

Ernesto d'Albergo, Giorgio Giovanelli

Politiche dell'intelligenza artificiale

Arene, strategie, poteri

STUDI

Politica



FrancoAngeli



OPEN ACCESS la soluzione FrancoAngeli

Il presente volume è pubblicato in open access, ossia il file dell'intero lavoro è liberamente scaricabile dalla piattaforma **FrancoAngeli Open Access** (<http://bit.ly/francoangeli-oa>).

FrancoAngeli Open Access è la piattaforma per pubblicare articoli e monografie, rispettando gli standard etici e qualitativi e la messa a disposizione dei contenuti ad accesso aperto. Oltre a garantire il deposito nei maggiori archivi e repository internazionali OA, la sua integrazione con tutto il ricco catalogo di riviste e collane FrancoAngeli massimizza la visibilità, favorisce facilità di ricerca per l'utente e possibilità di impatto per l'autore.

Per saperne di più: [Pubblica con noi](#)

I lettori che desiderano informarsi sui libri e le riviste da noi pubblicati possono consultare il nostro sito Internet: www.francoangeli.it e iscriversi nella home page al servizio "[Informatemi](#)" per ricevere via e-mail le segnalazioni delle novità.

Ernesto d'Albergo, Giorgio Giovanelli

Politiche dell'intelligenza artificiale

Arene, strategie, poteri

FrancoAngeli 

Questo volume è stato pubblicato con un contributo del Dipartimento di Scienze Sociali ed Economiche di Sapienza, Università di Roma, a valere sulla ricerca “Le conoscenze scientifiche nelle policy e pratiche locali basate sull’Intelligenza Artificiale” – Seed PNR 2021.

Isbn: 9788835180098

Isbn e-book Open Access: 9788891731630

Copyright © 2025 by FrancoAngeli s.r.l., Milano, Italy.

Pubblicato con licenza *Creative Commons
Attribuzione-Non Commerciale-Non opere derivate 4.0 Internazionale*
(CC-BY-NC-ND 4.0).

Sono riservati i diritti per Text and Data Mining (TDM), AI training e tutte le tecnologie simili.

*L'opera, comprese tutte le sue parti, è tutelata dalla legge sul diritto d'autore.
L'Utente nel momento in cui effettua il download dell'opera accetta tutte le condizioni
della licenza d'uso dell'opera previste e comunicate sul sito
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.it>*

Stampa: Litogi, Via Idro 50, 20132 Milano.

A Tommaso

Indice

Introduzione		pag.	9
1. Rischi, potenzialità e poste in gioco nelle agende dell'intelligenza artificiale			
1.1. I rischi pubblici dell'intelligenza artificiale	»		17
1.2. L'intelligenza artificiale e i processi democratici: rischi e potenzialità	»		31
1.3. Le poste in gioco economiche e politiche	»		43
2. Governare l'intelligenza artificiale: fra mercato e arene internazionali			
2.1. Le forme della regolazione: autoregolazione, <i>soft law</i> e <i>hard law</i>	»		56
2.2. Il potere delle imprese e la cornice dell'etica	»		69
2.3. Le arene politiche internazionali	»		83
3. Le strategie politiche			
3.1. Agende e strategie politiche nazionali: i “giganti”	»		98
3.2. La politica dell'Unione europea: strategie di regolazione e innovazione	»		114
3.3. Dall'intelligenza artificiale europea alla “via italiana”	»		129
4. L'intelligenza artificiale nelle politiche locali			
4.1. Dalle Smart City all'intelligenza artificiale urbana: agende e strategie	»		143
4.2. Governare i <i>trade-off</i> dell'intelligenza artificiale urbana	»		157
4.3. La politica dei gemelli digitali urbani	»		168
Bibliografia			181

Introduzione^{*}

Questo libro si occupa di uno dei fenomeni che, a partire dagli ultimi dieci anni, sta più profondamente e rapidamente trasformando le caratteristiche economiche e culturali della società in pressoché tutto il mondo e occupando prepotentemente le agende di attori e ambiti di azione di vari tipi. L'intelligenza artificiale (di seguito e in tutto il volume: IA) è oggetto di molte definizioni, in parte fra loro contrastanti, provenienti da tutti i mondi – economico, tecno-scientifico, politico e della società civile no-profit, comprese le istituzioni religiose – che saranno considerati nei prossimi capitoli. L'IA e le sue concrete utilizzazioni sono non solo un insieme di metodi e operazioni computazionali e di dispositivi tecnici ma anche, in senso lato, il prodotto di processi di costruzione sociale, un insieme stratificato e interdipendente di «tecnologia, istituzioni e ideologia» (McQuillan, 2022, p. 1). Per capire perché, consideriamo una definizione costruita in un ambiente politico: secondo l'articolo 3 della legge dell'UE sull'IA in vigore dal 2025¹ «per “sistema di IA” si intende un sistema basato su macchine progettato per operare con diversi livelli di autonomia e che può mostrare capacità di adattamento dopo la sua implementazione e che, per obiettivi esplicativi o impliciti, deduce, a partire dagli *input* che riceve, come generare *output* quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali». L'IA, quindi, è capace di percepire l'ambiente (con strumenti audio, visivi, testuali e tattili), di estrarre automaticamente conoscenze e riconoscere modelli a partire da dati e basandosi su correlazioni statistiche, di ragionare logicamente, di comunicare interattivamente, di decidere azioni da compiere sulla base di previsioni accurate delle loro possibili conseguenze (Skaug e Sætra, 2021). Per questi motivi l'uso e lo sviluppo dell'IA sono

* Il lavoro è frutto della ricerca e delle riflessioni congiunte degli autori. Tuttavia, a Ernesto d'Albergo vanno attribuiti i paragrafi 1.2, 1.3, 2.3, 3.1, 3.2 e 4.1, mentre a Giorgio Giovannelli i paragrafi 1.1, 2.1, 2.2, 3.3, 4.2 e 4.3. L'Introduzione è stata scritta insieme.

¹ Fonte: <https://artificialintelligenceact.eu/article/3/>

considerati decisivi per il presente e il futuro dell’umanità. Una percezione collettiva mediata da “immaginari”, ossia rappresentazioni socialmente prodotte che contribuiscono, svolgendo un ruolo cognitivo e normativo, a strutturare le relazioni fra i sottosistemi della società coinvolti nell’IA. La pervasività della “questione IA” è testimoniata anche dalla sua presenza in un grande numero di eventi politici, culturali e financo religiosi. Papa Leone XIV ha motivato la scelta del suo nome ricollegandosi al suo predecessore Papa Leone XIII, il quale «con la storica enciclica *Rerum Novarum* affrontò la questione sociale nel contesto della prima grande rivoluzione industriale. E oggi la Chiesa offre a tutti il suo patrimonio di dottrina sociale per rispondere a un’altra rivoluzione industriale e agli sviluppi dell’intelligenza artificiale, che comportano nuove sfide per la difesa della dignità umana, della giustizia e del lavoro»². In questo modo, ha dato seguito all’attivismo del pontefice precedente, che era intervenuto in un vertice dei G7 e aveva istituito una commissione per studiare i problemi dell’IA e proporre soluzioni.

Infatti, sono ampiamente diffusi e consolidati immaginari circa le sfide che essa può porre, i rischi e le minacce che può generare. I rischi dell’IA possono essere involontari, ad esempio in scenari spesso evocati in cui gli agenti umani potrebbero perderne il controllo, sebbene si ritenga che in fondo siano le scelte umane su come progettare e utilizzare l’IA a determinarne gli effetti (Simons, 2023). Invece, le minacce ci riconducono ai suoi usi intenzionalmente malevoli, ad esempio in ambito militare o per diffondere notizie false. Ma è opinione diffusa che l’IA offra anche una grande varietà di potenzialità a vantaggio della società, o di sue specifiche componenti, di volta in volta diverse, inclusive o esclusive, a seconda delle applicazioni. La consapevolezza della necessità di un equilibrio fra i suoi rischi e potenzialità, che nel libro è più volte richiamato come *trade-off* dell’IA, è il *frame* – ossia la cornice di significato condivisa – prevalente nella costruzione sociale e politica dell’IA in tutto il mondo. La ricerca di questo equilibrio ha strutturato il modo in cui l’IA è entrata nelle agende e nelle strategie di tutti gli attori, i sistemi e i livelli del potere politico considerati in questo volume.

Noi ci occupiamo in particolare delle politiche dell’IA, esaminandole da una prospettiva di sociologia dei fenomeni politici, che li interpreta alla luce delle costanti interazioni tra la politica e le diverse componenti della società (Viviani, 2017). Pertanto, abbiamo cercato di capire, nel caso dell’IA, come fattori economici e culturali e forze sociali influenzino e siano influenzati da

² La Repubblica 10 05 2025 (https://www.repubblica.it/cronaca/2025/05/10/video/papa_prevost_il_mio_nome_scelto_perche_leone_xiii_affronto_la_questione_sociale-424178681/)

attori e sistemi politici, secondo l’ambizione di provare a collocare il fenomeno politico nello studio della realtà sociale intera (Pennati, 1963). Le domande poste da questa angolazione riguardano quindi le modalità con le quali un fenomeno tecnico-scientifico come l’IA, caratterizzato dalla particolare rapidità con le quali si rinnova e si trasforma, oltre che dalla radicalità dei suoi impatti sulla vita degli individui e della società, diviene oggetto di azioni e decisioni pubbliche.

Nel libro confluiscono evidenze empiriche e riflessioni emerse da un lavoro di studio e ricerca sui rapporti tra IA e politica condotto dagli autori, con la collaborazione del nostro collega Tommaso Fasciani, nel periodo 2021-2025. Sono state svolte attività di: (i) analisi *desk* di letteratura scientifica e “grigia”, documenti politici e istituzionali, media e artefatti discorsivi prodotti da imprese di consulenza e fornitrice di tecnologie IA; (ii) mappatura di casi di applicazione dell’IA su scala urbana; (iii) approfondimento con metodi qualitativi di casi di studio. Sono sintetizzati e aggiornati anche risultati di ricerca discussi in occasione di convegni scientifici nazionali e internazionali e condivisi con gli studenti e le studentesse in insegnamenti nel settore di sociologia della politica presso Sapienza, Università di Roma.

In inglese, il libro potrebbe essere intitolato *The politics of artificial intelligence policy*. Un’espressione che, tradotta letteralmente in italiano – “la politica delle politiche dell’IA” – non suonerebbe bene. Tuttavia, è di questo che il volume si occupa, poiché nel significato condiviso nella comunità scientifica internazionale questa espressione indica l’insieme dei processi, delle strategie e degli episodi di conflitto, negoziazione e cooperazione, di contrapposizione e di convergenza di visioni e rappresentazioni, valori, credenze e interessi che caratterizzano le azioni pubbliche. Nel nostro caso, esse hanno per oggetto questo specifico sviluppo tecnologico.

L’IA è stata introdotta nelle agende politiche con tre tipi di finalità, che ritroveremo nei prossimi capitoli: (i) favorire la sua adozione nelle attività sociali e il suo sviluppo tecnologico ed economico, anche per incrementare la competitività di sistemi e sotto-sistemi nazionali (ad esempio: industriali, scientifici, di servizi pubblici, bellici); (ii) regolarne gli usi privati (ad esempio nelle diverse fasi della produzione di beni e servizi e, quindi, nel lavoro, o nella comunicazione); (iii) favorirne gli usi a fini pubblici, nelle politiche e nelle organizzazioni coinvolte nei processi di governance, ad esempio per migliorare la pianificazione delle azioni di governo e le prestazioni delle amministrazioni. Quest’ultima finalità poggia sulla convinzione di poter incrementare la razionalità delle decisioni e ridurre le limitazioni cognitive dei processi di azione, passando da processi decisionali affidati solo ad agenti umani a processi condivisi tra agenti umani e l’IA (Bullock *et al.*, 2022b). In queste circostanze, alla definizione delle modalità di uso pubblico si accom-

pagna la necessità di regolarle, con finalità e accorgimenti non molto diversi da quelli relativi agli usi privati.

Questa pluralità di finalità riflette e, a sua volta, può innescare dei conflitti, che la politica cerca di affrontare e risolvere e dagli esiti di questi processi politici dipende l'allocazione di risorse importanti. Possiamo quindi chiederci: quali sono gli interessi sociali economici e politici coinvolti nell'uso e nello sviluppo di IA? Chi guadagna o perde a seconda di quali decisioni politiche vengono prese (o meno) sull'IA? Quali valori della convivenza sociale e politica sono chiamati in causa, perché potenzialmente minacciati o, comunque, come riferimento per prendere decisioni collettive che riguardano l'IA? In che modo l'insieme di questi fattori si incontra con le caratteristiche dell'azione e delle strutture politiche e istituzionali? Provando ad articolare e a rispondere a queste domande, questo libro cerca di contribuire a processi conoscitivi che, ovviamente, sono già stati avviati e riguardano “come vengono adottate le decisioni collettive e le politiche sull'IA, come sono stabilite le priorità e scelti gli obiettivi nel contesto di una pluralità di attori, punti di vista, idee e valori” e “chi è incluso e chi è escluso dal prendere e influenzare le decisioni collettive e le politiche pubbliche” (Ulnicane e Erkkilä 2023, p. 613).

In altri termini, partendo dagli attori e includendo nell'osservazione il ruolo dei contesti in cui avvengono le loro interazioni, i nostri obiettivi principali sono capire chi, quanto e come esercita potere nelle politiche dell'IA – in particolare, i governi, le grandi imprese, i portatori di competenze scientifiche e tecniche o i cittadini? – e come reagiscono le istituzioni e i sistemi politici di fronte a un'innovazione tecnologica dalle implicazioni sociali così dirompenti.

Per rispondere a queste domande sui poteri nelle politiche dell'IA – anche se in modo per forza di cose provvisorio, poiché esaminiamo processi che si stanno svolgendo e trasformando rapidamente – abbiamo quindi cercato di ricostruire le caratteristiche degli attori coinvolti, delle arene in cui interagiscono e delle strategie politiche cui questi processi stanno mettendo capo. Intendiamo per arene nelle quali si svolgono attività di governo, producendo politiche pubbliche, degli ambienti strutturati che possono presentare caratteristiche di conflitto o consenso e che sono plasmate dai costi e dai benefici percepiti dai soggetti coinvolti (Heinelt, 2007), oltre che da regole formali e informali che influenzano il corso dei processi decisionali (John e Cole, 2000). Queste arene possono avere diverse caratteristiche, a partire dalle scale, ossia l'estensione geografica che le caratterizza e le relazioni fra “livelli” di potere politico e sovranità implicati. Nei capitoli seguenti sono presentate le caratteristiche di arene politiche internazionali dell'IA, come le Nazioni Unite (UN), il gruppo dei G7 e altri forum internazionali, dell'U-

nione europea (UE), con la sua natura ibrida, che include poteri sopra-nazionali, e di alcune arene nazionali. Sono esaminate in modo distinto le arene dei “giganti” – gli Stati Uniti d’America (Usa), la Cina e il Regno Unito (UK) – e quella italiana, un’arena nazionale fra le molte possibili, che abbiamo voluto approfondire per ovvie ragioni. Sono inoltre considerate le arene in cui si definiscono l’uso e il governo locali dell’IA, più specificamente in ambito urbano.

