni o formulare associazioni con categorie sociali e dati biometrici, oppure che siano adibiti alla generazione o manipolazione di contenuti⁵⁰⁹. Tra i vari sistemi annoverati in questo sottogruppo destinatario di obblighi più stringenti vi sono anche i sistemi in grado di produrre *deepfakes*⁵¹⁰.

Al fine di mitigare le potenziali problematiche associate alla diffusione di *deepfake*, specialmente per la loro capacità di manipolare il pubblico su vasta scala, chiunque produca tali contenuti è tenuto a fornire un chiaro e inequivocabile avviso agli utenti. Questa comunicazione deve avvenire attraverso informazioni facilmente comprensibili non solo da parte delle macchine, ma anche da parte degli utenti, indicando che il contenuto in questione è generato mediante l'uso di IA⁵¹¹.

Nelle bozze dell'*AI Act* che hanno succeduto la Versione I, è stata inclusa la regolamentazione di sistemi che hanno suscitato un ampio impatto nell'opinione pubblica dopo il lancio di ChatGPT nel 2022, in particolare i sistemi di intelligenza artificiale generativa⁵¹². I sistemi che utilizzano IA generativa sono stati oggetto di

⁵⁰⁷ Relazione di accompagnamento, *AI Act*, (Versione I), 16.

⁵⁰⁸ Art. 52(4), *AI Act*, (Versione I).

⁵⁰⁹ Relazione di accompagnamento, *AI Act*, (Versione I), 16. *Vedi anche* Art. 52, *AI Act*, (Versione I).

⁵¹⁰ Schepisi (2023), 68-69. Per *deepfakes* si intendono quelle tecnologie che consentono di produrre foto, audio e video falsificati bensì fortemente realistici e convincenti, a tal punto da essere idonei ad influenzare e manipolare l'opinione pubblica, o ad essere sfruttati per ricatti o diffusione di notizie false. L'uso deregolamentato di una tecnologia simile pone un rischio non indifferente nei confronti della stabilità della democrazia e dello stato di diritto, data la propria capacità di influenzare gli utenti. *Vedi* Risoluzione del Parlamento europeo del 20 gennaio 2021 sull'IA: *questioni relative all'interpretazione e applicazione del diritto internazionale nella misura in cui l'UE è interessata relativamente agli impieghi civili e militari e all'autorità dello Stato al di fuori dell'ambito della giustizia penale* (2020/2013(INI)). *Vedi anche* considerando 70a-70e, *AI Act*, (Versione IV).

⁵¹¹ Considerando 70a-70e, *AI Act*, (Versione IV).

⁵¹² *Vedi* la definizione di IA generativa, nota 2.

una vera e propria irruzione nel panorama dell'opinione pubblica, tra il 2022 e il 2023⁵¹³. Anche la regolamentazione di questi sistemi, come numerosi analizzati precedentemente, ha comportato degli scontri a livello politico tra Stati Membri⁵¹⁴, a causa della sua rilevanza e forte copertura mediatica che delle regole applicabili ai sistemi al momento più attenzionati dall'opinione pubblica possono aver generato. La definizione dell'IA generativa in sé era stata inserita nella Versione II dell'*AI Act*, “*sistemi di intelligenza artificiale specificamente destinati a generare, con vari livelli di autonomia, contenuti come testi complessi, immagini, audio o video.*”⁵¹⁵ Inizialmente, l'IA generativa doveva essere sottoposta di *default* ai requisiti previsti per i sistemi ad alto rischio: con l'evolversi del dibattito, è lentamente venuta meno questa ipotesi in quanto eccessivamente gravosa per coloro che sviluppano grandi modelli di IA generativa, come ad esempio gli LLMs⁵¹⁶.

A questo proposito, un modello di IA generativa può essere considerato “*high-risk*” secondo gli stessi parametri previsti per gli altri sistemi di IA, cioè secondo quella metodologia di classificazione prevista all'art. 6 dell'*AI Act*⁵¹⁷. A fronte di un'intensa attività di *lobbying* da parte di entità esterne al processo legislativo appartenente al settore dell'industria *tech*⁵¹⁸, l'assetto dei requisiti a cui sono soggetti i modelli di IA generativa è visibilmente cambiato e dotato di una complessità maggiore, anche solo dal punto di vista definitorio.

Innanzitutto, è recentemente nata una nuova nozione normativa nell'*AI Act*, la c.d. *General Purpose AI*, ascrivibile a quei sistemi di IA che vengono sviluppati per scopi vari e adattabili⁵¹⁹. All'interno di questa nozione rientrano sia l'IA generativa, che il relativo sottogruppo della stessa spesso menzionato, cioè gli LLMs. Un sottogruppo di IA generativa ancora più preciso è quello dei c.d. *foundational models*, cioè quei modelli che sono talvolta utilizzati come “fondamenta” per costituire delle sovrastrutture di sistemi di IA più specifici ed

⁵¹³ *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 2: Technical performance, chapter highlights, 73.

⁵¹⁴ Bertuzzi, L., (24 novembre 2023b). France, Germany, Italy push for ‘mandatory self-regulation’ for foundation models in EU’s AI law. [www.euractiv.com. https://www.euractiv.com/section/artificial-intelligence/news/france-germany-italy-push-for-mandatory-self-regulation-for-foundation-models-in-eus-ai-law/](https://www.euractiv.com/section/artificial-intelligence/news/france-germany-italy-push-for-mandatory-self-regulation-for-foundation-models-in-eus-ai-law/).

⁵¹⁵ Vedi Art. 28b (4), *AI Act*, (Versione II).

⁵¹⁶ Per la definizione di LLMs, vedi nota 92.

⁵¹⁷ Novelli, C., Casolari, F., Hacker, P., Spedicato, G. A., & Floridi, L. (2024). Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity. *Productivity, Innovation & Technology eJournal*, 3 <https://doi.org/10.2139/ssrn.4694565>.

⁵¹⁸ Byte by byte: How Big Tech undermined the AI Act. *Corporate Europe Observatory* (17 novembre 2023) <https://corporateeurope.org/en/2023/11/byte-byte>.

⁵¹⁹ Novelli *et al.* (2024), 3-4.

articolati, spesso programmati da soggetti diversi dai primi sviluppatori dei *foundational models*. La sovrapposizione, ancorchè confusa, di queste diverse categorie di IA è stata vista criticamente da alcuni, a causa della scarsa chiarezza sia nei confronti degli Stati membri che per le imprese, ma anche per la difficoltà di coordinamento con altri testi normativi⁵²⁰.

Secondo quanto previsto dai novellati artt. 52(a) e ss. della IV versione dell'*AI Act*, i modelli di GPAI possono essere anche porre un “*systemic risk*”⁵²¹ (che viene individuato diversamente rispetto il rischio “alto”). In ultimo, sono considerati GPAI che pongono un rischio sistemico di *default* tutti quei modelli di IA che dimostrino di avere un potere di calcolo pari a 10^{25} FLOPs⁵²² (*floating point operations per second*, un’unità di misura della capacità computazionale). I sistemi di IA così caratterizzati devono presentare una serie di ulteriori accortezze *ex ante* sia di sicurezza informatica che di trasparenza, nonché sulle informazioni relative al *training* algoritmico svolto prima dell’immissione sul mercato. Il parametro dei 10^{25} FLOPs è tuttavia un parametro modificabile dall’*AI Office*. La scelta di fare affidamento ad un valore numerico di tale sorta è stato tuttavia ritenuto, secondo alcuni, un indicatore limitante: modelli dotati di capacità computazionali leggermente al di sotto della soglia sono in grado di porre allo stesso modo un rischio sistemico⁵²³.

3.3.4 Le misure a sostegno dell’innovazione e governance dell’IA

In ultimo, occorre analizzare due elementi fortemente innovativi della regolamentazione europea, cioè la presenza di misure a sostegno dell’innovazione nel settore dell’IA, in parallelo con l’istituzione di un modello di *governance* dettagliatamente articolato a livello comunitario, ma anche all’interno dei singoli Stati membri. L’*AI Act*, infatti, oltre a prevedere il sistema di obblighi di conformità secondo l’approccio *risk-based* appena analizzato, tende ad implementare una serie

⁵²⁰ *Id.*

⁵²¹ Secondo l’art. 52(a) dell’*AI Act*, un modello pone un rischio “sistemico quando:” (a) *ha un’elevata capacità di impatto valutata sulla base di strumenti tecnici e metodologie adeguate, compresi indicatori e parametri di riferimento*” oppure “[...] (b) *sulla base di una decisione della Commissione, d’ufficio o a seguito di una segnalazione qualificata da parte del gruppo di esperti scientifici, che un modello generale di IA ha capacità o impatto equivalenti a quelli di cui al punto a)*”. Art. 52(a)(1), *AI Act*, (Versione IV).

⁵²² Art. 52(a)(2), *AI Act*, (Versione IV). Vedi anche considerando 60(n), *AI Act*, (Versione IV). Vedi anche Zorloni, L. (23 gennaio 2024) *AI Act, abbiamo letto l’ultima versione della legge europea sull’intelligenza artificiale. Wired Italia*. <https://www.wired.it/article/ai-act-testo-ultima-versione-gennaio-divieti-riconoscimento-facciale/#tre> Vedi anche Novelli et al. (2024), 3-4.

⁵²³ Novelli et al. (2024), 3-4.

di novità per mantenere quell'approccio incentrato alla creazione di un ecosistema di fiducia e, in ultimo, di innovazione nel campo dell'IA. A questo proposito, si comprende l'istituzione da parte della Commissione di strumenti di supporto per le PMI e le start-up, che si vedono ridotti i propri oneri in forza di quanto previsto all'art. 55(2)⁵²⁴.

Inoltre, nel Titolo V della proposta di regolamento, viene incentivata la creazione delle c.d. “*regulatory sandboxes*”, cioè degli spazi controllati di sperimentazione normativa per lo sviluppo di nuovi modelli di IA. Il concetto di spazio delimitato per la sperimentazione non è nuovo al settore della tecnologia. I *sandboxes* regolatori nascono infatti immediatamente subito la crisi finanziaria del 2008, con lo scopo di essere impiegati per sperimentare la tenuta di nuove tecnologie applicate nel settore finanziario⁵²⁵. Sebbene inizialmente non inseriti in quel più ampio disegno della Commissione UE sull'IA fornito con il Libro Bianco del 2020, essi sono stati integrati con il primo disegno di legge sull'IA (Versione I)⁵²⁶. Questi *sandboxes* consistono dunque in spazi regolamentari tutelati adibiti alla sperimentazione della tenuta di determinati sistemi secondo un piano di prova controllato e con la previsione di appropriate salvaguardie⁵²⁷. Le autorità nazionali competenti nel settore di IA sono investite del compito di creare i *sandboxes*: vige infatti un obbligo generale per tutti gli Stati membri di creare almeno uno spazio di sperimentazione normativa nel periodo immediatamente successivo l'entrata in vigore dell'*AI Act*⁵²⁸. L'obbligo ad istituire almeno un *sandbox* è stato tempestivamente recepito dagli Stati Membri: la Spagna, ad esempio, è stato il primo paese dell'UE a predisporre l'apertura dell'ambiente di sperimentazione

⁵²⁴ Art. 55, *AI Act*, (Versione I).

⁵²⁵ Yordanova, K., Bertels, N. (2023), *Regulating AI: Challenges and the Way Forward Through Regulatory Sandboxes*, in Antunes, H. S., Freitas, P. M., Oliveira, A. L., Pereira, C. M., De Sequeira, E. V., & Xavier, L. B. (2023). *Multidisciplinary perspectives on artificial intelligence and the law*. Springer Nature, 441-454.

⁵²⁶ Yordanova & Bertels (2023), 442.

⁵²⁷ Vedi la definizione di “*regulatory sandbox*” fornita dal Consiglio in una comunicazione del 2020. Essi sono “*quadri concreti che, fornendo un contesto strutturato per la sperimentazione, consentono, se del caso, di testare in un ambiente reale tecnologie, prodotti, servizi o approcci innovativi - al momento soprattutto nel contesto della digitalizzazione - per un periodo di tempo limitato e in una parte circoscritta di un settore o di un'area sottoposta a supervisione regolamentare, assicurando l'esistenza di adeguate salvaguardie.*” Vedi Consiglio dell'Unione Europea (2020) *Council Conclusions on Regulatory Sandboxes and Experimentation Clauses as tools for an innovation-friendly, future-proof, and resilient regulatory framework that masters disruptive challenges in the digital age*, 2020/C 447/01,

⁵²⁸ Art. 53(1), *AI Act*, (Versione IV).

controllata⁵²⁹. Fra i paesi che si sono attivati in questo senso vi è anche l'Italia, che ha predisposto un disegno di legge relativo all'istituzione di una *regulatory sandbox*⁵³⁰. L'utilizzo di queste sperimentazioni stabilisce un ruolo guida e di riferimento per il regolatore, anziché considerarlo come un avversario⁵³¹.

Un ultimo accenno merita l'assetto di *governance* istituito con la Proposta di regolamento. Insieme al già menzionato *AI Office*, l'*AI Act* istituisce diverse entità per rafforzare la *governance* a livello sia comunitario che nazionale. In *primis*, l'art. 56 istituisce l'*European Artificial Intelligence Board*, formato da rappresentanti degli Stati Membri, allo scopo di facilitare l'applicazione efficace dell'*AI Act*⁵³². I compiti assegnati al suddetto *Board* ricordano quelli assegnati all'*European Data Protection Board* con l'entrata in vigore del GDPR: infatti, tra di essi vi sono la supervisione sulla corretta implementazione del contenuto dell'*AI Act*, il coordinamento con le autorità nazionali sull'IA, la supervisione e il contributo sull'armonizzazione delle varie pratiche amministrative degli Stati Membri, nonché, la possibilità di emettere pareri quando necessario⁵³³. Di fondamentale importanza sono, inoltre, anche le varie autorità nazionali sull'AI che l'*AI Act* istituisce e a cui adibisce una serie di scopi diversi. Infatti, in forza dell'art. 59, viene predisposta la nuova istituzione di un'autorità nazionale competente o altresì la nomina di un'autorità preesistente con conseguente conferimento di poteri di supervisione e controllo⁵³⁴.

3.4 Spunti di critica dell'*AI Act*

La normativa relativa all'IA così strutturata nella proposta di Regolamento dell'*AI Act* non è tuttavia priva di spunti di critica. Nonostante la redazione di una legislazione vincolante per regolare un settore tecnologico cruciale come l'IA rappresenti un notevole successo per l'Unione Europea⁵³⁵, il cui potere regolatore è stato più volte sottolineato, alcuni accademici si sono soffermati sulla concreta

⁵²⁹ Rivaya, J. F., & Vidal, A. (29 settembre 2023). Spain: The artificial intelligence regulatory "sandbox" has arrived. *Garrigues*. https://www.garrigues.com/en_GB/garrigues-digital/spain-artificial-intelligence-regulatory-sandbox-has-arrived.

⁵³⁰ Camera dei deputati n. 1084, Proposta di legge d'iniziativa dei deputati Centemero, Andreuzza, Bagnai, Barabotti, Cavandoli, Di Mattina, Gusmeroli, Toccalini, Billi, Bof, Coin, Dara, Furgiuele, Pretto Zinzi, *Disposizioni concernenti l'adozione di una disciplina temporanea per la sperimentazione dell'impiego di sistemi di intelligenza artificiale*, (11 aprile 2023).

⁵³¹ Yordanova & Bertels (2023), 446.

⁵³² Art. 58, *AI Act*, (Versione IV).

⁵³³ *Id.*

⁵³⁴ Art. 59, *AI Act* (Versione IV).

⁵³⁵ *Vedi* nota 407.

possibilità per il testo normativo di centrare gli (ancorchè ambiziosi) obbiettivi espliciti ed impliciti che l'UE gli ha attribuito. Infatti, sin dall'emanazione della prima versione dell'*AI Act*, ad aprile 2021, sono emersi dei profili di criticità che sono stati più volte sottolineati da vari esponenti del mondo accademico, e che vale la pena di riportare per comprendere la concreta portata della normativa in esame, anche nell'ottica di ridimensionarne l'effetto dirompente e innovativo che essa sembrerebbe implicare.

In *primis*, la Prof.ssa Lilian Edwards dell'*Ada Lovelace Institute* ha sottolineato come la scelta di un approccio alla regolamentazione mirato a garantire la sicurezza dei prodotti per la tutela dei consumatori non consente di considerare diversi profili di complessità strutturale dell'IA, che viene trattata nell'*AI Act* come un *on-off product*⁵³⁶. Un simile approccio comporta un'eccessiva semplificazione della tecnologia dell'IA, ridotta ad un mero "prodotto" di cui assicurare la sicurezza per il consumo⁵³⁷. Infatti, l'*AI Act*, con il suo approccio *risk-based*, individua quelle regole di sicurezza di cui sono destinatari i c.d. fornitori o *providers*, cioè coloro che si occupano di sviluppare i sistemi di IA e di immetterli sul mercato o in servizio nei confronti degli utenti⁵³⁸, alla stregua di quegli strumenti di diritto derivato dell'UE volti ad assicurare la protezione dei consumatori in materia di sicurezza dei prodotti⁵³⁹. L'IA, tuttavia, non può essere oggetto di una semplificazione alla stregua un singolo prodotto materiale ed autonomo: esso è piuttosto un sistema dinamico e multiforme, assolutamente *general-purpose*, scaturente da una serie di entità diverse che ne permettono il funzionamento in contesti diversi e con impatti diversi a seconda di chi ne fa utilizzo⁵⁴⁰.

Un altro aspetto oggetto della critica da parte di accademici e società civile è un altro dei fondamenti dell'*AI Act*, il quale ha nella sua ragione d'essere il

⁵³⁶ Edwards (2022), 5.

⁵³⁷ *Id.*

⁵³⁸ Art.3(2), *AI Act*, (Versione I).

⁵³⁹ Ad esempio, si pensi ai seguenti strumenti di diritto derivato dell'UE che si sono occupati di stabilire regole e standard per la tutela dei consumatori: il Regolamento UE 2023/988, che ha abrogato la precedente direttiva 2001/95/CE, in materia di sicurezza generale dei prodotti in commercio dell'UE; allo stesso modo, si pensi anche agli strumenti di regolamentazione di settori particolarmente rilevanti, ad esempio la regolamentazione sulle sostanze chimiche. *Vedi* Regolamento (UE) 2023/988 del Parlamento europeo e del Consiglio del 10 maggio 2023 relativo alla sicurezza generale dei prodotti, che modifica il regolamento (UE) n. 1025/2012 del Parlamento europeo e del Consiglio e la direttiva (UE) 2020/1828 del Parlamento europeo e del Consiglio, e che abroga la direttiva 2001/95/CE del Parlamento europeo e del Consiglio e la direttiva 87/357/CEE del Consiglio (Testo rilevante ai fini del SEE); *vedi anche* il Regolamento REACH sulla delle sostanze chimiche (nota 319).

⁵⁴⁰ Edwards (2022), 5-6.

preventivo bilanciamento tra diritti fondamentali (sia con le esigenze di mercato, sia tra diritti fondamentali più o meno “cedevoli”⁵⁴¹). Secondo ONG impegnate nel settore, come Access Now⁵⁴², il presupposto del bilanciamento tra interessi economici e diritti fondamentali per valutare il rischio di determinate tecnologie è irrimediabilmente fallace: i diritti fondamentali in quanto tali non sono negoziabili e devono essere rispettati a prescindere dal livello di rischio associato a fattori esterni⁵⁴³.

Un altro dato strutturale dell’*AI Act* che ha suscitato perplessità è l’obiettivo, spesso richiamato, della realizzazione di un “ecosistema di fiducia.” È proprio la creazione di un’IA che fosse “*trustworthy*”⁵⁴⁴ ad aver spinto idealmente la Commissione a formulare una disciplina omnicomprensiva e ben strutturata per questi sistemi, fin dai tempi del Libro Bianco sull’IA del 2020⁵⁴⁵. Tuttavia, alcuni hanno richiamato l’inadeguatezza dell’*AI Act* a generare di per sé un’IA che ispiri “fiducia” nei confronti degli utenti finali⁵⁴⁶. L’Unione Europea, secondo Laux *et al.*, ha erroneamente equiparato l’accettabilità percepita del rischio alla capacità di stimolare fiducia nell’utilizzo dei sistemi di IA (soprattutto quando questi sono impiegati nel settore pubblico)⁵⁴⁷. Secondo l’analisi empirica del concetto di “fiducia” (concetto non di natura legale bensì sociologica, e pertanto complesso da inquadrare in termini rigorosi⁵⁴⁸) e di come questa si genera a livello sociale nell’ambito della elaborazione di politiche pubbliche, essa non può consistere in una mera accettabilità del rischio⁵⁴⁹. La sostenibilità degli eventuali rischi, infatti, costituisce solo uno dei vari elementi che possono stimolare maggiore fiducia a livello collettivo⁵⁵⁰.

In base a queste conclusioni, l’Unione Europea starebbe sopravvalutando le proprie ambizioni normative e semplificando eccessivamente un insieme estremamente complesso ed eterogeneo di concetti strettamente correlati: l’*AI Act*

⁵⁴¹ Schepisi (2023), 60-70.

⁵⁴² Access Now official website, *Access Now* <https://www.accessnow.org/>.

⁵⁴³ Hidvegi, F., Daniel, L., & Masse, E. (2021), The EU should regulate AI on the basis of rights, not risks, *Access Now*, <https://www.accessnow.org/eu-regulation-ai-risk-based-approach/>. Vedi anche Benbouzid *et al.* (2022), 26-27.

⁵⁴⁴ Meritevole di fiducia.

⁵⁴⁵ Vedi il paragrafo 2 e 3.1

⁵⁴⁶ Laux, J., Wachter, S., & Mittelstadt, B. (2022). Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4230294>.

⁵⁴⁷ Laux *et al.* (2022).

⁵⁴⁸ Mantelero (2022), 189.

⁵⁴⁹ Laux *et al.* (2022).

⁵⁵⁰ *Id.*

potrebbe essere dunque inadeguato nel raggiungere uno degli scopi principali per i quali è stato creato⁵⁵¹. Alcuni hanno ricollegato questa inadeguatezza ad un altro aspetto critico: l'*AI Act*, così come strutturato, non rispecchia a pieno quell'intento di imprimere quell'approccio "etico" all'utilizzo dell'IA che era emerso dopo gli intensi lavori dei gruppi di esperti come l'HLEGAI. Dopo diversi anni di discussioni sulla dimensione etica dell'IA⁵⁵², la visione che risulta aver prevalso con l'attuale struttura dell'*AI Act* sembra un approccio che delega le questioni etiche ad altre iniziative al di fuori della regolamentazione *hard*⁵⁵³.

Questo a primo impatto potrebbe rappresentare un dato allarmante; altri però ritengono che questa mancanza risvolti prettamente etici nell'*AI Act*, non sia necessariamente una criticità⁵⁵⁴. A questo proposito, gli strumenti di *soft law* come le linee guida, le raccomandazioni da parte di istituzioni, le *best practices* aziendali possono porre rimedio a questo *deficit*, introducendo nella pratica comportamenti ed esiti desiderabili: essi non sono sostitutivi bensì meramente integrativi della regolamentazione, e si occupano di elevare gli standard oltre la semplice conformità alla legge⁵⁵⁵. Questi strumenti collettivamente considerati rientrano nel concetto di c.d. "*etica soft*," cioè strumenti assolutamente post-*compliance* rispetto la legge, che chiariscono le implicazioni morali più "giuste" in un determinato settore come l'IA, che presenta una serie questioni prettamente etiche⁵⁵⁶. D'altronde, come spesso sostenuto da Floridi "*una volta che si è rispettata la legge, si può desiderare di fare qualcosa di più che seguirla.*"⁵⁵⁷

Ancora, molti hanno manifestato la propria preoccupazione in merito all'assoluta mancanza di diritti azionabili da parte dei diretti interessati che entrano in contatto con un dato sistema di IA. L'intera regolamentazione assume come destinatario principale i *providers*, mentre ignora in gran parte il potenziale ruolo degli utenti finali⁵⁵⁸.

⁵⁵¹ *Id.*

⁵⁵² *Vedi* paragrafo 2.

⁵⁵³ Mantelero (2022), 173.

⁵⁵⁴ *Vedi* Floridi (2021) e Floridi (2022).

⁵⁵⁵ Floridi (2021), 620.

⁵⁵⁶ Floridi (2022), 130-138.

⁵⁵⁷ Floridi (2021), 620. *Vedi anche* Floridi (2022), 131.

⁵⁵⁸ Gli utenti, intermediari o finali, ma comunque intesi come coloro che interagiscono con il sistema di IA, erano denominati "*users*" nella Versione I dell'*AI Act*; ad oggi, hanno assunto la denominazione di "*deployers*" a partire dalla Versione II della proposta di regolamento, in quanto più confacente al loro reale ruolo. I *deployers*, infatti, impiegano i sistemi di IA con cui interagiscono anche nei confronti di altri soggetti, in forza di quella continua catena di rapporti che talvolta identifica questo genere di sistemi. Tutto ciò porta a sfocare, talvolta, i confini concettuali fra i

Innanzitutto, essi non sono titolari di particolari diritti nei confronti delle potenziali conseguenze negative causate dai sistemi di IA⁵⁵⁹. Tuttavia, questo apparente vuoto normativo dell'*AI Act* circa i rimedi giurisdizionali a disposizione degli individui che subiscono le conseguenze negative di determinati sistemi di IA, sembra essere parzialmente colmato, o comunque in gran parte affrontato, con le due proposte di direttive sulla responsabilità dell'IA, di cui al paragrafo 4. Ad ogni modo, è sembrato paradossale che all'interno della proposta di regolamento siano assenti meccanismi rimediali o di reclamo a disposizione degli utenti finali, dal momento che il testo normativo annovera in uno degli obiettivi primari proprio la tutela dei diritti fondamentali di quest'ultimi⁵⁶⁰.

Allo stesso modo, gli utenti non hanno modo di partecipare alle procedure di valutazione di conformità da parte dei *providers*, processo complesso ad alto contenuto tecnico e inaccessibile ai più per svariati motivi (dall'asimmetria informativa tra utenti e *providers* alla presenza di tutela della proprietà intellettuale⁵⁶¹). Inoltre, se ad esempio, i fornitori di un dato sistema di IA commercializzano soluzioni personalizzabili dagli utenti, che hanno dunque l'opportunità di modificare autonomamente il tipo di rischio ai quali si espongono, essi comunque non sono destinatari di alcun obbligo in merito a ciò: il loro ruolo nei processi di gestione del rischio è marginalizzato anche nel caso in cui la stessa entità del rischio dipendesse da loro⁵⁶².

La sostanziale assenza di diritti e facoltà attribuita a coloro che utilizzano i sistemi di IA è uno dei tratti che maggiormente contraddistinguono l'*AI Act* da un'altra regolamentazione che ha significativamente rivoluzionato il settore del digitale e, per la precisione, della tutela dei dati personali: il GDPR. Il GDPR, a

“fornitori” di un sistema (i *providers*, cioè le prime aziende produttrici) e gli utenti finali (che comunque vengono considerati “*deployers*”). Vedi il confronto tra Art. 3(1)(4), *AI Act*, (Versione I) e Art. 3(1)(4), *AI Act*, (Versione II, III, IV).

⁵⁵⁹ *Ex multis*, Mantelero (2022), 160, 173; vedi anche Catanzariti, M. (2023b). Rischio e vulnerabilità nel modello europeo di intelligenza artificiale. *Società Mutamento Politica: Rivista Italiana Di Sociologia*, 13(25), 81 <https://doi.org/10.36253/smp-13804>; vedi anche Wendehorst, C. (2022) AI Liability in Europe: Anticipating the EU Liability Directive, *Ada Lovelace Institute*, 7 <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/09/Ada-Lovelace-Institute-Expert-Explainer-AI-liability-in-Europe.pdf>; vedi anche Smuha, N. A., Ahmed-Rengers, E., Harkens, A., Li, W., Maclaren, J., Piselli, R., & Yeung, K. (2021). How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act. *Social Science Research Network*, 45-46 <https://doi.org/10.2139/ssrn.3899991>; vedi anche Alfieri, C., Carocchia, F., & Inverardi, P. (2022). AI Act and Individual Rights: A Juridical and Technical Perspective. *IAIL@HHAI*, 4, <https://www.semanticscholar.org/paper/AI-Act-and-Individual-Rights%3A-A-Juridical-and-Alfieri-Carocchia/96b164abc9e27c8945c5133511ee1429aef69043>.

⁵⁶⁰ Smuha *et al.* (2021), 45-46.

⁵⁶¹ Catanzariti (2023b), 81

⁵⁶² Mantelero (2022), 172-173.

differenza dell'*AI Act*, identifica nella figura degli “interessati al trattamento”⁵⁶³ i desinari dell’intera sovrastruttura di diritti e facoltà relativi alla tutela dei dati personali⁵⁶⁴: il regolamento sulla protezione dei dati, infatti, elenca chiaramente sia i diritti dell’interessato (cap. 3, artt. 12-23) sia i rimedi e le sanzioni (artt. 77-84) in caso di condotta illecita. Una sovrastruttura tanto dettagliata anche per coloro che interagiscono con gli strumenti di IA non è prevista, mettendo dunque in dubbio la capacità di proteggere realmente i *deployers*⁵⁶⁵.

3.5 Conclusioni sull’*AI Act*: il rapporto tra *AI Act* ed innovazione; la concreta possibilità di un *Effetto Bruxelles*

A fronte dell’analisi finora condotta è d’uopo, in primo luogo, interrogarsi sul rapporto tra regolamentazione dell’IA ed innovazione, precisamente in relazione al modello scelto dall’Unione Europea con l’*AI Act*.