In tutte queste arene, sulla base dei sistemi di regole che ne disciplinano il funzionamento, vengono prese delle decisioni sull’IA e si cerca di attuarle, con processi il più delle volte tutt’altro che semplici e lineari. In queste decisioni e nei loro esiti abbiamo cercato di rintracciare delle strategie, intese come dei piani di azione deliberati e coordinati adottati per raggiungere degli specifici obiettivi nelle politiche dell’IA. I loro principali elementi sono: (i) un orientamento che può privilegiare, ai due estremi di un continuum, finalità di regolazione “social-oriented”, ossia volta a tutelare i diritti umani e mitigare i rischi pubblici e sociali o, all’opposto, “market-oriented”, per consentire lo sviluppo maggiore e più rapido delle tecnologie di IA, considerate un ingrediente ormai imprescindibile della competitività dei sistemi di impresa, specialmente nazionali, sui mercati globali; (ii) gli strumenti giuridici e organizzativi che vengono attivati; (iii) le risorse mobilitate, di tipo sia materiale (come gli stanziamenti per investimenti pubblici), sia immateriale, consistenti in regolazioni, norme e immaginari veicolati da rappresentazioni e narrazioni. Questi ultimi, come abbiamo visto, possono produrre paura, oppure consenso e fiducia nei confronti dell’IA da parte dei loro potenziali utenti, come le imprese che acquistano e usano sistemi di IA e i loro clienti, o i consumatori, o coloro che usufruiscono di servizi privati o pubblici, direttamente o indirettamente trasformati dall’IA.

Il primo capitolo mette a fuoco la dimensione problematica con la quale l’IA entra nelle agende politiche in tutte queste arene, in particolare i rischi pubblici e le minacce ad essa associati, oggetto di interesse in primo luogo per l’analisi giuridica. Ad esempio, le discriminazioni sociali o i danni per la privacy e per la libertà delle persone potenzialmente provocati da decisioni autonome dell’IA, o dal suo uso a supporto di decisioni, in particolare nel campo della sicurezza e della sorveglianza. Rischi che aumentano con il crescere della opacità di modelli e sistemi di IA, i quali possono occultare i pregiudizi sociali riprodotti nel loro addestramento e nei dati utilizzati a questo fine. Questi rischi sfidano non solo la capacità della politica di produrre regolazioni adeguate, ma anche l’adeguatezza di quelle esistenti, ad esempio in materia di privacy. Ovviamente i rischi sono ancora più gravi quando i danni sono provocati intenzionalmente. Si tratta non solo del *phishing* commerciale, ma anche di usi che caratterizzano il funzionamento dei processi

comunicativi (*fake news, filter bubbles, echo chambers*), della politica e della democrazia. In questo ambito, il *trade-off* fra potenzialità e rischi, diretti e indiretti, è particolarmente delicato, anche perché interviene in un campo – la democrazia – di per sé attualmente esposto a pericoli e minacce particolarmente evidenti, provocati dall’ascesa dei populismi di destra e dal conseguente diffondersi dell’autoritarismo. Le applicazioni di IA, in particolare generativa, rischiano quindi di amplificare alcune preesistenti fragilità dei sistemi democratici.

Il secondo capitolo offre risposte a interrogativi circa i rapporti fra le forme che la regolazione dell’IA può assumere, tenendo conto della complessità e della rilevanza degli interessi economici e sociali coinvolti. Da un lato emerge il ruolo delle imprese, a partire dalle Big Tech, che cercano di preservare la propria autonomia nella ricerca e lo sviluppo di IA sempre nuova attraverso forme di autoregolazione, o *soft law*, che spesso mirano a prevenire l’adozione di norme autoritative e stringenti da parte di autorità pubbliche. Dall’altro emergono dei tentativi di configurare arene internazionali in cui concordare principi e strumenti per la governance dell’IA, che includono forme di *hard law, soft law* e loro combinazioni. Sono visibili attori e ambienti istituzionali trainanti, i quali hanno introdotto i principali *frame* cognitivi e normativi delle politiche dell’IA. Questi, pur non dando luogo necessariamente a politiche omogenee o convergenti, hanno ispirato e strutturato i comportamenti di altri attori e istituzioni e le loro strategie politiche. I protagonisti sono stati l’OECD, che ha riconfermato il suo ruolo di indirizzo cognitivo e normativo nel contesto internazionale, e l’UE, con il suo tentativo di giustapporre strategie social-oriented e market-oriented.

Il terzo capitolo approfondisce e mette a confronto le strategie europee con quelle perseguitate dagli Stati i cui sistemi economici e tecnologici sono così potenti da indirizzare e condizionare lo sviluppo dell’IA in tutto il mondo, gli Usa, la Cina e, seppure in misura minore, lo UK. Su questa scala nazionale emerge con particolare evidenza il ruolo svolto da due fattori nell’indirizzare e cambiare gli orientamenti delle strategie politiche dell’IA: da un lato i rapporti fra la leadership politica e le maggiori imprese tecnologiche, di fatto in grado di controllare mercati oligopolistici dell’IA; dall’altro fattori strettamente politici come i cambiamenti elettorali e l’avvicendamento delle élite di governo. A questi si può ricondurre, in particolare, l’oscillazione fra due polarità strategiche, con una tendenza all’affermazione di visioni e misure liberiste favorita dai cambiamenti elettorali che, negli ultimi due anni, hanno rafforzato in molti casi leader e partiti di destra.

Il quarto capitolo, infine, concentra l’attenzione sulla scala locale delle politiche dell’IA, mettendo a fuoco le continuità e le trasformazioni nel passaggio dalle smart city a una sorta di nuovo paradigma dell’“IA urbana”. Le

città sono ambienti privilegiati per l'utilizzazione non solo privata, ma anche pubblica dell'IA. Anche in questo caso si pongono problemi relativi a governarne i *trade-off*, ma i margini di autonomia regolativa dei sistemi politici locali sono differenziati, più elevati ad esempio negli Usa, minori in Europa. In tutti questi contesti, abbiamo potuto ricostruire dinamiche diverse ma spesso caratterizzate dal ruolo importante dei contesti politici e del grado di istituzionalizzazione dei rispettivi sistemi di governo nel dare spunto, ma anche nel vincolare le modalità di ingresso in agenda e la conduzione di politiche dell'IA. In particolare, evidenziamo l'importanza del ruolo dei contesti istituzionali. Ciò emerge non solo nelle dinamiche all'interno dei sistemi politici nazionali, ma anche nel funzionamento delle arene internazionali. Le UN, l'OECD e il gruppo dei G7 inseriscono l'IA nelle rispettive agende e la processano poi con modalità predefinite dalle prerogative piuttosto deboli del loro status politico e istituzionale. Quelle dell'UE, invece, sono più forti e strutturate, tanto da consentirle di produrre la regolazione "dura" al momento più avanzata al mondo. L'esistenza di regolazioni e interventi politici preesistenti in ambiti contigui di governance delle tecnologie digitali – si pensi alla tutela della privacy o della concorrenza – rende dappertutto più facile attingere a repertori di soluzioni normative e relativi alle routine di azione. Questi due aspetti permettono di capire perché, anche a fronte della necessità di affrontare sfide di portata inedita come quelle dell'IA, non siano sinora emerse innovazioni nel funzionamento dei sistemi politici e istituzionali, se non di carattere incrementale. Ad esempio, per quanto riguarda l'UE, possiamo rilevare l'applicazione della consolidata sequenza "documento della Commissione, gruppo di esperti, libro bianco, consultazione ad ampio spettro nella società, presentazione di una proposta di regolamento, esame da parte del Parlamento e del Consiglio, trilogo", con gli interventi e le pressioni di lobby e società civile, organizzazioni no profit, e il coinvolgimento di esperti. Anche a fronte della dirompenza del suo oggetto, dalla nostra ricostruzione emerge come la "politica delle politiche dell'IA" condivida modalità di azione, orientamenti strategici alternativi, conflittualità e stili di gestione dei conflitti con altre arene come, in epoca recente, quelle delle politiche per lo sviluppo sostenibile e del contrasto ai cambiamenti climatici.

1. Rischi, potenzialità e poste in gioco nelle agende dell'intelligenza artificiale

1.1. I rischi pubblici dell'intelligenza artificiale

“Il problema dell’IA”. Questo titolo, conciso e immediato, apriva le riflessioni di un articolo sulla regolazione dei sistemi di IA pubblicato nel 2016 dalla rivista di diritto delle tecnologie della Harvard Law School. Sin dalle prime righe M. U. Scherer (2016), esperto di diritto del lavoro, descrive il problema dell’IA in riferimento alla categoria dei rischi pubblici¹ (*public risks*) associati agli sviluppi tecnologici. L’applicazione di questa categoria era motivata dalla crescente ubiquità delle tecnologie di IA che oggi riscontriamo in molteplici ambiti della vita quotidiana, nelle transazioni di mercato e financo nell’elaborazione e conduzione delle politiche pubbliche. Le innovazioni di IA pongono le stesse sfide di qualsiasi tecnologia emergente, di cui è necessario stabilire i requisiti di conoscenza circa i rischi ad essa associati, in altre parole conoscere i fatti che creano un pericolo, la sua natura e la portata che lo rende irragionevole (Stein, 2022).

Un importante punto di vista è quello giuridico, dal quale è possibile analizzare fatti e situazioni e le loro conseguenze concrete per interessi e diritti. Si pensi che tra i primi ambiti di osservazione rientrano i casi di incidenti provocati dalle automobili a guida autonoma (es. Roe, 2019; Kaminski, 2023), con studi che hanno mostrato la complessità nell’imputare responsabilità e colpe dei danni provocati. Se invece osserviamo fatti e situazioni da un punto di vista sociologico, è necessario operare una distinzione tra le categorie di pericolo, rischio e minaccia, per evitare di creare confusione. In tal senso ci viene in aiuto l’efficace sintesi proposta da F. Battistelli (2016)

¹ Secondo la definizione originaria fornita da P. Huber nella seconda metà degli anni Ottanta i rischi pubblici: «sono prodotti a livello centrale o di massa, ampiamente distribuiti e in gran parte al di fuori della comprensione e del controllo diretto del singolo portatore del rischio» (1985: 277).

basata sulle variabili della genesi e dell'intenzionalità dei danni provocati. La prima categoria, i pericoli, consiste nei danni che provengono dall'esterno della società, quindi esogeni e inintenzionali, di cui le calamità naturali sono esempio. I rischi, invece, rientrano nella categoria dei danni provocati da decisioni di individui, gruppi o organizzazioni che non hanno una intenzionalità negativa e che non intendono arrecare un danno. Sono quindi danni provocati dall'attività umana, dalle attività produttive e dalle decisioni di policy che si pongono obiettivi positivamente orientati, ma falliscono. Secondo Battistelli, infine, la categoria di minaccia comprende i danni pienamente intenzionali. In accordo con tale distinzione parleremo di rischi o minacce dell'IA, attribuendo ai possibili danni dell'IA un carattere endogeno ai processi e ai sistemi sociali, accantonando quindi l'idea – pur presente nel discorso pubblico – che un'IA pienamente autonoma si renda indipendente dalle decisioni e intenzionalità degli attori sociali.

Quindi in questo paragrafo ci occupiamo del rapporto tra IA, (nuovi) rischi pubblici e minacce. Infatti, il bilanciamento efficace, il *trade-off*, tra potenzialità e rischi rappresenta la principale sfida dell'IA per la società e dunque anche per la politica, i suoi poteri e le sue istituzioni. Esplicitare i principali rischi connessi con lo sviluppo e l'applicazione trasversale delle tecnologie di IA implica, tuttavia, tenere conto del dualismo che le connota in termini sia di effetti negativi, sia di potenzialità e benefici. Lo *High-level Advisory Body on AI*, organo consultivo delle UN istituito nel 2023, ha recentemente sostenuto che stilare un elenco esaustivo dei rischi dell'IA sia un'impresa “da pazzi” (UN, 2024, p. 28), considerata la rapidità nella sua evoluzione e nella diffusione dei suoi utilizzi. Tuttavia, la letteratura scientifica sul tema ha identificato alcuni degli approcci più efficaci nell'identificazione dei tipi di rischio dell'IA. Le tassonomie dei rischi che ne derivano si diversificano sulla base di diversi criteri di classificazione. Esse possono riguardare, ad esempio, la scala del rischio (Critch e Russell, 2023), i fattori causali e i settori di attività interessati dai rischi dominio (Slattery *et al.*, 2024) e financo le vulnerabilità, che l'IA può aggravare nelle categorie sociali più fragili (UN, 2024). L'idea dell'esistenza di diversi destinatari dei rischi origina da una semplice domanda: di chi è il problema dei rischi dell'IA? Il framework elaborato dal *National Institute of Standards and Technology* (NIST), istituito dal governo degli Usa, fornisce la seguente risposta: i danni dell'IA possono essere subiti da persone, organizzazioni e interi sistemi (NIST, 2022). Riguardo i rischi che interessano le persone il NIST distingue a sua volta tra danni per gli individui – lesione delle libertà e dei diritti individuali – danni subiti da gruppi e comunità, come nel caso della discriminazione di sottocategorie di popolazione e danni per la società nel suo complesso se, ad esempio, venisse ostacolato e manipolato l'accesso e

la partecipazione ai sistemi politici. Le organizzazioni, profit e no, rappresentano la seconda categoria di destinatari dei danni identificata dal NIST. L'IA può impattare negativamente, ad esempio, sulla reputazione di un'impresa o di un altro tipo di organizzazione, o sulla conduzione delle sue attività. La terza categoria di destinatari ha una scala più ampia e si riferisce ai danni che possono riguardare sistemi che operano su scala globale (es. sistema finanziario, filiere globali di approvvigionamento di beni) e, dal punto di vista ambientale, interi ecosistemi, ponendo così il problema della sostenibilità dell'IA. Si tratta con tutta evidenza di un approccio utile a "scalare" il problema dell'IA e a mostrare le sue implicazioni sulla vita degli individui e sul funzionamento di sistemi complessi che travalicano i confini nazionali.

Tuttavia, non abbiamo ancora definito i tipi di rischi pubblici associati all'IA. A questo scopo ci viene in aiuto un poderoso lavoro di categorizzazione, tuttora in corso, condotto da ricercatori del Massachusetts Institute of Technology (MIT) che, al gennaio 2025, ha identificato 1.518 rischi associati all'IA. Il lavoro del MIT è stato condotto attraverso una revisione sistematica della letteratura scientifica e "grigia" sul tema, finalizzata all'individuazione di estratti di testo indicativi dei possibili danni derivanti dall'impiego di sistemi di IA². Ne sono emerse due distinte, ma complementari, tassonomie del rischio: *Domain Taxonomy of AI Risks*; *Causal Taxonomy of AI Risks*. La prima identifica sette ambiti (*domains*) del rischio, a loro volta distinti in 23 sotto-ambiti (*subdomains*): 1. discriminazione e tossicità; 2. privacy e sicurezza; 3. disinformazione; 4. usi malevoli; 5. interazione uomo-macchina; 6. danni ai sistemi economici e danni ambientali; 7. sicurezza e fallimenti dei sistemi di IA (Slattery *et al.*, 2024). La seconda tassonomia comprende tre variabili: 1. entità, intesa come l'autore del danno – umani, sistema di IA, altro (quando il danno non è attribuibile in modo univoco); 2. intento, inteso come intenzionalità o meno del danno rispetto all'obiettivo iniziale del sistema; tempo – quando il rischio, in termini di danno occorso, si presenta nel ciclo di vita del sistema di IA, in fase di progettazione o dopo essere entrato in fase d'uso. La tassonomia causale evidenzia un dato rilevante: il 51% dei rischi classificati è da attribuire alle decisioni o, meglio, alle azioni, dei sistemi di IA. Un dato che riflette le preoccupazioni e le evidenze della comunità scientifica circa il rapporto tra l'autonomia, caratteristica costitutiva dell'IA, e i rischi che ne possono derivare, tema su cui torneremo.

Da un lato, il lavoro del MIT mostra la proliferazione di studi sugli effetti avversi dello sviluppo dell'IA e, parafrasando A. Giddens (1994), evidenzia

² Il lavoro dei ricercatori MIT ha considerato un corpus iniziale di 1.700 documenti da cui sono stati estratti 56 paper scientifici riguardanti altrettanti framework analitici dei rischi dell'IA. Il Database completo e aggiornato è disponibile al link: <https://airisk.mit.edu/>

la presenza di un “monitoraggio riflessivo” da parte delle diverse comunità epistemiche coinvolte; dall’altro, rende manifesta la necessità di velocizzare la produzione di conoscenze circa il panorama dei rischi con evidenti ripercussioni sulla prevenzione di condizioni di anomia provocate dall’IA. I sette ambiti identificati dai ricercatori MIT, forniscono comunque una pista da seguire per esplorare i diversi impatti negativi associati all’IA, nella loro dimensione sia individuale, sia di sistema. In particolare, ci occuperemo di tre ambiti di rischio che a nostro avviso restituiscano la complessità dei rischi pubblici e delle minacce dell’IA su cui si è strutturata un’agenda pubblica e politica: discriminazioni; privacy; usi malevoli.

Il primo ambito si riferisce alle discriminazioni nei confronti di singoli individui o di particolari categorie sociali. Le discriminazioni mediate dai sistemi di IA, o pregiudizi algoritmici, rappresentano una delle principali preoccupazioni che si sono consolidate nel dibattito scientifico sui sistemi di decisione autonoma (es. Pasquale, 2015; Danks e London, 2017). I sistemi di IA possono essere ricondotti a due tipi principali, distinti sulla base dell’autonomia e delle modalità in cui sono impiegati nei processi decisionali (Candrian e Scherer, 2022): algoritmi di consulenza e IA performativa. I primi servono a potenziare le decisioni umane, contrastando le condizioni di razionalità limitata; il secondo tipo si riferisce a sistemi che agiscono in autonomia. Essi si estendono a una molteplicità di attività in cui la decisione finale è delegata a sistemi IA che operano senza il controllo da parte di supervisori umani. Nel settore della comunicazione e dei consumi mediati da anni siamo entrati in contatto quotidiano con le decisioni autonome dell’IA in merito, ad esempio, alla selezione delle notizie che potrebbero interessarci (Carlson, 2018). Si pensi che le decisioni autonome dell’IA sono diventate parte integrante dei processi aziendali in riferimento alla definizione delle strategie di impresa, l’identificazione nel mercato di possibili attori concorrenti, clienti potenziali (Huang e Rust, 2021) o nei sistemi di reclutamento del personale (Black e van Esch, 2020). Risulta evidente come il processo di automazione delle decisioni abbia implicazioni sociali di diversa natura a seconda degli ambiti della vita sociale e dei destinatari degli *output* decisionali. Un esempio efficace riguarda la decennale applicazione nei processi di giustizia penale negli Usa del sistema di predizione del rischio di recidiva *Correctional Offender Management Profiling for Alternative Sanction*³ (COMPAS). Sebbene esso rientri fra i sistemi che potenziano la decisione umana, il giudice mantiene autonomia decisionale. Ciò nonostante, il suo

³ Il sistema COMPAS consiste in un software di proprietà privata che fornisce il calcolo della probabilità di recidiva di un imputato sulla base della combinazione di 137 item sulle sue caratteristiche socioeconomiche.

impiego ha attirato numerose polemiche circa il rischio di favorire decisioni discriminatorie sulla base di caratteristiche ascritte dell'imputato. Si pensi in tal senso all'iniziativa di ProPublica, organizzazione no-profit di giornalismo investigativo statunitense, che ha monitorato il caso di diecimila condannati in una corte penale della Florida mostrando gli errori nella previsione di recidive violente che avrebbero penalizzato gli imputati neri⁴. I dati resi disponibili da ProPublica hanno aperto un acceso dibattito scientifico sulle discriminazioni da attribuire all'applicazione di COMPAS nei processi (es. Washington, 2018; Engel *et al.*, 2024), una condizione che si estende ad altri ambiti della vita sociale.

Tali discriminazioni deriverebbero dai possibili *bias* dell'IA: risultati (pregiudizievoli) dovuti a pregiudizi umani che distorcono i dati di training originali, o l'algoritmo⁵. Dunque, l'origine di tali discriminazioni può essere imputata al processo di addestramento dell'algoritmo, se esso è basato su decisioni passate che sono a loro volta, espressione di pregiudizi sociali, o all'uso improprio dell'algoritmo stesso, se, ad esempio, viene impiegato su categorie di popolazione diverse da quella per cui è stato costruito (Ferrer *et al.*, 2021). L'individuazione dei pregiudizi algoritmici è però particolarmente difficile, sia se si prova a ricostruire i processi di addestramento degli algoritmi, sovente sottoposti a segreto commerciale, sia che si osservino gli *output* decisionali (van Nuenen, 2020). Il problema dell'opacità delle decisioni dei sistemi di IA è stato oggetto del famoso saggio *The Black Box Society - The Secret Algorithms That Control Money and Information* del giurista F. Pasquale, pubblicato nel 2015 dalla Harvard University Press. L'autore discute le nuove forme di potere, esercitato attraverso il controllo delle conoscenze su cui si basano complessi sistemi analisi dati, algoritmi e dunque decisioni automatizzate. In tal senso esplora le strategie per tenere segreti, offuscare o al contrario decriptare i contenuti degli algoritmi, applicando la metafora della "scatola nera": «le aziende cercano di scoprire dettagli intimi sulla vita di potenziali clienti e dipendenti, ma forniscono alle autorità di regolamentazione il minor numero possibile di informazioni sulle proprie statistiche e procedure» (Pasquale, 2015, pp. 3-4).

Come vedremo, il controllo delle conoscenze, la gestione di enormi patrimoni di dati e il mantenimento della segretezza sono alcune delle poste in gioco nei rapporti di forza tra attori pubblici e grandi imprese tecnologiche nella regolazione dell'IA. La metafora della scatola nera è stata ripresa in uno studio sui rischi di discriminazione nel decision-making algoritmico

⁴ I risultati dello studio sono disponibili al link: <https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm>

⁵ Fonte: <https://www.ibm.com/it-it/think/topics/ai-bias>

commissionato nel 2018 dalla Direzione Generale della Democrazia del Consiglio d’Europa: «spesso non è chiaro a qualcuno perché un sistema prenda una certa decisione nei suoi confronti; a causa dell’opacità di tali decisioni, è difficile per le persone valutare se sono state discriminate sulla base di un’informazione o di un’altra» (Zuiderveen Borgesius, 2018, p. 10). Lo studio riportava una serie di ambiti in cui era già stato possibile rinvenire episodi di discriminazione mediata dall’IA: attività di prevenzione del crimine, processi di selezione del personale o dei candidati all’ingresso a scuole e università, pubblicità, acquisti e ricerche online, sistemi di traduzione automatica. L’individuazione dei *bias* rappresenta quindi un campo fecondo per gli studi delle scienze sociali che hanno l’obiettivo di analizzare la riproduzione di stereotipi e di forme di discriminazione attraverso la loro reificazione negli algoritmi. Sono innumerevoli le ricerche che hanno esplorato le discriminazioni nei confronti delle caratteristiche protette⁶ con particolare riferimento al genere e all’origine etnica (es. Collett e Dillon, 2019; Nuseir *et al.*, 2021; Lee, 2018; Adams, 2021; Andrews Bucher, 2022; Hu, 2023; Franzoni, 2023; Broussard, 2023). Un problema spesso derivante da una sottorappresentanza di tali caratteristiche sia tra gli sviluppatori del settore, si pensi al gender gap che tuttora connota il settore dell’IA (Young *et al.*, 2023), che nelle basi dati impiegate per l’addestramento degli algoritmi. Ne è un esempio la mancanza di un’adeguata considerazione delle caratteristiche protette nello sviluppo degli algoritmi applicati alle pratiche mediche e diagnostiche che si traduce sulla loro inefficacia (Giovanola e Tiribelli, 2023), questione recentemente denunciata dal *World Economic Forum* rispetto, ancora una volta, ai *bias* su base razziale⁷. I. Ulnicane (2024) ha mostrato come stia emergendo un’agenda di studio sulle discriminazioni dell’IA basata sul concetto sociologico di intersezionalità, che evidenzia la necessità di comprendere le interdipendenze e le sovrapposizioni tra discriminazioni che riguardano identità multiple, caratteristiche meritevoli di protezione, come il genere, l’etnia, l’orientamento sessuale e la classe sociale.

Il secondo ambito dei rischi dell’IA categorizzato dai ricercatori MIT riguarda la violazione della privacy delle persone. Privacy è un termine di uso comune nella quotidianità e nel dibattito pubblico che, tuttavia, necessita di

⁶ Con questa espressione ci riferiamo alla gergalità dei diritti umani. Si pensi in tal senso all’articolo 21 della Carta dei Diritti Fondamentali dell’UE vieta qualsiasi forma di discriminazione in base alle seguenti caratteristiche protette: «il sesso, la razza, il colore della pelle o l’origine etnica o sociale, le caratteristiche genetiche, la lingua, la religione o le convinzioni personali, le opinioni politiche o di qualsiasi altra natura, l’appartenenza ad una minoranza nazionale, il patrimonio, la nascita, gli handicap, l’età o le tendenze sessuali», fonte: https://www.europarl.europa.eu/charter/pdf/text_it.pdf

⁷ Fonte: <https://www.weforum.org/stories/2024/09/racial-bias-healthcare-data-equity/>

una definizione. I giuristi statunitensi K. Manheim e L. Kaplan (2019) riconducono alla fine dell’Ottocento l’origine del principio della privacy, declinato da S. Warren e L. Brandeis (1890) come il diritto delle persone ad essere lasciate in pace. A questo principio sono associati almeno quattro tipi di violazione: l’intrusione negli affari privati, la divulgazione di fatti privati, la diffamazione (*false light publicity*) e l’appropriazione del nome o dell’immagine di una persona (*ibidem*)⁸. È opinione diffusa che lo sviluppo e l’impiego dei sistemi di IA possa amplificare i rischi di incorrere, intenzionalmente e non, in ognuna di queste violazioni. Uno studio recente ha costruito una tassonomia che ha identificato 12 tipi di rischi per la privacy associati ai sistemi di IA (Lee *et al.*, 2024). Secondo H. Lee e colleghi essi riguardano, in primo luogo, la massiva raccolta di dati personali, attività che conduce ad un “aggravamento della sorveglianza” delle attività degli individui (*ibidem*), intesa come l’osservazione, ascolto e registrazione delle attività personali. Essa rappresenta, dunque, uno dei principali strumenti di intrusione negli affari privati delle persone su cui si concentrano gli studi sui rischi dei sistemi di decisione automatica. Da un lato, i sistemi di IA necessitano di una considerevole mole di dati per l’addestramento degli algoritmi, un bisogno che trova soddisfazione nel web, divenuto progressivamente un “*all you can eat*” dei dati personali⁹. Dall’altro lato, i sistemi di IA si servono di una raccolta permanente di informazioni (*data gathering*) per svolgere le proprie funzioni decisionali, ottenuta attraverso la cessione volontaria da parte di individui e consumatori, o l’acquisizione delle tracce lasciate dalle attività personali online (Puntoni *et al.*, 2021; Martin e Zimmerman, 2024). Essi si “nutrono” di *data privacy*, un particolare tipo di privacy relativo al rapporto tra individuo e macchine¹⁰. Secondo S. Sharma (2020) l’IA aggiorna e amplia i confini dei

⁸ Come ricordava S. Rodotà (2004), la “rivoluzione pacifica della privacy” si è consolidata nell’ordinamento italiano con l’approvazione della Legge 675 del 1996 che ha attribuito a ciascuno il potere di governo delle informazioni che lo riguardano. Ad essa è seguita la costruzione a livello nazionale ed europeo di una complessa infrastruttura normativa e organizzativa culminata a, circa venti anni di distanza, con l’approvazione del Regolamento Generale sulla Protezione dei Dati - GDPR 2016/679 (UE). Fonte: <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/1093776>

⁹ Fonte: <https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/10034090>.

¹⁰ Un esempio riportato da J. Silverman (2017: 157) mostra le implicazioni di data privacy del consumo delle foto di Instagram come oggetti di dati: «la foto viene analizzata da programmi di riconoscimento facciale e degli oggetti; gli esperti di marketing la scansionano per vedere come appaiono i prodotti dei loro clienti; i metadati rivelano agli inserzionisti dove è stata scattata la foto; le forze dell’ordine e le agenzie di *intelligence* passano i commenti della foto attraverso un software di *sentiment analysis*, alla ricerca di attività illegali o di segni di radicalizzazione; dei bot loschi usano la foto come avatar; un numero incalcolabile di computer nei centri dati e negli hub Internet di tutto il mondo sminuzzano la foto e la trasportano in giro come pacchetti di informazioni, producendo nel frattempo registrazioni sul suo transito».

contesti tradizionali della privacy (es. la propria abitazione, la porta chiusa di una sala riunioni) e le modalità di condivisione e appropriazione delle informazioni che divengono meno appariscenti e meno riconducibili a violazioni esplicite. Tali appropriazioni sono fondative del “capitalismo della sorveglianza”, il cui sistema di accumulazione si basa sulla mercificazione dei comportamenti umani (online e non) che diventano “materie prime” e sulla possibilità di elaborare previsioni finalizzate alla creazione di mercati dei comportamenti futuri (Zuboff, 2019). In tal senso, gli algoritmi di IA sono strumenti imprescindibili per potenziare la capacità di tracciare, prevedere e influenzare i comportamenti delle persone a fini privati (Jones, 2025).

I meccanismi di appropriazione delle informazioni personali caratteristici del capitalismo della sorveglianza possono essere ritrovati anche nel settore pubblico (Stahl *et al.*, 2023). L'esempio più immediato è il Sistema di Credito Sociale (SCS), sviluppato e adottato in Cina dal 2014, un insieme di meccanismi che forniscono ricompense o punizioni come feedback alle persone, basati non solo sulla legalità, ma anche sulla moralità delle loro azioni, che riguardano la condotta economica, sociale e politica (Creemers, 2018). Il SCS si basa sulle logiche e le infrastrutture del capitalismo della sorveglianza rivolte non a fini commerciali ma per incoraggiare comportamenti socialmente, politicamente ed economicamente desiderabili (Aho e Duffield, 2020). Gli algoritmi di IA forniscono gli strumenti per l'elaborazione della grande mole dei dati del SCS (Backer, 2019) e per attribuire ricompense e punizioni. Sebbene recenti studi e riflessioni ne abbiano discusso e ridimensionato la portata¹¹, il SCS rappresenta una complessa infrastruttura di sorveglianza di Stato abilitata da Big Data¹² e IA (Liang *et al.*, 2018), uno scenario contrastato dal sistema istituzionale dell'UE e dalla sua risposta di regolazione pubblica. Tuttavia, l'aggiornamento degli strumenti di sorveglianza con le tecnologie di IA è un processo consolidato anche nei Paesi occidentali per motivi di sicurezza pubblica. Le tecnologie di riconoscimento facciale sono entrate negli spazi pubblici, negli aeroporti, nelle stazioni e negli stadi di calcio attraverso sperimentazioni e progetti pilota, talvolta falliti, che stanno progressivamente socializzando i cittadini a questa forma di controllo. Come vedremo, l'AI Act dell'UE consente alle autorità pubbliche l'impiego delle tecniche di identificazione biometrica per il contrasto alla criminalità organizzata, la prevenzione delle azioni di terrorismo e la ricerca delle persone scomparse. Questa intrusione, pur legittimata

¹¹ Fonte: <https://www.technologyreview.com/2022/11/22/1063605/china-announced-a-new-social-credit-law-what-does-it-mean/>

¹² Per Big Data intendiamo una grande e complessa mole di dati raccolti quotidianamente da diverse fonti, elaborati tramite AI e impiegati per ottenere nuove informazioni e scenari predittivi. Fonte: <https://www.europarl.europa.eu/topics/it/article/20210211STO97614/big-data-definizione-benefici-e-sfide-infografica>

pubblicamente, è oggetto di numerosi interventi da parte di organizzazioni della società civile. Particolarmente significativo in tal senso è il movimento europeo *Reclaim your face*, nato per aggregare la *voice* pubblica sui rischi del riconoscimento facciale sia da parte delle imprese nei contesti lavorativi che dei governi, che si è espresso negativamente rispetto al compromesso politico raggiunto dalle istituzioni europee sul tema della sorveglianza biometrica di massa¹³.