Infatti, una delle critiche principali rivolte all’Unione Europea riguardo alla sua strategia regolamentare rispetto l’IA e il digitale complessivamente considerato, consiste nel fatto che, mentre gli Stati Uniti e la Cina sono attivi nella produzione di nuova tecnologia, l’UE si limiti solamente a regolarla, senza tuttavia contribuire significativamente alla sua creazione⁵⁶⁶. Tuttavia, non bisogna dimenticare che l’Unione Europea è comunque vincolata costituzionalmente a aderire ai valori fondanti europei, come sancito dall’art. 2 del TUE, senza dunque poter permettere che i principi fondanti della comunità siano compromessi anche se sulla scorta di regolamentazione più lassista nei confronti dell’IA, le cui capacità di compromettere i principi di cui all’art. 2 del TUE sono state abbondantemente esaminate⁵⁶⁷.

La recente regolamentazione europea ha dunque sollevato dubbi riguardo la concreta possibilità di costituire un ostacolo all’innovazione in futuro. In via generale, alcuni ritengono che le normative eccessive dell’UE possano generare un aumento dei costi di conformità e, in ultima analisi, limitare lo sviluppo di nuove tecnologie⁵⁶⁸. Tuttavia, come più volte considerato, l’adozione di chiara regolamentazione per i sistemi di IA potrebbe portare a una maggiore certezza

⁵⁶³ Art. 4(1) GDPR.

⁵⁶⁴ Solove, D. J., & Schwartz, P. M. (2020). *EU data protection and the GDPR*. Aspen Publishing, 77-83.

⁵⁶⁵ Alfieri *et al.* (2022), 5.

⁵⁶⁶ Bradford (2023), 136.

⁵⁶⁷ Smuha (2023), 9.

⁵⁶⁸ Bradford (2023), 137.

legale, prevedibilità e fiducia nel settore, potendo, auspicabilmente, comportare un vantaggio commerciale per i consumatori, nell'ottica di promuovere un'IA sicura, etica e affidabile⁵⁶⁹.

Eppure, secondo l'analisi di Bradford, la scelta di regolamentare il mercato del digitale in generale, ancorchè quello dell'IA, non costituisce, di per sé, quell'insormontabile ostacolo per innovazione di cui spesso si parla. Un dato è comunque da sottolineare: il ritardo nel settore dell'innovazione tecnologica da parte dell'UE, specialmente per quanto riguarda i sistemi di IA, può anche essere totalmente slegato dalla regolazione⁵⁷⁰. Nel momento in cui si scrive, in cui l'*AI Act* non è ancora entrato in vigore, la situazione delle *start-up* europee che producono IA è molto meno meritevole di ottimismo rispetto alle equivalenti statunitensi o cinesi, anche senza alcuna cornice normativa⁵⁷¹.

Secondo Bradford sono ben altri fattori a frenare la spinta innovativa dell'UE nel settore del digitale, e non il ricorso all'iper-regolamentazione: tra questi fattori figurano l'eccessiva frammentazione del mercato unico digitale, i mercati finanziari decisamente sottosviluppati rispetto agli altri paesi, le leggi sulla disciplina del fallimento punitive che non incoraggiano gli investimenti ancorchè rischiosi, e l'assenza di una politica migratoria proattiva che consenta all'UE di sfruttare il capitale umano nel settore della tecnologia e dell'IA⁵⁷².

In secondo luogo, è opportuno discutere della concreta possibilità che si verifichi nuovamente di un “*effetto Bruxelles*” anche per quanto riguarda l'*AI Act* e qualora ciò accada, in che misura. Il mondo accademico sembra essere diviso su questo tema, portando argomentazioni pertinenti ad entrambe le cause.

Secondo un recente studio, il verificarsi nuovamente di un “*effetto Bruxelles*”, non è necessariamente da escludersi⁵⁷³. Infatti, è auspicabile che si verifichino sia un *effetto Bruxelles* “*de facto*” che “*de iure*”⁵⁷⁴. Il primo mira ad indicare le conseguenze che si verificheranno a livello del mercato. Le imprese extra-UE, infatti, con l'entrata in vigore dell'*AI Act*, dovranno affrontare un *trade-off*: da un lato, la possibilità di creare diverse versioni dei propri prodotti di IA (una

⁵⁶⁹ *Id.*

⁵⁷⁰ *Id.*, 138, 370.

⁵⁷¹ Bradford (2023), 137-138.

⁵⁷² *Id.*, 369-376.

⁵⁷³ Siegmann, C., & Anderljung, M., (2022), *The Brussels Effect and Artificial Intelligence: How the EU regulation will impact the global AI market*, arXiv (Cornell University). <https://doi.org/10.48550/arxiv.2208.12645>.

⁵⁷⁴ Siegmann & Anderljung (2022), 3-5.

conforme agli standard dell'UE, e una non conforme), praticando una c.d. “*biforcazione precoce*” (*early forking*, cioè modificare caratteristiche fondamentali del prodotto nelle prime fasi del processo di sviluppo)⁵⁷⁵, finendo così per mantenere due pacchetti tecnologici separati in parallelo, ispirati a diversi standard; oppure, dall'altro, offrire un unico prodotto a livello globale, con un alta probabilità di ridurre i costi di sviluppo grazie alla non differenziazione dei prodotti⁵⁷⁶. Un siffatto effetto *de facto* ritrova la sua ragione d'essere in una serie di fattori quali ad esempio, l'attrattività di una non differenziazione dei costi a causa dell'essenza stessa del mercato dell'IA, che è comunque strutturato come un oligopolio costituito da poche società multinazionali che servono larghe porzioni di mercato⁵⁷⁷, che sarebbero più inclini a sopportare dei costi fissi di *compliance* per continuare a servire il mercato europeo, piuttosto che abbandonarlo totalmente⁵⁷⁸.

Tuttavia, è più probabile che uno effetto di influenza dell'*AI Act de facto* sul mercato globale dell'IA si verifichi per quei sistemi che sono considerati dalla normativa “ad alto rischio”⁵⁷⁹, che sono soggetti ad una serie di requisiti di sicurezza⁵⁸⁰. Dato il parallelo scopo di salvaguardia dei diritti fondamentali più volte rimarcato, salvaguardia che si esplica a mezzo dell'individuazione dei livelli di rischio posto dai sistemi di IA, ciò potrebbe portare i fornitori di sistemi considerati nel novero di quelli che pongono un rischio “alto” di adeguarsi agli standard europei, senza creare alcuna differenziazione di prodotti conformi e non, che potrebbero scoraggiare i consumatori finali. D'altronde, come è stato più volte sottolineato, la corrispondenza ad un utilizzo sicuro e *trustworthy* dei prodotti che pongono rischi maggiori comporta un significativo vantaggio competitivo sul mercato⁵⁸¹.

Il secondo, *de iure*, descrive l'attitudine a adottare regolamentazioni simili da parte di altre potenze globali, soprattutto quei paesi che tessono intense relazioni economiche con l'UE⁵⁸². Se infatti questi decidessero di introdurre una cornice normativa per i sistemi che sia assolutamente incompatibile con quella dell'UE, si creerebbero indubbiamente delle frizioni nel mercato⁵⁸³. L'UE, inoltre, potrebbe

⁵⁷⁵ *Id.*, 4.

⁵⁷⁶ *Id.*

⁵⁷⁷ *Id.*, 30-31.

⁵⁷⁸ *Id.*, 31.

⁵⁷⁹ Per i requisiti stabiliti nell'*AI Act* per i sistemi ad alto rischio, *vedi* il sottoparagrafo 3.3.2.

⁵⁸⁰ Siegmann & Anderljung (2022), 3-5.

⁵⁸¹ *Vedi* Smuha (2021b),

⁵⁸² *Id.*, 5.

⁵⁸³ *Id.*

promuovere la propria iniziativa di regolamentazione dell'*AI Act* anche a livello interazionale, partecipando alle negoziazioni e a conferenze internazionali⁵⁸⁴. Un approccio di questo tipo si sta osservando nel caso della Convenzione internazionale oggetto dei recenti lavori del Consiglio d'Europa⁵⁸⁵. Il Consiglio d'Europa⁵⁸⁶ è infatti impegnato nello strutturare la prima Convenzione internazionale per l'IA, la cui adozione si prevede come minimo per il prossimo anno, e tra i cui potenziali firmatari si annoverano, *ex multis*, anche gli Stati Uniti⁵⁸⁷. La struttura della futura Convenzione, ad oggi, ricalca per gran parte contenuto dell'*AI Act* soprattutto per quanto riguarda la scelta di adottare un approccio *risk based*⁵⁸⁸. Se si verificasse questo effetto Bruxelles *de iure*, sarebbe realizzato quel proposito europeo di creare una convergenza normativa in tema di IA. A questo proposito, Smuha ha rimarcato che vi è comunque un certo tasso di desiderabilità di siffatta convergenza di regole e standard, a livello globale.⁵⁸⁹

Altri, tuttavia, ritengono che vi siano scarse possibilità del verificarsi di un siffatto effetto Bruxelles. Tra le argomentazioni più convincenti, vi sarebbe ad esempio il fatto che l'*AI Act* e il GDPR non possano essere paragonabili in quanto, al di là delle loro differenze strutturali e contenutistiche, il regolamento sulla protezione dei dati comunque non rappresentava il primo approccio alla regolamentazione dei dati personali nell'Unione Europea⁵⁹⁰. A ragione, l'Unione Europea aveva provveduto a fornire una cornice normativa alla protezione dei dati personali dei residenti nell'Unione già nel 1995, con la Direttiva 95/46/CE⁵⁹¹. È quindi meno probabile il verificarsi di un effetto dirompente come quello Bruxelles

⁵⁸⁴ *Id.*

⁵⁸⁵ Tiulkanov, A. (18 dicembre 2023). *Council of Europe is Progressing on the First International AI Treaty – in the Newsletter AI, Data & Digital Policy Thoughts about current and emerging digital regulation and policy*. LinkedIn. <https://www.linkedin.com/pulse/council-europe-progressing-first-international-ai-treaty-tiulkanov-xp88e%3FtrackingId=2fKarTumQ46R6uT86HPujw%253D%253D/?trackingId=2fKarTumQ46R6uT86HPujw%3D%3D>.

⁵⁸⁶ Il Consiglio d'Europa è un'organizzazione internazionale indipendente, separata dal Consiglio dell'Unione Europea e dalle altre istituzioni dell'Unione. Il suo ruolo principale è stabilire standard nei settori dei diritti umani, democrazia e stato di diritto per 46 Stati, compresi tutti i 27 Stati membri dell'UE, oltre a nazioni al di fuori dell'Unione come Regno Unito e Turchia. Parte delle importanti convenzioni da esso emanate coinvolge anche paesi extraeuropei, come gli Stati Uniti, l'Australia, l'Argentina e il Messico. *Vedi Id. Vedi anche* Conforti, B. (2018). *Diritto Internazionale* (11^a ed.). Editoriale Scientifica.

⁵⁸⁷ Tiulkanov (2023).

⁵⁸⁸ *Id.*

⁵⁸⁹ Siegmann & Anderljung (2023), 30-31, 59-60.

⁵⁹⁰ Mantelero (2022), 189.

⁵⁹¹ Direttiva 95/46/CE del Parlamento Europeo e Del Consiglio del 24 ottobre 1995 relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati G.U. L 281 (*Data protection directive*).

nel caso di una normativa estremamente giovane, che non ha precedenti né a livello europeo né a livello globale⁵⁹².

Secondo altri ancora, l'effetto Bruxelles in quanto diretta conseguenza dell'adozione dell'*AI Act* è di natura esogena rispetto alla scelta di regolamentare l'IA a livello europeo: infatti, a determinare l'impatto di una siffatta normativa a livello globale, che renderebbe l'UE pioniera nella legislazione in questo settore, sono principalmente gli accordi bilaterali di mutuo riconoscimento che l'UE può concludere con altri paesi, nonché gli standard in materia riconosciuti dalle agenzie dedicate⁵⁹³. Questi fattori comporterebbero non tanto il venir meno dell'intera possibilità del verificarsi di un effetto Bruxelles, ma comprometterebbero tuttavia la genuinità alla base dello stesso⁵⁹⁴.

4. *Le direttive sulla responsabilità civile dei sistemi di IA*

Per contribuire allo sviluppo di sistemi di IA che siano sicuri, affidabili, e rispettosi dei diritti fondamentali nei confronti degli individui, l'Unione Europea ha recentemente affiancato al disegno di legge dell'*AI Act* anche una proposta di direttiva sulla responsabilità civile per danno causato da un sistema di IA, denominata *AI Liability Directive* (AILD)⁵⁹⁵. La direttiva, in combinazione all'aggiornamento di una preesistente direttiva relativa alla sicurezza dei prodotti⁵⁹⁶, la *Product Liability Directive* (PLD⁵⁹⁷), ha lo scopo di precisare delle disposizioni in tema della responsabilità civile derivante dai danni che i sistemi *AI-powered* sono in grado di causare, per i più svariati motivi: da semplici malfunzionamenti di sistema che possono causare eventuali danni (regolati dalla

⁵⁹² Mantelero (2022), 189.

⁵⁹³ Pagallo, U. (2023) *Dismantling Four Myths in AI & EU Law Through Legal Information 'About' Reality*, in Antunes et al. (2023). *Multidisciplinary perspectives on artificial intelligence and the law*. Springer Nature, 251-261, 255-256.

⁵⁹⁴ *Id.*

⁵⁹⁵ Proposta di Direttiva del Parlamento Europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale) (Testo rilevante ai fini del SEE) {SEC(2022) 344 final} - {SWD(2022) 318 final} - {SWD(2022) 319 final} - {SWD(2022) 320 final}, 28 settembre 2022. Questa proposta di direttiva ha lo scopo di aggiornare la previgente direttiva 85/374/CEE,

⁵⁹⁶ Direttiva sulla responsabilità da intelligenza artificiale, 13.

⁵⁹⁷ Proposta di Direttiva del Parlamento Europeo e del Consiglio sulla responsabilità per danno da prodotti difettosi (Testo rilevante ai fini del SEE) {SEC(2022) 343 final} - {SWD(2022) 315 final} - {SWD(2022) 316 final} - {SWD(2022) 317 final}, 28 settembre 2022.

revisionata PLD)⁵⁹⁸, a decisioni algoritmiche discriminatorie e lesive dei diritti fondamentali degli individui⁵⁹⁹. Le due direttive sono intese come un pacchetto unitario: mentre la PLD riguarda la responsabilità oggettiva del produttore per prodotti difettosi, determinando un risarcimento per alcuni tipi di danni subiti principalmente da persone fisiche, la AILD contempla le azioni di responsabilità a livello nazionale, principalmente per colpa di un unico soggetto, nell'ottica di risarcire qualsiasi tipo di danno e qualsiasi tipo di danneggiato⁶⁰⁰.

Nonostante la proposta di regolamento dell'*AI Act* sia più volte indicata come una disciplina sostanzialmente omnicomprensiva dell'IA, essa lascia ad oggi irrisolta l'annosa questione della responsabilità extracontrattuale derivante dai danni causati dall'IA, necessitando quindi un'integrazione da parte del legislatore europeo⁶⁰¹. A questo proposito, l'UE ha dovuto necessariamente selezionare uno strumento normativo diverso da un regolamento come l'*AI Act*, in quanto l'armonizzazione dei regimi di disciplina della responsabilità civile (estremamente compositi a livello di tutti gli Stati Membri⁶⁰²) e il risarcimento del danno non rientrano nelle competenze specifiche dell'UE⁶⁰³ per le quali è consentito l'utilizzo di un regolamento, atto normativo *erga omnes*, dotato delle sue tipiche caratteristiche di diretta e immediata applicabilità⁶⁰⁴. L'utilizzo di una direttiva consente dunque inoltre agli Stati Membri di implementare quanto previsto al suo interno secondo le proprie necessità, senza che si vengano a creare attriti con i regimi nazionali di responsabilità civile⁶⁰⁵.

La combinazione tra l'*AI Act* e le direttive sulla responsabilità e la sicurezza dei prodotti appena menzionate porterebbe, in linea generale, a delineare una cornice normativa non solo *ex ante* (dalle fasi di *training* fino all'immissione sul mercato), ma anche *ex post*, che possa, in ultimo, tutelare a tutto tondo coloro che interagiscono con i sistemi di IA. La definizione normativa di questo particolare

⁵⁹⁸ Commissione Europea. (28 settembre 2022). *Domande e risposte sulla revisione della direttiva sulla responsabilità per danno da prodotti difettosi*. https://ec.europa.eu/commission/presscorner/api/files/document/print/it/qanda_22_5791/QANDA_22_5791_IT.pdf.

⁵⁹⁹ Ziosi, M., Mökander, J., Novelli, C., Casolari, F., Taddeo, M., & Floridi, L. (2023). The EU AI Liability Directive: shifting the burden from proof to evidence, *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4470725>.

⁶⁰⁰ Direttiva sulla responsabilità da intelligenza artificiale, 3-4.

⁶⁰¹ Ziosi *et al.*, (2023), 2.

⁶⁰² Wendehorst (2022) 4

⁶⁰³ Schepisi (2023), 74.

⁶⁰⁴ Strozzi & Mastroianni (2018), 290-295.

⁶⁰⁵ Direttiva sulla responsabilità da intelligenza artificiale, 7-8.

aspetto legato all'IA risulta essere un compito tutt'altro che agevole, potenzialmente dimostrandosi più intricato della redazione stessa dell'*AI Act*, a causa della presenza dei diversi regimi di responsabilità vigenti nei vari ordinamenti⁶⁰⁶. A questo proposito, il Garante Europeo della Protezione dei Dati (GEPD) ha recentemente fornito delle proprie indicazioni sulla direttiva sulla responsabilità da IA, e, nello specifico, sulla gestione del rapporto tra l'AILD e l'*AI Act*⁶⁰⁷. Dalla relazione del GEPD emerge che la previsione di un regime di responsabilità per gli sviluppatori di sistemi di IA non deve essere interpretata come una struttura alternativa alle attuali disposizioni dell'*AI Act*, ma come una solamente disciplina integrativa di quanto proposto nel Regolamento⁶⁰⁸.

Alcuni obiettivi perseguiti dalla proposta di direttiva (AILD) rispecchiano le finalità precedentemente analizzate che l'*AI Act* persegue⁶⁰⁹: in primo luogo l'armonizzazione dei regimi legali per incrementare la certezza del diritto in questo settore⁶¹⁰; in secondo luogo, evitare il verificarsi di vuoti di responsabilità tra i fornitori e gli utenti, stimolando l'implementazione di meccanismi adeguati che possano garantire una più chiara attribuzione delle responsabilità; in terzo luogo, facilitare il processo di risarcimento dei danni per le parti danneggiate dall'utilizzo di sistemi di IA⁶¹¹. Tuttavia, anche nel caso della disciplina prevista dalla AILD, così come nell'*AI Act*⁶¹², vi sono delle finalità secondarie, non esplicitate nel testo normativo: tra queste figurano lo scongiurare risarcimenti insufficienti per le parti danneggiate⁶¹³, nonché, l'incrementare la fiducia del pubblico nei confronti dell'uso dell'IA per favorire un ambiente imprenditoriale che non impedisca l'innovazione⁶¹⁴. In ultimo, la ricerca di generare un “*effetto Bruxelles*”⁶¹⁵ anche in questo caso rappresenta senza dubbio una priorità per il legislatore europeo⁶¹⁶ nella ricerca di una maggiore

⁶⁰⁶ Wendehorst (2022), 6.

⁶⁰⁷ *EDPS Opinion 42/2023 on the Proposal for two Directives on AI liability rules*. (11 ottobre 2023). European Data Protection Supervisor. https://edps.europa.eu/data-protection/our-work/publications/opinions/2023-10-11-edps-opinion-422023-two-directives-ai-liability-rules_en.

⁶⁰⁸ *Id.*

⁶⁰⁹ *Vedi* il paragrafo 3.2.

⁶¹⁰ Analogamente a quanto rischiava di accadere con l'*AI Act*, l'inerzia del legislatore europeo poteva portare una proliferazione tra i vari Stati membri di profili di responsabilità civile diversi, ancorché incompatibili. *Vedi* Direttiva sulla responsabilità da intelligenza artificiale, 1.

⁶¹¹ Ziosi *et al.* (2023), 2.

⁶¹² Per gli obiettivi principali dell'*AI Act*, sia espliciti che non, *vedi* il paragrafo 3.2.

⁶¹³ Wendehorst (2022), 6-9, 10-13.

⁶¹⁴ Wendehorst (2022), 6-9.

⁶¹⁵ *Vedi* nota 319.

⁶¹⁶ Wendehorst (2022), 8-9.

armonizzazione a livello comunitario in materia di responsabilità civile degli strumenti di IA.

La direttiva, basata anch'essa sull'art. 114 del TFUE⁶¹⁷, focalizza la maggior parte dei propri obiettivi mediante l'alleggerimento dell'onere della prova a carico del soggetto ricorrente, il quale ha subito un danno tramite un sistema che utilizza IA. A questo proposito, l'art. 4 dell'AILD prevede che, qualora i danneggiati dimostrino che qualcuno, nell'intera catena di produzione del sistema di IA, ha agito con colpa, trascurando un preciso obbligo correlato al danno subito, e che esiste una ragionevole probabilità di connessione con l'*output* del sistema di IA, il giudice può liberamente presumere che tale mancato adempimento sia collegato da un nesso logico-causale con il danno stesso⁶¹⁸.

Insieme a ciò, la direttiva prevede un ulteriore strumento per rafforzare la posizione dei danneggiati a livello prettamente processuale, incrementando la facilità con cui essi possono raccogliere degli elementi di prova rilevanti. I danneggiati, infatti, in forza di quanto prevede l'art. 3, potranno chiedere al giudice di ordinare la divulgazione di informazioni sui sistemi di IA "*ad alto rischio*"⁶¹⁹. Ciò potrebbe consentire ai danneggiati, ad esempio, di identificare la persona che potrebbe essere ritenuta responsabile nella catena di sviluppo del sistema di IA, così da ridurre sia il divario di responsabilità (il c.d. *liability gap*⁶²⁰) che si crea a causa della difficoltà, tipica dei sistemi di IA, di risalire ai soggetti responsabili. La possibilità di richiedere accesso alle informazioni da parte di coloro che producono e mettono sul mercato sistemi di IA è comunque controbilanciata dal fatto che la divulgazione delle informazioni è limitata a quanto sia "necessario" e "proporzionato" per supportare una richiesta di risarcimento⁶²¹. Una limitazione di questo tipo funge da contrappeso per proteggere i segreti commerciali o informazioni riservate di interesse nazionale⁶²².

Ciò che tuttavia la proposta di direttiva sembra non prendere in considerazione, secondo Ziosi *et al.*, è la sostanziale asimmetria informativa che sussiste tra la parte attrice, l'utente danneggiato, e la parte convenuta: spesso e

⁶¹⁷ Direttiva sulla responsabilità da intelligenza artificiale, 5-6.

⁶¹⁸ Art. 4, Direttiva sulla responsabilità da intelligenza artificiale.

⁶¹⁹ La direttiva, dunque, utilizza la classificazione utilizzata nello stesso *AI Act* per caratterizzare i vari sistemi di IA, consentendo un collegamento normativo che aumenta l'unitarietà delle discipline volte a regolare l'IA a tutto tondo. Art. 3, Direttiva sulla responsabilità da intelligenza artificiale.

⁶²⁰ Ziosi *et al.* (2023), 2-3.

⁶²¹ Art. 3, Direttiva sulla responsabilità da intelligenza artificiale.

⁶²² Ziosi *et al.* (2023), 5.

volentieri, il ricorrente possono non essere addirittura a conoscenza di star interagendo con un sistema di IA, men che meno di essere i destinatari di un *output* lesivo, perché ad esempio discriminatorio⁶²³. Allo stesso tempo, gli utenti finali non hanno accesso alle informazioni relative agli output dei sistemi di IA utilizzati, non avendo dunque modo di sostenere adeguatamente in giudizio le proprie ragioni a causa della suddetta asimmetria informativa⁶²⁴. Per questo motivo, una delle raccomandazioni fornite da Ziosi *et. al.*, è quella di incentivare meccanismi di condivisione delle informazioni ma soprattutto di notifica diretta all'utente nel caso di violazioni danni subiti⁶²⁵. In ultimo, un'altra delle raccomandazioni più generali fornite da parte degli accademici, è in relazione alla creazione di incentivi per sviluppare sistemi di IA che siano meno complessi ed opachi, *by design* e *by default*, incrementando la trasparenza *ex ante* e non solo facendo un eccessivo affidamento ad obblighi di trasparenza *ex post*⁶²⁶.

⁶²³ *Id.*, 4.

⁶²⁴ *Id.*

⁶²⁵ *Id.*, 6.

⁶²⁶ *Id.* Vedi anche Rudin & Radin (2019).

Capitolo 3: le altre potenze mondiali e la regolamentazione ideale dell'IA

Il tema della regolamentazione dell'IA non si esaurisce unicamente con l'analisi delle normative europee che si prefiggono l'obiettivo di offrire un quadro disciplinare omnicomprensivo nei confronti di questa tecnologia. La natura assolutamente transnazionale dello sviluppo dei sistemi di IA, in un mondo indissolubilmente interconnesso, non permette di valutare in maniera isolata il solo regolamento in via di approvazione da parte dell'Unione Europea, sebbene questo, ad oggi, sia l'unico testo con l'ambizione di offrire una struttura di *hard law* corredato di sanzioni, e con un approccio basato sul rischio. Sarebbe dunque un errore limitarsi ad un'analisi strettamente legata alla legislazione europea, senza tener conto di quei paesi che hanno cercato di individuare una strategia regolamentare per l'IA, seppure in maniera parzialmente o totalmente divergente da quella dell'Unione. L'approccio comparatistico tra paesi con diverse culture legislative, nonché con diversi approcci all'IA e al progresso tecnologico in generale, consente di completare quel quadro generale sulla regolamentazione dell'IA che si è cercato di delineare, almeno a livello europeo. Ciò che è tuttavia certo, nonostante il panorama composito relativo alla *race to AI regulation*⁶²⁷ correntemente in corso, è che ad oggi si rende necessaria la predisposizione di un quadro di regole secondo le più varie modalità (che siano di *hard* o *soft law*) per limitare quegli effetti negativi che i sistemi di IA sono in grado di portare con sé in qualsiasi campo di applicazione, soprattutto quelli in cui vi è una maggiore possibilità di intaccare i diritti fondamentali degli individui e della collettività.

A questo proposito, l'apparente dicotomia tra regolamentazione e innovazione è stata superata, grazie alla previsione di una serie di ipotesi graduate di legislazione in questa materia (più o meno tendenti alla *soft law*, *hard law*, o *co-regulation*⁶²⁸ che integra entrambi i modelli), fissando dei principi "eticici" di utilizzo, o di parametri normativi che possano meglio indirizzare l'innovazione stessa. L'idea per la quale l'imposizione di qualsiasi tipo di regolamentazione sull'IA costituisca semplicemente un impedimento all'innovazione è (abbastanza

⁶²⁷ Per la definizione di *Race to AI regulation*, vedi nota 312.

⁶²⁸ Mantelero (2022), 142.

pacificamente) superata⁶²⁹. La questione ora al centro dell'attenzione sia nel settore pubblico che nel privato riguarda le modalità di regolamentazione dell'IA, piuttosto che il quesito se questa debba essere regolamentata o meno.

I diversi approcci scelti dai vari paesi rispecchiano dunque quelle tendenze legislative tipiche in relazione al settore digitale totalmente considerato. Infatti, a questo proposito, si è scelto di portare avanti un'analisi considerando due *super-players* del mercato dell'IA: gli Stati Uniti d'America e la Repubblica Popolare Cinese. Questa scelta si pone in continuità con quanto analizzato nel paragrafo 1 del precedente Capitolo 2, in cui si è comparato l'orientamento delle due potenze globali (insieme alla terza, l'UE) che Bradford, denomina come “imperi digitali.”⁶³⁰ Secondo quanto sostenuto dall'ideatrice del concetto di “*Brussels effect*”, l'Unione Europea, la Cina e gli USA sono, in modo diverso tra loro e in modalità differenti “imperi digitali,” che rappresentano tre diverse versioni di come si può diversamente esplicitare un sistema capitalistico digitale fortemente avanzato in un mondo iper-connesso⁶³¹.

Nonostante la scelta di soffermarsi su queste tre entità sovranazionali, non bisogna tuttavia concludere che non vi siano interessanti spunti di principi relativi all'utilizzo di IA che sono stati recentemente introdotti da altre nazioni. Nel corso degli ultimi cinque anni, in diversi paesi sono stati adottati principi concernenti lo sviluppo, l'applicazione e l'utilizzo di questi sistemi, con un *focus* più o meno specifico su settori particolari. A questo proposito, sono da menzionare i lavori del Brasile, del Giappone, e dell'Australia.

In Brasile, è stata recentemente proposto un disegno di legge che mira a costruire una cornice normativa per i sistemi di IA secondo un approccio basato su una matrice cumulativa di rischi⁶³²: a differenza dell'*AI Act* dell'UE, la proposta di legge brasiliana considera separatamente i rischi per i diritti umani e l'impatto su

⁶²⁹ Vedi Weaver, J. F., (2018) *Regulation of artificial intelligence in the United States*, in Barfield, W. & Pagallo, U., *Research Handbook on the Law of Artificial Intelligence* (2018), Elgar, 155. Vedi anche Smuha (2021b). Vedi anche Bradford (2023), 351.

⁶³⁰ Gli “imperi digitali” di cui parla Bradford sono gli USA, la Cina e l'UE. Nel definirli “imperi,” Bradford non si limita a considerare unicamente gli aspetti di crescita e sviluppo dell'*industry* dell'IA: queste potenze sono considerate imperi in forza del potere tecnologico, economico e regolamentare che essi hanno, in considerazione con la loro capacità di influenzare l'ordine digitale mondiale verso l'una o l'altra tendenza. Bradford (2023), *Introduction*, 7.