L’intrusione nelle attività delle persone non è tuttavia l’unico processo mediato dall’IA che impatta sulla privacy. Il “movimento” della violazione può essere rivolto anche verso l’esterno, attraverso la diffusione accidentale di informazioni personali sensibili (Slattery *et al.*, 2024). Le maggiori preoccupazioni degli esperti riguardano i *Large Language Models* (LLMs) (Hagendorff, 2024), che impiegano nelle fasi di addestramento nomi, indirizzi mail, conversazioni che, come mostrato in diversi studi, possono essere recuperate ed eventualmente divulgare (es. Carlini *et al.*, 2021; Huang *et al.*, 2022; Wang *et al.*, 2023).

Dunque, lo sviluppo dell’IA pone nuove sfide ai preesistenti strumenti di regolazione pubblica nel campo della privacy, con particolare riferimento alla anonimizzazione dei dati personali¹⁴. Si pensi che solo nel dicembre 2024 il Comitato europeo per la protezione dei dati (EDPB) ha provato a districare le condizioni per le quali i dati per l’addestramento dei modelli di IA possono essere considerati anonimi e dunque non sottoponibili alle prescrizioni del GDPR¹⁵. I rischi per la privacy sono stati una delle porte di accesso per la definizione di un’agenda pubblica, politica e istituzionale sulle innovazioni provenienti dal settore dell’IA. In Italia gli esempi più eclatanti riguardano le istruttorie aperte dal Garante della Privacy nei confronti di OpenAI per il servizio “ChatGPT” e delle aziende cinesi che forniscono la *chatbot* “DeepSeek”¹⁶. Le due iniziative hanno condotto a limitazioni (temporanee) nell’uso dei due servizi sul territorio italiano e ad altre azioni sanzionatorie basate sulla strumentazione giuridica fornita agli stati membri dell’UE dal GDPR¹⁷, a dimostrazione della stretta relazione tra preoccupa-

¹³ Il testo del parere è disponibile al link: <https://reclaimyourface.eu/eu-ai-act-will-fail-commitment-to-ban-biometric-mass-surveillance/>

¹⁴ Il GDPR non si applica a dati anonimi, intesi all’ Art. 26 come: «informazioni che non si riferiscono a una persona fisica identificata o identificabile o [...] dati personali resi sufficientemente anonimi da impedire o da non consentire più l’identificazione dell’interessato».

¹⁵ Il testo integrale del parere 28/2024 è disponibile al link: https://www.edpb.europa.eu/system/files/2024-12/edpb_opinion_202428_ai-models_en.pdf

¹⁶ Per una ricostruzione si rimanda a: <https://www.garanteprivacy.it/home/docweb-/docweb-display/docweb/10098369>

¹⁷ Il *General Data Protection Regulation*, ovvero il Regolamento 2016/679 dell’Unione europea in materia di trattamento dei dati personali e di privacy.

pazioni per la privacy e azione istituzionale nel campo dell'IA. Tali preoccupazioni non riguardano tuttavia i soli servizi di IA generativa¹⁸, piuttosto risultano essere trasversali ai diversi settori di uso dell'IA. Ne è esempio il settore sanitario, dove sistemi di IA sono usati in molteplici attività diagnostiche, pratiche di cura e di formazione del personale medico-sanitario. Nel 2023 l'Organizzazione Mondiale della Sanità (OMS) ha pubblicato le linee guida sull'IA in ambito sanitario, dedicando un intero capitolo alla privacy e alla protezione dei dati. Come ogni strumento di *soft law*, le linee guida dell'OMS non forniscono sanzioni o prescrizioni vincolanti per gli attori del settore, ma definiscono un percorso desiderabile per lo sviluppo di applicazioni e sistemi di IA rispettosi delle diverse giurisdizioni nazionali e sovranazionali in ambito di protezione dati e privacy degli individui. In estrema sintesi, le indicazioni dell'OMS riguardano l'anonimizzazione dei dati e la de-identificazione della fonte dei dati (es. pazienti), la protezione degli stessi da cyberattacchi o da altre fonti di violazione che possono avere rilevanti implicazioni per la dignità delle persone, favorire azioni discriminatorie o di stigmatizzazione.

Il terzo ambito dei rischi di cui ci occupiamo riguarda i danni provocati intenzionalmente, che identificano, quindi, le minacce propriamente procurate attraverso l'IA. La relazione tra sicurezza e IA è stata oggetto, già nel 2018, di un importante studio curato da esperti delle università di Cambridge e Oxford, organizzazioni no-profit e think tank nel campo dei diritti e la sicurezza¹⁹. Il rapporto poneva l'attenzione sugli sviluppi e gli usi malevoli dei sistemi di IA intesi come: «le pratiche volte a compromettere la sicurezza di individui, gruppi o società» (Brundage *et al.*, 2018, p. 9). Secondo gli esperti, a essere a rischio sarebbero state la sicurezza digitale, fisica e politica, quest'ultimo un tipo di sicurezza che implica rischi per la tenuta dei sistemi democratici. Sebbene non sia possibile in questa sede fornire una rassegna esaustiva delle minacce dell'IA, è opportuno dotarsi, ancora una volta, di una tassonomia di partenza che consente di esplorarne le molteplici implicazioni sociali. Particolarmenete efficace in tal senso è il lavoro di T. F. Blauth e colleghi (2022) che collocano le minacce degli usi malevoli dell'IA negli

¹⁸ L'IA generativa consiste di algoritmi (come ChatGPT) che possono essere utilizzati per creare nuovi contenuti, tra cui audio, codice, immagini, testo, simulazioni e video. Gli strumenti di IA generativa possono produrre un'ampia varietà di scritti credibili in pochi secondi e rispondere alle critiche per renderli più adatti allo scopo. Ciò ha implicazioni per un'ampia gamma di settori. In breve, tutte le organizzazioni che hanno bisogno di produrre materiale scritto chiaro possono trarne vantaggio. Fonte: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>

¹⁹ Gli autori del rapporto: Future of Humanity Institute; University of Oxford; Centre for the Study of Existential Risk; University of Cambridge; Center for a New American Security; Electronic Frontier Foundation; OpenAI.

ambiti, rispettivamente, dell’ingegneria sociale (*social engineering*), delle attività di *hacking*, dello sviluppo e uso di sistemi di arma autonomi e della disinformazione. Il primo ambito riguarda le note azioni di *phishing* online – sottrazione fraudolenta di informazioni sensibili – condotte attraverso il sempre più massiccio ricorso a *social bot*, ovvero finti profili basati su algoritmi che simulano il comportamento umano e che possono interagire, anche con finalità malevole, con gli utenti umani. L’IA generativa può rendere i vettori di attacco più realistici e convincenti attraverso contenuti testi, immagini e video basati e “appresi” su dati esistenti (Schmitt e Flechais, 2024). Le azioni di *phishing* sono solo uno degli esempi di uso criminale dell’IA. L. Floridi (2022) ha analizzato il ruolo dell’IA in diverse aree illegali che sono o possono essere interessate dai crimini dell’IA: reati che potrebbero non verificarsi se non grazie all’uso dell’IA. Da un lato, questi crimini possono interessare il funzionamento di sistemi complessi come il mercato finanziario attraverso azioni di distorsione delle informazioni o fenomeni criminali come il contrabbando di sostanze illecite realizzato con veicoli a guida autonoma. Dall’altro lato, essi possono avere come bersaglio singoli individui, nel caso dei reati contro la persona che possono comprendere molestie, furti d’identità o persino forme di tortura psicologica quando l’IA è impiegata nella conduzione di interrogatori (*ibidem*). Non si può non accennare alle minacce di danni fisici e letali alle persone causati dall’uso dei sistemi d’arma autonomi, tecnologie in grado di identificare, selezionare e attaccare il bersaglio senza l’intervento umano (Blanchard e Taddeo, 2024). Sebbene l’autonomia di tali sistemi comporti evidenti preoccupazioni circa il rispetto dei principi del diritto internazionale di proporzionalità, discriminazione e necessità (*ibidem*), l’uso dell’IA in contesti di guerra si sta diffondendo rapidamente. Un’inchiesta del giornalista israeliano Yuval Abraham²⁰, pubblicata da numerose testate giornalistiche, ha documentato l’utilizzo a Gaza di un sistema di identificazione automatica dei potenziali bersagli cui indirizzare i pesanti bombardamenti dell’esercito israeliano che, secondo dati ritenuti attendibili dalle UN, hanno condotto all’uccisione di oltre 45 mila palestinesi. Risulta evidente come tali strumenti pongano serissimi rischi pubblici per l’incolumità delle persone, tanto che tale questione è stata inserita nell’agenda istituzionale delle Nazioni Unite con l’approvazione nel dicembre 2024 della risoluzione 77/L. 79²¹.

L’ultimo ambito considerato riguarda il settore dell’informazione e le minacce di disinformazione e manipolazione dell’opinione pubblica anche a

²⁰ Il testo completo è disponibile al link: <https://www.972mag.com/lavender-ai-israeli-army-gaza/>

²¹ La risoluzione ONU è disponibile al link: <https://reachingcriticalwill.org/images/documents/Disarmament-fora/1com/1com24/resolutions/L77.pdf>

fini elettorali. Il complesso rapporto tra IA e informazione non può essere riassunto in poche righe ma è comunque opportuno accennare a due importanti fenomeni che da circa due decenni minacciano e alterano il mercato delle notizie e i consumi informativi: le bolle di filtraggio (*filter bubbles*) e camere d'eco (*echo chambers*); la disinformazione associata alle *fake news*.

Secondo E. Pariser (2011), giornalista e attivista, le *filter bubbles* definiscono un universo personalizzato di informazioni che alterano il modo con cui incontriamo notizie e idee che riproducono dunque contesti informativi chiusi, *echo chambers*, dove le proprie credenze vengono amplificate (Sunstein, 2017). In particolare, le *echochambers* sono ambienti comunicativi in cui gli individui ricevono in modo ripetuto informazioni finalizzate a rafforzare i loro preesistenti punti di vista in base alle preferenze informative dedotte dai comportamenti online (Jungherr, 2023). A causa di questa trasmissione e ritrasmissione, visioni e interpretazioni divergenti finiscono per essere omesse, mentre idee o credenze più o meno veritieri vengono amplificate (Treccani, 2017), non lasciando spazio per una pluralità di significati e per il cambiamento delle proprie opinioni da parte dei partecipanti alle interazioni. Gli utenti dei social media e di piattaforme di intrattenimento (non specialistiche sulla politica) risultano così aggregati in cluster omofili (Cinelli *et al.*, 2021), ad esempio *antiwax* con *antiwax*, *brexiters* con *brexiters*, con la conseguente polarizzazione delle opinioni politiche e delle preferenze elettorali. Le *filter bubbles* sono “bolle” in cui le convinzioni individuali sono rafforzate e non esposte a punti di vista opposti e oggetto di *confirmation bias* (König e Wenzelburger, 2020; Arsenault *et al.*, 2022; Bullock *et al.*, 2022; Coeckelberg, 2022; Duberry, 2022). Entrambe contribuiscono all’isolamento informativo degli individui, ma le *filter bubbles* si formano attraverso algoritmi che personalizzano i contenuti sulla base delle interazioni passate dell’utente in modo inconsapevole, mentre le *echochambers* si creano prevalentemente attraverso scelte consapevoli delle persone, che preferiscono interagire con gruppi e fonti di informazione che confermano le loro idee. Le bolle di filtraggio e le camere d’eco sono divenuti concetti largamente applicati nelle analisi sul funzionamento degli algoritmi di profilazione degli utenti, selezione e suggerimento dei contenuti informativi (es. Flaxman *et al.*, 2016; Rhodes, 2022; Figà Talamanca e Arfini, 2022). Essi hanno consentito di osservare e verificare, talvolta smentire, i fenomeni di isolamento informativo, che, quando presenti conducono alla radicalizzazione e polarizzazione delle opinioni politiche, due processi che minacciano la democrazia e la coesistenza civile (Rodiloso, 2024).

Particolarmente rilevante in tal senso è il fenomeno delle *fake news* online. Il problema delle notizie false, generate o meno con l’IA, è un fenomeno monitorato a livello globale da istituzioni, organizzazioni indipendenti e im-

prese attraverso diversi strumenti che ne tracciano l'andamento e le caratteristiche²². Ancora una volta quindi l'IA non deve essere intesa come un elemento esterno ai processi sociali preesistenti, quanto piuttosto un fattore che ne consente il potenziamento e perfezionamento. Si pensi che al 2024 il solo sistema di tracking della disinformazione di NewsGuard, impresa privata nel campo delle tecnologie, ha identificato a livello globale 1.150 siti di notizie inaffidabili generate dall'IA. I dati del Digital News Reports 2024 del Reuters Institute e dell'Università di Oxford evidenziano la crescente preoccupazione delle persone circa la veridicità delle notizie online (sentimento condiviso dal 59% dei rispondenti a livello globale e dal 72% negli Stati Uniti). Tali preoccupazioni sono amplificate dalla diffusione dei *deepfake*²³, la comunicazione di notizie false, diffuse in modo automatizzato su *social media*, sotto forma di immagini, video e audio falsi, di facile produzione attraverso algoritmi di IA allenati su grandi quantità di dati per simulare situazioni o persone in modo artificiale. Per gli utenti può essere difficile distinguerli da quelli creati dall'uomo e di queste comunicazioni, particolarmente efficaci nel produrre disinformazione, raramente si può rintracciare la fonte (EPRS, 2023). Il termine “*deepfake*” è un neologismo che deriva dalla combinazione delle parole “*deep learning*” (un ramo dell'IA) e “*fake*” (falso). I *deepfake* sempre più impiegati nelle competizioni elettorali o nella narrazione dei conflitti bellici, tipi di contenuti a cui le giovani generazioni pensano di essere regolarmente esposte (Reuters Institute, 2024). Allo stesso tempo, gli usi malevoli dei *deepfake* possono essere rivolti alla realizzazione di truffe sofisticate: in Italia nel 2024 la nota conduttrice televisiva Mara Venier ha denunciato all'Autorità Garante della Concorrenza e del Mercato (AGCM) l'utilizzo malevolo di un *deepfake* video in cui la sua replica digitale incoraggiava le persone a fare investimenti in *bitcoin* attraverso una finta intervista con E. Musk²⁴. Dunque, le immagini, gli audio e i filmati falsificati attraverso gli algoritmi possono rappresentare un potente strumento di manipolazione intenzionale dell'opinione pubblica (Gambin *et al.*, 2024), non solo a fini politici, che si aggiunge al lungo elenco delle minacce mediate dall'IA.