⁶³¹ *Id.*

⁶³² Vedi per la legislazione brasiliana, Emenda -CTIA (Substitutivo) ao Projeto de Lei 2.338/2023.

ogni singolo diritto/libertà⁶³³. Inoltre, il Brasile è stato uno dei primi paesi ad approvare l'apertura di una pubblica consultazione sponsorizzata dall'agenzia di protezione dei dati brasiliana, con la quale facilitare l'istituzione di una *regulatory sandbox*, in cui sperimentare collaborativamente l'IA all'interno di uno schema normativo predefinito, attraverso una metodologia strutturata di cooperazione tra vari *stakeholders*⁶³⁴.

Il Giappone, analogamente, ha in un primo momento approcciato l'IA con un criterio regolamentare fondato su dei principi (i *Social principles of Human-Centric AI*, anche detti *Social Principles*, redatti nel 2019⁶³⁵), che sembrano ricalcare in gran parte le linee guida europee dell'HLEGAI⁶³⁶ e comunque quel modello tipicamente *rights-driven* europeo⁶³⁷. All'emanazione di questi principi, tuttavia, non è stato dato seguito con un testo di legge di *hard law*, ma solo linee guida a adesione volontaria⁶³⁸, combinando quell'attitudine tutta europea alla tutela dei diritti e delle libertà fondamentali con l'approccio *market-driven*, tendenzialmente più lassista, degli USA. I+

Infine, nel 2019 anche l'Australia ha stilato delle linee guida simili a quelle europee⁶³⁹, focalizzandosi sull'utilizzo dell'IA secondo dei principi "etici" che assicurassero la sicurezza e l'affidabilità di questi sistemi⁶⁴⁰. L'Australia sembra invece replicare *in toto* l'approccio dell'Unione Europea nel fornire un set di regole all'IA, il che sarebbe comprensibile tenendo conto che entrambe hanno concluso un accordo di cooperazione in aree critiche come l'IA e la *blockchain*⁶⁴¹.

Tutti questi orientamenti, seppur interessanti per quanto riguarda il tema più generale della *governance* dell'IA a livello globale, non provengono da una vera e propria fonte del diritto come invece accade con l'*AI Act*, ma rappresentano comunque un valido indicatore del fatto che l'attenzione per la regolamentazione

⁶³³ Mantelero, A. (30 novembre 2023). Alessandro Mantelero on LinkedIn: PL 2338-2023 (Inteligência artificial) - Emenda 1. <https://www.linkedin.com/feed/update/urn:li:activity:7136124845989257216/>

⁶³⁴ Mari, A. (13 ottobre 2023). Brazil consults on regulatory sandbox for AI and data protection. *Forbes*. <https://www.forbes.com/sites/angelicamarideoliveira/2023/10/13/brazil-consults-on-regulatory-sandbox-for-ai-and-data-protection/>.

⁶³⁵ *Regulation of artificial intelligence in Europe and Japan*. (24 agosto 2020). White & Case LLP International Law Firm, Global Law Practice. <https://www.whitecase.com/insight-our-thinking/regulation-artificial-intelligence-europe-and-japan>.

⁶³⁶ Vedi il paragrafo 2 del Capitolo 2.

⁶³⁷ Bradford (2023), 349-350.

⁶³⁸ *Id.*

⁶³⁹ Le *Ethics Guidelines* emanate dell'Unione Europea nel 2019. Vedi nota 352.

⁶⁴⁰ *Id.*, 350-351.

⁶⁴¹ *Id.*, 351.

dell'IA non è appannaggio della sola Unione Europea, nonostante il suo ben noto potere regolatore nel settore digitale⁶⁴².

PARTE I – GLI STATI UNITI D'AMERICA

I.1 La posizione statunitense nel mercato dell'IA e l'approccio alla regolamentazione

Gli Stati Uniti d'America possono vantare di essere uno dei Paesi *leader* e degli innovatori più influenti all'interno del mercato globale dell'IA, il cui volume di mercato auspicabilmente raggiungerà la cifra di 106,50 miliardi di dollari nel corso del 2024⁶⁴³. Si prevede che le dimensioni del suddetto mercato mostreranno un tasso di crescita annuale del 14,27%, con un volume di 237,10 miliardi di dollari entro il 2030⁶⁴⁴. Gli Stati Uniti ricoprono le posizioni apicali in diversi indici di analisi del ruolo di un dato paese nel settore dell'IA, come ad esempio nei *rankings* globali sull'entità degli investimenti pubblici per l'IA, il numero di brevetti complessivi relativi all'IA depositati (oltre 58.000 tra il 2016 e il 2021⁶⁴⁵), il numero di accordi che coinvolgono aziende di IA e la presenza di istituzioni influenti che promuovono il capitale umano all'interno del settore. Inoltre, sei delle c.d. “*Big Nine*” che dominano il panorama dell'IA hanno sede negli Stati Uniti: Google, Amazon, Apple, Meta, IBM e Microsoft⁶⁴⁶. Queste aziende hanno avuto un impatto significativo sullo sviluppo di IA, spesso non unicamente limitato ai meri profitti aziendali (comunque ingenti nel caso di tutte le aziende precedentemente menzionate), quanto piuttosto alla qualità e alla direzione dell'innovazione: esse hanno la possibilità, tramite l'investimento di ingenti capitali, di sfruttare il potere dell'IA per creare soluzioni tecnologiche innovative. Ancora, negli USA è prevista una crescita dei ricavi complessivi dagli attuali 42,7 miliardi di dollari, ad una cifra superiore a 128,8 miliardi di dollari entro il 2028⁶⁴⁷.

⁶⁴² Vedi in generale Smuha (2021b).

⁶⁴³ *Artificial Intelligence - United States*. Statista. [Grafico] (n.d.) <https://www.statista.com/outlook/tmo/artificial-intelligence/united-states?currency=USD>.

⁶⁴⁴ *Id.*

⁶⁴⁵ Davies, N., (10 novembre 2021). Index shows US is winning the AI race – but for how long? *Investment Monitor*. <https://www.investmentmonitor.ai/ai/ai-index-us-china-artificial-intelligence/>

⁶⁴⁶ *Id.*

⁶⁴⁷ Revenues from the artificial intelligence market in North America from 2017 to 2028 (in billion U.S. dollars). [Grafico]. Statista. (19 gennaio 2022) <https://www.statista.com/statistics/721748/north-america-artificial-intelligence-market/>.

In questo panorama di florida innovazione, gli Stati Uniti hanno dunque esitato nell'implementazione di schemi normativi sull'IA, lasciando il mercato sostanzialmente libero da parte di interventi governativi mirati alla disciplina di questa tecnologia per un lungo periodo. Il modello regolamentare adottato dagli USA ha inciso nel rafforzare le capacità di espansione del mercato dell'IA statunitense e dell'intero settore della tecnologia ad esso riconnesso. In un effetto di questo tipo, si può individuare il tradizionale approccio *market-driven*⁶⁴⁸ da parte delle istituzioni americane, principalmente ispirato alle logiche di mercato e alla massimizzazione dei profitti aziendali, che si riconosce in tutto il settore del digitale statunitense totalmente considerato. La base ideologica di questo modello regolamentare è da ritrovarsi in quel tecno-ottimismo di matrice tendenzialmente cyber-libertaria⁶⁴⁹ particolarmente diffuso nei grandi *hub* di innovazione come la Silicon Valley⁶⁵⁰. La strategia statunitense di ritardare il più possibile la creazione di cornici normative *ad hoc* per l'IA è direttamente collegata ad un retaggio di fine secolo scorso di "*permissionless innovation*" (cioè innovazione che avviene "senza permessi" o comunque senza l'approvazione di autorità governative o altri enti regolatori di alcuna sorta) che era stata attivamente implementata per consentire lo sviluppo di Internet⁶⁵¹. Alcuni hanno argomentato che l'espansione e il consolidamento di Internet nell'economia digitale sono stati possibili solamente grazie al suddetto contesto largamente permissivo e privo di regolazione precoce⁶⁵².

Un'argomentazione del genere riconosce nella coppia concettuale "innovazione-regolamentazione" un chiaro *trade-off*, una dicotomia assoluta nella

⁶⁴⁸ Vedi paragrafo 1 Capitolo 2.

⁶⁴⁹ Chiamato anche tecno-libertarismo, il cyberlibertarismo è una delle cinque grandi teorie generali dell'*Internet governance*: esso comporta una visione assolutamente non interventista nei confronti della regolamentazione della tecnologia e di Internet, con la ferma convinzione che lo stato non ha alcuna autorità per imporre leggi nel *cyberspace*, sulla scorta di un pensiero ispirato al *laissez-faire* tipico del capitalismo di primo '900. Il cyberlibertarismo e la tradizionale dottrina libertaria concordano per quanto riguarda i valori di massimizzazione della libertà individuale e di sfiducia nel governo e nel suo intervento economico *tout court*. Vedi Koenig, T., & Rustad, M. (2021). *Global Information Technologies: Ethics and the Law*, West Academic Publishing, 80-81. Vedi anche Rustad, M. L. (2021). *Global Internet Law in a nutshell*, West Academic Publishing, Nutshell Series, 34-36.

⁶⁵⁰ Vedi Bradford (2023), 33-40.

⁶⁵¹ Si vedano a questo proposito, le iniziative come la *Framework for Global Electronic Commerce* emanata dall'amministrazione Clinton nel 1996, in cui si prevedeva che "*the Internet should develop as market driven arena and not a regulated industry.*" Vedi *Read the Framework for Global Electronic Commerce*, White house <https://clintonwhitehouse4.archives.gov/WH/New/Commerce/read.html>.

⁶⁵² Thierer, A. (2016). *Permissionless Innovation: the continuing case for comprehensive technological freedom*. *Mercatus Center at George Mason University*, 12-16.

quale dove prevale l'una, l'altra viene irrimediabilmente compressa⁶⁵³. Il primo approccio regolamentare negli USA, dunque, è stato mutuato dalla precedente esperienza relativa al successo di Internet: è stato considerato altresì allettante proseguire con una strategia simile nel campo dell'IA, consentendo la ricerca e lo sviluppo iniziali con una struttura regolamentare molto limitata e creandone una solo in un secondo momento, quando il settore si sarà ulteriormente sviluppato⁶⁵⁴.

Questa generale ritrosia, tuttavia, non è indice di una mancanza di interesse generale da parte del legislatore statunitense nei confronti di eventuali leggi per disciplinare l'IA: il 116° Congresso (quel periodo di legislatura che va dal 1° gennaio 2019 al 3 gennaio 2021) è stata la sessione congressuale più incentrata sull'IA della storia, con più di 486 menzioni dell'espressione "IA" nei lavori parlamentari rispetto a quanto emerge nei rapporti pubblicati dalle commissioni congressuali e nei rapporti del servizio di ricerca del Congresso degli Stati Uniti⁶⁵⁵.

Alcuni, contrariamente, hanno voluto interpretare la prolungata assenza di un quadro normativo per l'IA da parte degli Stati Uniti come una mancata opportunità di generare un c.d. "*Washington effect*", che avrebbe potuto permettere di influenzare positivamente l'ecosistema digitale globale, dato il potere economico e il peso specifico del Paese nel settore dell'IA⁶⁵⁶.

Ad ogni modo, anche il modello di regolazione statunitense *market-driven*, ha nel tempo subito delle attenuazioni, adattatosi con la crescente consapevolezza dei rischi intrinseci individuali e collettivi che l'IA è in grado di portare con sé, e con l'aumento di casistiche pratiche in cui l'IA ha causato danni concreti. Infatti, se in principio era possibile sostenere che non vi fosse alcuna necessità di disporre una cornice regolamentare di alcun genere⁶⁵⁷, il crescente avanzamento tecnologico

⁶⁵³ *Id. Vedi anche* Khanna, D. (15 settembre 2015). Regulations stifle innovation. *The Hill*. <https://thehill.com/blogs/congress-blog/technology/253625-regulations-stifle-innovation/>

⁶⁵⁴ Weaver (2018), 165.

⁶⁵⁵ Il report dell'università di Stanford non presenta (ancora) i dati relative ai congressi successivi al 116°: la previsione che si può avanzare è tuttavia che il dato di "menzioni" dell'IA nei lavori parlamentari possa solamente aumentare. *Vedi* Number of times when artificial intelligence (AI) was mentioned in congressional record in the United States from 2001 to 2020 [Grafico]. *Statista*. (5 marzo 2021). <https://www.statista.com/statistics/1233574/united-states-congress-artificial-intelligence-mention/>

⁶⁵⁶ Mökander, J., Juneja, P., Watson, D.S., Floridi, L. (2022). The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other? *Minds & Machines* 32, 757 <https://doi.org/10.1007/s11023-022-09612-y>.

⁶⁵⁷ Si vedano ad esempio i risultati della consultazione pubblica avutasi nel giugno 2016 negli USA nell'ambito dell'iniziativa dell'*Executive Office of the (US) President* e del *National Science and Technology Council Committee on Technology* denominata "*Preparing for the future of Artificial Intelligence*." Gran parte dei commentatori coinvolti nella pubblica consultazione hanno sostenuto che "*una regolamentazione ampia della ricerca o della pratica dell'IA sarebbe sconsigliabile in*

degli ultimi anni in questo settore ha progressivamente aumentato la necessità di predisporre almeno delle linee guida che informassero meglio l'utilizzo e lo sviluppo dei sistemi *AI-powered*. Il cambiamento di approccio è particolarmente visibile analizzando come hanno operato le diverse amministrazioni presidenziali che si sono avvicinate nell'ultimo decennio.

1.2 Le iniziative durante la Presidenza Donald Trump (2016-2020)

Durante l'amministrazione Trump, l'obiettivo scelto era quello di garantire e mantenere l'eccellenza e la *leadership* americana nell'economia digitale⁶⁵⁸; pertanto, la creazione di un quadro giuridico per l'IA era considerata un potenziale ostacolo all'innovazione e un limite alle possibilità di raggiungere tale scopo. Tuttavia, lo sviluppo di IA è stato oggetto di dibattito e ha portato all'emanazione dell'*Executive Order* No. 13859, che ha predisposto la cosiddetta "*American Initiative on AI*", un'iniziativa governativa che ha stabilito una serie di principi guida. Tra questi, tre principi riguardavano direttamente la necessità che l'America si ponesse alla guida dell'innovazione con scoperte tecnologiche, creazione di standard di sviluppo volti a ridurre le "*barriere alla sperimentazione e all'impiego sicuro dell'IA*" e alla promozione di un ambiente internazionale di innovazione che sostenesse la ricerca e le industrie del *made in USA*⁶⁵⁹.

Nell'affrontare queste priorità, l'*Executive Order* No. 13859 ha taciuto sull'opportunità di una regolamentazione vera e propria, delegando al *Director of the Office of Management and Budget* statunitense (l'"OMB") la redazione di un documento integrativo che informasse "*lo sviluppo di approcci normativi e non normativi da parte di [tutte le] agenzie federali, relativamente alle tecnologie e ai settori industriali, potenziati o abilitati dall'IA.*"⁶⁶⁰ L'*Executive Order* No. 13859 ha dunque predisposto un approccio prettamente di settore, trattando unicamente l'implementazione dell'IA nel settore pubblico (precisamente da parte delle agenzie federali), senza però fare alcun riferimento alla sua diffusione nel settore privato⁶⁶¹.

questo momento". Vedi Executive Office of the President & National Science and Technology Council Committee on Technology. (Ottobre 2016). *Preparing for the Future of Artificial Intelligence*.

https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.

⁶⁵⁸ Vedi in generale Exec. Order No. 13859, 84 Fed. Reg. 3967 § 1 (14 febbraio 2019).

⁶⁵⁹ Exec. Order No. 13859, 84 Fed. Reg. 3967 § 1 (a), (b), (c), (d), (e).

⁶⁶⁰ Exec. Order No. 13859, 84 Fed. Reg. 3967 § 6 (i)

⁶⁶¹ *Id.*

L'OMB ha pubblicato il suddetto documento integrativo nel novembre 2020. Anche in questo caso, si è mantenuto un approccio sulla stessa lunghezza d'onda dell'*Executive Order* No. 13859, non dando alcuno spazio al tema relativo ai i danni potenziali dell'IA e la loro prevenzione. Questo documento, con una marcata impostazione deregolamentare, presenta una serie di linee guida delle quali le agenzie federali devono tener conto nell'adozione di qualsiasi cornice regolamentare di settore relativa ai sistemi che le agenzie federali possano implementare nelle loro attività⁶⁶². Tuttavia, alcuni hanno criticato la scelta di stabilire queste linee guida, non solo perché sembrano ribadire il *laissez-faireism* all'interno dell'industria dell'IA⁶⁶³, ma anche perché sembrano suggerire di escludere quadri normativi più esigenti che le agenzie stesse potrebbero voler applicare⁶⁶⁴.

In questo contesto, il Congresso ha approvato una legge con appoggio bipartisan denominata *National Artificial Intelligence Initiative Act* nel gennaio 2021, pochi giorni dopo il cambio dell'amministrazione presidenziale in favore dell'attuale presidente Joe Biden⁶⁶⁵. La legge stabilisce un piano di investimenti coordinato e accelerato verso l'utilizzo dell'IA nel settore pubblico; crea un ufficio specifico che supervisiona questo processo, il *National Artificial Intelligence Initiative Office* (NAIIO); definisce un quadro generale per rafforzare la ricerca sull'IA tra le istituzioni, con la collaborazione dell'industria, delle organizzazioni *no profit* e del mondo accademico. Nel momento della sua emanazione, questa legge era lo strumento legislativo più vicino a una strategia nazionale sull'IA approvato dal Congresso, anche se i suoi effetti erano ancora una volta limitati unicamente al settore pubblico, in quanto non era coperta in nessuno modo l'IA impiegata nel settore privato⁶⁶⁶.

⁶⁶² Vedi in generale Office. Of Management & Budget, Executive Office of The President, (17 novembre 2020). *Memorandum For the Heads of The Executive Departments and Agencies: Guidance for Regulation of Artificial Intelligence Applications*.

⁶⁶³ MacCarthy, M., (9 marzo 2020). AI needs more regulation, not less, *Brookings*. <https://www.brookings.edu/research/ai-needs-more-regulation-not-less/>.

⁶⁶⁴ Engler, A., (8 dicembre 2020) New White House guidance downplays important AI harms, *Brookings*. <https://www.brookings.edu/blog/techtank/2020/12/08/new-white-house-guidance-downplays-important-ai-harms/>.

⁶⁶⁵ National Artificial Intelligence Initiative Act of 2020, H.R. 6216, 116th Cong (2020).

⁶⁶⁶ Samp, T. (26 dicembre 2020). Landmark artificial intelligence legislation should become law, *The Hill* <https://thehill.com/opinion/technology/531680-landmark-artificial-intelligence-legislation-should-become-law/>.

I.3 La Biden Administration: dall'AI Bill of Rights all'Executive Order No. 14110

Con il cambio di amministrazione Trump-Biden, si è definitivamente osservato quel già menzionato cambiamento di orientamento in relazione a come approcciare il tema della regolamentazione dell'IA. Innanzitutto, è da sottolineare il tentativo, seppur vano, di approvare una proposta di legge presentata da parte del partito democratico statunitense, denominata l'*Algorithmic Accountability Act* (AAA)⁶⁶⁷. L'AAA, introdotto nel 2022, proponeva un impianto normativo mirato ai c.d. "*automated-decision systems*" (anche detti ADSs, sistemi di decisione automatizzata), accantonando dunque l'ambizioso progetto di fornire una definizione unica e omnicomprensiva ai sistemi di IA in generale⁶⁶⁸. L'ambito di applicazione dell'AAA avrebbe interessato solo le aziende con un fatturato annuo superiore a 50 milioni di dollari statunitensi o che elaboravano dati da più di un milione di utenti⁶⁶⁹. L'applicazione così delimitata del disegno di legge evitava di porre eccessivi vincoli agli ADSs utilizzati a livello governativo ed al contempo metteva al riparo le PMI dalla previsione di stringenti requisiti, spesso forieri di costi di *compliance* eccessivi⁶⁷⁰.

Tuttavia, nonostante queste interessanti caratteristiche della proposta di legge, questa non ha superato il vaglio parlamentare a causa della mancanza di un appoggio bipartisan necessario⁶⁷¹. A ragione, si era dubitato della possibilità dell'AAA di essere approvato fin dal momento della sua introduzione, in quanto il testo mancava di elementi che controbilanciassero i tratti regolamentari *tout court* con degli stimoli per l'innovazione, tali da stimolare l'appoggio da parte del partito repubblicano strumentale al raggiungimento della maggioranza necessaria⁶⁷². Allo stesso tempo, Mökander *et al.* ricordano comunque il clima sfavorevole per la regolamentazione della tecnologia in generale⁶⁷³, rappresentato dal fallimento, tra

⁶⁶⁷ Algorithmic Accountability Act of 2022, H.R.6580, 117th Cong. (2022) (d'ora in avanti abbreviata in "H.R.6580").

⁶⁶⁸ H.R.6580, § 2(2).

⁶⁶⁹ H.R.6580, § 2(7)(A)(i).

⁶⁷⁰ Mökander *et al.* (2022), 753.

⁶⁷¹ *Digital Policy Alert - United States of America: Rejected Algorithmic Accountability Act of 2022* (3 gennaio 2023). <https://digitalpolicyalert.org/event/8384-rejected-algorithmic-accountability-act-of-2022>.

⁶⁷² Mökander *et al.* (2022), 757.

⁶⁷³ Kang, C. (23 aprile 2022). As Europe approves new tech laws, the U.S. falls further behind. *The New York Times*. <https://www.nytimes.com/2022/04/22/technology/tech-regulation-europe-us.html>.

gli altri, dell'*American Innovation and Choice Online Act* del 2022 (un disegno di legge per certi versi paragonabile al *Digital Service Act* dell'Unione Europea), e del *Data Protection Act* del 2021 (anche questo, analogo al GDPR⁶⁷⁴).

Come accennato in precedenza, il panorama della regolamentazione dell'IA ha subito un cambiamento di passo con l'ultima amministrazione presidenziale. La pubblicazione nell'ottobre 2022 di un libro bianco emanato dall'Ufficio della Politica Scientifica e Tecnologica della Casa Bianca denominato il "*Blueprint for an AI Bill of Rights: making automated systems work for the American people*"⁶⁷⁵ è stata accolta come il primo reale tentativo di regolamentazione sostanziale dell'IA negli USA⁶⁷⁶. Il *Blueprint* prevede una serie di principi e relative pratiche per informare meglio il procedimento di *design* e impiego di sistemi automatizzati, in modo tale da fornire una serie di protezioni nei confronti dei diritti fondamentali della popolazione americana⁶⁷⁷.

Il cambiamento di politica normativa all'interno del governo degli Stati Uniti è palpabile dal tono adottato dal *Blueprint*. Infatti, se da un lato l'*Executive Order* No. 13859 menzionava la capacità dell'IA di "*migliorare la qualità della vita*" degli americani nelle prime righe dell'*incipit*⁶⁷⁸, il *Blueprint* invita invece a considerare i sistemi automatizzati "*tra le grandi sfide poste oggi alla democrazia [...] in modi che minacciano i diritti dei cittadini americani.*"⁶⁷⁹ Contrariamente all'*Executive Order* precedentemente menzionato, e alla *Guidance for Regulation of AI* dell'OMB, questo libro bianco inizia ad esporre i danni che hanno già colpito la popolazione americana. Questi sistemi automatizzati hanno, tra l'altro, negato

⁶⁷⁴ Mökander *et al.* (2022), 757.

⁶⁷⁵ The White House Office of Science and Technology Policy, *Blueprint for an AI Bill of Rights: making automated systems work for the American people* (Ottobre 2022), (d'ora in avanti abbreviato in "*Blueprint*").

⁶⁷⁶ Wiley, M., (4 ottobre 2022). *AI Bill of Rights Blueprint Can Protect Individuals from Bias and Discrimination, Implementation Will Be Critical. The Leadership Conference on Civil and Human Rights*. <https://civilrights.org/2022/10/04/ai-bill-of-rights-blueprint-can-protect-individuals-from-bias-and-discrimination-implementation-will-be-critical/>.

⁶⁷⁷ *Blueprint*, 14.

⁶⁷⁸ Exec. Order No. 13859, 84 Fed. Reg. 3967 § 1.

⁶⁷⁹ *Blueprint*, 3.

un'adeguata assistenza medica nei confronti dei pazienti, ovvero hanno effettuato valutazioni circa il *credit score*⁶⁸⁰ di cittadini americani non chiare e distorte⁶⁸¹.

Il libro bianco ha una struttura particolare: pur presentando i cinque capisaldi fondanti nelle sue pagine iniziali, esso contiene un c.d. “*technical companion*” che supporta ogni principio con delle raccomandazioni approfondite rivolte direttamente a coloro che sviluppano i modelli di IA⁶⁸².

Il primo principio prevede l'introduzione di “sistemi sicuri ed efficaci” per evitare danni tramite test preliminari all'implementazione del sistema e valutazioni indipendenti⁶⁸³. Nell'ambito di questo principio, il *Blueprint* è simile alla proposta di legge sull'IA dell'Unione Europea: il *technical companion*, infatti, propone valutazioni del rischio da svolgere prima di distribuire il sistema sul mercato e prescrive la qualità dei dati in ogni fase di sviluppo⁶⁸⁴, approcci che ricordano molto il modello *risk-based* dell'*AI Act*. Il secondo, attiene la “protezione dalla discriminazione algoritmica”, richiedendo ai progettisti di IA di aumentare la loro cautela nei confronti di eventuali rischi discriminatori⁶⁸⁵. Il *Blueprint* non si limita ad affrontare il problema della potenziale discriminazione algoritmica a livello teorico, ma fornisce piuttosto esempi pratici e reali di casi di discriminazione⁶⁸⁶. Tra gli altri, il *Blueprint* cita anche il celebre l'algoritmo difettoso del sistema di valutazione dei *curricula* di Amazon⁶⁸⁷. Il terzo principio del *Blueprint*, “privacy dei dati”, riguarda principalmente la *governance* dei dati utilizzati nell'addestramento algoritmico attraverso protezioni in relazione alla qualità e alla sicurezza dei *dataset* integrate nei sistemi di IA⁶⁸⁸. Questo principio non è strettamente riconducibile al mero sviluppo o all'implementazione di sistemi di IA sul mercato: tuttavia, come più volte sottolineato soprattutto nel Capitolo 1, la

⁶⁸⁰ Per *credit scoring* si intende quell'indicatore, solitamente un'unica cifra numerica, con cui banche, istituti di credito o società che offrono servizi vari, effettuano una previsione del comportamento creditizio dei loro clienti, come ad esempio la probabilità di rimborsare un prestito in tempo, basata sulle informazioni contenute nei rapporti di credito. Talvolta il punteggio ottenuto dipende anche da una serie di dati esterni, le cui computazioni sono affidate ad algoritmi anche *AI-powered*. *What is a credit score? | Consumer Financial Protection Bureau*. (28 agosto 2023). Consumer Financial Protection Bureau. <https://www.consumerfinance.gov/ask-cfpb/what-is-a-credit-score-en-315/>.

⁶⁸¹ *Blueprint*, 4.

⁶⁸² *Blueprint*, 12.

⁶⁸³ *Id.*, 5.

⁶⁸⁴ *Id.*, 18.

⁶⁸⁵ *Id.*, 5.

⁶⁸⁶ Come, al contrario, altri documenti precedenti, come l'Ordine Esecutivo 13859, in cui non compaiono le parole chiave “discriminazione”, “pregiudizio” o “danno.”

⁶⁸⁷ *Blueprint*, 24. Per le vicende dell'algoritmo di Amazon utilizzato nello *screening* iniziale dei *curricula*, vedi paragrafo 3.2 del Capitolo 1.

⁶⁸⁸ *Blueprint*, 6.

qualità, la quantità, e la sicurezza dei dati hanno un legame indissolubile con le capacità dell'IA di performare, e, talvolta, di generale dei risultati che impattano negativamente nei confronti degli utenti⁶⁸⁹. In questo caso, il *Blueprint* suggerisce l'adozione di procedure di raccolta e di gestione dei dati da utilizzare per i *training* dell'IA⁶⁹⁰, che ricordano in parte quanto previsto dal GDPR europeo: ad esempio, il principio di *privacy by design*⁶⁹¹, la limitazione della raccolta di dati a seconda dello scopo (*purpose limitation*), o ancora l'importanza di una manifestazione chiara e non ingannevole del consenso degli utenti al trattamento dei dati, per citarne alcuni. Il quarto principio, denominato “Avviso e spiegazione”, richiede che i progettisti e gli sviluppatori forniscano una documentazione estesa e comprensibile all'utente, per ridurre le decisioni o i risultati inspiegabili dei sistemi che operano come *black-box*⁶⁹². L'ultimo principio stabilisce “alternative umane, considerazione e ripiego.”⁶⁹³ Si tratta di fornire agli utenti un portavoce umano che possa “*considerare e risolvere rapidamente i problemi riscontrati.*”⁶⁹⁴ La definizione di un'alternativa umana consente agli utenti di scegliere di non essere costretti a confrontarsi con soli sistemi automatizzati artificiali.

L'emanazione del *Blueprint for an AI Bill of Rights* ha rappresentato un passo significativo verso una maggiore attenzione ai potenziali danni dell'IA e ai diritti civili e digitali degli americani. L'amministrazione Biden, con la redazione del *Blueprint*, ha contribuito ad instaurare ad una presenza governativa più incisiva nel campo dello sviluppo tecnologico⁶⁹⁵. Tuttavia, questo genere di approccio per certi versi più “interventista” ha ricevuto critiche sia da coloro che sono maggiormente favorevoli al mantenimento di un ambiente di innovazione deregolamentato e ispirato unicamente alle logiche di mercato, così come da parte di coloro che in preferiscono l'introduzione di regole, qualsiasi esse siano, tale da controllare lo sviluppo e l'introduzione sul mercato dei sistemi *AI-powered*.

Questi ultimi, infatti, hanno più volte lamentato che il *Blueprint* non sia uno strumento di *policy* governativa ad effetto vincolante⁶⁹⁶. La sua mancanza di esecutività lo rende inadatto a incidere sostanzialmente nei confronti dei produttori

⁶⁸⁹ Vedi il sottoparagrafo 2.2.2 del Capitolo 1.

⁶⁹⁰ *Blueprint*, 30.

⁶⁹¹ Per la definizione del principio di *privacy by design*, vedi nota 95.

⁶⁹² *Blueprint*, 41. Per la definizione di effetto *black-box*, vedi il sottoparagrafo 2.2.1 del Capitolo 1.

⁶⁹³ *Id.*, 7.

⁶⁹⁴ *Id.*, 46.

⁶⁹⁵ Kovacevich, A. (2 agosto 2022). The ‘Biden Tech Doctrine’ — one year in. *The Hill*. <https://thehill.com/opinion/technology/593238-the-biden-tech-doctrine-one-year-in/>.

⁶⁹⁶ *Blueprint*, 2.

di matrice non governativa di IA, soprattutto per i grandi *players* del settore⁶⁹⁷. Altri⁶⁹⁸, invece, hanno sottolineato come il libro bianco predisponga un'eccezione nel caso dell'utilizzo di sistemi algoritmici da parte delle forze dell'ordine⁶⁹⁹. Questa deroga rappresenta un'occasione mancata per fornire un quadro giuridico alle agenzie federali che attualmente utilizzano, tra gli altri, strumenti di riconoscimento facciale basati su algoritmi di IA che dovrebbero essere regolamentati in modo adeguato in quanto in grado di portare con sé risultati altamente discriminatori⁷⁰⁰.

Dall'altra parte del dibattito, la *U.S. Chamber of Commerce*⁷⁰¹ ha espresso apertamente le sue preoccupazioni sul *Blueprint*, affrontando l'effetto controproducente di elevare le barriere all'innovazione dell'IA, attraverso una lettera del suo Vicepresidente esecutivo della Camera indirizzata al Direttore dell'*Office of Science and Technology Policy* della Casa Bianca⁷⁰². Nella lettera, il Vicepresidente della Camera di Commercio statunitense ha osservato, *inter alia*, che il processo di stesura del Programma non ha comportato un contributo sostanziale da parte degli *stakeholder*⁷⁰³. Un discorso di questo tipo scaturisce dall'approccio assolutamente *top-down* del contenuto del libro bianco, il quale nasce a partire dalle valutazioni discrezionali dei membri dell'Ufficio in questione. La mancanza di coinvolgimento delle imprese e dei loro rappresentanti a livello

⁶⁹⁷ Johnson, K. (4 ottobre 2022). Biden's AI Bill of Rights is toothless against big tech. *Wired*. <https://www.wired.com/story/bidens-ai-bill-of-rights-is-toothless-against-big-tech>. *Contra* Engler, A. (7 ottobre 2022). The AI Bill of Rights Makes Uneven Progress on Algorithmic Protections. *Lawfare*. <https://www.lawfareblog.com/ai-bill-rights-makes-uneven-progress-algorithmic-protections>.

⁶⁹⁸ *Vedi* Engler (2022).

⁶⁹⁹ *Blueprint*, 2.

⁷⁰⁰ Ad esempio, negli USA le forze dell'ordine traggono enormi vantaggi dall'utilizzo di sistemi di rilevazione biometrica come Clearview AI, un sistema di riconoscimento facciale basato sull'IA, che consente di identificare gli individui anche in tempo reale, tramite le immagini raccolte su Internet e la creazione di una sorta di "carta di identità biometrica" presente nei *database* proprietari di questa società. *Vedi in generale Law Enforcement | Clearview AI*. (n.d.). Clearview AI. <https://www.clearview.ai/law-enforcement>.

⁷⁰¹ La U.S. Chamber of Commerce costituisce una delle più note, nonché più antiche organizzazioni a tutela delle imprese americane. Fondata oltre un secolo fa, si impegna a sostenere la crescita economica degli Stati Uniti, promuovendo politiche favorevoli alla creazione di posti di lavoro e alla prosperità delle aziende del *Made in USA*. Nel fare ciò, la Camera fornisce analisi di politica, collega i suoi membri con *leaders* degli affari e del governo, mentre agisce come alleata delle imprese e degli interessi aziendali in diverse sedi, compreso Capitol Hill. *Vedi About us*. (n.d.). <https://www.uschamber.com/about>.

⁷⁰² U.S Chamber of Commerce, *Letter to Director Prabhakar (OSTP) on "A Blueprint for an AI Bill of Rights"* (10 ottobre 2022).

⁷⁰³ *Id.*, 1.

politico ha delegittimato il contenuto del *Blueprint*, riducendo la sua generale incisività politica nei confronti di una futura regolamentazione.

Tuttavia, l'*input* di regolamentare l'IA con un mezzo di *soft law* come il libro bianco di uno specifico ufficio della Casa Bianca è stato utile per saggiare il terreno per l'eventuale redazione di altri testi normativi di natura più incisiva. A questo proposito, infatti, esattamente un anno dopo l'emanazione del *Blueprint*, l'amministrazione Biden ha emanato l'*Executive Order No. 14110 "on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence,"*⁷⁰⁴ con il quale si è definitivamente accantonata quell'idea di un'IA completamente *permissionless*, a favore di un approccio parzialmente più ispirato a delle logiche di regolamentazione.

Prima di trattare il suo contenuto è d'uopo analizzare la tipologia di strumento normativo scelto dall'amministrazione Biden, per fornire il quadro di indicazioni per un'IA "sicura ed affidabile," come indicato nel titolo del decreto⁷⁰⁵. Gli *Executive Orders* in via del tutto generale sono provvedimenti normativi paragonabili nell'ordinamento italiano in parte ai decreti-legge, ed in parte ai decreti del Presidente del Consiglio dei ministri⁷⁰⁶. Essi hanno un effetto vincolante sulle sole amministrazioni pubbliche (tra cui le agenzie federali), mentre non hanno una particolare presa sui privati⁷⁰⁷. Il testo, inoltre, non attribuisce diritti o pretese agli individui, essendo comunque indirizzato al settore pubblico. Inoltre, gli ordini esecutivi non costituiscono strumenti legislativi a disposizione delle istituzioni parlamentari, ma sono frutto unicamente di un potere discrezionale amministrativo e non godono di alcuna legittimazione democratica diretta (ma solo indiretta, in quanto provengono dall'emanazione del Presidente degli Stati Uniti d'America, democraticamente eletto dai cittadini americani). L'*Executive order*, in ultimo, non ha una sua intrinseca stabilità legislativa in quanto, una volta emanato, è soggetto ad eventuali *revoche*, modifiche o integrazioni da parte dei presidenti delle

⁷⁰⁴ Exec. Ord. No. 14110.

⁷⁰⁵ *Id.*

⁷⁰⁶ Servizio Studi del Senato della Repubblica, Servizio degli Affari Internazionali del Senato della Repubblica, & Ufficio Rapporti con l'Unione Europea della Camera dei deputati. (2024). Conferenza sull'intelligenza artificiale - Bruxelles, 28-29 gennaio 2024. In *Dossier Europei XIX Legislatura (65/DE)*, 77, <https://www.senato.it/service/PDF/PDFServer/BGT/01403117.pdf>. (d'ora in avanti abbreviato in "Dossier – Conferenza sull'intelligenza artificiale").

⁷⁰⁷ *Id.*

amministrazioni successive⁷⁰⁸. È quindi evidente la sua diversità da uno strumento normativo di matrice parlamentare come ad esempio l'*AI Act*⁷⁰⁹. Ad ogni modo, il tempismo in cui gli USA hanno redatto il predetto *Executive Order* lascia immaginare la volontà statunitense di fondo di voler inserire un proprio vitale contributo nella *race to AI regulation*⁷¹⁰.

Le basi su cui si fonda l'intera emanazione dell'*Executive Order* è quella di fornire un “*cronoprogramma per la creazione di un ambiente normativo e amministrativo entro cui il progresso dell'intelligenza artificiale sia ben stimolato ma anche regolato.*”⁷¹¹ Il testo innanzitutto prevede la presentazione di 8 finalità principali che fungano da punti cardinali dell'ordine esecutivo: sicurezza, innovazione responsabile, promozione e protezione dei lavoratori americani, tutela dell'uguaglianza e dei diritti umani, protezione di utenti e consumatori nell'uso quotidiano di IA, protezione dei dati personali, reclutamento di personale qualificato nelle agenzie federali, l'assunzione di un ruolo guida a livello globale in termini sia di innovazione che di *governance* dell'IA⁷¹².

Oltre ai principi guida, l'*Executive Order* ha introdotto una serie di definizioni, 33 per la precisione, in grado di cristallizzare alcuni concetti tecnici e creare degli standard definatori per la legislazione statunitense, ma anche rilevanti a livello internazionale⁷¹³: tra le nozioni adottate vi è proprio quella di IA, già precedentemente oggetto di questa trattazione nel Capitolo 1 paragrafo 2.3, insieme ad altri concetti ed attività innovative, tra cui la definizione l'IA generativa⁷¹⁴ o la pratica di “*red-teaming,*” cioè una di quelle attività di *cybersecurity* che consentono di testare il livello di resistenza di un dato sistema agli attacchi informatici⁷¹⁵,

⁷⁰⁸ Sibilla, F., Bruschi, B.S., (19 gennaio 2024). *Cenni e confronti tra le due principali normative a livello mondiale sull' AI: AI Act europeo e Executive Order americano*. Rivista Diritto Di Internet – Osservatorio Dati & Imprese.

⁷⁰⁹ *Id.*

⁷¹⁰ Dossier – Conferenza sull'intelligenza artificiale (2024), 77, 79.

⁷¹¹ *Id.*

⁷¹² Exec. Ord. No. 14110 § 2.

⁷¹³ Floridi (2023).

⁷¹⁴ Definita come “*la classe di modelli di IA che emulano la struttura e le caratteristiche dei dati in ingresso per generare contenuti sintetici derivati. Questi possono includere immagini, video, audio, testi e altri contenuti digitali.*” Exec. Ord. No. 14110. § 3 (c).

⁷¹⁵ Questo genere di pratica consente a degli esperti operanti nella cybersicurezza di simulare le condotte “*antagoniste*” (vale a dire quelle dei potenziali malintenzionati, intenti ad attaccare un sistema di IA per manipolare i risultati e impossessarsi dei suoi dati), contrastandole attivamente. *Vedi* Dossier – Conferenza sull'intelligenza artificiale (2024), 80.

nonché la definizione della pratica “*watermarking*”, particolarmente di rilievo per quanto riguarda i contenuti multimediali visivi prodotti dall’IA generativa⁷¹⁶.

Oltre alla definizione delle otto finalità, l’*Executive Order* predispone una dettagliata *timeline* di attività (ad esempio pubblicazione di documenti e predisposizione di particolari iniziative) che impegnano una molteplicità di dipartimenti governativi statunitensi, nel porre in essere specifiche azioni secondo scadenze temporali. Tra queste, quelle che necessitano menzione sono le seguenti. In *primis*, è prevista l’emanazione di un documento che contenga le linee guida per l’identificazione di buone pratiche per la sicurezza dei sistemi di IA (tra queste l’uso del *watermarking*) da parte del *National Institute of Standards and Technology* (NIST), sotto a supervisione del Ministro del Commercio⁷¹⁷. In secondo luogo, è interessante il coinvolgimento del Ministro per l’Energia, che è tenuto a portare avanti delle valutazioni in relazione all’impatto dell’uso dell’IA nei confronti dell’ambiente, nell’ottica di mitigare gli effetti negativi del cambiamento climatico⁷¹⁸, e di perseguire obiettivi di innovazione responsabile⁷¹⁹. Ancora, il Segretario di Stato e il Ministro dell’Interno sono collettivamente investiti di obblighi attinenti alla promozione di un ambiente in grado di sfruttare correttamente il capitale umano, anche straniero, nel campo dell’IA (il c.d. *AI talent*): ad esempio tramite l’istituzione di programmi di scambio con paesi esteri, o con l’accelerazione delle procedure di esame delle domande di ingresso negli USA per motivi professionali riconnessi con le discipline attinenti all’IA⁷²⁰. Questo aspetto relativo all’*AI talent* non è da sottovalutare: si ricordi infatti come, nell’analisi di quei fattori che hanno impedito e tutt’ora impediscono all’Unione Europea di sviluppare un’ambiente di florida innovazione nel settore del digitale è (tra gli altri) la mancanza di uno sfruttamento del capitale umano europeo e di capacità di attrattiva di talento estero da parte dell’Unione⁷²¹. Tra le altre iniziative predisposte

⁷¹⁶ Il *watermarking* è una pratica utilizzata nei confronti dei contenuti multimediali visivi prodotti dall’IA generativa, che consente di incorporare nell’output delle informazioni codificate, ancorché a livello grafico, per garantire l’autenticità del risultato *AI-generated* e per identificare la provenienza, eventuali modifiche o trasmissioni del contenuto. Il *watermarking* è utile per distinguere immagini e video creati con IA generativa piuttosto che reali. Esso rientra in una delle strategie da impiegare per limitare la diffusione di informazioni false e disinformazione basate su contenuti realistici, ma tutta via non effettivamente. *Vedi* Exec. Ord. No. 14110. § 3 (gg), § 4.5 (a) (ii). *Vedi anche* Dossier – Conferenza sull’intelligenza artificiale (2024), 80-81.

⁷¹⁷ Exec. Ord. No. 14110. § 4.1 (a).

⁷¹⁸ Exec. Ord. No. 14110. § 5.2 (g).

⁷¹⁹ Dossier – Conferenza sull’intelligenza artificiale (2024), 81.

⁷²⁰ Exec. Ord. No. 14110. § 5.1.

⁷²¹ *Vedi* il paragrafo 3.5 del Capitolo 2, a pag. 103.

dall'*Executive Order* vi è il potenziamento della ricerca, delle università, e delle procedure di brevetto nel campo dell'IA, vitale per assicurare protezione a chi produce e a chi investe⁷²².

Un'altra iniziativa di interesse, soprattutto se analizzata alla luce dell'ottica europea della regolamentazione sull'IA secondo il modello *rights-driven*, è la particolare attenzione alla funzione di tutela dei diritti umani e dell'uguaglianza che l'*Executive Order* prevede, attribuendo all'*Attorney General* (Ministro della Giustizia) il compito di predisporre un convegno a stretto giro con i rappresentanti di funzionari federali incaricati della tutela dei diritti⁷²³; a questo, si darà seguito con un documento di relazione da presentare al presidente degli USA per quanto riguarda i seguenti temi che interessano particolarmente la materia dei diritti umani e la loro intersezione con l'IA: la quantificazione delle pene, le misure alternative alla detenzione; le cauzioni e le misure cautelari; la valutazione dei rischi nella concessione dei benefici penitenziari; o la prognosi criminologiche e la attività di polizia predittiva⁷²⁴. Insieme alla tutela dei diritti fondamentali, particolare attenzione viene conferita al settore del lavoro, anche a seguito dei numerosi scioperi dei lavoratori dello spettacolo di Hollywood, mossi anche dalle rivendicazioni circa i pericoli posti dall'IA (soprattutto quella generativa) in ambito di produzione di contenuti artistici multimediali⁷²⁵. D'altronde, il settore lavorativo è quello che ha causato maggiori preoccupazioni a livello dell'intera opinione pubblica statunitense, in base all'ultimo report condotto dall'Università di Stanford precedentemente menzionato⁷²⁶.

Inoltre, viene affidato il compito di occuparsi della protezione dei dati e della privacy nell'ambito dello sviluppo e dell'utilizzo di IA al Direttore dell'Ufficio per il bilancio del Presidente degli stati Uniti (il *Director* dell'OMB, precedentemente menzionato⁷²⁷). L'*Executive Order* infine, predisporre delle nuove

⁷²² Exec. Ord. No. 14110. § 5.2 (c).

⁷²³ Exec. Ord. No. 14110. § 7.1 (a)(ii).

⁷²⁴ Si noti la differenza tra USA e UE in relazione all'utilizzo dell'IA nel *law enforcement*: mentre nei primi esso è oggetto di attente valutazioni per permetterne l'utilizzo con lo scopo, tra gli altri, di razionalizzare il sistema penitenziario e giudiziario *in toto*, l'UE guarda a queste pratiche con molto più scetticismo, arrivando a vietare totalmente (salvo eccezioni precisamente definite dalla legge) alcune di queste pratiche, come ad esempio proprio la polizia predittiva. *Vedi in generale* art. 5, *AI Act* (Versione IV).

⁷²⁵ Dossier – Conferenza sull'intelligenza artificiale (2024), 81.

⁷²⁶ *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 8: public opinion, 331 (2023)

⁷²⁷ Exec. Ord. No. 14110. § 9. È da sottolineare la peculiarità di affidare la disciplina della protezione dei dati ad un ufficio non settoriale, non dotato di uno specifico *expertise* nel campo della privacy. Nonostante l'OMB abbia, in precedenza, operato come un'autorità dalle competenze diversificate

entità, quali il *Chief AI Officer*, responsabile dei risvolti pratici dell'uso dell'IA e della diffusione dei sistemi *AI-powered* nei pubblici uffici⁷²⁸, e il *White House AI Council*, all'interno del più ampio Ufficio esecutivo del Presidente degli Stati Uniti, dotato di attribuzioni quali il coordinamento generale dell'*AI governance*, che possa fornire ausilio al Presidente stesso⁷²⁹.

È necessaria una considerazione finale alla luce di quanto appena trattato in relazione all'*Executive Order* No. 141110, nell'ottica di operare un confronto con l'*AI Act* europeo. I due testi normativi certamente riflettono gli approcci differenziati nella regolamentazione dell'IA nelle due entità sovranazionali. Tuttavia, mentre l'*AI Act* si propone di disciplinare in modo completo ed orizzontale l'IA con un approccio graduato basato sul rischio, vietando sistemi a rischio elevato e stabilendo requisiti e certificazioni, l'*Executive Order* statunitense definisce standard per sicurezza e privacy senza prevedere meccanismi sanzionatori diretti nei confronti dei soggetti non conformi⁷³⁰. Vi è anche da sottolineare come l'*Executive Order* sia indirizzato alle agenzie federali, e non ai privati. Inoltre, l'UE adotta uno strumento normativo stabile e legalmente vincolante, il regolamento europeo, mentre l'ordine esecutivo è un atto unilaterale del Presidente USA, suscettibile di modifiche con il cambio di amministrazione, e dunque non dotato di particolare stabilità sul lungo periodo⁷³¹. Nonostante ciò, i due testi convergono su alcuni punti come nella definizione di IA, segnando una possibile a livello globale, benché le divergenze nella legittimazione e stabilità dell'ordine esecutivo sollevano interrogativi sulla certezza del diritto e sull'applicabilità a lungo termine degli standard normativi ivi previsti⁷³².

anche in un settore come quello della privacy, (come emerge dalle precedenti decisioni dell'OMB riportate dall'*Executive Order*, in cui si fa riferimento alle nozioni attinenti ai temi della privacy precedentemente fornite da questo particolare ufficio, *vedi* Exec. Ord. No. 14110. § 3 (y), (z) etc.), esso viene dunque investito di questa particolare competenza in quanto negli USA non è stata istituita una *data protection authority* (DPA) governativa, intesa secondo quell'accezione tipicamente europea indicata all'art. 51 del GDPR. Gli USA non hanno infatti alcuna normativa generale che possa regolare la disciplina della privacy e della protezione dei dati a livello federale e generale; diverse istituzioni hanno dunque assunto quei comportamenti tipici di un DPA, tra cui l'OMB, ma molto più frequentemente la *Federal Trade Commission* (FTC), soprattutto durante l'amministrazione Biden. *Vedi* Solove & Schwartz (2020), 113-114.

⁷²⁸ Exec. Ord. No. 14110. § 10 (b).

⁷²⁹ Exec. Ord. No. 14110. § 12. *Vedi anche* Dossier – Conferenza sull'intelligenza artificiale (2024), 84.

⁷³⁰ Sibilla & Bruschi (2024). *Vedi anche* Ryan-Mosley, T. (30 ottobre 2023). Three things to know about the White House's executive order on AI. *MIT Technology Review*. <https://www.technologyreview.com/2023/10/30/1082678/three-things-to-know-about-the-white-houses-executive-order-on-ai/>.

⁷³¹ *Id.*

⁷³² *Id.*

Ad ogni modo, l'*AI Bill of Rights* emanato nel 2022, e l'*Executive Order* No. 14110 che lo ha succeduto nel 2023, consentono di evidenziare un dato: l'assenza generale di iniziative di regolamentazione da parte del Congresso. Tutto ciò che è stato fino ad ora prodotto in termini di disciplina dell'IA deriva da agenzie federali ed uffici esecutivi del Presidente, nonché da istituti dedicati come il NIST⁷³³. Con tale vuoto nella creazione di politiche federali, governi statali e locali si sono affrettati a colmare quei vuoti normativi effettivamente percepibili anche a livello regionale. Infatti, durante il corso del 2023, i legislatori dei parlamenti statali hanno presentato oltre il 440% in più di proposte di legge relative all'IA rispetto al 2022⁷³⁴. Tra questi, ad esempio, il Texas e il Connecticut hanno approvato legislazioni *ad hoc*⁷³⁵, mentre i governatori in California e Pennsylvania che hanno emesso ordinanze esecutive⁷³⁶; persino le amministrazioni cittadine di Seattle e della città di New York hanno creato proprie normative sull'IA a livello locale⁷³⁷. Tuttavia, il quadro composito venutosi a creare a causa delle iniziative normative degli enti locali è suscettibile di generare un elevato grado di incertezza circa gli standard e le regole da seguire nello sviluppo e utilizzo di IA a livello delle singole giurisdizioni statali⁷³⁸.

PARTE II – LA REPUBBLICA POPOLARE CINESE

II.1 Il modello “state-driven” cinese nell’economia digitale e lo stato del mercato dell’IA in Cina

Le capacità di produrre, sviluppare e commercializzare applicazioni dell'IA, nonché le dimensioni del mercato cinese, certamente non passano in secondo piano rispetto a quelle statunitensi. Per questo motivo, si è scelto di trattare il tema della

⁷³³ Turner Lee, N. & Malamud, J., (25 gennaio 2024) How Congress can secure Biden’s AI legacy. *Brookings*. <https://www.brookings.edu/articles/how-congress-can-secure-bidens-ai-legacy/>.

⁷³⁴ *Id.*

⁷³⁵ Williams, Z. (28 agosto 2023). AI Regulation Ramping Up in States from Texas to Connecticut. *Bloomberg Law*. <https://news.bloomberglaw.com/in-house-counsel/ai-regulation-ramping-up-in-states-from-texas-to-connecticut>

⁷³⁶ BCLP - Bryan Cave Leighton Paisner. (n.d.). US state-by-state AI legislation snapshot. BCLP - Bryan Cave Leighton Paisner - US State-by-state AI Legislation Snapshot. <https://www.bclplaw.com/en-US/events-insights-news/2023-state-by-state-artificial-intelligence-legislation-snapshot.html>.

⁷³⁷ Davidson, N. (4 dicembre 2023). MAP: How are state and local governments navigating AI regulation? *GovTech*. <https://www.govtech.com/biz/data/how-are-state-and-local-governments-navigating-ai-regulation>

⁷³⁸ Turner Lee & Malamud (2024).

regolamentazione dell'IA in Cina, per completare la comparazione con l'approccio europeo e quello statunitense.

In generale, la Cina rappresenta uno dei *players* più rilevanti nel settore del digitale complessivamente considerato: essa, tra l'altro, detiene otto delle maggiori venti aziende nel settore di Internet (tra cui i giganti Tencent, Alibaba, Meituan e JD⁷³⁹); essa, inoltre, ha una posizione centrale nell'ambito della *governance* delle risorse attinenti alla produzione di *hardware* e *software* da impiegare per lo sviluppo di IA, nonché un ruolo di assoluto rilievo nelle tecnologie innovative e nella ricerca⁷⁴⁰. Lo sviluppo del settore del digitale in Cina è stato da sempre facilitato da interventi e investimenti di matrice governativa, che potessero consentire con facilità l'emersione di un'industria non solo ben strutturata e solida a livello nazionale, ma che potesse agevolmente competere anche a livello internazionale, soprattutto con le *Big tech* americane⁷⁴¹. La Cina, infatti, fin dagli anni immediatamente successivi alla crisi del 2008, ha facilitato la creazione di un ambiente di innovazione nel settore dello sviluppo digitale grazie ad un approccio sostanzialmente deregolamentato, frutto della combinazione degli interessi di consolidamento del potere da parte del Partito Comunista Cinese (PCC), ma anche di quelli appartenenti alle aziende private emergenti, che hanno potuto crescere a livello esponenziale, in maniera sostanzialmente incontrollata⁷⁴².

La strategia economica cinese è sempre stata (e, per certi versi, è ancora)⁷⁴³ volta ad incentivare il maggior sviluppo possibile delle imprese della tecnologia, aumentando l'influenza economica del paese a livello globale, e rafforzando contemporaneamente proprio l'autorità cinese. I fitti legami tra le grandi società del *digital* e il PCC, non si esauriscono unicamente in investimenti e nel sostegno governativo alle aziende. È infatti noto che, in cambio della suddetta regolamentazione permissiva, soprattutto per quanto riguarda la normativa antitrust, le grandi aziende cinesi hanno operato come surrogati del PCC per quanto concerne attività di censura e sorveglianza della società civile, in modo tale da

⁷³⁹ Bradford (2023), 71.

⁷⁴⁰ *Id.*, 197, 210. Tra l'altro, la Cina nel 2023 è stata il paese con il maggior numero di articoli scientifici in materia di IA, redigendo circa il 40% delle pubblicazioni mondiali in questo settore. *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 1: research and development, 34 (2023). Allo stesso modo, la top 10 delle istituzioni che hanno effettuato il maggior numero di pubblicazioni

⁷⁴¹ *Vedi in generale id.*, Chapter 2: the Chinese State-Driven Regulatory Model, 69-104.

⁷⁴² *Id.*, 73.

⁷⁴³ *Id.*, 94-99.

mantenere un saldo controllo sul tessuto sociale cinese⁷⁴⁴. Infatti, la mancanza di regole per disciplinare le aziende del digitale è controbilanciata dalle restrizioni poste dall'autorità cinese all'accesso ad Internet da parte della popolazione, sia dal punto di vista delle informazioni disponibili online, che da quello dei servizi, le piattaforme, e le app, prodotte da paesi stranieri⁷⁴⁵. A differenza di quanto accade, ad esempio, negli USA e nell'UE, in cui vigono principi generali di *net neutrality*⁷⁴⁶, la Cina ha scelto di restringere l'accesso alla rete⁷⁴⁷, nonché ad alcuni servizi su essa disponibili, nell'ottica di mantenere un saldo monopolio sulle informazioni e sui contenuti online a disposizione dei consumatori cinesi⁷⁴⁸. A questo proposito, nel mercato digitale cinese, alcune piattaforme online *made in USA* come YouTube o Whatsapp sono censurate da molti anni, sostituite da prodotti social simili offerti da aziende locali cinesi, come Youku Tudou (una delle più note piattaforme di *streaming* video) o WeChat (il social più diffuso di messaggistica istantanea⁷⁴⁹).

Un approccio al digitale di così delineato è stato denominato da Bradford come c.d. *state-driven*⁷⁵⁰, cioè ispirato alle logiche di autoritarismo digitale dello stato. La deregolamentazione di un determinato settore non è ispirata a logiche di mercato e fiducia nella *self-regulation* dello stesso, come accade nel caso del modello *market-driven* americano; l'assenza di norme in un campo convive perfettamente con l'incisiva censura di Internet e dei servizi digitali in generale, perché a dettare la linea politica in questo settore rimane l'autorità centrale⁷⁵¹.

⁷⁴⁴ *Id.*, 71-72.

⁷⁴⁵ *Id.*, 75-77.

⁷⁴⁶ La neutralità della rete è un principio che stabilisce che tutti i dati su Internet devono essere trattati in modo equo e senza discriminazioni da parte degli *Internet Service Providers* (ISP). Ciò significa che gli stessi ISP non possono rallentare, bloccare o discriminare l'accesso a determinati siti web o servizi online in base al contenuto, all'origine o alla destinazione del traffico dati. In generale, la neutralità della rete può essere manipolata non solo da parte degli ISP, ma anche da parte dei governi nazionali stessi: per questo motivo, all'altra estremità dello spettro, rispetto al principio generale di *net neutrality* (vigente in misura differenziata sia negli USA che in UE) vi sono le pratiche restrittive e di censura dell'accesso a determinate informazioni o servizi online da parte delle autorità governative, ispirate alla capacità di sorveglianza da parte dello Stato (come quello cinese), in grado di bloccare al di fuori dei confini nazionali determinati contenuti. *Vedi* Contissa (2017), 125. *Vedi anche* Marsden, C. T. (2017). *Network neutrality: From policy to law to regulation*. Manchester University Press, 9-10.

⁷⁴⁷ *Inter alia*, la Cina è in grado di censurare le suddette piattaforme e altri servizi disponibili online a mezzo di progetti come il c.d. *Great Firewall of China*, un sistema composito di tecnologie e di norme che restringono l'accesso a determinate informazioni da parte di qualsiasi individuo in tutto il territorio cinese. Bradford (2023), 78-79.

⁷⁴⁸ *Id.*, 77-85.

⁷⁴⁹ *Id.*, 154-155.

⁷⁵⁰ *Vedi* il paragrafo 1 del Capitolo 2, in cui viene anticipata una descrizione sommaria del modello *state-driven* teorizzato da Bradford.