Riassumendo, lo sviluppo dell'IA porta potenzialmente a più estreme conseguenze le riflessioni sulla società del rischio teorizzata dal sociologo

²² Di particolare interesse è il *Database of Known Fakes*, strumento di *fact-checking* finanziato dalla Commissione Europea, basato anche sull'utilizzo di tecnologie di IA (progetto vera.ai).

²³ La tecnologia dei *deepfake* non ha necessariamente usi e scopi malevoli ma è uno strumento diffuso nel settore dell'intrattenimento, nell'educazione e nell'industria cinematografica (Westerlund, 2019).

²⁴ Per una ricostruzione dell'accaduto si rimanda a:

<https://www.open.online/2024/01/23/truffa-deepfake-venier-musk-trading-fc/>

U. Beck alla fine degli anni Ottanta del secolo scorso. Secondo Beck (2000) l’evoluzione del processo di modernizzazione ha implicato l’emergere di nuove situazioni e conflitti sociali connessi con la distribuzione di nuovi rischi, i rischi della modernizzazione, prodotti dall’industrializzazione e dal progresso tecnologico. Dunque, il processo di modernizzazione può implicare nuove forze distruttive che associano la produzione sociale di ricchezza alla produzione sociale di rischi (*ibidem*). I rischi della modernità analizzati da Beck si riferivano in particolare agli sviluppi dell’industria chimica e nucleare. Essi erano caratterizzati dall’essere spesso invisibili, negoziabili da processi sociali di definizione e trasversali nella loro distribuzione. Sebbene l’esposizione ai rischi sia da mettere in relazione con strutture sociali di potere e con le diseguaglianze, i rischi della modernizzazione travalicano i confini nazionali, recidono in parte il legame con il luogo di “produzione” e possono colpire anche chi ne è produttore. Lo schema interpretativo di Beck, dunque, regge alla prova del tempo e consente di interpretare anche la questione dei rischi dell’IA, il loro bilanciamento nel *trade-off* con le sue potenzialità positive, con particolare riferimento alle sfide per le diverse forme di azione istituzionale e alle questioni di *meta-governance*. In tal senso, facciamo nostre e riproponiamo le parole di F. Fischer (1998, p. 113): «(la società del rischio dell’IA è) intrappolata da un repertorio di risposte politiche obsolete e inadeguate alle catastrofi moderne [...], ci troviamo di fronte a un paradosso – nel momento stesso in cui le minacce e i pericoli sembrano diventare più pericolosi ed evidenti, essi scivolano contemporaneamente attraverso le reti di prove, responsabilità e risarcimenti con cui i sistemi giuridici e politici cercano di catturarli».

L’elenco dei rischi – che, come abbiamo detto, in questa sede è anche incompleto – definisce una ricca agenda pubblica di questioni che interessano diversi fenomeni politici e sociali: ampliamento o contrazione della partecipazione politica, espansione delle diseguaglianze tra gruppi sociali, riproduzione di forme di discriminazione. Da un lato, tali fenomeni coinvolgono una molteplicità di attori che sul *trade-off* tra potenzialità e rischi dell’IA propongono una “variegata” domanda sociale di *input* rivolti ai sistemi politici (Easton, 1957). Allo stesso tempo gli esperti, le organizzazioni non governative, le agenzie pubbliche e le stesse imprese del settore rivendicano la propria compartecipazione alla definizione delle risposte pubbliche in accordo con un assetto di governance multiscalare che “governa”, mentre usa, l’IA. Dall’altro lato, l’agenda del *trade-off* può essere soddisfatta senza il ricorso a “decisioni autoritative” ma attraverso soluzioni di autoregolazione delle imprese che hanno aderito a interessanti esercizi di riflessività circa le possibili conseguenze dell’IA. Ne è un esempio la firma, nel 2023, da parte di imprenditori e amministratori delegati di imprese leader nel settore

re dello “*statement*” lanciato nel 2023 dall’organizzazione no-profit *Center of AI Safety*: «mitigare il rischio di estinzione causato dall’IA dovrebbe essere una priorità globale, insieme ad altri rischi su scala sociale come le pandemie e la guerra nucleare»²⁵ (trad. nostra). Iniziative che rendono pubblica, più o meno strumentalmente, la consapevolezza degli impatti attesi o imprevisti del lavoro di sviluppatori, imprenditori e investitori nel settore, ma che lasciano aperta una domanda: chi e come se ne deve occupare?

1.2. L’intelligenza artificiale e i processi democratici: rischi e potenzialità

Il 13 giugno 2023 Eva Maydell, parlamentare europea del gruppo PPE, *rapporteur* della legge sull’IA ha sintetizzato così il senso degli oltre due anni di lavoro del Parlamento Europeo per regolare gli usi dell’IA: “quindi no, non è una questione tecnica. Si tratta del futuro della nostra democrazia”. Quest’affermazione richiede di esplorare le trasformazioni che l’uso di algoritmi può indurre, non solo per le istituzioni, rappresentative e non, della democrazia (Dent, 2020), ma anche per la stessa sovranità dei poteri democratici sulla società (Innerarity, 2022). Soprattutto nel momento in cui, conquistando prepotentemente posizioni di governo, la c.d. *Tech Right*²⁶, cerca di imporre cambiamenti epocali, che non si limitano alle agende di specifiche politiche, come quella ambientale o economica, ma riguardano anche le forme del potere politico e dei suoi rapporti con la società. In questi cambiamenti l’elevato potenziale trasformativo delle tecnologie digitali e dell’IA svolge un ruolo centrale, legittimando l’idea che siamo in un punto di svolta critico, uno snodo in cui saranno definite traiettorie future (Coeckelbergh e Saetra, 2023). Per questo, nelle agende di tutti gli attori coinvolti nel governo dell’IA sono entrati quei *trade-off* fra potenzialità e rischi che caratterizzano

²⁵ La dichiarazione è stata firmata da circa 350 ricercatori accademici, “altre figure di rilievo”, imprenditori e amministratori delegati di imprese leader nel settore. L’elenco completo dei firmatari è disponibile al link: <https://www.safe.ai/work/statement-on-ai-risk#open-letter>

²⁶ Si intende per *Tech Right* – denominazione coniata da Elon Musk – il connubio fra la destra radicale americana, che ha portato due volte alla Presidenza degli Usa il suo imprenditore-leader Donald Trump, e l’élite tecno-economica delle Big Tech della Silicon Valley. Il suo programma consiste in una miscela di populismo e sovranismo autoritario e di neo-liberismo radicale che testimonia di una nuova divaricazione fra il capitalismo ipertecnologico e liberal-democrazia: riduzione al minimo delle regolazioni pubbliche dell’economia e della società, lotta alla visione “woke” dei diritti civili, svuotamento delle istituzioni e delle procedure democratiche fondate sulla divisione dei poteri a favore della capacità di governo del leader designato dal popolo, assunzione di poteri politici da parte dei manager delle Big Tech e dei loro tecnici.

specificamente i processi politici democratici e le pratiche della democrazia. Ciò permette anche di accertare se e con quali modalità le istituzioni e i sistemi politici sono (o meno) in grado di trattare queste implicazioni dell'IA, sottoponendola a sua volta alle procedure e ai poteri della democrazia.

Come evidenziato da molti studi recenti della sociologia politica e della scienza politica, la democrazia liberale del mondo occidentale, le sue istituzioni, le sue pratiche e i suoi rapporti con la società sono ammalati, minati da inefficacia e delegittimazione. Le sue principali sindromi riguardano i processi e le forme di legittimazione sociale del potere politico. Prima, la disaffezione elettorale resa evidente dalla crescita del fenomeno dell'astensionismo. Seconda, la crisi dei partiti politici, sostituiti da forme di raccordo fra politica e società meno partecipate, non fondate sulla militanza, le ideologie, o la condivisione di interessi simili, ma su strumenti tecnici e apparati professionali a sostegno di candidati e liste. Terza sindrome, l'affermarsi di formazioni e candidati apertamente ostili alla democrazia stessa, che si pongono come esponenti di aspirazioni e interessi del “popolo”, non bisognosi di mediazioni e accordi fra interessi e valori sociali diversi. Di fronte a questo scenario, ci si può chiedere da un lato se e come l'IA possa aiutare ad affrontare alcuni elementi della crisi delle democrazie liberali e a curarne le sindromi; dall'altro, se e come l'IA possa somministrare alla democrazia dosi di ulteriori veleni, radicalizzandone la crisi. Come vedremo, entrambi i possibili decorsi e, in alcuni casi, la prescrizione di alcuni antidoti, sono già in molte agende, sia politiche, sia di ricerca scientifica.

In questo paragrafo sono presentate le interpretazioni proposte dagli studi sul rapporto tra IA e democrazia, che classifichiamo combinando due variabili applicate alla polarità degli immaginari del *trade-off* dell'IA, ossia le potenzialità, gli antidoti, e i rischi, i veleni, per i sistemi democratici²⁷. La prima variabile riguarda le modalità con cui si verificano gli effetti dell'IA sulle democrazie, che possono essere *diretti* – dovuti, cioè all'impiego di sistemi e modelli di IA in procedimenti e pratiche delle democrazie (elezioni, lavoro dei parlamenti e dei governi, pratiche partecipative), oppure *indiretti* – ossia le conseguenze sulla democrazia delle trasformazioni alimentate dall'IA nella società. La seconda variabile considera invece le diverse fasi e i momenti con cui questi effetti, positivi e negativi, si manifestano. Essi possono interessare, infatti, componenti diverse dei rapporti fra il sistema sociale e il sottosistema politico (Easton, 1965; Schmidt, 2010) e diverse fasi del ciclo di vita delle politiche pubbliche. In tal senso distinguiamo tra:

²⁷ Sugli immaginari dell'IA si veda il numero monografico di *Im@go. A Journal of the Social Imaginary* 23/2024 dedicato a *The Imaginary of Artificial Intelligence* (<https://cab.unime.it/journals/index.php/IMAGO/issue/view/385>).

- i flussi di *input* che la società indirizza verso i sistemi politici attraverso le richieste e il sostegno di vari tipi di attori sociali (come cittadini, imprese, associazioni). Questi comprendono i processi elettorali, le consultazioni condotte dalle istituzioni, la pressione degli interessi e le azioni delle lobby per strutturare le agende politiche o influenzare le decisioni politiche, ma anche varie forme di partecipazione popolare per esprimere richieste e pareri prima delle decisioni;
- i flussi di *output*, che vanno dalla politica verso la società, ossia le decisioni prodotte e il modo in cui vengono prodotte, il cosiddetto *throughput*, i processi all'interno dei sistemi politici nelle fasi della formazione delle agende, della decisione e dell'implementazione di politiche pubbliche²⁸.

Consideriamo in primo luogo gli effetti diretti dell'impiego di IA nei processi democratici e nelle pratiche di democrazia. A tutte le tecnologie digitali impiegate in processi e pratiche democratici – la comunicazione attraverso internet e i social media, le piattaforme e le applicazioni per analizzare dati e soppesare alternative decisionali – è già possibile applicare sistemi di IA. Pressoché tutti sono oggetto di immaginari sia di potenzialità, sia di rischio. Le *potenzialità* riguardano processi elettorali (*input*) e non elettorali (*throughput e output*) dei sistemi democratici. In primo luogo, le potenzialità dell'IA alimentano aspettative di miglioramento dei sistemi democratici riguardo alla fase di *input*, consentendo di prevedere effetti positivi sui processi elettorali e la partecipazione politica.

La prima promessa dell'IA riguarda sia l'aiuto agli elettori a prendere decisioni informate su politiche e candidati²⁹, sia il supporto alle campagne elettorali. Ne sono esempio le applicazioni di consulenza agli elettori che possono aiutare i cittadini a comprendere meglio la politica, fornendo conoscenze su *issue* specifiche e sulle posizioni di singoli attori politici (Kouroupis, 2023). Esse sono basate su algoritmi di profilazione politica e di *matching*, che forniscono suggerimenti a ciascun elettore in merito all'intenzione di voto migliore per lui combinando il suo profilo, desunto dalle preferenze dichiarate, come i *like e dislike* (EPRS, 2023), con informazioni sui candidati e i loro post sui social. Ad esempio, da tempo (2003) gli elettori svizzeri possono usare un algoritmo (*Smartvote*)³⁰ in occasione delle elezioni e dei frequenti referendum, un supporto che finora sembra avere influito nell'orientare le scelte sui candi-

²⁸ Due studiosi hanno cercato di capire gli effetti dell'integrazione di dati digitali nei processi politici dell'UE (*data driven form of governance*) distinguendo proprio fra *input*, *throughput e output* in una survey sui cittadini europei (Starke, Lünich 2022).

²⁹ Queste ed altre possibilità e limiti dell'uso politico dell'IA sono discussi nel blog "Democracy on Mars. Musings about AI, governance and philosophy" (<https://tantum.stack.com>).

³⁰ Fonte: <https://www.smartvote.ch/en/home>

dati, più che sui partiti (Benesch *et al.*, 2023). L'IA può anche sostituire virtualmente i candidati nel rispondere ai sondaggi loro rivolti dalle stesse applicazioni di consulenza agli elettori (Mancera Andrade J.A. e Terán L., 2024). Si possono inoltre immaginare sistemi di IA che votino per conto degli individui, gemelli digitali degli elettori addestrati attraverso le loro preferenze sociali, economiche e politiche, inferite ascoltandoli parlare e osservando le loro azioni *online* e *offline*. Per essere affidabile potrebbe essere progettato da un'entità pubblica e l'elettore dovrebbe poter sempre correggere il voto del suo gemello digitale, quando cambia opinione. I governi potrebbero sondare l'opinione dei cittadini sulle *policy* sondando i loro gemelli (Duberry, 2021). Gli effetti non sono certi: da un lato potrebbero elevare i tassi di partecipazione al voto, dall'altro dissuadere i cittadini dall'impegno nel capire la politica e dalla partecipazione democratica (Schneier, 2023).

Un processo di particolare interesse è anche il supporto dell'IA alle campagne elettorali, che può avvenire attraverso strategie di *micro-targeting*³¹ individuale. Sui dati degli elettori vengono realizzate pratiche note come *opinion mining*, *sentiment analysis* (analisi semantica ed emozionale dei contenuti), *digital listening* e profilazione psicométrica di specifici tratti di personalità, ad esempio attraverso i *like* apposti dall'utente di social media. Una volta capito come gli utenti percepiscono specifiche questioni pubbliche, gli analisti e i consulenti politici possono condurre campagne di comunicazione e costruire “brand” mirati di candidato (Duberry, 2022). Ricevendo aggiornamenti costanti che riguardano i temi che l'algoritmo riconosce come i loro preferiti (Viehoff, 2022), i singoli elettori possono diventare destinatari di comunicazioni che li indirizzano verso un partito, o candidato. A loro volta esse possono essere inviate da *political bots*³² automatizzati, account spam che pubblicano autonomamente, utilizzando script pre-programmati (Howard *et al.*, 2018). I loro effetti comunicativi sono più potenti se uniti a *micro-targeting* e amplificati dalla centralizzazione del potere sulla generazione e moderazione dei contenuti in un piccolo numero di grandi piattaforme.

Le promesse dell'IA interessano anche altri flussi di *input* che dalla società arrivano ai sistemi politici, come le pratiche di democrazia partecipa-

³¹ Il *micro-targeting* consiste nell'indirizzare messaggi personalizzati ai singoli elettori, identificandone il profilo attraverso le piattaforme di social media come Instagram o TikTok, analizzando Big Data (comportamentali, emozionali, con) e tracciando le conversazioni online relative a specifici argomenti.