⁷⁵¹ *Vedi id.*

Sebbene il governo cinese abbia, nel tempo, modificato il proprio approccio nei confronti della normativa antitrust⁷⁵², conscio delle grandi dimensioni e influenza dei giganti dell'*industry* precedentemente menzionati⁷⁵³, l'approccio *state-driven* ha permesso alle aziende di crescere fino alle dimensioni e al potere che esse detengono oggi. Inoltre, per comprendere il modello *state-driven* è fondamentale riconoscere l'importanza di un altro fattore determinante, rappresentato dalle intense attività di *lobbying* a favore del PCC: i rapporti che intercorrono tra il Partito e le grandi aziende della tecnologia divergono infatti dal tipo di influenza che sono in grado di generare le *Big tech* dei paesi occidentali, dal momento che gli stessi esponenti di punta di queste aziende cinesi sono spesso membri attivi del PCC⁷⁵⁴. Le trame di connessione tra Stato e industria digitale sono dunque molto fitte, per assicurare la supremazia tecnologica cinese a livello globale⁷⁵⁵. Queste caratteristiche, in combinazione con una politica di protezionismo digitale e all'utilizzo di strumenti come la sorveglianza e la censura, sono gli indicatori dell'adozione del modello economico e di regolamentazione *state-driven* tipico di Pechino⁷⁵⁶. Tuttavia, non vanno tralasciati quei casi di attività di regolamentazione che, seppur recentemente, hanno interessato l'autorità cinesi. Infatti, in tempi recenti, sono state varate leggi relative alla privacy e alla protezione dei dati⁷⁵⁷ (molto simili, ancorchè a livello formale, al GDPR⁷⁵⁸) e inoltre degli aggiornamenti di natura più stringente alla disciplina antitrust⁷⁵⁹. Per cui, il modello di regolamentazione *state-driven* si può riassumere considerando che le decisioni prese dal governo riguardo alla regolamentazione di settori come l'economia

⁷⁵² Zhang, A. Z. (2021). Agility over stability: China's great reversal in regulating the platform economy. *Harvard International Law Journal*, Vol. 63, No. 2, 2022, <https://doi.org/10.2139/ssrn.3892642>. Vedi anche Bradford (2023), 92.

⁷⁵³ Bradford (2023), 71.

⁷⁵⁴ *Id.*, 73.

⁷⁵⁵ *Id.*, 71.

⁷⁵⁶ Vedi in generale *id.*, Chapter 2: the Chinese State-Driven Regulatory Model, 69-104.

⁷⁵⁷ Standing Committee of the National People's Congress, Geren Xinxi Baohu Fa (个人信息保护法) (*Personal Information Protection Law*, abbreviata in "PIPL") (20 agosto 2021).

⁷⁵⁸ La somiglianza della PIPL al GDPR è tuttavia di matrice prettamente formale. Sebbene i principi presenti in entrambi le normative sembrano astrattamente sovrapponibili e dotati di una simile robusta struttura (corredata di diritti applicabili da parte dei singoli individui), uno sguardo più accurato rivela, contrariamente, una serie di scappatoie, di *loopholes* ed eccezioni varie che consentono all'autorità cinese, in ultimo, di aggirare in gran parte le tutele previste per la privacy per poter svolgere indisturbata l'attività di sorveglianza di massa nei confronti della società civile cinese. Roberts, H., Cowls, J., Morley, J., Taddeo, M. R., Wang, V., Floridi, L. (2020) The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & Society* 36, 59–77, 72 <https://doi.org/10.1007/s00146-020-00992-2>. Vedi anche Bradford (2023), 91-93.

⁷⁵⁹ Bradford (2023), 91-93.

digitale sono principalmente basate sulle esigenze e sulle direttive governative, piuttosto che su altre considerazioni.

Dopo aver fornito un panorama sulla situazione della *digital industry* cinese in combinazione con l'approccio regolamentare finora adottato, è necessario delineare la situazione del mercato dell'IA, in rapida espansione anche in questo paese. Infatti, dal 2016 al 2023, le dimensioni del mercato dell'IA sono aumentate da 15,4 miliardi di yuan, agli attuali 304,3 miliardi⁷⁶⁰. Allo stesso modo, le proiezioni per il 2026 vedono l'*AI core industry* (l'industria che si occupa unicamente di IA, in contrapposizione con il settore delle tecnologie/dispositivi che integrano l'IA, le c.d. *AI-driven industries*) crescere fino ad un valore complessivo di 605 miliardi di yuan⁷⁶¹. Dal punto di vista degli investimenti privati nell'IA nello scorso anno, la Cina si trova al secondo posto dei paesi con una cifra maggiore di risorse allocate per questa finalità, per un valore di 13,4 miliardi di dollari⁷⁶². Inoltre, sebbene gli Stati Uniti siano ancora *leader* nel settore della creazione di startup nell'IA, la Cina presenta ad oggi il maggior numero di domande di brevetto relativi alla creazione di sistemi *AI powered* al mondo⁷⁶³. Inoltre, sembrerebbe che la società cinese sia incline ad accettare le implementazioni dell'IA, in ambito sia pubblico che privato, in misura maggiore rispetto all'opinione pubblica americana⁷⁶⁴. Non solo, la Cina si trova tra i primi paesi a presentare un alto livello di volontà di accettare la maggiore presenza di sistemi *AI-powered* in ogni settore, così come un'elevata tendenza ad affidarsi alle sue valutazioni e ai suoi risultati⁷⁶⁵.

Tuttavia, nel riportare lo sviluppo del mercato dell'IA in Cina, non si può prescindere dal menzionare che una delle motivazioni principali per le quali il paese è in grado di competere a livello mondiale nella c.d. *race to AI*, soprattutto contro gli USA⁷⁶⁶, è proprio l'adozione di un modello di autoritarismo digitale. Infatti, le aziende cinesi specializzate nello sviluppo dell'IA hanno a disposizione una vastissima quantità di dati provenienti dai consumatori del paese, i quali sono, come più volte ricordato, soggetti ad una sorveglianza online estensiva e pervasiva, a

⁷⁶⁰ China: AI market size 2023 [Grafico] Statista. (30 gennaio 2024). <https://www.statista.com/statistics/1262377/china-ai-market-size/>.

⁷⁶¹ China: scale of AI industry by segment 2026. [Grafico] Statista. (3 gennaio 2024) <https://www.statista.com/statistics/1256546/china-scale-of-ai-industry-by-segment/>

⁷⁶² *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 4: the economy, 189 (2023).

⁷⁶³ Bradford (2023), 210.

⁷⁶⁴ *The Artificial Intelligence Index Report 2023, AI Index*, Chapter 8: public opinion, 334 (2023).

⁷⁶⁵ Trust and acceptance of AI worldwide 2022. [Grafico] Statista (10 agosto 2023) <https://www.statista.com/statistics/1369185/trust-and-acceptance-of-ai-worldwide/>

⁷⁶⁶ Bradford (2023), *Chapter 5: the Battle for Technological Supremacy*, 183- 220.

scapito del proprio diritto alla privacy⁷⁶⁷. Inoltre, i sistemi di IA sviluppati dalle aziende cinesi sono essi stessi strumenti da impiegare per perseguire una sorveglianza di matrice autoritaria, in grado di compromettere le libertà individuali e i diritti così come concepiti da altre entità sovranazionali come l'Unione Europea e dagli USA⁷⁶⁸. Per questa ragione, la Cina è, *inter alia*, sede di diverse aziende che stanno sfruttando l'IA all'interno di progetti di *smart cities* per sviluppare strumenti di riconoscimento facciale, da utilizzare per scopi di pubblica sicurezza e ordine pubblico⁷⁶⁹. L'autorità nazionale ha implementato una molteplicità di sistemi di sorveglianza invasiva in tutto il paese (ad esempio con il progetto c.d. “*Sharp Eyes*,” che prevede la creazione di un unico, onnipresente, sistema di sorveglianza a livello nazionale⁷⁷⁰) giustificandoli con l'obiettivo di promuovere una società più armoniosa e ordinata. Tuttavia, voci critiche hanno evidenziato che l'obiettivo reale di tali pratiche di sorveglianza è piuttosto quello di instaurare un controllo autoritario attraverso l'uso della tecnologia stessa nei confronti dell'intera società cinese.

Per questo motivo, la lotta per la supremazia nel campo della produzione di sistemi *AI-powered* tra Cina e USA non si esaurisce in una sfida a livello economico, ma diventa può essere interpretata alla stregua di uno scontro tra le concezioni ideologiche di impiego degli strumenti tecnologici di entrambi i paesi.

II.2 La regolamentazione dell'IA in Cina

La Cina, ormai da diversi anni, sta gettando le basi intellettuali e burocratiche per redigere una normativa nazionale completa concernente l'IA, dalla portata epocale per la *governance* globale di questa tecnologia e, potenzialmente, al pari dell'*AI Act* dell'Unione Europea⁷⁷¹. Le normative cinesi sull'IA, sebbene spesso trascurate poiché considerate non propriamente rilevanti o un mero strumento politico da parte dei commentari occidentali, colpevoli di assumere una miope prospettiva eccessivamente euroatlantica su questo tema, hanno in realtà

⁷⁶⁷ *Id.*, 85, 208.

⁷⁶⁸ *Id.*, 209.

⁷⁶⁹ *Id.*, 85.

⁷⁷⁰ Gershgorn, D. (2 marzo 2021). China's 'Sharp Eyes' program aims to surveil 100% of public space. *Medium*. <https://onezero.medium.com/chinas-sharp-eyes-program-aims-to-surveil-100-of-public-space-ddc22d63e015>.

⁷⁷¹ Sheehan, M. (2023). *China's AI regulations and how they get made* (No. 1), 7 https://carnegieendowment.org/files/202307-Sheehan_Chinese%20AI%20gov.pdf.

un'importanza fondamentale⁷⁷². È infatti cruciale analizzarle per comprendere appieno il loro impatto sullo sviluppo dell'IA in Cina e per estrapolarne insegnamenti preziosi per la regolamentazione a livello globale, e comprendere ed analizzare le strutture sottostanti e la fattibilità tecnica di diversi approcci regolatori, sebbene vi possano essere possibili disaccordi sulle modalità adottate⁷⁷³. Inoltre, siccome questo paese ricopre (e ricoprirà) un ruolo d'importanza progressivamente crescente nel settore, come uno dei maggiori produttori delle variegate applicazioni di IA esistenti, non soffermandosi su di essa e sui suoi potenziali effetti a livello anche internazionale, potrebbe essere una leggerezza.

I piani di sviluppo dell'IA che la Cina ha negli anni progettato hanno comportato anche il *drafting* di diversi documenti normativi che hanno, sostanzialmente, fornito una cornice regolamentare all'IA. A questo proposito, è stato redatto il piano c.d. “*New Generation Artificial Intelligence Development Plan*” (AIDP)⁷⁷⁴ nel 2017, nell'ottica di fissare un cronoprogramma degli obiettivi strategici per rendere la Cina *leader* nel settore dell'IA entro il 2030⁷⁷⁵. L'AIDP, contenente una serie di strategie (geopolitiche, economiche e legali/etiche) per raggiungere questo ambizioso obiettivo di supremazia, prevedeva, *inter alia*, che l'autorità cinese avrebbe ampliato e codificato in legge gli standard etici per l'IA entro il 2025⁷⁷⁶.

Per questo motivo, negli anni successivi all'emanazione dell'AIDP, il Ministero della Scienza e della Tecnologia, nel marzo del 2019, ha predisposto dei principi di *governance* per l'IA⁷⁷⁷, da applicare a tutti coloro che intendono sviluppare sistemi *AI-powered*. Il documento propone otto principi, tra cui figurano il rispetto della privacy individuale, la sicurezza e la robustezza, e la *governance*

⁷⁷² Il fatto che le leggi di uno stato autoritario come la Cina possano essere impiegate come *proxy* per dispiegare misure di sorveglianza di massa nei confronti della società civile in generale non dovrebbe portare a sottovalutare completamente gli sforzi di regolamentazione né a ignorare l'entità e l'impatto di tali leggi. *Id. Vedi anche* Blomquist, K. (2022). *In pursuit of (soft) power: Chinese artificial intelligence governance in an age of great power competition*, [Tesi nel programma *MSc in Social Science of the Internet*]. University of Oxford, Oxford Internet Institute, 50.

⁷⁷³ Sheehan (2023), 10.

⁷⁷⁴ *New Generation Artificial Intelligence Development Plan*, 新一代人工智能发展计划, (luglio 2017).

⁷⁷⁵ Roberts *et al.* (2020), 72. *Vedi anche* Roberts *et al.*, (2020), 60-61.

⁷⁷⁶ Roberts *et al.* (2020), 61.

⁷⁷⁷ *Translation: Chinese expert group offers “Governance principles” for “Responsible AI” - DigiChina*. (1 ottobre 2021). *DigiChina*. <https://digichina.stanford.edu/work/translation-chinese-expert-group-offers-governance-principles-for-responsible-ai/>. *Vedi anche* Bradford (2023), 92. *Vedi anche* Roberts *et al.*, (2020), 68.

agile⁷⁷⁸. Le indicazioni contenute in questo documento presentano principi assolutamente generali e di alto livello, senza tuttavia predisporre alcuni meccanismi di concreta implementazione⁷⁷⁹. Nonostante, dunque, lo slancio a livello di *drafting* regolamentare presente già diversi anni fa, vi sono delle considerazioni da fare per quanto riguarda il contenuto di questi principi di *governance*. Infatti, ad una lettura prettamente formale, la disciplina delineata dall'autorità cinese può apparire molto simile a quanto l'UE ha predisposto con i suoi strumenti regolamentari prodromici all'*AI Act* (tra cui, ad esempio, le *Ethical Guidelines* emanate dall'HLEGAI nel 2019). In realtà, sebbene i principi sembrano piuttosto sovrapponibili con quelli europei, la situazione effettiva è ben diversa, dal momento che l'autorità cinese, come più volte ricordato, non si esime dal condurre attività di sorveglianza invasiva e totalitaria nei confronti della società, anche a mezzo di applicazioni di IA come il riconoscimento facciale⁷⁸⁰.

Il governo cinese ha, successivamente, provveduto ad integrare il proprio assetto di regolamentazione, queste indicazioni di principio, con strumenti di normativa più incisiva e di dettaglio. Così operando, ha utilizzato un approccio di settore, dedicando più fonti del diritto a disciplinare aspetti selezionati e particolare dell'IA. In *primis*, nel 2021 sono state emanate le *Provisions on the Management of Algorithmic Recommendations in Internet Information Services*⁷⁸¹. Questo documento è il primo esempio di fonte del diritto vincolante emanato dal legislatore cinese, volto a disciplinare una precisa applicazione degli algoritmi impiegati nella predisposizione di contenuti “raccomandati” agli utenti⁷⁸². Questo testo normativo è nato a causa della crescente preoccupazione per la quale i sistemi *AI-powered* di raccomandazione di contenuti potessero diffondere informazioni online in maniera non conforme alle linee guida del PCC. In realtà, il Partito ha fortemente ampliato i suoi obiettivi non limitandosi a questo *input* iniziale. La normativa, infatti, include non solo alcune misure concrete per il controllo dei contenuti online (come l'obbligo per le piattaforme di intervenire manualmente negli elenchi dei “temi caldi” sui social media per garantire che questi riflettano le priorità del governo),

⁷⁷⁸ Sheehan (2023), 10. Per quanto riguarda un'indicazione generale dell'IA applicata ai sistemi di raccomandazione di contenuti, si veda il paragrafo 2 del Capitolo 1.

⁷⁷⁹ Roberts *et al.*, (2020), 72.

⁷⁸⁰ Bradford (2023), 92.

⁷⁸¹ *Provisions on the Management of Algorithmic Recommendations in Internet Information Services* 互联网信息服务算法推荐管理规定. *China Law Translate*. (4 gennaio 2022) <https://www.chinalawtranslate.com/en/algorithms/>.

⁷⁸² Sheehan (2023), 12.

ma anche ricomprendendo delle previsioni interessanti a tutela del diritto dei lavoratori le cui mansioni sono gestite da parte di sistemi automatizzati di IA⁷⁸³. A questo proposito, infatti, sono state accolte positivamente quelle previsioni che, seppur in termini vaghi, indicano alle piattaforme che si avvalgono di lavoratori le cui mansioni sono stabilite da algoritmi di IA, di assicurarsi che questi forniscano un compenso e un riposo adeguati e che si facciano carico di garantire i diritti e gli interessi dei lavoratori stessi⁷⁸⁴.

Oltre, dunque, quelle previsioni di puro autoritarismo digitale cinese, la legge in analisi predispone delle salvaguardie, aggiungendo disposizioni che impediscono ai fornitori di sistemi di AI di utilizzare gli algoritmi per pratiche commerciali anticoncorrenziali e scorrette⁷⁸⁵. In ultimo, questa prima legge predispone obblighi nei confronti degli sviluppatori di IA, che sono tenuti a fornire informazioni dettagliate sul *training* e la distribuzione degli algoritmi di IA, insieme a un rapporto di *self-assessment* sulla loro sicurezza, e istituisce un “registro per i sistemi di IA” (*AI registry*, 算法备案系统), che è stato successivamente aggiornato nelle leggi successive, in modo tale da poter operare da pubblico archivio per quei sistemi che hanno “*proprietà di [influenzare] l’opinione pubblica o di mobilitazione sociale.*”⁷⁸⁶

Successivamente, nel 2022, è stata emanata un’altra legge volta a disciplinare gli usi di algoritmi di IA per generare o alterare sinteticamente i contenuti online, come voci, testi, immagini e video (un’attività sostanzialmente equiparabile alla creazione di *deepfake*⁷⁸⁷). La regolamentazione, intitolata *Provisions on the Administration of Deep Synthesis Internet Information Services*⁷⁸⁸ mira a garantire che tali contenuti siano conformi ai controlli sulle informazioni, etichettati come generati sinteticamente e che siano adottate misure per mitigarne l’abuso, inclusa la necessità di un’etichetta evidente quando

⁷⁸³ *Id.*

⁷⁸⁴ *Id.* Vedi anche Sheehan, M., & Du, S. (2 novembre 2022). *How food delivery workers shaped Chinese algorithm regulations*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/2022/11/02/how-food-delivery-workers-shaped-chinese-algorithm-regulations-pub-88310>.

⁷⁸⁵ *Id.*

⁷⁸⁶ *Id.*, 13.

⁷⁸⁷ Per la definizione di *deepfake*, vedi nota 510.

⁷⁸⁸ *Provisions on the Administration of Deep Synthesis in Internet-Based Information Services*, 互联网信息服务深度合成管理规定. *China Law Translate* (1 febbraio 2023). <https://www.chinalawtranslate.com/deep-synthesis/>.

potrebbero causare confusione o indurre in errore il pubblico⁷⁸⁹. Così come era accaduto nel caso precedente, l'autorità cinese ha basato la propria attività regolamentare sulla preoccupazione derivante dalla diffusione dei *deepfakes*, in grado anch'essi di manipolare l'opinione pubblica⁷⁹⁰. Anche in questo caso, è evidente l'approccio assolutamente settoriale di una disciplina diretta unicamente ai sistemi in grado di manipolare testi, video e immagini. Curiosamente, la normativa indirizzata a regolare queste attività è entrata in vigore solo cinque giorni prima dell'uscita di ChatGPT⁷⁹¹. Dopo la messa a disposizione sul mercato del noto strumento di IA conversazionale, e dopo aver valutato le conseguenze da esso generato nell'intero settore, la Cina ha ritenuto necessario proporre un'ulteriore integrazione alla disciplina dei sistemi di c.d. *deep synthesis*.

Per questo motivo, il legislatore cinese ha presentato nel 2023 una prima bozza di disciplina per quei sistemi di IA generativa che ricomprendano anche le applicazioni di chatbot conversazionali come ChatGPT, denominata *Measures for the Management of Generative Artificial Intelligence Services*⁷⁹². Come osservato da Sheehan, l'utilizzo di etichette per permettere l'individuazione di contenuti multimediali generati dall'IA non è applicabile allo stesso modo ai testi generati dagli LLMs, né una pratica simile si adatta a tutti i tipi di IA, che è progressivamente sempre più *general-purpose*⁷⁹³. In queste previsioni, non ancora entrate in vigore ma ancora allo stato di mera proposta, si estende in *primis* l'obbligo di inserire i sistemi di IA generativa utilizzati sul mercato all'interno dell'*AI registry*⁷⁹⁴; in secondo luogo, vi sono dei requisiti specifici in materia di qualità e gestione dei dati dei *training* dei sistemi di IA generativa, che, *inter alia*, non devono presentare alcun contenuto in grado di violare i diritti di proprietà intellettuale⁷⁹⁵. Il

⁷⁸⁹ Sheehan (2023), 12-13. Per quanto riguarda la pratica dell'etichettatura dei contenuti *AI-generated*, si sottolinea come l'approccio cinese non differisce in gran parte con quello degli USA nell'*Executive Order* 14110, nella previsione del *watermarking* così come definito nella nota 716. Questo approccio in via generale non differisce da quanto previsto anche dalla disciplina europea contenuta nell'*AI Act*, in cui vi è esplicitata la necessità di indicare la matrice artificiale di un dato contenuto con mezzi intellegibili sia dalle macchine che dagli utenti. Vedi il paragrafo 3.3.3. del Capitolo 2.

⁷⁹⁰ Sheehan (2023), 12-13.

⁷⁹¹ *Id.*, 13.

⁷⁹² *Measures for the Management of Generative Artificial Intelligence Services (Draft for comment)* (国家互联网信息办公室关于《生成式人工智能服务管理办法（征求意见稿）》公开征求意见的通知) – *DigiChina*. (19 aprile 2023). *DigiChina*. <https://digichina.stanford.edu/work/translation-measures-for-the-management-of-generative-artificial-intelligence-services-draft-for-comment-april-2023/>.

⁷⁹³ Sheehan (2023), 14.

⁷⁹⁴ *Id.*

⁷⁹⁵ *Id.*

regolamento, inoltre, stabilisce che l'IA generativa non debba discriminare sulla base della razza o del sesso e che i contenuti generati debbano essere necessariamente veritieri e accurati⁷⁹⁶. In conclusione, la nuova bozza di regole indirizzate alla disciplina dell'IA generativa, emanate nel 2023, consente di sottolineare come, anche in Cina, le preoccupazioni relative alla regolamentazione di questa particolare applicazione dell'IA siano diffuse, allo stesso modo che negli USA e in UE.

Il quadro generale di *governance* dell'IA cinese è dunque variegato, fornito di diversi testi normativi che si occupano di disciplinare alcune caratteristiche o particolari applicazioni dell'IA. Tra gli obiettivi dell'autorità governativa cinese non vi è tuttavia unicamente l'esplicazione di forme di controllo e sorveglianza sulla popolazione, come spesso viene ricordato dai commentatori occidentali, secondo i principi di autoritarismo digitale più volte menzionato⁷⁹⁷. Infatti, vi sono altre prerogative che guidano l'inclinazione alla regolamentazione: tra queste, vi è l'affrontare l'impatto sociale, etico ed economico che l'IA sta avendo sul popolo cinese, nonostante il controllo politico predominante del PCC⁷⁹⁸. Inoltre, altri obiettivi rilevano ai fini di questa analisi, come ad esempio la volontà di ottenere il *first mover advantage* più volte menzionato nel corso di questo elaborato⁷⁹⁹.

L'uso di questo approccio sostanzialmente verticale (in contrapposizione con quello orizzontale adottato dall'UE con l'*AI Act*, che mira ad essere una disciplina omnicomprensiva di questa tecnologia) è strumentale, secondo l'autorità cinese, a mantenere una dinamicità normativa. Nel metodo di regolamentazione cinese, la verticalità si fonde con l'iteratività: nel caso in cui la tecnologia dovesse avanzare troppo in fretta rispetto alla legge, che solitamente fatica a stare al passo con l'innovazione, nulla vieta l'emanazione di nuove regole per il particolare settore oggetto di questa accelerazione innovativa⁸⁰⁰.

⁷⁹⁶ *Id.*

⁷⁹⁷ *Id.*, 17.

⁷⁹⁸ *Id.*

⁷⁹⁹ Per la definizione di *first mover advantage* si veda il paragrafo 3.2 del Capitolo 2.

⁸⁰⁰ *Id.*, 16. *Vedi anche* Blomquist (2022), 6.

III. 1 Comparazione, convergenze a livello internazionale e standardizzazione

L'approccio comparatistico adottato in questa analisi offre una panoramica strumentale a comprendere le varie ipotesi di struttura regolamentare a disposizione dei legislatori a livello internazionale. Un dato è tuttavia evidente: il tema del regolamento dell'IA come intera disciplina tecnologica, così come inquadrata nel Capitolo 1 di questo elaborato⁸⁰¹, è divenuto di rilevanza assoluta, soprattutto negli ultimi mesi⁸⁰². Molti paesi, al di là di quelli precedentemente oggetto di analisi, sono in procinto di redigere, valutare, delineare, dei principi applicabili all'IA, in quanto sono evidenti i rischi ad essa connessi⁸⁰³ (nonché le possibilità di progresso ed innovazione se adeguatamente sfruttate⁸⁰⁴). La corsa alla disposizione di una cornice normativa tra i vari paesi non è mai stata così concitata; tuttavia, nell'evitare una c.d. *race to the bottom*⁸⁰⁵, cioè un'infruttuosa gara al ribasso a fronte della necessità di emettere celermente principi, norme imperative o linee guida di qualsiasi tipo, gli stati devono necessariamente interrogarsi su quale sia l'approccio regolamentare che maggiormente risponde alle proprie prerogative. Per questo è stato cruciale sottolineare non solo i punti di forza dei tre approcci oggetto della precedente trattazione, quanto le criticità che determinati assetti normativi possono generare. Tutto ciò, considerando gli effetti su quella coppia concettuale di regolamentazione-innovazione, più o meno concepibili in antitesi tra di loro a seconda del modello regolamentare adottato.

Ad esempio, alcuni ritengono che la scelta di una regolamentazione omnicomprensiva, orizzontale, che ricomprenda la nozione in generale di IA come quella europea riportata nell'*AI Act*, sia prematura dal punto di vista dello sviluppo tecnologico ed eccessivamente ambiziosa⁸⁰⁶; così configurata, la regolamentazione

⁸⁰¹ Vedi Benbouzid *et al.*, (2022), 5.

⁸⁰² Smuha, N. A. (2023c). Biden, Bletchley, and the emerging international law of AI. *Verfassungsblog*. <https://doi.org/10.17176/20231115-113402-0>.

⁸⁰³ Per un efficace visualizzazione dei paesi attualmente impegnati nella redazione di una qualsivoglia attività di regolamentazione dell'IA, vedi *Global AI Regulation Tracker*. (n.d.). <https://www.techieray.com/GlobalAIRegulationTracker>. Questo Global AI Regulation Tracker consiste in una risorsa completa che traccia le regolamentazioni globali riguardanti l'IA. Questo sito fornisce un'analisi dettagliata delle leggi, delle politiche e delle iniziative normative in tutto il mondo, offrendo agli utenti un'ampia panoramica sul tema.

⁸⁰⁴ Floridi (2022), 280-281.

⁸⁰⁵ Smuha (2021b), 5.

⁸⁰⁶ Vedi il commento della Casa Bianca nel 2020, in relazione ai potenziali rischi di iperregolamentazione che gli USA intendevano segnalare alle istituzioni europee che si

rischia di frenare eccessivamente la spinta propulsiva dell'innovazione a causa, *inter alia*, degli elevati costi di *compliance* che essa è in grado di imporre, soprattutto nei confronti delle PMI⁸⁰⁷. Da questo punto di vista, alcuni hanno ritenuto più efficaci iniziative di settore come quelle dell'autorità cinese, volte a disciplinare particolari applicazioni dell'IA, e non l'intera categoria in maniera trasversale⁸⁰⁸. La Cina, come precedentemente analizzato, ha infatti mantenuto un approccio più attendista per quanto riguarda l'emanazione di una disciplina orizzontale omnicomprensiva che fosse applicabile a tutti i dispositivi riconducibili all'IA⁸⁰⁹: una legislazione simile è nei piani futuri del governo cinese, ma non di immediata emanazione come suggerisce al contrario la *timeline* dell'*AI Act*⁸¹⁰. Al contrario, si potrebbe ugualmente inferire, che l'approccio cinese rischia di creare vuoti normativi consistenti, dal momento che si concentra maggiormente sull'emanazione di diversi impianti di discipline settoriali a seconda dei diversi sistemi di IA e delle loro relative caratteristiche peculiari⁸¹¹.

Anche l'approccio regolamentare statunitense non è privo di criticità che, complessivamente considerate, possono incidere sulla dicotomia regolamentazione-innovazione. Infatti, nonostante l'approccio generalmente *market-driven*, che lascia maggiore spazio alle aziende di sviluppare e implementare le più varie applicazioni dell'IA, una simile attitudine non è certo priva di rischi. La sotto-regolamentazione dell'IA può infatti portare all'esplicazione di quei rischi ampiamente osservati nel Capitolo 1, che possono avere come effetto, in ultimo, quello di danneggiare considerevolmente non solo gli individui e la collettività, ma che possano anche avere un effetto controproducente nei confronti di quell'ambiente di fervida innovazione che pone gli USA in testa nello sviluppo e utilizzo di sistemi di IA a livello globale⁸¹².

A questo proposito, i paesi che intendono fornire una qualsiasi regolamentazione all'IA nonostante l'assoluta polifunzionalità e multiformità di questa tecnologia, tenderanno (auspicabilmente) ad una maggiore convergenza a

apprestavano alla redazione della prima bozza dell'*AI Act*. Shepardson, D. (7 gennaio 2020). White House proposes regulatory principles to govern AI use. *Reuters*. <https://www.reuters.com/article/idUSKBN1Z60GL/>.

⁸⁰⁷ *Vedi in generale* Mueller, B. (2021). How much will the Artificial Intelligence Act cost Europe? In *Center for Data Innovation*. <https://www2.datainnovation.org/2021-aia-costs.pdf>

⁸⁰⁸ Sheehan (2023), 15-18.