³² Un bot è un programma automatizzato, programmato per svolgere determinate azioni in maniera regolare o reattiva e adattare i suoi messaggi anche senza intervento umano simulano di essere utenti umani sui social network, in modo tale da essere difficili da individuare. Gli altri utenti finiscono per confondersi e danno per scontato di trovarsi di fronte a utenti umani. In questo modo possono influenzare le discussioni pubbliche alimentandole di credenze, o anche diffondendo, il più delle volte con effetto di rinforzo, dei punti di vista già condivisi dai singoli utenti.

tiva e deliberativa, basate sullo scambio informato di opinioni e sul confronto dialogico (Moini e Nupieri, 2024) e le forme di pressione degli interessi. Le tecnologie digitali, i Big Data e l'IA consentono in particolare di sviluppare e potenziare le forme di partecipazione online, già da tempo note come *e-participation* (Medaglia, 2012; Wirtz *et al.*, 2018; Starke, 2020), aiutando, se non sostituendo, il ruolo dei facilitatori umani. In particolare, nelle piattaforme online in cui si formano ed elaborano opinioni sui temi oggetto di decisione pubblica durante consultazioni, l'IA generativa permette di capire più facilmente le opinioni degli altri attraverso sintesi discorsive di intuizioni e “intelligenza collettiva”. Facilitando l'*input* attraverso *citizen-sourcing*³³ e la comprensione e la validazione delle preferenze (Hallin e Lipka, 2023), l'IA può ridurre la complessità quantitativa e discorsiva delle informazioni (Cortez, 2023) e costruire ponti tra persone in disaccordo, aiutandole a raggiungere il consenso (Schneier *et al.*, 2023). Ad esempio, raggruppando automaticamente i commenti pubblicati in *cluster*, per votare poi su queste alternative (Duberry, 2022). O, al contrario, sostituendo le votazioni con conversazioni dense su scala ampia³⁴.

Le potenzialità dei modelli di IA generativa possono essere dispiegate anche nel rendere più efficace la relazione tra imprese, organizzazioni di rappresentanza degli interessi e decisori politici. In altre parole, l'IA può aiutare le pratiche di *lobbying* sintetizzando proposte di nuova legislazione, valutando se essa è rilevante per un'impresa e persino redigendo lettere a esponenti di governi e parlamenti per chiedere modifiche normative (EPRS, 2023). Inoltre, i lobbisti possono utilizzare l'IA per prevedere la probabilità di successo delle proposte da presentare ai legislatori (Jungherr, 2023). La cosiddetta “microlegislazione IA” indica piccole proposte di legge che soddisfano interessi ristretti. Gli algoritmi sono in grado di prevedere il probabile destino degli emendamenti legislativi proposti, nonché i percorsi attraverso i quali i lobbisti possono ottenere con maggiore efficacia i risultati desiderati. A questo fine l'IA può individuare piccole modifiche da apportare a un disegno di legge o a una legge esistente e che avrebbero il massimo impatto su un interesse sociale particolare. L'IA può creare milioni, persino miliardi di possibili emendamenti e identificare i pochi potenzialmente più vantaggiosi, producendo un sistema automatico per generare microlegislazione redditizia per gli interessi promossi. La produzione legislativa “omnibus” (leggi contenenti moltissime norme e approvate tutte insieme o con

³³ L'espressione, per derivazione dal crowdsourcing, identifica pratiche di coinvolgimento di cittadini nella progettazione e nello svolgimento di attività di interesse pubblico e la valorizzazione della loro intelligenza collettiva per conoscere i problemi e trovare soluzioni, fonte: https://en.wikipedia.org/wiki/Citizen_sourcing

³⁴ Fonte: <https://tantum.substack.com>

scadenze temporali molto strette) rende questi sistemi più facili da realizzare e usare (Sanders e Schneier, 2023).

Peraltro, così come gli attori più influenti si affidano spesso a professionisti per sviluppare i loro interventi politici, l'IA può fornire supporto discorsivo alle persone “qualsiasi” (Schneier *et al.*, 2023). In questo senso L'IA potrebbe anche essere impiegata per contrastare le tendenze populiste, promuovendo al contempo la legittimità democratica (Cavaliere e Romeo, 2022).

Per quanto riguarda gli *output*, le potenzialità di calcolo, elaborazione e simulazione dei sistemi di IA offrono ai decisori pubblici nuove risorse per ampliare la razionalità su cui si basano i processi decisionali e la formulazione delle politiche pubbliche. In tal senso, l'IA potrebbe fornire un efficace supporto alle decisioni e al miglioramento dei rapporti fra istituzioni e destinatari delle decisioni. L'IA potrebbe aiutare le democrazie non sostituendo gli umani nel prendere decisioni, ma rendendoli più capaci di gestire le informazioni, configurando una “*AI-augmented democracy*” (Coeckelbergh e Saetra, 2023). I primi studi su questo tema sostengono che risolvere compiti cognitivi e di coordinamento nel *throughput* con l'aiuto di IA può contribuire a incrementare la trasparenza (Starke e Lünich, 2020), la rapidità, reattività (Paltiel, 2023), la accountability (König e Wenzelburger, 2020) e l'efficacia (EPRS, 2023) delle azioni pubbliche. Le applicazioni di IA possono essere usate per delineare azioni pubbliche o per simularne gli effetti. Come nei contesti organizzativi del business, dove l'IA aiuta il management attraverso raccomandazioni e decisioni basate sull'elaborazione di un immenso *input* di informazioni (König e Burger, 2020), l'IA può fornire ai *decision maker* elaborazioni degli *input* derivanti dalle opinioni e dalle preferenze dei cittadini e dalle prese di posizione di attori collettivi sociali ed economici sui temi presenti nelle agende politiche.

I parlamentari possono poi usare l'IA per scrivere e modificare disegni di legge, per scrivere interventi con l'IA generativa e per comunicare con il pubblico (Schneier, 2023). Si pensi che il Parlamento italiano utilizza già sistemi di IA per produrre ed elaborare i dati per la pubblicità dei suoi lavori, per la classificazione automatica dei documenti parlamentari, con sistemi di riconoscimento vocale automatico per i resoconti di seduta³⁵, con il supporto alla stesura dei testi legislativi, all'analisi degli emendamenti e dei riferimenti normativi e alla verifica della *compliance* dei nuovi testi rispetto al quadro normativo preesistente. Inoltre, è in corso di sviluppo una *chatbot* per fornire risposte ai cittadini sulle attività dei singoli parlamentari³⁶. Tuttavia,

³⁵ Fonte: <https://comunicazione.camera.it/eventi/g7verona/intelligenza-artificiale>

³⁶ Per maggiori informazioni si rimanda a <https://www.altalex.com/documents/news/2024/09/16/intelligenza-artificiale-in-parlamento>

è stato dimostrato come le comunicazioni con gli elettori automatizzate con IA risultino più efficaci se sottoposte a controllo umano, mentre quelle più formulate dalla sola IA rischiano di generare sfiducia nelle istituzioni rappresentative (Kreps e Jakesch, 2023).

Riassumendo, se agenti umani e agenti-macchina prendono decisioni insieme perché l'IA aggiunge modelli analitici e sintetici in grado non solo di memorizzare e recuperare informazioni, ma anche di definire i problemi, valutare le soluzioni e prendere decisioni, vengono ridisegnati i confini della razionalità limitata, assunta per decenni come un presupposto (Simon, 2001) nell'analisi dell'azione politica e del comportamento amministrativo (Boix, 2022). Queste qualità migliorate, possono arricchire anche le interazioni fra cittadini e istituzioni e la loro legittimazione attraverso il *throughput* e l'*output* (Schmidt, 2010).

Finora abbiamo ricostruito i possibili effetti diretti e positivi dell'applicazione dell'IA alle pratiche dei sistemi democratici. Tuttavia, i cambiamenti sociali indotti dall'IA possono riflettersi negativamente sulle istituzioni democratiche e sulle conquiste politico-sociali su cui si basano le democrazie liberali anche attraverso effetti indiretti. La preoccupazione più radicale riguarda l'*output*, con lo svuotamento delle istituzioni democratiche e la concentrazione nelle mani di poche compagnie Big Tech³⁷ di potere non solo economico e cognitivo, ma anche politico, in modo opposto ai "pesi e contrappesi" delle democrazie moderne. I giganti tecnologici e i governi stessi potrebbero infatti delegare in misura crescente decisioni umane alle macchine, finendo per rimpiazzare la democrazia con forme di "governance algoritmica" o "algocrazia" (Landemore, 2023). Sono scenari lontani – ma da tenere d'occhio – in cui l'IA potrebbe sostituire gli umani nel coordinare e amministrare gli affari dello Stato e la regolazione. Prenderebbe così forma il futuro politico del neoliberismo tecnologico (Schröder, 2024), una variante radicale di tecnocrazia alimentata dall'IA (Coeckelbergh e Saetra, 2023).

Una seconda e più prossima preoccupazione riguarda l'*input*. I cambiamenti economici alimentati dall'IA che approfondiscono le disuguaglianze sociali (Inglehart e Norris, 2017), con la sostituzione di lavoro umano, la gestione automatizzata del ciclo della forza-lavoro (dal reperimento di manodopera all'assunzione, al controllo della produttività, al licenziamento), possono provocare l'ulteriore indebolimento delle già fragili tutele, della rappresentanza e dell'azione collettiva dei lavoratori dipendenti. Le forme sindacali di intermediazione, non solo fra lavoratori e datori di lavoro, ma anche

³⁷ Con questa espressione intendiamo le aziende più grandi e influenti, principalmente con sede negli Usa, protagoniste in settori tecnologici, tra cui l'IA, il cloud computing, i social media e l'e-commerce.

fra lavoratori e decisori politici, sono componenti storicamente importanti delle democrazie pluraliste. Al loro declino è associato anche quello della partecipazione politica attraverso la militanza e l'adesione a partiti o altre forme organizzate di impegno civile e cittadinanza e la stessa connessione tra elettori e rappresentanti (Raniolo, 2024). Inoltre, se l'IA produce nuove disuguaglianze fra chi è in grado di avvalersene e chi ne subisce solo gli effetti, possono ridursi le aspettative di mobilità sociale e in alcune fasce sociali può aumentare la percezione di privazione relativa, ingiustizia ed esclusione. Sinora, l'insieme di queste percezioni ha alimentato prevalentemente risposte politiche populiste (Levy, 2018), che delegittimano la separazione dei poteri in nome di un'autorità dei governi sottratta a controlli e contrappesi, fino alle cosiddette “democrazie illiberali”, caratterizzate dall'uso autoritario dei poteri statali e, in misura crescente, anche delle tecnologie digitali.

Queste preoccupazioni ci introducono all'altra faccia della medaglia relativa agli immaginari di rischio che definiscono in modo complementare la relazione tra IA e sistemi democratici. Essi riguardano sia *input*, sia *throughput* e *output* dei sistemi democratici.

Sul lato degli *input* i principali rischi, di tipo intenzionale e non intenzionale, come anticipato nell'introduzione, sono riferiti ai processi elettorali. Gli immaginari di rischio *intenzionali*, o minaccia, raffigurano l'IA come mezzo nelle mani di chi intende alimentare la polarizzazione politica, sia nel discorso pubblico, sia nelle scelte elettorali. Questo tipo di uso ha fatto configurare come ad alto rischio nell'EU AI Act i sistemi di IA che possono essere usati per influenzare il comportamento degli elettori in occasione di competizioni elettorali o di referendum (Malaschini, 2024).

Minacciare l'integrità di questi processi significa manipolare gli elettori (Coeckelbergh e Saetra, 2023) attraverso l'iper-personalizzazione automatizzata dei rapporti con la politica (Cupać *et al.*, 2024). Gli strumenti principali sono i *political bots* e il *microtargeting*. Come abbiamo visto precedentemente, le risultanti *echochambers* e *filter bubbles* producono effetti di isolamento informativo, poiché l'utente riconferma e proprie posizioni e di polarizzazione politica, fino a compromettere la coesione sociale (Helbing *et al.*, 2017). Con *bots* e *targeting* si può infatti combinare la generazione intenzionale di disinformazione basata sull'IA³⁸, usando *deepfake* che possono persuadere gli elettori a comportarsi in modo anche contrario ai loro interessi (Bueno de Mesquita *et al.*, 2024). Le vittime possono essere non solo gli

³⁸ La disinformazione è un'informazione falsa o fuorviante generata con l'intento esplicito di ingannare o manipolare un pubblico. Per un approfondimento si rimanda a (Standford University, 2024).

elettori, ma anche esponenti politici, mentre questi strumenti possono essere usati sia da individui, sia da “*truppe-cyber*”, o da imprese commerciali “ombra” (già dozzine nel 2020³⁹), come nei casi di campagne di influenza straniera sulle elezioni. Bombardare gli utenti con contenuti generati dall’IA può finire per minacciare la fiducia nei media stessi, oltre che nei processi e nelle istituzioni della democrazia (Schoch *et al.*, 2022).

Le azioni di disinformazione alimentate da IA sono considerate dal *World Economic Forum* (2024)⁴⁰ il maggiore rischio nel breve periodo. Non a caso, insieme a quelli per il lavoro, hanno aggiunto urgenza all’ingresso dell’IA in agende di regolazione (Ulnicane *et al.*, 2022). La disinformazione mediata dall’IA non è uno scenario futuro lontano, piuttosto è già stata mostrata la sua azione distorsiva e manipolativa nei sistemi democratici più consolidati. Nelle elezioni presidenziali Usa del 2016 sono state rilevate molte iniziative finalizzate a orientare e polarizzare l’elettorato attraverso falsi messaggi sui social media, mirati, generati da IA e diffusi da falsi account di sostenitori di Trump gestiti da soggetti cinesi, russi e iraniani che promuovevano teorie cospirazioniste⁴¹. Tuttavia, anche se i tentativi sono stati molti – ad esempio da parte di partiti di estrema destra francesi⁴², non ci sarebbero prove che disinformazione e *deepfake* abbiano avuto un impatto significativo sui risultati delle elezioni in UK, Francia e UE nel 2024. Uno studio dell’Alan Turing Institute ha rilevato che nelle elezioni in UK e del Parlamento europeo del 2024 la disinformazione abilitata dall’IA c’è stata e con modalità simili alle tendenze osservate in altre recenti elezioni, ma ha avuto un impatto minore di quanto temuto e non ci sono prove che il contenuto dell’IA abbia influenzato i risultati elettorali. Per questo il rischio per la democrazia dovuto all’uso malevolo di IA sarebbe moderato (CETAS, 2024). L’analisi di 11 elezioni svolte nel 2023 ha portato a concludere che più che di una “apocalisse informativa” è corretto parlare di “inquinamento dell’ecosistema dell’informazione pubblica” (Labuz e Nehring, 2024), arginato a sua volta da sistemi automatici: nelle presidenziali Usa del 2024 Chat GPT si è rifiutata di generare immagini dei candidati 250.000 volte⁴³. La facilità con cui gli utenti possono creare contenuti sintetici sofisticati utilizzando l’IA

³⁹ Fonte:

<https://www.isdglobal.org/explainers/commercial-disinformation-product-service/>

⁴⁰ Fonte: https://www3.weforum.org/docs/WEF_The_Global_Risks_Report_2024.pdf

⁴¹ Fonte:

<https://www.nytimes.com/2024/01/09/business/media/election-disinformation-2024.html>

⁴² Fonte:

<https://www.politico.eu/newsletter/digital-bridge/finally-someone-used-gen-ai-in-a-western-election/>

⁴³ Fonte:

https://www.italian.tech/2024/11/09/news/elezioni_usa_2024_chatgpt_deep-fake_intelligenza_artificiale-423607877/

generativa ha creato ulteriori fonti di rischio, ma l'esposizione si è limitata a una minoranza di utenti con convinzioni politiche già allineate alle narrazioni ideologiche incorporate in tali contenuti (Benaich e Chalmers, 2024; Stockwell, 2024). In altri termini, sono confermati i risultati di studi che documentano una bassa esposizione degli elettori a contenuti falsi, che si concentra in una frangia ristretta con forti motivazioni a ricercare tali informazioni (Budack *et al.*, 2024).