⁸⁰⁹ *Id.*

⁸¹⁰ *Id.*

⁸¹¹ *Id.*

⁸¹² *Vedi* il paragrafo I.1.

livello internazionale nella scelta di regole da implementare. Come è stato osservato in precedenza, sono stati già individuati degli indicatori di questa convergenza a livello internazionale su temi quali ad esempio la definizione normativa di IA (in questo caso tra USA e UE⁸¹³). Una simile tendenza è motivata in *primis* dalla direzione che la società digitale sta complessivamente prendendo: infatti, è auspicabile che nel futuro si verifichi un sempre crescente collegamento digitale tra paesi, individui, istituzioni e aziende, tramite gli innumerevoli nodi di interconnessione che la tecnologia ha messo a disposizione negli anni; l'IA non farà che accelerare questo processo che diverrà di “iperconnessione” sia a livello generalizzato (ad esempio tramite l'implementazione maggiore di applicazioni dell'IA nella pubblica amministrazione), sia a livello localizzato (con la creazione di ambienti in cui è disseminata una molteplicità di *devices* che utilizzano l'IA per raccogliere dati rispetto l'ambiente circostante, e utilizzarli per le proprie computazioni). Inoltre, la natura assolutamente transnazionale dello sviluppo e utilizzo dei sistemi di IA, i quali possono essere prodotti e forniti in luoghi assolutamente diversi, in paesi diversi, e presentare una forte frammentazione territoriale lungo tutta la catena di valore⁸¹⁴, impone una visione necessariamente *cross-border* di una serie di principi, che devono essere fissati il più possibile in maniera condivisa dalla comunità internazionale.

Per questi motivi, è importante ricordare uno dei tentativi in corso di fornire un quadro di regolamentazione per l'IA con uno strumento di diritto internazionale⁸¹⁵: il disegno di Convenzione sulla disciplina dell'IA (*Draft Framework Convention On Artificial Intelligence*) oggetto dei lavori della *Committee on Artificial Intelligence* (CAI) istituita presso il Consiglio d'Europa⁸¹⁶. Come esposto nel paragrafo 3.5 del Capitolo 2, lo sforzo di convergenza di principi per le fondamenta dello sviluppo e l'utilizzo di applicazioni dell'IA a livello globale, è al centro dell'intento del Consiglio di Europa, tenendo comunque in considerazione anche quell'intrinseca volontà di fornire “per primi” una cornice condivisa a livello internazionale⁸¹⁷. Le negoziazioni in seno al Consiglio d'Europa

⁸¹³ Floridi (2023). *Vedi anche* il paragrafo 2.3 del Capitolo 1.

⁸¹⁴ Thomas *et al.* (2023).

⁸¹⁵ Smuha, (2023c).

⁸¹⁶ *Vedi nota* 586.

⁸¹⁷ Smuha (2023c).

sono ad uno stadio abbastanza avanzato⁸¹⁸ ed è già stata pubblicata una prima bozza di convenzione a dicembre 2023⁸¹⁹: tuttavia, essa ha già in parte deluso le iniziali aspettative che erano state riposte in relazione alla capacità del trattato di incidere nel panorama internazionale della regolamentazione dell'IA, come si vedrà appena seguito.

Ad ogni modo, la convenzione prevede l'introduzione di una serie di principi definitivi, a seguito dei quali vengono indicate delle obbligazioni generali quali il rispetto dei diritti umani, dell'integrità dei processi democratici e dello stato di diritto⁸²⁰, nonché principi che devono essere rispettati durante tutta la durata del c.d. *AI lifecycle*⁸²¹. La convenzione, inoltre, indica una serie di rimedi generali per gli individui a cui non è stata assicurata la protezione dei diritti fondamentali come previsto dalla convenzione stessa⁸²², nonché delle indicazioni integrative per quanto riguarda le attività di valutazione e mitigazione dei rischi inerenti ai sistemi di IA⁸²³.

Sulla base della bozza di trattato pubblicata lo scorso dicembre e delle successive notizie riportate dalle testate europee in relazione ai conseguenti negoziati ancora in corso, il trattato sembrerebbe prevedere una forte limitazione di applicabilità per quanto riguarda l'intero settore privato⁸²⁴. Il Consiglio d'Europa avrebbe accettato un campo di applicazione ridotto a fronte di una sottoscrizione del trattato il più ampia e condivisa possibile a livello globale: tuttavia, è evidente l'influenza di paesi come gli USA nella negoziazione proprio dal punto di vista di

⁸¹⁸ Infatti, la proposta dell'*AI Act* e la bozza di convenzione internazionale del Consiglio di Europa sono state pensate per poter procedere su binari sostanzialmente paralleli per quanto riguarda le tempistiche dei diversi procedimenti legislativi e di negoziazione del contenuto del trattato. Vedi Open letter to COE AI Convention negotiators: Do not water down our rights (2024). *European Center for Not-for-profit Law*. <https://ecnl.org/news/open-letter-coe-ai-convention-negotiators-do-not-water-down-our-rights>.

⁸¹⁹ *Committee On Artificial Intelligence* (CAI) istituita presso il Consiglio d'Europa, *Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy And The Rule Of Law*, 18 dicembre 2023, CAI(2023)28 (d'ora in avanti denominate come *Draft Framework Convention On Artificial Intelligence*).

⁸²⁰ Art. 4 e ss., *Draft Framework Convention On Artificial Intelligence*.

⁸²¹ Per la definizione di *AI lifecycle*, vedi Edwards (2022), 6. Vedi anche art. 6 e ss., *Draft Framework Convention On Artificial Intelligence*.

⁸²² Art. 14 e ss., *Draft Framework Convention On Artificial Intelligence*. Vedi anche Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2023). Global AI governance: barriers and pathways forward. *Social Science Research Network*, 4 <https://doi.org/10.2139/ssrn.4588040>.

⁸²³ art. 16 e ss., *Draft Framework Convention On Artificial Intelligence*.

⁸²⁴ Bertuzzi, L., (6 Giugno 2023b). EU's AI ambitions at risk as US pushes to water down international treaty. www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/eus-ai-ambitions-at-risk-as-us-push-to-water-down-international-treaty/>. Vedi anche Bertuzzi, L., (12 gennaio 2024). EU prepares to push back on private sector carve-out from international AI treaty. www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/eu-prepares-to-push-back-on-private-sector-carve-out-from-international-ai-treaty/>.

questo indebolimento dell'*enforcement* della Convenzione⁸²⁵, che risulterebbe applicabile solo nell'ambito dell'IA utilizzata nel settore pubblico, lasciando ai Paesi firmatari la facoltà di decidere se optare per ricomprendere anche il settore privato, oppure limitarsi a disciplinare solo il pubblico⁸²⁶. Un assetto simile, ancorchè ancora in via di definizione dal momento che la terza ed ultima lettura del testo del trattato presso il Consiglio è prevista per marzo 2024, sarebbe foriero di un presupposto particolarmente preoccupante: i diritti fondamentali e le libertà individuali, attivamente oggetto della protezione della bozza di Convenzione, risultano non meritare una tutela nel settore privato alla stregua di quello pubblico, lasciando scoperti di tutela gli individui nei confronti dei sistemi di IA che possono essere utilizzati dalle aziende private⁸²⁷.

In ultimo, occorre menzionare un'ulteriore strategia di regolamentazione dell'IA a livello ancorchè internazionale che tuttavia non coinvolge direttamente un'attività normativa *tout court*: l'elaborazione di standard tecnologici di settore da parte degli istituti specializzati. I processi di c.d. standardizzazione consentono di introdurre dei modelli a cui fare riferimento in caso di particolari applicazioni dell'IA o generalizzate. Un esempio dell'utilizzo della standardizzazione proprio nel settore dell'IA è l'elaborazione della norma tecnica ISO/IEC 42001:2023, denominata "*Information Technology Artificial intelligence - Management system*" (AIMS)⁸²⁸. Questa norma tecnica, sviluppata dall'*International Organisation of Standardisation* (ISO), ha l'obiettivo di indicare alle organizzazioni che intendono implementare sistemi che utilizzano l'IA i requisiti per gestirli responsabilmente, nell'ottica di una sapiente gestione dei rischi connessi, con delle specifiche attenzioni alla tutela dei dati personali, nonché alla sicurezza, alla trasparenza,

⁸²⁵ A questo proposito, alcuni hanno sottolineato la problematicità di aver sostanzialmente esautorato dal processo di *drafting* del trattato la società civile, in quanto le negoziazioni si sono svolte sostanzialmente a porte chiuse, contravvenendo con i principi di trasparenza predisposti dalle policy interne al Consiglio d'Europa. Bertuzzi, L., (17 gennaio 2023a) US obtains exclusion of NGOs from drafting AI treaty. *www.euractiv.com*. <https://www.euractiv.com/section/digital/news/us-obtains-exclusion-of-ngos-from-drafting-ai-treaty/>.

⁸²⁶ Bertuzzi (2023b).

⁸²⁷ *Id.*

⁸²⁸ ISO/IEC 42001:2023. (18 dicembre 2023). ISO. <https://www.iso.org/standard/81230.html>. Questo standard non è il primo inerente all'IA e alle sue modalità di utilizzo. L'ISO aveva già provveduto in precedenza a fornire degli standard inerenti al settore dell'IA, all'inizio del 2023 e nel 2022; tuttavia, quest'ultimo è il *framework* più aggiornato.

prevedendo una struttura armonizzata di *governance* facilmente integrabile e scalabile⁸²⁹.

L'elaborazione di regole da parte enti di standardizzazione è uno dei possibili strumenti di *governance* c.d. "ibrida,"⁸³⁰ dal momento che coinvolge entità sia del settore pubblico che di quello privato. Una strategia simile è anche quella utilizzata dall'Unione Europea per quanto riguarda le regolamentazioni sulla sicurezza dei prodotti⁸³¹ (tra cui figura anche l'*AI Act*⁸³²). Modelli simili consentono di superare agevolmente la natura normalmente localizzata dell'*hard law*, rendendosi accessibili ad un'ampia varietà di soggetti che possono beneficiare degli orientamenti e, soprattutto, delle standardizzazioni provenienti da enti qualificati

III. 2 Alla ricerca di un equilibrio regolamentare ottimale: dalle valutazioni d'impatto sui diritti fondamentali e sull'etica (HRESIA) all'AI literacy

È dunque evidente che, a fronte delle considerazioni appena ultimate nel paragrafo precedente, nonché di tutto l'elaborato complessivamente considerato, per regolamentare in maniera efficace e trasversale l'IA, l'elaborazione di sole cornici normative a sé stanti non è sufficiente, né deve essere la sola strategia da utilizzare nella regolamentazione di questa tecnologia estremamente all'avanguardia. Sebbene l'utilizzo di fonti del diritto *stricto sensu* (internazionale, europeo o interno che siano) possa essere un *input* utile per delineare i contorni di legittimità di operatività di sistemi *AI-powered*, ad esempio grazie alla previsione di particolari applicazioni o specifiche modalità di utilizzo vietate⁸³³, tuttavia un tale approccio può beneficiare di una serie di ulteriori considerazioni di corollario, sia per quanto riguarda gli elementi che possono essere inseriti all'interno delle leggi a disciplina dell'IA, sia all'utilizzo di strumenti di regolamentazione paralleli di *soft law*, o a pratiche del tutto scollegate da una disciplina legislativa in senso stretto. Ciò occorre per far sì che l'intera classificazione dei rischi inerenti all'IA,

⁸²⁹ Vedi ISO/IEC 42001:2023. Vedi anche Perego, M. (4 gennaio 2024). Il sistema di gestione dell'intelligenza artificiale: la norma ISO/IEC 42001:2023. *Federprivacy*. <https://www.federprivacy.org/informazione/primo-piano/il-sistema-di-gestione-dell-intelligenza-artificiale-la-norma-iso-iec-42001-2023>.

⁸³⁰ Prifti, K., Morley, J., Novelli, C., Floridi, L., (2024). Regulation by design: features, practices, limitations, and governance implications. *Centre for Digital Ethics (CEDE) Research Paper Series*, 18 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4724454.

⁸³¹ Prifti (2024), 18.

⁸³² Per la definizione di *AI Act* come *product safety regulation*, vedi il paragrafo 3.4 del Capitolo 2.

⁸³³ Vedi il sottoparagrafo 3.3.2. del capitolo 2.

così come delineati nella tabella 1 del Capitolo 1⁸³⁴, siano in qualche modo affrontati, trasversalmente o direttamente, da un set di regole e principi applicabili, nell'ottica di implementare quel concetto di *broad AI assurance*⁸³⁵, così come delineato da Thomas *et al.* Per cui, l'obbiettivo del presente paragrafo è quello di indicare ulteriori spunti di natura regolamentare, all'interno del più ampio spettro di possibilità che vede alle proprie estremità, la *hard law* da un lato, e le *best practices* auto-imposte e standard dell'industria che possono anch'essi servire a regolamentare l'IA su base sostanzialmente volontaria.

Innanzitutto, occorre considerare quali elementi possono essere inseriti all'interno della legislazione nazionale e internazionale che, secondo alcuni studiosi, possono incidere sulla qualità dell'IA prodotta in un dato settore. Questo è il caso della predisposizione delle valutazioni di impatto di diverso tipo, tra cui spicca quella relativa ai diritti fondamentali. In generale, nel contesto dell'IA, gli *assessment* assumono varie forme e possono essere di diversi tipi, ma la loro importanza può anche esorbitare le mere valutazioni tecniche, essendo in realtà suscettibili di proteggere anche diritti e valori fondamentali⁸³⁶. Questo genere di valutazioni non è nuovo nel campo normativo della tecnologia complessivamente considerata, dal momento che sono previsti degli *impact assessment* sia per quanto riguarda la disciplina europea in materia di protezione dei dati⁸³⁷, ma anche proprio nell'*AI Act*, specificatamente in materia di salvaguardia dei diritti fondamentali⁸³⁸.

Le valutazioni di impatto sui diritti fondamentali dei sistemi di IA, soprattutto quelle impiegate nel settore pubblico, sono cruciali per mitigare quanto più possibile quei rischi inerenti allo sviluppo di pregiudizi integrati nelle applicazioni, che si propagano e riverberano in ogni computazione, valutazione o predizione della macchina⁸³⁹. Come infatti sottolineato da Mantelero, “*proprio come i clienti non devono controllare la sicurezza delle auto che acquistano, allo stesso modo gli utenti finali dei sistemi di intelligenza artificiale non dovrebbero*

⁸³⁴ Thomas *et al.* (2023), 4.

⁸³⁵ Si veda lo stesso titolo dell'articolo citato redatto da Thomas *et al.*: “*The case for a broader approach to AI assurance: addressing “hidden” harms in the development of Artificial intelligence.*”

⁸³⁶ Mantelero (2022), 13-14.

⁸³⁷ Vedi Art. 35 GDPR (che prevede le c.d. valutazioni d'impatto sulla protezione dei dati, anch'esse richieste quando un certo trattamento dei dati “*può presentare un rischio elevato per i diritti e le libertà delle persone fisiche*”).

⁸³⁸ Vedi Art. 29(a), *AI Act*, (Versione IV).

⁸³⁹ Mantelero (2022), 191.

dover verificare se i loro diritti e la loro libertà sono tutelati,” essendo questo piuttosto un dovere in capo a chi li produce, allena o mette sul mercato⁸⁴⁰.

A questo proposito, Mantelero propone il modello del c.d. *Human Rights, Ethical and Social Impact Assessment* (HRESIA), una valutazione di impatto che combina la dimensione ampia del rispetto dei diritti umani, con quella più contenuta dell’analisi sull’impatto di un dato sistema in relazione ai valori etici di una società⁸⁴¹. Tale modello non si limita ad essere una mera *checklist* di requisiti da completare per rendere *compliant* il sistema di IA, ma piuttosto costituisce un processo articolato di analisi e valutazione che mira a rendere i sistemi di IA orientati positivamente nei confronti dell’umanità (*human-centered*), fin dalle prime fasi di sviluppo⁸⁴². Ciò implica dunque l’idea che gli utenti finali non debbano essere costretti a verificare autonomamente la tutela dei propri diritti e delle proprie libertà, ma che tali aspetti siano intrinsecamente incorporati nei sistemi stessi a partire dalla propria progettazione.

Un assetto come quello previsto dalla HRESIA prevede una vicendevole integrazione tra valutazioni di impatto *ex ante* unicamente attinenti al rispetto dei diritti umani fondamentali⁸⁴³ e valutazioni di tipo etico e sociale, utili ad essere modulate a seconda dei contesti regionali e che consentono un approccio flessibile all’*assessment* delle applicazioni dell’IA⁸⁴⁴. Una simile integrazione è necessaria: infatti, l’analisi del semplice impatto a livello sociale (ancorchè regionalmente localizzato) di un dato sistema di IA è molto più complesso rispetto a quella che si limita ad analizzare l’influenza generalizzata sui diritti umani, concettualmente più determinati come punti di riferimento per un simile *assessment*⁸⁴⁵. Al contrario, valori etici e standard socialmente determinati sono tuttavia meno precisi e più complessi da focalizzare correttamente per portare a termine una valutazione ben determinata. Questo è uno dei motivi per cui il settore generale dell’etica in relazione all’IA coinvolge una varietà amplissima di contributi prettamente teorici, così come una proliferazione di linee guida, in alcuni casi parzialmente influenzate

⁸⁴⁰ *Id.*, 11.

⁸⁴¹ *Id.*, 16.

⁸⁴² Mantelero (2022), 16.

⁸⁴³ Come un tipo di *assessment* denominato HRIA (*Human Right Impact Assessment*), al cui genere appartiene la procedura più volte citata in precedenza, denominata FRIA e prevista nell’*AI Act*, art. 29(a), *AI Act*, (Versione IV). Vedi anche Magrani, E., & Guedes Fernandes da Silva, P., (2023), *the Ethical and Legal Challenges of Recommender System Driven by Artificial Intelligence*, 157, in Antunes et al. (2023). *Multidisciplinary perspectives on artificial intelligence and the law. Springer Nature*.

⁸⁴⁴ Mantelero (2022), 16.

⁸⁴⁵ *Id.*, 130-132.

da quell'attività di “*bluwashing*”⁸⁴⁶ o semplicemente che riflettano i valori aziendali anziché quelli realmente etici⁸⁴⁷. Per questo motivo, sarà necessario prevedere un processo di valutazione *multistakeholder*, che coinvolga in determinate fasi anche la società civile, proprio per scongiurare il rischio di insinuarsi di pratiche etiche di sola facciata⁸⁴⁸. Infine, l'utilizzo di valutazioni di impatto come l'HRESIA, non deve essere percepita come un ostacolo allo sviluppo dei sistemi IA, o come un eccessivo fardello burocratico per raggiungere un certo livello di accettabilità di un sistema: al contrario, modelli di *assessment* come l'HRESIA sono opportunità per far venir meno quel grado di incertezza che caratterizza l'innovazione tecnologica, incrementandone la sicurezza nei confronti del rispetto dei diritti fondamentali, ma anche dell'etica e delle norme sociali in un dato contesto⁸⁴⁹.

Le valutazioni di impatto non sono tuttavia l'unico strumento possibilmente integrato all'interno di testi normativi (e non) che consente una regolamentazione intelligente dell'IA, anche nell'ottica di prevenire e non solo mitigare i rischi di danni a livello individuale, collettivo e sociale. Per assicurare un livello alto di quella che è stata definita sicurezza dell'IA (*AI assurance*) lungo tutta la catena di valore dell'IA, sempre più complessa ed estesa a livello territoriale⁸⁵⁰, Thomas *et al.* forniscono una serie di proposte contestualmente rivolte sia ai i regolatori (anche denominati “responsabili politici”) che ai soggetti dell'industria dell'IA⁸⁵¹. Secondo le loro raccomandazioni, per i responsabili politici, è imprescindibile promuovere lo sviluppo di quadri normativi e di valutazione specifici per gestire in modo proattivo l'impatto sociale e anche quello ambientale dell'IA, oltre a intervenire efficacemente per colmare le lacune nell'attuale approccio del settore attraverso l'implementazione di soluzioni politiche mirate⁸⁵². Per gli attori dell'industria, invece, è di vitale importanza integrare le strategie di garanzia dell'IA all'interno della gestione generale dei rischi digitali lungo l'intera catena di fornitura, con un particolare *focus* sull'adozione di un approccio di investimento

⁸⁴⁶ Per la definizione di *bluwashing*, vedi Floridi (2022), 112. Vedi anche nota 372.

⁸⁴⁷ Mantelero (2022), 130-132.

⁸⁴⁸ *Id.*

⁸⁴⁹ *Id.*, 50.

⁸⁵⁰ Thomas *et al.* (2023), 5.

⁸⁵¹ *Id.*, 10-16.

⁸⁵² *Id.*, 10-13.

orientato alla sostenibilità, il quale potrebbe svolgere un ruolo chiave nel promuovere uno sviluppo responsabile e sostenibile dell'IA⁸⁵³.

Queste pratiche sono suggerite nell'ottica di produrre un'IA non solo accurata e precisa, ma che sia anche propriamente indicata come appunto "sostenibile," secondo quella concezione di sostenibilità ampia proposta da Heilinger *et al.*, non limitata alle sole implicazioni prettamente ecologiche⁸⁵⁴. A questo proposito, per IA sostenibile si deve piuttosto intendere un generale "movimento per promuovere il cambiamento nell'intero ciclo di vita dei prodotti di IA [...] verso una maggiore integrità ecologica e [soprattutto] di giustizia sociale."⁸⁵⁵

Tuttavia, anche prevedendo l'inserimento degli elementi appena considerati in un'ipotetica disciplina dell'IA, ciò non basterebbe a creare, ad avviso di chi scrive, un ecosistema di fiducia e di sviluppo nel rispetto dei diritti fondamentali e degli individui, nonché della società collettivamente considerata e nel rispetto della democrazia e dello stato di diritto, senza la previsione di un ultimo (fondamentale) tassello. La c.d. "AI literacy," intesa come "alfabetizzazione" all'IA, è infatti un'ulteriore implementazione che può fornire un apporto fondamentale in quell'inevitabile processo di progressiva integrazione dell'IA nella vita quotidiana. Il concetto di *AI literacy* è in *primis* emerso nella Versione II dell'*AI Act* (il mandato del Parlamento Europeo) ed era definita come quell'insieme di competenze che consentono ai fornitori, agli utenti e alle persone interessate, di utilizzare con cognizione di causa i sistemi di IA, nonché di essere consapevoli delle opportunità e dei rischi dell'IA e dei possibili danni che può causare, promuovendo così la sua democratizzazione in maniera orizzontale⁸⁵⁶. Sebbene questa definizione non sia stata mantenuta nell'ultima versione dell'*AI Act*, il concetto di alfabetizzazione all'IA viene comunque efficacemente riassunto nel considerando 9(b), che comunque rimanda all'importanza dell'*empowerment* di tutti i soggetti in relazione all'IA complessivamente considerata. Tra le altre cose, lo sviluppo ed utilizzo di questa maggiore consapevolezza nell'uso dell'IA viene utile anche nell'ottica di quegli obiettivi non solo di *compliance* previsti dalla normativa, ma in generale

⁸⁵³ *Id.*, 13-16.

⁸⁵⁴ Heilinger, J.C., Kempt, H. & Nagel, S. Beware of sustainable AI! Uses and abuses of a worthy goal (2023), *AI Ethics*. <https://doi.org/10.1007/s43681-023-00259-8>

⁸⁵⁵ Heilinger *et al.* (2023), 5.

⁸⁵⁶ Considerando 9(b), *AI Act*, (Versione II). Il concetto di *AI literacy*, *tuttavia*, è presente in più parti dell'*AI Act*, essendo menzionata anche nei considerando 58, 81, e negli artt. 4(b), 58 e 69.

per un contributo di importanza vitale per sostenere il consolidamento di un percorso di innovazione caratterizzato da un'IA *trustworthy* in tutta l'UE⁸⁵⁷.

Anche questo elemento, in congiunzione con tutto ciò che è stato precedentemente menzionato, può essere efficacemente utilizzato per completare il complesso e composito quadro della regolamentazione dell'IA, se correttamente bilanciato da un approccio assolutamente multidisciplinare e trasversale all'interno di tutti i settori⁸⁵⁸. Mai come in questo momento l'interdisciplinarietà assume un'importanza cruciale, dal momento che una corretta comprensione dell'IA e del suo funzionamento non è compatibile con una visione eccessivamente compartimentalizzata⁸⁵⁹.

III. 3 Conclusioni

Durante gli albori di Internet all'inizio degli anni '90, l'intero mondo accademico e politico si poneva dei quesiti pressoché simili a quelli che vengono oggi riferiti all'IA, sebbene in relazione alla regolamentazione di Internet. Il tema della potenziale, eventuale, normazione del cyberspazio in generale veniva diversamente affrontata dai commentatori, che avevano fornito diversi punti di vista in materia. Un passaggio importante in questo dibattito è la celebre Dichiarazione di Indipendenza del Cyberspazio redatta da John Perry Barlow nel 1996, in cui si è ribadito quell'approccio della dottrina cyberlibertaria già precedentemente menzionata quando si è trattato il modello statunitense di regolamentazione del digitale in generale⁸⁶⁰. Nella dichiarazione, Barlow ricordava come Internet non potesse essere in alcun modo considerato come uno spazio da regolamentare o governare da parte dei governi nazionali ed internazionali⁸⁶¹, ma che piuttosto doveva rimanere al di sopra di imposizioni normative di alcuna sorta, per tutelare a tutto tondo le libertà percepite come assolutamente incompressibili dei cittadini del cyberspazio⁸⁶². L'argomentazione proposta da Barlow prevede che una qualsiasi autorità legittima debba necessariamente scaturire in ultima analisi dal consenso

⁸⁵⁷ Considerando 9(b), *AI Act* (Versione IV).

⁸⁵⁸ Magrani & Guedes Fernandes da Silva (2023), 156.

⁸⁵⁹ *Id.*

⁸⁶⁰ Per una definizione completa di cyberlibertarismo, vedi nota 649.

⁸⁶¹ Vedi Barlow, J. P., *A declaration of the independence of cyberspace*. (1996). Electronic Frontier Foundation, <https://www.eff.org/it/cyberspace-independence>.

⁸⁶² *Id.* Vedi anche Koenig & Rustad (2021), 82-83.

degli utenti di Internet di tutto il mondo, e non piuttosto da legislatori, giudici o dalle grandi aziende del settore⁸⁶³.

Allo stesso modo, anche esperti del settore giuridico avevano avanzato simili conclusioni, che contrastavano con la creazione di una regolamentazione per Internet. *Ex multis*, il magistrato americano Frank H. Eastbrook ritenne sostanzialmente inutile la creazione di una regolamentazione specifica di Internet⁸⁶⁴, dal momento che non si crea una branca del diritto specifica per ogni questione particolare attinente all'esperienza umana. Eastbrook, con una provocazione, sottolinea come nessuno si sognerebbe mai di costituire una branca del diritto attinente ai cavalli, sebbene esistano tutta una serie di casi che coinvolgono questi animali (atti di compravendita, illeciti civili relativi ad incidenti, contratti di licenza⁸⁶⁵). La “legge del cavallo” non ha ragione di esistere, poiché a regolare questi rapporti sono sufficienti le categorie del diritto che esistono in via generale⁸⁶⁶.

A questa corrente di pensiero contraria alla regolamentazione di Internet, si pone uno dei più importanti pensatori del nostro tempo in relazione all'intersezione tra diritto e informatica: Lawrence Lessig⁸⁶⁷. Lessig riteneva che le derive cyberlibertarie che gridavano alla deregolamentazione dell'Internet fossero eccessivamente *naïf*⁸⁶⁸: la branca del diritto di Internet, al contrario, era in grado di cambiare ogni altro settore del diritto, fornendo degli elementi ulteriori all'esperienza giuridica e regolamentare, soprattutto su temi cruciali quali la globalizzazione, la proprietà intellettuale o la regolamentazione a livello privato⁸⁶⁹. Dal canto suo, Lessig indicava già all'inizio degli anni 2000 la necessità di incidere nella creazione di un'efficace Internet *governance* ispirata a quattro elementi che permettono di costituire delle limitazioni, regole, o principi ispiratori per quanto riguarda l'istituzione di una cornice disciplinare variegata e basata sulla combinazione di quattro “modalità:” in *primis*, l'utilizzo del codice sorgente e di architetture di sistema, secondo quel principio per il quale “*code is law*”⁸⁷⁰; in secondo luogo, le dinamiche economiche di mercato; in terzo luogo, la legge,

⁸⁶³ Koenig & Rustad (2021), 83.

⁸⁶⁴ Eastbrook, F. H. (1996), *Cyberspace and the Law of the Horse*, *University of Chicago Legal Forum*. Vedi anche Rustad (2021), 27-29.

⁸⁶⁵ Eastbrook (1996).

⁸⁶⁶ Rustad (2021), 27-29.

⁸⁶⁷ *Id.*

⁸⁶⁸ Koenig & Rustad (2021), 84.

⁸⁶⁹ Rustad (2021), 29.

⁸⁷⁰ *Vedi in generale* Lessig, L. (2006), *Code and other laws of Cyberspace* (Version 2.0).

complessivamente considerata; e in ultimo le norme sociali applicate da una data comunità⁸⁷¹. Ognuno di questi elementi non è ordinato secondo una gerarchia: sono tutti astrattamente idonei a fornire un tassello del composito *puzzle* regolamentare di Internet. A questo proposito, lo stesso Lessig critica fortemente l’approccio dei cyberlibertari nei confronti della *governance* di Internet, il quale viene da questi percepito come totalmente scollegato dal sistema legale tradizionale: una fallacia simile potrebbe portare alla compromissione di valori fondamentali come il rispetto della *privacy* individuale e di gruppo o la libertà di parola e di espressione⁸⁷².

Un discorso simile, ad avviso di chi scrive, può essere trasposto sull’attuale tema relativo alla necessaria regolamentazione dell’IA nell’avvenire. L’IA, allo stesso modo di Internet, possiede quella capacità trasformativa ed assolutamente rivoluzionaria, per la quale sono stati sollevati quesiti non dissimili da quelli che interessavano gli studiosi di Internet negli anni ’90, riguardo la preoccupante diminuzione dei posti di lavoro, nonché delle libertà e dei diritti degli individui a causa dell’incremento della sorveglianza individuale e sociale.