Anche il ruolo delle piattaforme nel plasmare il dibattito pubblico e polarizzare le opinioni potrebbe essere sopravvalutato. Sia nei messaggi più tradizionali, sia *online*, infatti, specialmente gli utenti meno informati e critici cercano le informazioni più affini alle proprie posizioni e ignorano quelle in contrasto con esse. Nelle piattaforme è difficile che qualcuno riesca a far cambiare idea a un utente (Quattrociocchi, 2024).

Il target di disinformazione e *deepfake* possono essere i politici stessi, che possono non riuscire a distinguere fra comunicazioni ricevute da esseri umani, o da IA generativa. Potendo generare facilmente migliaia di comunicazioni orientate a senso unico, un attore malintenzionato può distorcere le percezioni dei decisori politici su quali questioni sono più importanti per gli elettori e sui loro sentimenti in proposito. Si tratta di “astroturfing politico”, ossia campagne di disinformazione coordinate in cui l'emittente finge di essere dei normali cittadini che agiscono in modo indipendente l'uno dall'altro.

In sistemi politici autocratici l'uso strumentale della tecnologia e dei dati necessari per indirizzare la competizione politica può essere più facile a causa del maggiore controllo centralizzato dei dati (Jungherr, 2023). Tuttavia, anche nei sistemi democratici la moderazione del discorso da parte dell'IA può riuscire a plasmare il discorso pubblico (Elkin-Koren e Perel, 2023). Questo attribuisce a chi controlla le piattaforme e fa grande uso di IA grande potere di influenza. Questo è un esempio della possibile inefficacia di una disciplina dell'uso politico di IA affidata alla sola autoregolazione da parte delle piattaforme. Un altro è la repentina decisione di Meta (Facebook, Instagram, WhatsApp) di sostituire i dispositivi indipendenti di controllo delle *fake news* adottati volontariamente, anche per prevenire eventuali regolamentazioni istituzionali, con segnalazioni degli utenti per allineare la linea aziendale al nuovo corso politico e al “discorso dominante” dopo l'elezione di D. Trump.

Vi sono inoltre rischi *non intenzionali* derivanti dall'uso di IA nei processi elettorali. In primo luogo, nei cittadini l'idea stessa di azione politica collettiva può essere sostituita dall'accettazione di essere semplicemente destinatari di meccanismi di distribuzione delle informazioni (Viehoff, 2022). Inoltre, è possibile un “rischio epistemico”, ossia l'erosione della capacità degli individui di formare e rivedere autonomamente le proprie credenze politiche. La dipendenza dagli *input* dei social media mina la *agency* politica delle persone (Coeckelbergh, 2023). D'altra parte, anche partiti e candidati

sono indotti a conformarsi alle preferenze degli elettori così come rilevate o, a loro volta, precedentemente indotte attraverso l'uso di IA. Una simile riduzione dell'autonomia umana può essere dannosa per la democrazia, che richiede invece deliberazione e scelta autonoma (Corsi e d'Albergo, 2024). L'esito può essere aggravato dai *bias* imposti all'IA generativa dai *dataset* usati per l'addestramento e apprendimento, che possono essere plasmati dalle disuguaglianze e discriminazioni esistenti (Jungherr, 2023).

Rischi intenzionali e non riguardano anche l'uso di IA in varie forme di partecipazione pubblica. Nelle consultazioni è possibile utilizzare l'IA generativa per inondare le piattaforme istituzionali di messaggi artificiali a vantaggio di interessi specifici (Kreps e Kriner, 2023). Nelle pratiche di democrazia partecipativa e deliberativa vi è invece un rischio di opacità se i partecipanti non hanno consapevolezza critica circa il funzionamento di sistemi di IA modellati da esperti e fornitori privati, su chi li sviluppa e li gestisce, sulla trasparenza e la responsabilità del trattamento dei dati e sulla *cybersecurity* (Duberry, 2022). La partecipazione sostenuta da sistemi di IA potrebbe avere addirittura l'effetto di spostare il potere dalla sfera sociale a quella degli attori politici, tecnici ed economici, a partire dai fornitori di algoritmi (d'Albergo *et al.*, 2024). L'uso di IA potrebbe quindi contribuire allo svuotamento sociale di queste pratiche in cui le tecnologie vengono impiegate per dare l'impressione di inclusività senza un reale trasferimento di potere decisionale ai cittadini, riducendole a forme simboliche e inefficaci, ovvero forme di "tokenism" (Arnestein, 1969), se non a legittimare decisioni già prese (Timcke, 2024).

Se si guarda alle componenti di throughput e *output* possono essere identificati tre immaginari di rischi dell'IA producibili da comportamenti non intenzionali. Il primo riassume i rischi relativi alla perdita di controllo umano sui processi decisionali, che può avere molteplici implicazioni sull'accentuarsi di processi di depoliticizzazione, erosione dell'*agency* umana, opacità delle decisioni e sulle relazioni di potere tra pubblico e privato.

Il supporto cognitivo alle azioni pubbliche attraverso IA può favorirne la depoliticizzazione se le preferenze consapevoli ed esplicitate di cittadini e *policy maker* e la loro stessa facoltà di giudizio sono sostituite dalla standardizzazione dei processi cognitivi. La riduzione della *agency* epistemica umana (cfr. sopra) ha come corollario l'induzione alla delega nei confronti della tecnologia, ossia di chi è in grado di controllarla. Ciò può finire per attribuire potere politico ai fornitori di software e hardware (Saetra, 2020), oltre che a quegli esperti che si affidano a loro volta a modelli supportati da IA artificiale per produrre i loro consigli (Jungherr, 2023). I fornitori da un lato hanno le competenze per installare, mantenere e sviluppare sistemi di decisione o istruttoria automatica, dall'altro considerano gli algoritmi e l'IA forniti "i loro segreti commerciali più custoditi" (Paltiel, 2022). Possono

quindi crescere le asimmetrie di potere dovute al controllo di risorse informative fra chi le padroneggia e le elabora attraverso algoritmi – siano essi attori politici di governo, amministrativi o imprese private – e chi, come i cittadini, o la stessa opposizione politica, non sono capaci di raggiungerle e contrastarle (König e Wenzelburger, 2020; Rosso, 2023).

Rischio epistemico e potenziali deficit di controllo democratico e *accountability* dei sistemi di IA (Kouroupis, 2023) sono dunque problemi causati soprattutto dall’opacità: per gli utenti – i decisori di politiche pubbliche e i loro destinatari – può infatti essere difficile capire e controllare il design e la gestione di piattaforme e algoritmi, come siano processati i dati e la logica complessa e spesso imperscrutabile con cui l’IA produce previsioni, raccomandazioni o decisioni algoritmiche (Starke e Lünich, 2020; Arsenault e Kreps, 2022). Per questo è preoccupante l’uso di IA soprattutto nei processi decisionali che riguardano questioni su cui è importante il giudizio umano (Boix, 2022). Più in generale, l’opacità minerebbe le norme di responsabilità democratica, trasparenza (Arsenault *et al.*, 2022; Coeckelberg, 2022), accountability e legittimità delle decisioni. Soprattutto se funzionari amministrativi e attori politici accettassero in modo acritico decisioni automatiche affidate all’IA (Paltiel, 2022).

Il secondo immaginario riprende due dei principali rischi pubblici dell’IA ovvero i *bias* e le discriminazioni. Nel fare previsioni sulle domande di regolazione o di interventi pubblici provenienti dai cittadini l’IA può sopravvalutare il peso dei gruppi e degli interessi sociali che hanno maggiore visibilità a causa del loro attivismo, ignorando le richieste di chi si rende meno visibile (Jungherr, 2023). Nel sintetizzare discorsi e posizioni politiche l’IA può produrre una selezione delle posizioni statisticamente mediane, favorendo le posizioni più comuni, rendendo invisibili quelle di minoranza (*ibidem*). Questa preoccupazione può assumere anche caratteri paradossali come nel caso delle rivendicazioni dei gruppi tecnologicamente attivi di estrema destra, che individuano un *bias* negativo dell’IA nei loro confronti dovuto alla congiura di un “monopolio ebraico”. Quando questi gruppi tentano di sfruttare l’IA per fini malevoli, i loro sforzi possono infatti essere ostacolati dalle restrizioni automatiche sull’uso dell’IA da parte di imprese come OpenAI (Koblenz-Stenzler e Klempner, 2024).

Un calcolo basato su dati del passato può anche rafforzare modelli preesistenti di disuguaglianza ed esclusione rilevati statisticamente⁴⁴. L’effetto

⁴⁴ Ad esempio, genitori appartenenti a minoranze povere nelle aree urbane possono essere più facilmente presi come target da software algoritmici di protezione dei bambini, o si può essere erroneamente etichettati come truffatori, con conseguenze gravi come è accaduto in Olanda, quando le autorità fiscali hanno utilizzato un algoritmo di autoapprendimento per creare profili di rischio nel tentativo di individuare le frodi sui sussidi per l’assistenza

di questi *bias* è ancora più significativo quando all'IA si delegano direttamente decisioni di *policy* (Duberry, 2022) che rischiano di replicare le condizioni sociali esistenti. Anche le rappresentazioni delle relazioni di potere esistenti nella società possono essere percepite dall'IA come verità oggettive e, una volta consolidate, diventare senso comune, naturalizzate e non discutibili (Lindgren, 2023).

Il terzo immaginario dei rischi non intenzionali riguarda il controllo dei dati e dunque la violazione della privacy. L'uso che l'IA fa dei Big Data, sia per fini di consenso politico, sia per svolgere funzioni pubbliche in modo più rapido ed efficace può portare da un lato a violazioni sistematiche del diritto alla privacy e al controllo dei propri dati, dall'altro a una strana forma di partecipazione, in cui i cittadini influenzano la politica fornendo informazioni, ma senza l'intenzione di farlo, né sapendolo, perché il consenso per l'uso è stato dato, spesso superficialmente, per finalità generiche. E ciò nonostante che in alcune strategie nazionali per l'IA i dati vengano definiti un “bene comune”, il che dovrebbe comportare che il loro uso dipende da una decisione collettiva (Paltiel, 2023).

A causa di comportamenti intenzionali e non le applicazioni di IA, in particolare generativa, rischiano quindi non solo di favorire in modo asimmetrico attori, fazioni politiche o interessi sociali ristretti, ma anche di amplificare alcune preesistenti fragilità dei sistemi democratici. Nei casi peggiori, si possono porre le basi per forme di “autoritarismo digitale” (Pearson, 2024), particolarmente adatto a sistemi autoromatici, i cui leader possono trarre maggiori vantaggi rispetto quelli che operano in sistemi democratici (Junguerr, 2023). In questi ultimi, oltretutto, può ridursi la già calante fiducia dei cittadini nelle istituzioni e la legittimazione delle pratiche democratiche, con effetti dirompenti sull'efficacia e la stessa stabilità della democrazia.

1.3. Le poste in gioco economiche e politiche

Gli immaginari e le narrazioni che caratterizzano i discorsi sullo sviluppo dell'IA alimentano aspettative di enormi potenzialità di trasformazione, oltre che di rischio, da mettere a frutto in quasi tutti i settori di attività sociali, economiche e di ricerca scientifica. Raramente però questi vantaggi sono distribuiti equamente in termini sociali e geografici. Per questo, lo sviluppo dell'IA e le conseguenze dei suoi usi sono delle poste in gioco nelle relazioni economiche e politiche e nei rapporti fra interessi e poteri economici e politici.

all'infanzia; fonte: <https://www.politico.eu/article/dutch-scandal-serves-as-a-warning-for-europe-over-risks-of-using-algorithms/>

Nella teoria economica e nella teoria dei giochi il concetto di posta in gioco viene utilizzato per indicare ciò che gli attori cercano di massimizzare, soprattutto in termini materiali, attraverso comportamenti strategici nei mercati. In ambito sociologico P. Bourdieu (1972; 1986) ha definito le poste in gioco (*enjeux*) come le risorse materiali e simboliche contese all'interno di un determinato spazio sociale, ciò che gli attori cercano di acquisire o mantenere, gli interessi per i quali gli agenti si mobilitano e intessono rapporti di potere. Questo concetto molto ampio consente di analizzare i motivi per i quali, in svariati campi di azione sociale, ciascun attore si impegna e investe le risorse che controlla, orientando le proprie strategie per massimizzare le prospettive di guadagno, materiale, o simbolico. In questo modo possiamo individuare varie dimensioni sia della competizione economica nei mercati digitali che qui ci interessano, sia delle poste in gioco più specificamente politiche.

Nella sfera economica sono in gioco, direttamente o indirettamente, sia i rendimenti degli investimenti pubblici, sia i profitti delle imprese tecnologiche. Nella sfera politica l'IA può permettere a stati e governi di consolidare o espandere potenza, o influenza. Per farsi una prima idea di che si tratta possono essere utili alcuni esempi relativi ad eventi relativi allo sviluppo dell'IA verificatisi nelle prime settimane del 2025. Insieme, sono infatti rivelatori non solo delle poste in gioco in ciascuna delle due sfere, ma anche della stretta connessione e interdipendenza di queste ultime.

Il primo evento: il 10 gennaio l'azienda cinese DeepSeek ha reso disponibile una *chatbot* basata sul proprio modello di IA generativa DeepSeek-R1, che ha avuto un immediato successo di pubblico con milioni di download, suscitando però preoccupazioni fra i produttori e fra gli investitori in IA negli Usa. Infatti, questo modello – rilasciato in *open-source*⁴⁵, quindi ispezionabile, replicabile e modificabile – sembra offrire prestazioni potenti come ChatGpt di OpenAI, soprattutto in ambito matematico e di programmazione informatica, ma i suoi costi di addestramento (ammonterebbero a 6 milioni di dollari contro i 100 di Gpt-4) e i suoi tempi di calcolo sono molto inferiori. Inoltre, utilizza solo 2.000 *chip* Nvidia – semiconduttori indispensabili per sviluppare i modelli – contro i 16.000 degli altri modelli simili di IA. Il potenziale ingresso in campo di un simile rivale per le Big Tech Usa è stato interpretato come un “grande schiaffo in faccia”, che minaccia le prospettive di mercato delle imprese di IA americane ed europee e dei loro

⁴⁵ Con questa espressione si intende un modello di sviluppo dell'IA in cui il codice sorgente, i dati di addestramento e le architetture dei modelli sono accessibili, modificabili e condivisibili liberamente da chiunque. Questo consente a chiunque, comprese imprese diverse da quella che lo ha prodotto, di studiare come funziona il sistema e apportare miglioramenti.

grandi centri di calcolo⁴⁶. Il governo Usa ha ipotizzato che DeepSeek abbia “distillato” parti dei modelli della stessa OpenAI⁴⁷, mentre il suo uso è stato vietato da vari Paesi – fra cui l’Italia – che hanno espresso preoccupazioni sulle sue pratiche etiche, di privacy e di sicurezza. Il Congresso degli Usa ha valutato l’opzione di impedirne l’uso in ambienti amministrativi per motivi di sicurezza internazionale, sulla falsariga di quanto già avvenuto per Tik-Tok. L’aspetto più interessante è che nella discussione sulle poste in gioco che ne è scaturita fra attori di governo, esponenti delle Big Tech ed esperti, emerge una sostanziale identificazione fra gli interessi dell’industria americana dell’IA e quelli della nazione. Insieme, sono interpretati come un sistema interdipendente di economia e politica posizionato nello scenario competitivo globale. Nel caso di DeepSeek, devono affrontare un concorrente che testimonia dei rapidi progressi dell’analogo sistema economico e politico dell’IA cinese, che attrae anche capitali esteri, come il fondo Prosperity7 della società petrolifera saudita Aramco, che ha recentemente investito 300 milioni di dollari nella società cinese Zhipu AI⁴⁸.