Per alcuni punti di vista, dunque, il dibattito relativo alla potenziale regolamentazione dell’IA e delle sue possibili forme non è dissimile da quanto appena riportato in tema di *Internet governance*: al pari proprio di quanto teorizzato da Lessig, la regolamentazione dell’IA dovrebbe sfruttare sia il proprio scopo “protettivo,” cioè in grado di minimizzare gli impatti negativi su individui e collettività, ma anche secondo quel suo ruolo “abilitante” lo stimolo all’innovazione⁸⁷³. Dunque, si auspica comunque una necessaria commistione di quei dei quattro elementi summenzionati, proposti da Lessig nell’ambito della regolamentazione di Internet, ma ugualmente applicabili anche alla disciplina dell’IA. La previsione di sole leggi, cornici normative alla stregua dell’*AI Act*, non sono né devono costituire l’unico strumento a disposizione dei regolatori per un corretto sviluppo e sfruttamento dell’IA.

Proprio secondo la dottrina di Lessig, essa deve essere necessariamente affiancata da corretti meccanismi di regolazione e vincoli del mercato, ad esempio atti a non permettere l’assunzione di posizioni dominanti né comportamenti predatori da parte delle grandi aziende (scopo protettivo), ma anche grazie ad

⁸⁷¹ *Id.* Vedi anche Rustad (2021), 29-32. Vedi anche Smuha (2021), 5.

⁸⁷² *Id.*, 33.

⁸⁷³ La nozione di *protective regulation* ed *enabling regulation* recepita dalla dottrina di Lessig viene riportata da Smuha in Smuha (2021), 5-6.

incentivi e investimenti che possano sostenere il progresso, soprattutto per quanto riguarda il sostegno alle *start-up* innovative del settore (scopo abilitante). Ancora, la previsione di comportamenti sociali informati all'utilizzo e alla comprensione dell'IA, secondo i principi di *AI literacy*⁸⁷⁴, può fornire da orientamento aggiuntivo per chi entra a contatto con strumenti di IA in generale, e ne comprende la portata.

Tra l'altro, si sottolinea che l'assoluta novità del settore normativo in analisi non è da considerare in secondo piano nel momento in cui si valuta la scelta di fornire una cornice normativa di alcuna sorta. Anche testi come l'*AI Act*, o la futura regolamentazione cinese rivolta a disciplinare in maniera orizzontale l'IA (non ancora oggetto di alcuna bozza da parte del legislatore), non sono altro che i primi tentativi di fornire una disciplina a questi sistemi; essi non sono un punto di arrivo, bensì un punto di partenza. Non si esclude la possibilità di interventi successivi che andranno a implementare lo scheletro previsto dagli strumenti di normazione: un esempio di questa continua implementazione che consente di rimanere al passo con lo sviluppo tecnologico è rappresentato dalle direttive sulla responsabilità civile dell'IA correntemente oggetto dell'attività normativa dell'Unione Europea⁸⁷⁵.

In conclusione, come più volte ribadito durante tutto il presente elaborato, la necessità di regolamentazione dell'IA non è più argomento di discussione, quanto piuttosto la forma che essa deve assumere, dando per scontata l'urgenza di fornire una cornice di principi informativi, regole, secondo le varie possibili strategie oggetto di questa trattazione. Si potrebbe infatti inferire che, così come l'argomentazione relativa alla de-regolamentazione di Internet abbia nel tempo lentamente perso la propria presa sulle istituzioni globali (con l'approvazione di numerose normative nel settore della *governance* di Internet), anche per quanto riguarda l'IA si deve auspicare un mutamento simile: solo con la minimizzazione e mitigazione dei rischi relativi all'applicazione dell'IA sarà possibile sfruttarne al meglio le proprie potenzialità e ottenere quei risultati di progresso che essa è effettivamente in grado di generare.

L'assoluta novità di questo settore normativo, in ultimo, non deve né spiazzare i responsabili politici chiamati ad intervenire nella regolamentazione dell'IA, né portare a ritenere che, dal momento che essa consiste in una delle tecnologie più trasformative del nostro tempo, essa non debba essere oggetto di alcuna limitazione o indirizzamento per un utilizzo più sicuro e realmente incentrato

⁸⁷⁴ Vedi il paragrafo III.2.

⁸⁷⁵ Vedi il paragrafo 4 del Capitolo 2.

ad assicurare il più alto grado di rispetto dei diritti fondamentali a livello individuale, nonché dello stato di diritto e della democrazia. Per cui, un quadro regolamentare è ormai un obbligo a cui le istituzioni dei vari paesi dovranno dare seguito, non in maniera frettolosa, nell'ottica figurare tra i vincitori della *race to AI regulation*, ma piuttosto per indirizzare sulla giusta via l'intero processo di innovazione e renderlo il più possibile *human-centered*. Esso dovrebbe peraltro necessariamente coinvolgere quegli elementi aggiuntivi che sono stati delineati in quest'ultima parte III del presente Capitolo, relativi alle valutazioni d'impatto sui diritti fondamentali, strumentali allo sviluppo consapevole della tecnologia senza intaccare le libertà dei singoli individui e limitando i più possibili usi malevoli o dannosi della stessa.

Conclusion

Il presente elaborato ha discusso in maniera trasversale l'attuale panorama relativo alla regolamentazione dell'IA, proponendosi di valutare in maniera comparativa i diversi approcci adottati dai singoli paesi o entità sovranazionali che occupano posizioni apicali nel settore dal punto di vista dello sviluppo e dell'avanzamento tecnologico dell'IA. In *primis*, si è proposta una visione particolare dell'IA tramite un suo inquadramento filosofico, tecnico e normativo, nonché incentrato a rilevarne opportunità e rischi, nell'ottica di comprendere meglio le strategie regolamentari successivamente proposte.

A fronte di questa analisi prodromica al tema della regolamentazione, l'IA emerge come uno strumento che rappresenta solo un ulteriore avanzamento della tecnologia, il quale, sebbene sia in grado di modificare radicalmente l'intera esperienza quotidiana umana per quanto riguarda generazione di testi, contenuti, nonché la predisposizione di calcoli, organizzazione del lavoro, mansioni, valutazioni e *output* decisionali che coinvolgono un numero amplissimo di variabili e parametri, non deve essere considerato in maniera avulsa dal settore dalla quale essa è emersa: è ormai chiaro che, a parere di chi scrive, l'IA non è una panacea⁸⁷⁶, né un arma di distruzione di massa, ma uno strumento tanto in grado di eccellere nel fornire soluzioni concrete e arricchire l'esperienza umana complessivamente considerata, ma anche tanto potenzialmente preoccupante per la capacità di assumere, assimilare, finanche istituzionalizzare, pregiudizi di matrice sistemica o umana. Assumere un atteggiamento che non ne riconosce queste caratteristiche in parallelo all'intento di fornire una regolamentazione permissiva o restrittiva che sia, può rendere inefficace l'intera previsione di qualsiasi norma, regola o principio.

Successivamente, l'analisi delle varie normative proposte in questo settore da varie entità sovranazionali tra cui l'UE, gli USA e la Repubblica Popolare Cinese, ha consentito di evidenziare quali siano i punti di forza e di debolezza i relativi approcci alla regolamentazione. Scopo di quest'analisi non è stato certamente quello di indicare quale modello sia il migliore per fornire un quadro di gestione dell'IA: ciò sarebbe impossibile in quanto i risultati derivanti dalle scelte regolamentari saranno evidenti solo negli anni a venire.

⁸⁷⁶ Vedi Floridi (2022), 86-88, 146-147.

Le future normative dovranno quindi essere flessibili e adattabili, in grado di evolversi al passo con il rapido progresso tecnologico. Dovranno inoltre essere basate su principi di trasparenza, responsabilità e inclusione, per garantire che l'IA sia utilizzata in modo etico e responsabile, a beneficio di tutti.

In definitiva, il successo della regolamentazione dell'IA dipenderà dalla capacità di trovare un equilibrio tra innovazione e sicurezza, tra progresso tecnologico e tutela dei diritti umani. Solo con un impegno collettivo e una visione lungimirante potremo costruire un futuro in cui l'IA sia una forza positiva per il bene dell'umanità.

Bibliografia

Fonti normative nazionali e internazionali, documenti delle istituzioni e di organizzazioni internazionali

- ◇ Accordo di Parigi, così come ratificato dall'UE; EurLex GU L 282 del 19 ottobre 2016.
- ◇ AI HLEG, Policy Recommendations for Trustworthy AI (giugno 2019)
<https://digital-strategy.ec.europa.eu/en/library/policy-and-investment-recommendations-trustworthy-artificial-intelligence>.
- ◇ Algorithmic Accountability Act of 2022, H.R.6580, 117th Cong. (2022).
- ◇ Camera dei deputati n. 1084, Proposta di legge d'iniziativa dei deputati Centemero, Andreuzza, Bagnai, Barabotti, Cavandoli, Di Mattina, Gusmeroli, Toccalini, Billi, Bof, Coin, Dara, Furgiuele, Pretto Zinzi, *Disposizioni concernenti l'adozione di una disciplina temporanea per la sperimentazione dell'impiego di sistemi di intelligenza artificiale*, (11 aprile 2023).
- ◇ Carta dei diritti fondamentali dell'Unione Europea, 7 giugno 2016, 2012 O.J. (C 326/02) (anche conosciuta come "Carta di Nizza").
- ◇ Commissione Europea, 'Artificial Intelligence for Europe' (Comunicazione) COM(2018) 237 final.
- ◇ *Committee On Artificial Intelligence (CAI)* istituita presso il Consiglio d'Europa, *Draft Framework Convention On Artificial Intelligence, Human Rights, Democracy And The Rule Of Law*, 18 dicembre 2023, CAI(2023)28
- ◇ Comunicazione della Commissione Europea "Verso uno spazio comune europeo dei dati" COM 232 final, 25 aprile 2018.
- ◇ Comunicazione della Commissione Europea, "Una strategia europea per i dati", COM 66 final, 19 febbraio 2020.
- ◇ Consiglio dell'Unione Europea (2020) *Council Conclusions on Regulatory Sandboxes and Experimentation Clauses as tools for an innovation-friendly, future-proof, and resilient regulatory framework that masters disruptive challenges in the digital age*, 2020/C 447/01,
- ◇ Consiglio Europeo, *Riunione del Consiglio Europeo (19 ottobre 2017) – Conclusioni* EUCO 14/17, 2017, 8.
- ◇ Consiglio Europeo, *Riunione Straordinaria del Consiglio Europeo (1° e 2 ottobre 2020) – Conclusioni*, EUCO 13/20, 2020.
- ◇ *Declaration of Cooperation on Artificial Intelligence*, 4 (10 aprile 2018),
<https://ec.europa.eu/jrc/communities/en/node/1286/document/eu-declaration-cooperation-artificial-intelligence#:~:text=Declaration%20signed%20at%20Digital%20Day%20on%2010th%20April%202018.&text=This%20Declaration%20builds%20on%20the,of%20a%20Digital%20Single%20Market>.

- ◇ Direttiva 95/46/CE del Parlamento Europeo e Del Consiglio del 24 ottobre 1995 relativa alla tutela delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati G.U. L 281 (*Data protection directive*).
- ◇ *EDPS Opinion 42/2023 on the Proposal for two Directives on AI liability rules*. (11 ottobre 2023). European Data Protection Supervisor. https://edps.europa.eu/data-protection/our-work/publications/opinions/2023-10-11-edps-opinion-422023-two-directives-ai-liability-rules_en.
- ◇ Emenda -CTIA (Substitutivo) ao Projeto de Lei 2.338/2023 (legislazione brasiliana).
- ◇ *Ethics guidelines for trustworthy AI*. (8 aprile 2019) <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.
- ◇ Exec. Ord. No. 14110.
- ◇ Exec. Order No. 13859, 84 Fed. Reg. 3967 § 1 (14 febbraio 2019).
- ◇ Exec. Order No. 14110 “*on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence*”, Fed. Reg. 75191 (30 ottobre 2023).
- ◇ Executive Office of the President & National Science and Technology Council Committee on Technology. (Ottobre 2016). *Preparing for the Future of Artificial Intelligence*. https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf.
- ◇ *High-level expert group on artificial intelligence, Shaping Europe’s Digital Future*. <https://digital-strategy.ec.europa.eu/en/policies/expert-group-ai>
- ◇ *Innovative Technology Arrangements and Services Act, Legizlazzjoni Malta* (1° novembre 2018), <https://legislation.mt/eli/cap/592/eng/pdf>.
- ◇ *ISO/IEC 42001:2023*. (18 dicembre 2023). ISO. <https://www.iso.org/standard/81230.html>.
- ◇ *Libro Bianco della Commissione Europea sull’Intelligenza Artificiale – Un approccio europeo all’eccellenza e alla fiducia*, 11 COM (2020) 65 final, 19 febbraio 2020.
- ◇ *Measures for the Management of Generative Artificial Intelligence Services (Draft for comment)* (国家互联网信息办公室关于《生成式人工智能服务管理办法（征求意见稿）》公开征求意见的通知) – *DigiChina*. (19 aprile 2023). *DigiChina*. <https://digichina.stanford.edu/work/translation-measures-for-the-management-of-generative-artificial-intelligence-services-draft-for-comment-april-2023/>.
- ◇ National Artificial Intelligence Initiative Act of 2020, H.R. 6216, 116th Cong (2020).
- ◇ *New Generation Artificial Intelligence Development Plan*, 新一代人工智能发展计划, (luglio 2017).
- ◇ *New seal for IT-security and responsible data use is in its way*, Ministry of industry, business and financial affairs (31 ottobre 2019), <https://eng.em.dk/news/2019/oktober/new-seal-for-it-security-and-responsible-data-use-is-in-its-way/>.
- ◇ *OECD AI Principles* (OECD18)
- ◇ *OECD AI Principles* (OECD23).

- ◇ Office. Of Management & Budget, Executive Office Of The President, (17 novembre 2020). *Memorandum For the Heads Of The Executive Departments And Agencies: Guidance for Regulation of Artificial Intelligence Applications.*
- ◇ *Opinion of the data Ethics Commission (Data Ethik Kommission) – Executive Summary,*
https://assets.contentstack.io/v3/assets/blt3de4d56151f717f2/blt300ce23c9789e0f3/5e5cfe13fa08326331360f93/191023_DEK_Kurzfassung_en_bf.pdf.
- ◇ Proposta di Direttiva del Parlamento Europeo e del Consiglio relativa all'adeguamento delle norme in materia di responsabilità civile extracontrattuale all'intelligenza artificiale (direttiva sulla responsabilità da intelligenza artificiale) (Testo rilevante ai fini del SEE) {SEC(2022) 344 final} - {SWD(2022) 318 final} - {SWD(2022) 319 final} - {SWD(2022) 320 final}, 28 settembre 2022.
- ◇ Proposta di Direttiva del Parlamento Europeo e del Consiglio sulla responsabilità per danno da prodotti difettosi (Testo rilevante ai fini del SEE) {SEC(2022) 343 final} - {SWD(2022) 315 final} - {SWD(2022) 316 final} - {SWD(2022) 317 final}, 28 settembre 2022.
- ◇ Proposta di Regolamento del Parlamento Europeo e del Consiglio che stabilisce regole armonizzate sull'Intelligenza Artificiale e modifica alcuni atti legislativi dell'Unione (AI Act) {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final, 21 aprile 2021. (Versione I della Commissione); (Versione II del Parlamento Europeo); (Versione III del Consiglio); (Versione IV, non ufficiale).
- ◇ Proposta di Regolamento del Parlamento Europeo e del Consiglio riguardante norme armonizzate sull'accesso equo ai dati e sul loro utilizzo (Data Act), COM 68 final, 23 febbraio 2022.
- ◇ *Provisions on the Administration of Deep Synthesis in Internet-Based Information Services, 互联网信息服务深度合成管理规定. China Law Translate (1 febbraio 2023).* <https://www.chinalawtranslate.com/deep-synthesis/>.
- ◇ *Provisions on the Management of Algorithmic Recommendations in Internet Information Services 互联网信息服务算法推荐管理规定. China Law Translate. (4 gennaio 2022)* <https://www.chinalawtranslate.com/en/algorithms/>.
- ◇ Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (GDPR), 27 aprile 2016.
- ◇ Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724 (*Data Governance Act*), 30 maggio 2022.
- ◇ Regolamento (UE) 2023/988 del Parlamento europeo e del Consiglio del 10 maggio 2023 relativo alla sicurezza generale dei prodotti, che modifica il regolamento (UE) n. 1025/2012 del Parlamento europeo e del Consiglio.
- ◇ Regolamento 2022/1925 sui mercati contendibili ed equi nel settore digitale (*Digital Market Act*), GU (L 265).

- ◇ Regolamento 2022/2065 relativo al mercato unico dei servizi digitali e che modifica la direttiva 2000/31/CE (*Digital Service Act*) 2022 GU (L 277).
- ◇ Risoluzione del Parlamento Europeo del 20 ottobre 2020 recante raccomandazioni alla Commissione concernenti il quadro relativo agli aspetti etici dell'intelligenza artificiale, della robotica e delle tecnologie correlate (2020/2021(INL)).
- ◇ Risoluzione del Parlamento europeo del 20 gennaio 2021 sull'IA: *questioni relative all'interpretazione e applicazione del diritto internazionale nella misura in cui l'UE è interessata relativamente agli impieghi civili e militari e all'autorità dello Stato al di fuori dell'ambito della giustizia penale* (2020/2013(INI))
- ◇ Servizio Studi del Senato della Repubblica, Servizio degli Affari Internazionali del Senato della Repubblica, & Ufficio Rapporti con l'Unione Europea della Camera dei deputati. (2024). Conferenza sull'intelligenza artificiale - Bruxelles, 28-29 gennaio 2024. In *Dossier Europei XIX Legislatura (65/DE)*, 77, <https://www.senato.it/service/PDF/PDFServer/BGT/01403117.pdf>.
- ◇ Standing Committee of the National People's Congress, *Geren Xinxi Baohu Fa* (个人信息保护法) (*Personal Information Protection Law*, abbreviata in "PIPL") (20 agosto 2021).
- ◇ The White House Office of Science And Technology Policy, *Blueprint for an AI Bill of Rights: making automated systems work for the American people* (Ottobre 2022).
- ◇ *Translation: Chinese expert group offers "Governance principles" for "Responsible AI"* - DigiChina. (1 ottobre 2021). *DigiChina*. <https://digichina.stanford.edu/work/translation-chinese-expert-group-offers-governance-principles-for-responsible-ai/>.
- ◇ U.S Chamber of Commerce, *Letter to Director Prabhakar (OSTP) on "A Blueprint for an AI Bill of Rights"* (10 ottobre 2022).
- ◇ Versione consolidata del Trattato sul funzionamento dell'Unione europea, art. 288, 9 maggio 2008. 288, 9 maggio 2008, 2008 G.U. (C 115)
- ◇ Versione consolidata del Trattato sull'Unione europea, art. 2, GU C 202 del 7.6.2016.

Sentenze, provvedimenti e pareri giudiziali

- ◇ C 6/64, Sentenza della Corte di Giustizia dell'Unione Europea del 15 luglio 1964, *Flaminio Costa v. E.N.E.L.* ECLI:EU:C:1964:66.
- ◇ CGUE, Parere n. 2/00 del 6 dicembre 2001.
- ◇ Sentenza 25 novembre 2021, n. 7891 del Consiglio di Stato, sezione III.
- ◇ Provvedimento n. 112 del 30 marzo 2023, Autorità Garante per la protezione dei dati personali.

- ◇ Alfieri, C., Carocchia, F., & Inverardi, P. (2022). AI Act and Individual Rights: A Juridical and Technical Perspective. *IAIL@HHAI*, 4, <https://www.semanticscholar.org/paper/AI-Act-and-Individual-Rights%3A-A-Juridical-and-Alfieri-Carocchia/96b164abc9e27c8945c5133511ee1429aef69043>
- ◇ Alongi, A., & Pompei, F. (2019). *Il valore dei dati nell'ecosistema digitale: etica e diritti nell'era del web*. EditorialeNovanta.
- ◇ Barlow, J. P., *A declaration of the independence of cyberspace*. (1996). Electronic Frontier Foundation, <https://www.eff.org/it/cyberspace-independence>.
- ◇ Bartlett, J. (2018). *The People vs Tech: How the Internet is killing democracy (and how we save it)*. Penguin Random House UK.
- ◇ Bauman, Z. (2013). *Modernity and the Holocaust*. John Wiley & Sons
- ◇ Beam, A. L., et al. (2023). Artificial intelligence in medicine. (Editorial) *The New England Journal of Medicine*, <https://doi.org/10.1056/nejme2206291>
- ◇ Benbouzid, B., Meneceur, Y., & Smuha, N. A. (2022). Four shades of AI regulation: a cartography of normative and definitional conflicts, *Réseaux*, N° 232-233(2) <https://doi.org/10.3917/res.232.0029>.
- ◇ Blomquist, K. (2022). *In pursuit of (soft) power: Chinese artificial intelligence governance in an age of great power competition*, [Tesi nel programma MSc in Social Science of the Internet]. University of Oxford, Oxford Internet Institute.
- ◇ Bradford, A. (2019b). *The Brussels effect: How the European Union Rules the World*, In Oxford University Press eBooks.
- ◇ Bradford, A. (2023). *Digital empires: The Global Battle to Regulate Technology*. Oxford University Press.
- ◇ Carman, C. C., & Evans, J. (2014). On the epoch of the Antikythera mechanism and its eclipse predictor. *Archive for History of Exact Sciences*, 68(6) <https://doi.org/10.1007/s00407-014-0145-5>.
- ◇ Catanzariti, M. (2023b). Rischio e vulnerabilità nel modello europeo di intelligenza artificiale. *SocietàMutamentoPolitica: Rivista Italiana Di Sociologia*, 13(25), <https://doi.org/10.36253/smp-13804>
- ◇ Cheng, Y., Li, Y., Chen, T., & Yang, Q., (2020) Viewpoint Federated Learning for Privacy-Preserving AI, *Communication of the ACM*, vol. 63, <http://dx.doi.org/10.1145/3387107>
- ◇ Columbro, D. (2022). *Dentro l'algoritmo: Le formule che regolano il nostro tempo*. Effequ.
- ◇ Conforti, B. (2018). *Diritto Internazionale* (11^a ed.). Editoriale Scientifica.
- ◇ Contissa, G. (2017). *Information technology for the law.*, G. Giappichelli Editore.
- ◇ Damasio, A. & Damasio, H. (1992) Brain and Language, *Scientific American*, Vol. 267 No. 3 Special Issue: Mind and Brain.
- ◇ Dreyfus, H. L. (1974). Artificial Intelligence. *The Annals of the American Academy of Political and Social Science*, 412 <http://www.jstor.org/stable/1040396>
- ◇ Druzin, B.H. (2016) *Why Does Soft Law Have Power Anyway*, 7 *Asian JIL*.

- ◇ Eastbrook, F. H. (1996), *Cyberspace and the Law of the Horse*, *University of Chicago Legal Forum*.
- ◇ Edwards, L. (2022) *Regulating AI in Europe: Four Problems and Four Solutions*, *Ada Lovelace Institute* <https://www.adalovelaceinstitute.org/report/regulating-ai-in-europe/>.
- ◇ Floridi, L. (2020). AI and Its New Winter: from Myths to Realities. *Philos. Technol.* 33 <https://doi.org/10.1007/s13347-020-00396-6>
- ◇ Floridi, L. (2021) *The End of an Era: from Self-Regulation to Hard Law for the Digital Industry*, *Philos. Technol.*
- ◇ Floridi, L. (2022). *Etica dell'intelligenza artificiale: Sviluppi, opportunità, sfide*. Raffaello Cortina Editore.
- ◇ Floridi, L. (2023) L. On the Brussels-Washington consensus about the legal definition of Artificial Intelligence. *Philos. Technol.* <https://doi.org/10.1007/s13347-020-00396-6>.
- ◇ Frosini, T. E., *L'Ordine Giuridico del digitale*, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.
- ◇ Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1), 5-7 205395171989794. <https://doi.org/10.1177/2053951719897945>
- ◇ Grother, P., Ngan, M., Hanaoka, K., (2019) *NISTIR 8280: Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, <https://doi.org/10.6028/NIST.IR.8280>.
- ◇ Gupta, R. *et al.* (2021), Artificial intelligence to deep learning: machine intelligence approach for drug discovery. *Mol Divers* 25, 1315–1360, 1320-1323. <https://doi.org/10.1007/s11030-021-10217-3>
- ◇ Hager, G., Drobnis, A.W., Fang, F., Ghani, R., Greenwald, A., Lyons, T., Parkes, D.C., Schultz, J., Saria, S., Smith, S.F., & Tambe, M. (2019). Artificial Intelligence for Social Good. *ArXiv*, abs/1901.05406.
- ◇ Heiling, J.C., Kempt, H. & Nagel, S. Beware of sustainable AI! Uses and abuses of a worthy goal (2023), *AI Ethics*. <https://doi.org/10.1007/s43681-023-00259-8>
- ◇ Huyskes, D. (2022). *Resistere all'inevitabile*, - *Luiss University Press – le meraviglie del possibile*, <https://luissuniversitypress.it/resistere-allinevitabile/>.
- ◇ Jung, A. (2022). *Machine learning: The Basics*. Springer., vii.
- ◇ Klonoff, D., & Han, J. (2019). The First Recall of a Diabetes Device Because of Cybersecurity Risks. *Journal of diabetes science and technology*, <https://doi.org/10.1177/1932296819865655>.
- ◇ Koenig, T., & Rustad, M. (2021). *Global Information Technologies: Ethics and the Law*, West Academic Publishing.
- ◇ Laux, J., Wachter, S., & Mittelstadt, B. (2022). Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4230294>.
- ◇ Lessig, L. (2006), *Code and other laws of Cyberspace* (Version 2.0).

- ◇ Lingevious, J. Military artificial intelligence as power: consideration for European Union actorness. *Ethics Inf Technol* 25, 19 (2023). <https://doi.org/10.1007/s10676-023-09684-z>.
- ◇ Magrani, E., & Guedes Fernandes da Silva, P., (2023), *the Ethical and Legal Challenges of Recommender System Driven by Artificial Intelligence*, in Antunes et al. (2023). *Multidisciplinary perspectives on artificial intelligence and the law*. Springer Nature.
- ◇ Mantelero, A. (2022). Beyond data: Human Rights, Ethical and Social Impact Assessment in AI. *Springer Nature*.
- ◇ Marsden, C. T. (2017). *Network neutrality: From policy to law to regulation*. Manchester University Press.
- ◇ Mészáros, J., Preuveneers, D., Marquet, E., Biasin, E., Peter, I. E., Santos, I. R., Vranckaert, K., Belkadi, L., & Menéndez, N. (2023). Chatgpt: How Many Data Protection Principles Do You Comply with?, *Artificial Intelligence eJournal*. <https://doi.org/10.2139/ssrn.4647569>.
- ◇ Mökander, J., Juneja, P., Watson, D.S., Floridi, L. (2022). The US Algorithmic Accountability Act of 2022 vs. The EU Artificial Intelligence Act: what can they learn from each other? *Minds & Machines* 32, <https://doi.org/10.1007/s11023-022-09612-y>.
- ◇ Nemitz, P., (2018) Constitutional Democracy and Technology in the Age of Artificial Intelligence, *Phil. Trans. R. Sc. A*.
- ◇ Nomura, A., et al. Artificial Intelligence in Current Diabetes Management and Prediction. *Curr Diab Rep* 21, 61, 1 (2021). <https://doi.org/10.1007/s11892-021-01423-2>.
- ◇ Novelli, C., Casolari, F., Hacker, P., Spedicato, G. A., & Floridi, L. (2024). Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity. *Productivity, Innovation & Technology eJournal*, 3 <https://doi.org/10.2139/ssrn.4694565>.
- ◇ O’Neil, C. (2016). *Weapons of math destruction: How Big Data Increases Inequality and Threatens Democracy*. Penguin UK.
- ◇ Pagallo, U. (2023) *Dismantling Four Myths in AI & EU Law Through Legal Information ‘About’ Reality*, in Antunes et al. (2023). *Multidisciplinary perspectives on artificial intelligence and the law*. Springer Nature.
- ◇ Prifti, K., Morley, J., Novelli, C., Floridi, L., (2024). Regulation by design: features, practices, limitations, and governance implications. *Centre for Digital Ethics (CEDE) Research Paper Series*, 18 https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4724454.
- ◇ Richardson, R., Schultz, J., & Crawford, K. (2019). Dirty Data, Bad Predictions: How Civil Rights Violations Impact Police Data, Predictive Policing Systems, and Justice. 94 *NYU Law Review*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3333423#.
- ◇ Roberts, H., Cows, J., Morley, J., Taddeo, M. R., Wang, V., Floridi, L. (2020) The Chinese approach to artificial intelligence: an analysis of policy, ethics, and regulation. *AI & Society* 36, <https://doi.org/10.1007/s00146-020-00992-2>

- ◇ Roberts, H., Hine, E., Taddeo, M., & Floridi, L. (2023). Global AI governance: barriers and pathways forward. *Social Science Research Network*, 4 <https://doi.org/10.2139/ssrn.4588040>.
- ◇ Rudin, C., & Radin, J. (2019). Why Are We Using Black Box Models in AI When We Don't Need To? A Lesson From An Explainable AI Competition. *Harvard Data Science Review*, 1(2) <https://doi.org/10.1162/99608f92.5a8a3a3d>.
- ◇ Rustad, M. L. (2021). *Global Internet Law in a nutshell*, West Academic Publishing, Nutshell Series.
- ◇ Schepisi, C., *Brevi note sulla "dimensione europea" della regolamentazione dell'intelligenza artificiale*, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.
- ◇ Scopelliti, M.A., *È ancora possibile la neutralità tecnologica della normativa?*, in Falce, V. (2023) *Strategia dei dati e intelligenza artificiale*. Giappichelli.
- ◇ Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019). Fairness and Abstraction in Sociotechnical Systems. *Conference on Fairness, Accountability, and Transparency (FAT* '19)*. <https://doi.org/10.1145/3287560.3287598>.
- ◇ Sheehan, M. (2023). *China's AI regulations and how they get made* (No. 1), https://carnegieendowment.org/files/202307-Sheehan_Chinese%20AI%20gov.pdf.
- ◇ Siegmann, C., & Anderljung, M., (2022), *The Brussels Effect and Artificial Intelligence: How the EU regulation will impact the global AI market*, arXiv (Cornell University). <https://doi.org/10.48550/arxiv.2208.12645>.
- ◇ Smuha, N. A. (2021b). From a 'race to AI' to a 'race to AI regulation': regulatory competition for artificial intelligence. *Law, Innovation and Technology*, 13(1), <https://doi.org/10.1080/17579961.2021.1898300>.
- ◇ Smuha, N. A. (2023). Digital Sovereignty in the European Union: Five Challenges from a Normative Perspectives. *Working Paper - ERA Conference Proceedings*, 3.
- ◇ Smuha, N. A., Ahmed-Rengers, E., Harkens, A., Li, W., Maclaren, J., Piselli, R., & Yeung, K. (2021). How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act. *Social Science Research Network*, <https://doi.org/10.2139/ssrn.3899991>;
- ◇ Solove, D. J., & Schwartz, P. M. (2020). *EU data protection and the GDPR*. Aspen Publishing.
- ◇ Strozzi, G., & Mastroianni, R., (2019), *Diritto dell'Unione Europea: parte Istituzionale*, ottava edizione, G. Giappichelli Editore.
- ◇ Taddeo, M., Blanchard, A. A Comparative Analysis of the Definitions of Autonomous Weapons Systems. *Sci Eng Ethics* (2022). <https://doi.org/10.1007/s11948-022-00392-3>.
- ◇ Thierer, A. (2016). Permissionless Innovation: the continuing case for comprehensive technological freedom. *Mercatus Center at George Mason University*, 12-16.
- ◇ Thomas, C., Roberts, H., Mökander, J., Tsamados, A., Taddeo, M., Floridi, L., (2023). The case for a broader approach to AI assurance: addressing "hidden" harms

in the development of Artificial Intelligence, Disponibile su SSRN

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4660737.

- ◇ Tsamados, A., Aggarwal, N., Cows, J., Morley J., Roberts H., Taddeo, M., Floridi, L. (2021). The Ethics of Algorithms: Key Problems and Solutions. In: Floridi, L. (eds) Ethics, Governance, and Policies in Artificial Intelligence. *Philosophical Studies Series*, vol 144. Springer, Cham. https://doi.org/10.1007/978-3-030-81907-1_8.
- ◇ Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, LIX (236), <https://doi.org/10.1093/mind/lix.236.433>.
- ◇ Ward, J. (2022). *The loop: How Technology Is Creating a World Without Choices and How to Fight Back*. Hachette Books.
- ◇ Weaver, J. F., (2018) *Regulation of artificial intelligence in the United States*, in Barfield, W. & Pagallo, U., *Research Handbook on the Law of Artificial Intelligence* (2018), Elgar.
- ◇ Wendehorst, C. (2022) AI Liability in Europe: Anticipating the EU Liability Directive, *Ada Lovelace Institute*, 7 <https://www.adalovelaceinstitute.org/wp-content/uploads/2022/09/Ada-Lovelace-Institute-Expert-Explainer-AI-liability-in-Europe.pdf>;
- ◇ Wylie, C. (2020). *Il mercato del consenso: come ho creato e poi distrutto* Cambridge Analytica. Longanesi.
- ◇ Yordanova, K., Bertels, N. (2023), *Regulating AI: Challenges and the Way Forward Through Regulatory Sandboxes*, in Antunes, H. S., Freitas, P. M., Oliveira, A. L., Pereira, C. M., De Sequeira, E. V., & Xavier, L. B. (2023). *Multidisciplinary perspectives on artificial intelligence and the law*. Springer Nature.
- ◇ Zhang, A. Z. (2021). Agility over stability: China's great reversal in regulating the platform economy. *Harvard International Law Journal*, Vol. 63, No. 2, 2022, <https://doi.org/10.2139/ssrn.3892642>.
- ◇ Ziosi, M., Mökander, J., Novelli, C., Casolari, F., Taddeo, M., & Floridi, L. (2023). The EU AI Liability Directive: shifting the burden from proof to evidence, *Social Science Research Network*. <https://doi.org/10.2139/ssrn.4470725>.
- ◇ Zittrain, J. (2005). The generative Internet. *Social Science Research Network*. https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID921636_code177769.pdf?abstractid=847124&mirid=1&type=2.
- ◇ Zuboff, S. (2019) *Il capitalismo della sorveglianza: il futuro dell'umanità nell'era dei nuovi poteri*, LUISS University Press.

Articoli di giornale, report e risorse online

- ◇ *About us*. (n.d.). <https://www.uschamber.com/about>.
- ◇ *Accelerating Climate Action with AI*, Boston Consulting Group, 5 (novembre 2023). <https://web-assets.bcg.com/72/cf/b609ac3d4ac6829bae6fa88b8329/bcg-accelerating-climate-action-with-ai-nov-2023-rev.pdf>.
- ◇ Access Now official website, *Access Now* <https://www.accessnow.org/>.

- ◇ *AI Act threatens to make facial surveillance commonplace in Europe.* (16 gennaio 2024). Patrick Breyer. <https://www.patrick-breyer.de/en/ai-act-threatens-to-make-facial-surveillance-commonplace-in-europe/>
- ◇ *AlphaGo's next move.* (27 maggio 2017). Google DeepMind. <https://deepmind.google/discover/blog/alphagos-next-move/>.
- ◇ *Artificial Intelligence - United States.* Statista. [Grafico] (n.d.) <https://www.statista.com/outlook/tmo/artificial-intelligence/united-states?currency=USD>.
- ◇ *Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI | News | European Parliament.* (9 dicembre 2023). <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>.
- ◇ Ball, J. (16 agosto 2023). Will AI steal my job? Maybe – but here are some possible new opportunities. *The Guardian*. <https://www.theguardian.com/global-development/2023/aug/16/will-ai-steal-my-job-maybe-but-here-are-some-possible-new-opportunities>
- ◇ BCLP - Bryan Cave Leighton Paisner. (n.d.). US state-by-state AI legislation snapshot. BCLP - Bryan Cave Leighton Paisner - US State-by-state AI Legislation Snapshot. <https://www.bclplaw.com/en-US/events-insights-news/2023-state-by-state-artificial-intelligence-legislation-snapshot.html>.
- ◇ Bertuzzi, L. (22 gennaio 2024). *LEAK: Given the massive public attention to the #AIAct, I've taken the rather unprecedented decision to publish the final text. The agreed text is on the right-hand side for those unfamiliar with four-column documents. Enjoy the reading! Some context: 1/6.* Twitter. <https://twitter.com/BertuzLuca/status/1749326217612820558>
- ◇ Bertuzzi, L. (7 marzo 2023). *EU lawmakers set to settle on OECD definition for Artificial Intelligence.* www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/eu-lawmakers-set-to-settle-on-oecd-definition-for-artificial-intelligence/>.
- ◇ Bertuzzi, L. (8 novembre 2022). *Artificial Intelligence definition, governance on MEPs' menu.* www.euractiv.com. <https://euractiv.com/section/digital/news/artificial-intelligence-definition-governance-on-meps-menu/>.
- ◇ Bertuzzi, L., (12 gennaio 2024). *EU prepares to push back on private sector carve-out from international AI treaty.* www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/eu-prepares-to-push-back-on-private-sector-carve-out-from-international-ai-treaty/>.
- ◇ Bertuzzi, L., (14 dicembre 2023b). *European Union squares the circle on the world's first AI rulebook.* www.euractiv.com. <https://www.euractiv.com/section/artificial-intelligence/news/european-union-squares-the-circle-on-the-worlds-first-ai-rulebook/>.
- ◇ Bertuzzi, L., (17 gennaio 2023a) *US obtains exclusion of NGOs from drafting AI treaty.* www.euractiv.com. <https://www.euractiv.com/section/digital/news/us-obtains-exclusion-of-ngos-from-drafting-ai-treaty/>.

- ◇ Bertuzzi, L., (24 novembre 2023b). France, Germany, Italy push for ‘mandatory self-regulation’ for foundation models in EU’s AI law. *www.euractiv.com*.
<https://www.euractiv.com/section/artificial-intelligence/news/france-germany-italy-push-for-mandatory-self-regulation-for-foundation-models-in-eus-ai-law/>.
- ◇ Bertuzzi, L., (6 Giugno 2023b). EU’s AI ambitions at risk as US pushes to water down international treaty. *www.euractiv.com*.
<https://www.euractiv.com/section/artificial-intelligence/news/eus-ai-ambitions-at-risk-as-us-push-to-water-down-international-treaty/>.
- ◇ Byte by byte: How Big Tech undermined the AI Act. *Corporate Europe Observatory* (17 novembre 2023) <https://corporateeurope.org/en/2023/11/byte-byte>.
- ◇ Capgemini Research Institute, Why addressing ethical questions in AI will benefit organisations (2019) https://www.capgemini.com/wp-content/uploads/2021/02/AI-in-Ethics_Web-1.pdf.
- ◇ Capone, E. (30 maggio 2023). Lo strano appello contro le IA: “Rischiame l’estinzione”. Ma le aziende continuano a svilupparle. *La Repubblica*.
https://www.repubblica.it/tecnologia/2023/05/30/news/ia_rischio_estinzione_appello_altman_hinton_musk_regole-402536379/.
- ◇ China: AI market size 2023 [Grafico] *Statista*. (30 gennaio 2024).
<https://www.statista.com/statistics/1262377/china-ai-market-size/>.
- ◇ China: scale of AI industry by segment 2026. [Grafico] *Statista*. (3 gennaio 2024)
<https://www.statista.com/statistics/1256546/china-scale-of-ai-industry-by-segment/>
- ◇ Commissione Europea. (28 settembre 2022). *Domande e risposte sulla revisione della direttiva sulla responsabilità per danno da prodotti difettosi*.
https://ec.europa.eu/commission/presscorner/api/files/document/print/it/qanda_22_5791/QANDA_22_5791_IT.pdf.
- ◇ Copeland, B. (7 ottobre 2008). MYCIN | *Expert System, Medical Diagnosis & Treatment*. Encyclopedia Britannica.
<https://www.britannica.com/technology/MYCIN>.
- ◇ Coraggio, G. (9 Gennaio 2024). *The Approval of the EU AI Act and What to expect Next - with Mr. Brando Benifei*.
- ◇ Dastin, J. (10 ottobre 2018). Insight - Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*. <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>.
- ◇ Davidson, N. (4 dicembre 2023). MAP: How are state and local governments navigating AI regulation? *GovTech*. <https://www.govtech.com/biz/data/how-are-state-and-local-governments-navigating-ai-regulation>
- ◇ Davies, N., (10 novembre 2021). Index shows US is winning the AI race – but for how long? *Investment Monitor*. <https://www.investmentmonitor.ai/ai/ai-index-us-china-artificial-intelligence/>
- ◇ *Digital Policy Alert - United States of America: Rejected Algorithmic Accountability Act of 2022* (3 gennaio 2023). <https://digitalpolicyalert.org/event/8384-rejected-algorithmic-accountability-act-of-2022>.
- ◇ Do Americans think artificial intelligence will lead to there being more jobs for people or fewer jobs? [Grafico]. In *Statista*. (30 dicembre 2022)

- <https://www.statista.com/statistics/1357711/opinions-on-artificial-intelligence-impact-on-jobs-by-age-us/>.
- ◇ DreaMed. (31 luglio 2023). *DREAMED MAIN Home 3* - DreaMed. DreaMed - Cutting-edge Diabetes Technology for Insulin Treatment Optimization. <https://dreamed.ai/>.
 - ◇ Engler, A. (7 ottobre 2022). The AI Bill of Rights Makes Uneven Progress on Algorithmic Protections. *Lawfare*. <https://www.lawfareblog.com/ai-bill-rights-makes-uneven-progress-algorithmic-protections>.
 - ◇ Engler, A., (8 dicembre 2020) New White House guidance downplays important AI harms, *Brookings*. <https://www.brookings.edu/blog/techtank/2020/12/08/new-white-house-guidance-downplays-important-ai-harms/>
 - ◇ *Enhancing Jail Management and Prisoner Management System with Facial Recognition* | *Corsight*. (18 maggio 2023). Corsight. <https://www.corsight.ai/prisons/>.
 - ◇ European Council - Council of the European Union. (9 dicembre 2023) <https://www.consilium.europa.eu/en/press/press-releases/2023/12/09/artificial-intelligence-act-council-and-parliament-strike-a-deal-on-the-first-worldwide-rules-for-ai/>.
 - ◇ European Digital Rights (EDRI). (8 dicembre 2023). *Open letter to EU AI Act negotiators: Do not trade away our rights!* https://edri.org/wp-content/uploads/2023/12/Open-letter-to-EU-AI-Act-negotiators_Do-not-trade-away-our-rights-8-December-2023.pdf
 - ◇ European Digital Rights, Access Now, Algorithm Watch, Panoptikon Organization, Article 19, Electronic Frontier Finland, & Electronic Privacy Information Center. (2022). *Scope of the EU Artificial Intelligence Act (AIA): Military Purposes and National Security*. *European Center for Not-for-profit Law*. https://ecnl.org/sites/default/files/2022-03/ECNL%20Pagers%20on%20scope%20of%20AIA%20ECNL_FINAL.pdf
 - ◇ Fenton-Harvey, J. (13 dicembre 2023). The Gospel: Israel's controversial AI used in the Gaza war. *The New Arab*. <https://www.newarab.com/analysis/gospel-israels-controversial-ai-used-gaza-war>
 - ◇ Gershgorn, D. (2 marzo 2021). China's 'Sharp Eyes' program aims to surveil 100% of public space. *Medium*. <https://onezero.medium.com/chinas-sharp-eyes-program-aims-to-surveil-100-of-public-space-ddc22d63e015>.
 - ◇ *Global AI Regulation Tracker*. (n.d.). <https://www.techieray.com/GlobalAIRegulationTracker>.
 - ◇ *Google's AlphaGo "plays just like a human," says top ranked Go player*. (1° febbraio 2016). SiliconANGLE. <https://siliconangle.com/2016/02/01/googles-alphago-plays-just-like-a-human-says-top-ranked-go-player/>.
 - ◇ *GuardianTM Connect Continuous Glucose Monitor* | *Medtronic*. (27 novembre 2023). <https://www.medtronicdiabetes.com/products/guardian-connect-continuous-glucose-monitoring-system>.

- ◇ Hao, K. (5 aprile 2021). What is machine learning? *MIT Technology Review*. <https://www.technologyreview.com/2018/11/17/103781/what-is-machine-learning-we-drew-you-another-flowchart/>.
- ◇ Heaven, W. D. (31 agosto 2023). ChatGPT is everywhere. Here's where it came from. *MIT Technology Review*. <https://www.technologyreview.com/2023/02/08/1068068/chatgpt-is-everywhere-heres-where-it-came-from/>.
- ◇ Heikkilä, M. (13 aprile 2022). Dutch scandal serves as a warning for Europe over risks of using algorithms. *POLITICO*. <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>.
- ◇ Heikkilä, M. (14 novembre 2023). AI is at an inflection point, Fei-Fei Li says. *MIT Technology Review*. <https://www.technologyreview.com/2023/11/14/1083352/ai-is-at-an-inflection-point-fei-fei-li-says/>.
- ◇ Hidvegi, F., Daniel, L., & Masse, E. (2021), The EU should regulate AI on the basis of rights, not risks, *Access Now*, <https://www.accessnow.org/eu-regulation-ai-risk-based-approach/>.
- ◇ Hodson, H. (6 maggio 2016). Revealed: Google AI has access to huge haul of NHS patient data. *New Scientist*. <https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/>
- ◇ Hunt, E. (24 marzo 2016). Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter. *The Guardian*. <https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter>.
- ◇ Johnson, K. (4 ottobre 2022). Biden's AI Bill of Rights is toothless against big tech. *Wired*. <https://www.wired.com/story/bidens-ai-bill-of-rights-is-toothless-against-big-tech>.
- ◇ Kang, C. (23 aprile 2022). As Europe approves new tech laws, the U.S. falls further behind. *The New York Times*. <https://www.nytimes.com/2022/04/22/technology/tech-regulation-europe-us.html>.
- ◇ Kerner, S. M. (13 settembre 2023). *Large Language Models (LLMs)* <https://www.techtarget.com/whatis/definition/large-language-model-LLM>
- ◇ Khanna, D. (15 settembre 2015). Regulations stifle innovation. *The Hill*. <https://thehill.com/blogs/congress-blog/technology/253625-regulations-stifle-innovation/>.
- ◇ Kovacevich, A. (2 agosto 2022). The 'Biden Tech Doctrine' — one year in. *The Hill*. <https://thehill.com/opinion/technology/593238-the-biden-tech-doctrine-one-year-in/>.
- ◇ *Law Enforcement | Clearview AI*. (n.d.). Clearview AI. <https://www.clearview.ai/law-enforcement>.
- ◇ Lomas, N. NHS memo details Google/DeepMind's five year plan to bring AI to healthcare, *Techcrunch* (8 giugno 2016) <https://techcrunch.com/2016/06/08/nhs-memo-details-googleddeepminds-five-year-plan-to-bring-ai-to-healthcare/>. *Costi ambientali dei dispositivi di IA - Privacy Network*. (24 ottobre 2022). Privacy Network. <https://privacy-network.it/costi-ambientali-dei-dispositivi-di-ia/>.

- ◇ LuissGuidoCarli. (22 dicembre 2023). *Gli Stati Generali del Diritto di Internet e della Intelligenza Artificiale (Giorno 2)* [Video]. YouTube., min. 8:32:39- 8:46:48, min 5:19:11-5.44.54 YouTube. <https://www.youtube.com/watch?v=vEnAZFHokm8>.
<https://www.youtube.com/watch?v=vEnAZFHokm8>.
- ◇ MacCarthy, M., (9 marzo 2020). AI needs more regulation, not less, *Brookings*. <https://www.brookings.edu/research/ai-needs-more-regulation-not-less/>.
- ◇ Madung, O. (9 agosto 2023). AI hysteria is a distraction: algorithms already sow disinformation in Africa. *The Guardian*. <https://www.theguardian.com/global-development/2023/aug/09/ai-chatgpt-doomerism-threat-already-here-big-tech-algorithms-sow-disinformation>
- ◇ Mantelero, A. (30 novembre 2023). Alessandro Mantelero on LinkedIn: PL 2338-2023 (Inteligência artificial) - Emenda 1. <https://www.linkedin.com/feed/update/urn:li:activity:7136124845989257216/>
- ◇ Mari, A. (13 ottobre 2023). Brazil consults on regulatory sandbox for AI and data protection. *Forbes* <https://www.forbes.com/sites/angelicamarideoliveira/2023/10/13/brazil-consults-on-regulatory-sandbox-for-ai-and-data-protection/>.
- ◇ Milmo, D. (21 agosto 2023). Elon Musk launches AI startup and warns of a ‘Terminator future.’ *The Guardian*. <https://www.theguardian.com/technology/2023/jul/13/elon-musk-launches-xai-startup-pro-humanity-terminator-future>
- ◇ Milmo, D., & Stacey, K. (1° novembre 2023). ‘It’s not clear we can control it’: what they said at the Bletchley Park AI summit. *The Guardian*. <https://www.theguardian.com/technology/2023/nov/01/elon-musk-calls-ai-one-of-the-biggest-threats-to-humanity-at-summit>.
- ◇ MIT Technology Review, *The Big Data Conundrum: How to Define It?* (23 ottobre 2013) <https://www.technologyreview.com/2013/10/03/82990/the-big-data-conundrum-how-to-define-it/>
- ◇ Moltzatu, A. (22 dicembre 2019) *EU Strategy On Artificial Intelligence In 2018: A Closer Look at the Strategy by the European Union*, Medium. <https://medium.com/towards-data-science/the-eu-strategy-on-artificial-intelligence-ab51ce5ce5bd>.
- ◇ Mueller, B. (2021). How much will the Artificial Intelligence Act cost Europe? In *Center for Data Innovation*. <https://www2.datainnovation.org/2021-aia-costs.pdf>.
- ◇ Mukherjee, S., Coulter, M., & Chee, F. Y. (14 dicembre 2023) Explainer: What’s next for the EU AI Act? *Reuters*. <https://www.reuters.com/technology/whats-next-eu-ai-act-2023-12-14/>
- ◇ Nathalie Smuha | legal scholar and philosopher. (n.d.). *Nathaliesmuha*. <https://www.nathaliesmuha.com/>
- ◇ Number of times when artificial intelligence (AI) was mentioned in congressional record in the United States from 2001 to 2020 [Grafico]. *Statista*. (5 marzo 2021). <https://www.statista.com/statistics/1233574/united-states-congress-artificial-intelligence-mention/>

- ◇ Open letter to COE AI Convention negotiators: Do not water down our rights (2024). *European Center for Not-for-profit Law*. <https://ecnl.org/news/open-letter-coe-ai-convention-negotiators-do-not-water-down-our-rights>.
- ◇ Osservatorio Amministrazione automatizzata. *Privacy Network*. <https://privacy-network.it/osservatorio/>.
- ◇ Pause Giant AI Experiments: An Open Letter, *The Future of Life Institute*, (22 marzo 2023) <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>
- ◇ Perego, M. (4 gennaio 2024). Il sistema di gestione dell'intelligenza artificiale: la norma ISO/IEC 42001:2023. *Federprivacy*. <https://www.federprivacy.org/informazione/primo-piano/il-sistema-di-gestione-dell-intelligenza-artificiale-la-norma-iso-iec-42001-2023>.
- ◇ *Press release - Artificial intelligence act: Council and Parliament strike a deal on the first rules for AI in the world*. (9 dicembre 2023).
- ◇ *Read the Framework for Global Electronic Commerce*, White house <https://clintonwhitehouse4.archives.gov/WH/New/Commerce/read.html>.
- ◇ *Regulation of artificial intelligence in Europe and Japan*. (24 agosto 2020). White & Case LLP International Law Firm, Global Law Practice. <https://www.whitecase.com/insight-our-thinking/regulation-artificial-intelligence-europe-and-japan>.
- ◇ Reiff, B. (3 dicembre 2023). 'A mass assassination factory': Inside Israel's calculated bombing of Gaza. *+972 Magazine*. <https://www.972mag.com/mass-assassination-factory-israel-calculated-bombing-gaza/>.
- ◇ Revenues from the artificial intelligence market in North America from 2017 to 2028 (in billion U.S. dollars), [Grafico]. *Statista*. (19 gennaio 2022) <https://www.statista.com/statistics/721748/north-america-artificial-intelligence-market/>.
- ◇ *Ruha Benjamin on deep learning: Computational depth without sociological depth is 'superficial learning.'* (n.d.). <https://www.techregister.co.uk/ruha-benjamin-on-deep-learning-computational-depth-without-sociological-depth-is-superficial-learning/>.
- ◇ Ryan-Mosley, T. (30 ottobre 2023). Three things to know about the White House's executive order on AI. *MIT Technology Review*. <https://www.technologyreview.com/2023/10/30/1082678/three-things-to-know-about-the-white-houses-executive-order-on-ai/>.
- ◇ Samp, T. (26 dicembre 2020). Landmark artificial intelligence legislation should become law, *The Hill* <https://thehill.com/opinion/technology/531680-landmark-artificial-intelligence-legislation-should-become-law/>.
- ◇ Saujani, R. (5 dicembre 2023). We don't have to choose between ethical AI and innovative AI. *TIME*. <https://time.com/6342280/ai-paid-leave-social-good/>
- ◇ Semuels, A. (6 agosto 2020). Millions of Americans have lost jobs in the Pandemic—And robots and AI are replacing them faster than ever. *Time*. <https://time.com/5876604/machines-jobs-coronavirus/>
- ◇ Sheehan, M., & Du, S. (2 novembre 2022). *How food delivery workers shaped Chinese algorithm regulations*. Carnegie Endowment for International Peace.

- <https://carnegieendowment.org/2022/11/02/how-food-delivery-workers-shaped-chinese-algorithm-regulations-pub-88310>.
- ◇ Shepardson, D. (7 gennaio 2020). White House proposes regulatory principles to govern AI use. *Reuters*. <https://www.reuters.com/article/idUSKBN1Z60GL/>.
 - ◇ Sibilla, F., Bruschi, B.S., (13 dicembre 2023). *Il web scraping nell'Intelligenza Artificiale: il caso Clearview AI e l'indagine del Garante*. Rivista Diritto Di Internet – Osservatorio Dati & Imprese <https://dirittodiInternet.it/il-web-scraping-nellintelligenza-artificiale-il-caso-clearview-ai-e-lindagine-del-garante/>
 - ◇ Sibilla, F., Bruschi, B.S., (19 gennaio 2024). *Cenni e confronti tra le due principali normative a livello mondiale sull' AI: AI Act europeo e Executive Order americano*. Rivista Diritto Di Internet – Osservatorio Dati & Imprese.
 - ◇ Smuha, N. A. (2023c). Biden, Bletchley, and the emerging international law of AI. *Verfassungsblog*. <https://doi.org/10.17176/20231115-113402-0>.
 - ◇ Spagnuolo, E. (29 settembre 2023). Intelligenza artificiale: i lavori a rischio e quelli emergenti secondo il World Economic Forum. *Corriere Della Sera*. https://www.corriere.it/tecnologia/cards/intelligenza-artificiale-i-lavori-a-rischio-e-quelli-emergenti/il-lavoro-al-tempo-di-chatgpt_principale.shtml.
 - ◇ *Technology to solve the access to justice gap | Suffolk Legal Tech Class 2023*. Quinten Steenhuis, <https://suffolklitlab.org/legal-tech-class/docs/a2j/a2j-technology>.
 - ◇ *The Artificial Intelligence Index Report 2023, AI Index*, (2023) https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf.
 - ◇ Thompson, C. (11 maggio 2022). What the history of AI tells us about its future. *MIT Technology Review*. <https://www.technologyreview.com/2022/02/18/1044709/ibm-deep-blue-ai-history/>.
 - ◇ *Timeline - Artificial intelligence*. (n.d.). European Council - Council of the European Union. <https://www.consilium.europa.eu/en/policies/artificial-intelligence/timeline-artificial-intelligence/>
 - ◇ Tiulkanov, A. (18 dicembre 2023). *Council of Europe is Progressing on the First International AI Treaty – in the Newsletter AI, Data & Digital Policy Thoughts about current and emerging digital regulation and policy*. LinkedIn. <https://www.linkedin.com/pulse/council-europe-progressing-first-international-ai-treaty-tiulkanov-xp88e%3FtrackingId=2fKarTumQ46R6uT86HPujw%253D%253D/?trackingId=2fKarTumQ46R6uT86HPujw%3D%3D>.
 - ◇ Tiulkanov, A. (23 gennaio 2024). *According to the leaked #AIAct text, the AI system's definition has undergone a controversial change that may potentially narrow the scope of the regulation*. LinkedIn. https://www.linkedin.com/feed/update/urn:li:activity:7155481476115873792/?updateEntityUrn=urn%3Ali%3Afs_updateV2%3A%28urn%3Ali%3Aactivity%3A7155481476115873792%2CFEED_DETAIL%2CEMPTY%2CDEFAULT%2Cfalse%29.
 - ◇ Trust and acceptance of AI worldwide 2022. [Grafico] Statista (10 agosto 2023) <https://www.statista.com/statistics/1369185/trust-and-acceptance-of-ai-worldwide/>

- ◇ Turner Lee, N. & Malamud, J., (25 gennaio 2024) How Congress can secure Biden's AI legacy. *Brookings*. <https://www.brookings.edu/articles/how-congress-can-secure-bidens-ai-legacy/>.
- ◇ Volpicelli, G. (18 gennaio 2024). EU set to allow draconian use of facial recognition tech, say lawmakers. *POLITICO*. <https://www.politico.eu/article/eu-ai-facial-recognition-tech-act-late-tweaks-attack-civil-rights-key-lawmaker-hahn-warns/>
- ◇ Waltz, E. (24 giugno 2021). Elon Musk announces Neuralink Advance toward syncing our Brains with AI. *IEEE Spectrum*. <https://spectrum.ieee.org/elon-musk-neuralink-advance-brains-ai>
- ◇ *We always hear that AI will take our jobs. But what jobs will it create?* (6 ottobre 2023). World Economic Forum. <https://www.weforum.org/agenda/2023/09/jobs-ai-will-create/>.
- ◇ *What is a credit score? | Consumer Financial Protection Bureau.* (28 agosto 2023). Consumer Financial Protection Bureau. <https://www.consumerfinance.gov/ask-cfpb/what-is-a-credit-score-en-315/>.
- ◇ Wiley, M. (4 ottobre 2022). *AI Bill of Rights Blueprint Can Protect Individuals from Bias and Discrimination, Implementation Will Be Critical. The Leadership Conference on Civil and Human Rights.* <https://civilrights.org/2022/10/04/ai-bill-of-rights-blueprint-can-protect-individuals-from-bias-and-discrimination-implementation-will-be-critical/>.
- ◇ Williams, Z. (28 agosto 2023). AI Regulation Ramping Up in States from Texas to Connecticut. *Bloomberg Law*. <https://news.bloomberglaw.com/in-house-counsel/ai-regulation-ramping-up-in-states-from-texas-to-connecticut>
- ◇ Zorloni, L. (23 gennaio 2024) *AI Act, abbiamo letto l'ultima versione della legge europea sull'intelligenza artificiale. Wired Italia.* <https://www.wired.it/article/ai-act-testo-ultima-versione-gennaio-divieti-riconoscimento-facciale/#tre>