Il secondo evento: appena entrato in carica per il secondo mandato, il Presidente degli Usa D. Trump ha adottato l’*executive order Maintaining American Leadership in Artificial Intelligence* per rimuovere le “barriere regolative” al mercato imposte dalla precedente amministrazione democratica, basandosi sull’idea che, per mantenere la leadership statunitense, sia necessario sviluppare sistemi di IA “che siano liberi da pregiudizi ideologici o da agende sociali artificiali”⁴⁹. Con lo stesso fine, negli stessi giorni il Presidente degli Usa ha dichiarato di voler favorire 500 mld di dollari di investimenti privati in infrastrutture per l’IA (cfr. sotto). Precedenti presidenti degli Usa (Biden e lo stesso Trump nel mandato 2016-2020) avevano più volte dichiarato l’aspirazione alla leadership mondiale nell’IA. Ma gli Usa non sono soli: dal 2017, attraverso massicci investimenti pubblici e privati nell’innovazione (ISPI, 2024), la Cina si sta sforzando esplicitamente di divenire leader globale nell’IA entro il 2030, oltre che una forza trainante nella definizione di norme e standard etici per l’IA (Huw *et al.*, 2020). A loro si aggiunge l’UE, che si è candidata ad essere leader mondiale nell’etica dell’IA, ma è comunemente ritenuta indietro nello sviluppo competitivo di conoscenze e sistemi

⁴⁶ Fonte: <https://www.nytimes.com/2025/01/27/business/us-stock-market-deepseek-ai-sp500-nvidia.html>

⁴⁷ Fonte: <https://archive.is/D9whR#selection-2447.1-2447.151>

⁴⁸ Fonte: <https://computerweekly.it/2024/06/04/larabia-saudita-investe-in-cina-per-fare-concorrenza-a-chatgpt/>

⁴⁹ Il testo integrale è disponibile al link: <https://www.federalregister.gov/documents/2025/01/31/2025-02172/removing-barriers-to-american-leadership-in-artificial-intelligence>

di IA rispetto a Usa e Cina, con conseguenze che ambienti accademici e d'impresa ritengono non solo tecnologiche: “se tutto viene inventato dai cinesi e dagli americani, l'Europa continuerà a ristagnare economicamente e il ristagno economico genera una recessione geopolitica. Di conseguenza, se l'Europa è irrilevante nel campo dell'IA, sarà sostanzialmente meno rilevante anche dal punto di vista geopolitico”⁵⁰.

In altri termini, è in corso una competizione politica ed economica insieme per essere il sistema nazionale – o formato da più Paesi nel caso dell'UE – che nel mondo esercita più potere nell'IA. Questa competizione si svolge in due campi di gioco distinti, ma fra loro interdipendenti, in cui opera la politica: l'innovazione tecnologica e il campo delle strategie di sviluppo e regolazioni istituzionali. Nel primo gli investimenti e i saperi sono mobilitati soprattutto da attori e capitali privati, ma la politica cerca di fornire loro ambienti regolativi e infrastrutturali favorevoli. Nel secondo i protagonisti politici negoziano o concordano le scelte di sviluppo e regolazione con quelli economici. La strutturazione dei campi di gioco, attraverso scelte che riguardano la *metagovernance* dell'IA, è a sua volta una posta in gioco, perché queste scelte possono favorire scenari favorevoli a valori (ad esempio, i diritti umani) o interessi (ad esempio la profittabilità degli investimenti) diversi. I giocatori appartengono però a una cerchia tutto sommato ristretta di élite economiche e politiche che, con modalità approfondite in altri capitoli, alternano strategie di collaborazione, competizione e rivalità, tentativi di esercitare influenza, alleanze e rotture. Anche la società civile no-profit e l'accademia cercano di partecipare ma, come vedremo, hanno generalmente minori risorse per imporre la loro presenza e, soprattutto, per influenzare lo svolgimento delle partite.

Vediamo ora i principali aspetti delle poste in gioco economiche e politiche, come gli attori le traducono in strategie per aggiudicarsene e le asimmetrie che caratterizzano le rispettive capacità di influenzare i processi nel campo economico, nelle arene politiche e nelle loro intersezioni.

Iniziamo dalle poste in gioco economiche. Il business globale dell'IA è enorme, dinamico e in rapida crescita. In questo mercato le poste in gioco riguardano il primato nella ricerca, nello sviluppo e, quindi, nella commercializzazione di una varietà di prodotti, con i relativi profitti. Le applicazioni di IA interessano quasi tutti i settori di attività economiche e sociali – comprese quelle delle amministrazioni pubbliche – e vengono impiegate con il fine di incrementare efficienza e innovazione, anche dei modelli organizza-

⁵⁰ Intelligenza artificiale, come si muove l'Europa. Ross: “Pensa solo alle regole, va verso l'irrilevanza”, Quotidiano Nazionale, 29 01 2025;
<https://www.quotidiano.net/economia/intelligenza-artificiale-europa-deepseek-lbakil2z>

tivi e di business. Secondo la piattaforma globale di dati Statista, le dimensioni del mercato dell'IA dovrebbero avere un tasso di crescita annuale del 27,67%, raggiungendo un volume di mercato di 826,70 miliardi di dollari entro il 2030⁵¹. Secondo ISPI (2024) il mercato globale dell'IA dovrebbe avere un tasso di crescita annuo composto (il tasso di rendimento dell'investimento nel periodo) del 37,3% dal 2023 al 2030. Goldman Sachs (2023) prevede che gli investimenti nell'IA potranno incrementare il PIL globale di circa 7.000 miliardi di dollari nei prossimi dieci anni, con un aumento del 7%. Una crescita misurata e stimata in modi diversi, ma dovuta in primo luogo all'incremento della produttività lavorativa, perché l'IA agisce come "acceleratore" della logica produttivista "dell'attuale capitalismo neoliberista tecnologico e tecnocratico" (Puglia e Lamanna, 2024).

Di seguito sono esplorate alcune modalità di azione delle Big Tech, protagoniste indiscusse del business IA. Queste imprese monopolizzano investimenti e profitti, ispirano le narrazioni sulla digitalizzazione e l'IA, contribuendo a legittimarla nella discussione pubblica (Tafani, 2024), e condizionano l'azione dei governi attraverso le loro capacità di controllare dati, opinioni e comportamenti della popolazione e di incrementare la dipendenza degli stati dai vantaggi economici offerti dall'IA (Suleyman, 2023). Inoltre, questa influenza viene esercitata attraverso i milioni investiti, come le imprese più tradizionali, in attività di *lobbying*.

I costi per lo sviluppo dell'IA sono elevati, non solo per la ricerca e l'innovazione: ad esempio si stima che OpenAI, che nel febbraio 2025 aveva in corso una raccolta di finanziamenti che può portarla al valore di 300 miliardi di dollari, ne spende 700.000 al giorno per i costi d'esercizio della sua IA generativa (Puglia e Lamanna, 2024). Per questo, nelle società di IA investono miliardi sia società di capitale di rischio, prevalentemente Usa, sia di *private equity* (investimento in imprese non quotate ad alto potenziale di sviluppo effettuato con l'obiettivo di ottenere consistenti guadagni dalla vendita della partecipazione acquisita o dalla quotazione in borsa). Dopo il lancio di ChatGPT nel 2022, la capitalizzazione di mercato di OpenAI ha raggiunto nel 2024 i 3.000 miliardi di dollari, la terza azienda negli Usa a raggiungere questo traguardo dopo Microsoft e Apple (Benaich e Chalmers, 2024). Per il fabbisogno dell'IA generativa serve una grande quantità di *chips* (semiconduttori, unità di elaborazione grafica - GPU) e soprattutto Nvidia (Usa) controlla il 90% di questo specifico mercato, raccogliendo i frutti dell'investimento in ricerca sull'IA che lega hardware e software (Benaich e Hogarth, 2022). Statunitensi sono anche i maggiori fornitori di servizi cloud, a loro volta indispensabili per l'IA, come Amazon Web Services (AWS), Google

⁵¹ Fonte: <https://www.statista.com/outlook/tmo/artificial-intelligence/worldwide>

Cloud Platform (GCP) e Microsoft Azure. Sul mercato dell'IA sono presenti con sempre maggiore importanza le cinesi Baidu, Tencent, Alibaba, SenseTime, 4Paradigm e Yitu Technology. Alcune aziende di IA sono *spin-out* universitari, ma è il settore privato a guidare la maggior parte degli sviluppi e degli investimenti nell'IA (UN, 2024), anche dominando la ricerca sull'IA su larga scala, con una produzione di articoli e paper in cui, secondo dati del 2022, prevalgono gli Usa (aumento rispetto al 2021 dell'11%), seguiti però da una Cina in forte crescita (+ 24%) e con la Tsinghua University prima produttrice individuale mondiale (Benaich e Hogarth, 2022).

Con gli investimenti e l'innovazione, crescono anche l'ampiezza e le interdipendenze spaziali dei mercati. Questo non riguarda solo gli utenti, o consumatori di IA. Anche il fabbisogno di investimenti dei laboratori più innovativi cresce più di quanto possa investire il solo capitale di rischio tradizionale e con base nazionale. Per questo le imprese di IA statunitensi ricevono proposte di investimento e acquisto di azioni dall'estero, come quelle del fondo sovrano di Abu Dhabi, degli Emirati Arabi e di fondi dell'Arabia Saudita, talvolta respinte per motivi di sicurezza nazionale (Benaich, 2024). Lo sviluppo dell'IA non è però distribuito in modo equilibrato all'interno dei sistemi economici, né sul piano spaziale e tale distribuzione è essa stessa una posta in gioco. Le aziende più grandi e tecnologicamente avanzate possono permettersi di beneficiare dell'adozione di IA più delle piccole. A sua volta, l'adozione disomogenea di IA tende a incrementare le disparità economiche e, indirettamente, di influenza politica tra i sistemi economici tecnologicamente avanzati e quelli dei Paesi "in via di sviluppo". Nei primi l'IA viene più facilmente associata nel discorso pubblico con il *trade-off* fra i benefici derivanti dal suo sviluppo e i rischi di varia natura, anche esistenziali, che ne possono derivare. La maggior parte di coloro che hanno poca IA – a partire dai Paesi africani – individuano piuttosto un rischio esistenziale nel suo mancato utilizzo (Cameron *et al.*, 2025) e nella conseguente dipendenza dai sistemi economici e politici più ricchi e forti. Questo divario fra sistemi economici si accompagna anche a fenomeni di stampo neocoloniale, come la concentrazione nei Paesi più deboli di attività malpagate e sfruttamento del lavoro per estrarre dati e per addestrare i sistemi di IA e strutturare l'hardware necessario (Hao, 2022). Da una prospettiva di analisi critica dei processi economici questi fenomeni sono concretualizzati come "assemblaggi planetari di colonialità" che rinforzano un'architettura di potere diseguale e stratificata in cui sono contrapposte élites, o centri ricchi e influenti (Usa e Cina controllano più del 90% dei mercati), e periferie digitali (Hung, 2024). Nella crescente concorrenza globale, al momento l'Asia orientale è il motore manifatturiero ma – come abbiamo visto con il caso di DeepSeek – sfida il Nord-America anche per la ricerca e l'innovazione nell'IA. India, UK e Giappone emergono come terzi contendenti (ISPI, 2024),

mentre l’Europa viene spesso considerata indietro sul piano degli investimenti e dell’innovazione (Aresu, 2024).

Le poste in gioco relative all’appropriazione di ricchezza in questi mercati possono portare a strategie e comportamenti sia di collaborazione, sia di competizione – ossia una rivalità per raggiungere lo stesso obiettivo, ma nel rispetto delle regole – fino al conflitto aperto. Alcuni esempi di cooperazione economica di imprese dell’IA statunitensi che sanno di non poter fare da sole: Meta acquista GPU da Nvidia; il partner di Alphabet sui *Tensor Processor Unit* entra in Meta (Aresu, 2024); Apple stabilisce una partnership con la cinese Alibaba per l’IA negli iPhone⁵². Tutt’altro che cooperative sono invece le strategie e i comportamenti che impediscono la concorrenza e l’ingresso nel mercato dell’IA. I fattori che alimentano rischi per la concorrenza sono le barriere all’entrata nel mercato costituite dalla necessità di grandi e costose quantità di *chip* specializzati, di servizi *cloud*, di grandi volumi di dati, di competenze avanzate di *data science*, quindi di copiosi investimenti⁵³. In particolare, il passaggio a modelli di IA di grandi dimensioni sta concentrando il potere nelle mani di grandi aziende tecnologiche che godono di un accesso preferenziale alle risorse necessarie per progettarli e produrli. Inoltre, non solo le grandi imprese IA, ma spesso anche le maggiori piattaforme digitali acquisiscono start-up di IA promettenti, utilizzando le proprie posizioni di vantaggio competitivo sui rispettivi mercati per espandere il loro potere verso il mercato specificamente legato all’IA (Norton Rose Fulbright, 2024).

La competizione può anche diventare conflitto aperto, con comportamenti e pratiche extraeconomiche che, nel caso dell’IA, utilizzano azioni che investono anche il campo politico. Ne è un esempio il conflitto fra E. Musk e OpenAI, da cui il primo aveva disinvestito nel 2018, ufficialmente a causa di divergenze sulle prospettive e i rischi insiti nello sviluppo dell’IA ma, secondo molti osservatori, anche per l’attenzione di OpenAI alla massimizzazione dei profitti per Microsoft.

In questo quadro si inserisce l’appello per la moratoria sviluppo dei sistemi di IA avanzati, proprio come ChatGPT, co-firmato nel 2023 da E. Musk e molti dirigenti e ingegneri di Big Tech come Apple e Google). In realtà l’obiettivo principale dell’appello sarebbe stato quello di indebolire la competitività di OpenAI e Microsoft, proprio attraverso l’introduzione di quella regolazione politica che la cultura imprenditoriale americana cerca

⁵² Fonte: <https://www.reuters.com/technology/apple-partners-with-alibaba-develop-ai-features-iphone-china-information-reports-2025-02-11/>

⁵³ Come è stato notato ad esempio dall’Autorità per la concorrenza francese (28 giugno 2024, <https://www.autoritedelaconcorrence.fr/en/press-release/generative-artificial-intelligence-autorite-issues-its-opinion-competitive>)

sempre di contrastare. Il conflitto fra questi imprenditori è proseguito nel 2025 dopo l’insediamento di D. Trump, anche se E. Musk e S. Altman erano nel frattempo divenuti entrambi suoi sostenitori, il primo anche come componente influente dello staff della Casa Bianca. L’oggetto è stato l’annuncio (21 gennaio 2025) di una *joint venture* americana sull’IA, sostenuta politicamente dal Presidente, relativa allo “Stargate Project”. Il progetto, creato da SoftBank, OpenAI, Oracle e MGX con Arm, Microsoft, NVIDIA, Oracle, e OpenAI come partner tecnologici chiave⁵⁴, ha l’obiettivo di realizzare infrastrutture di IA negli Usa entro il 2029. L’effettiva disponibilità delle ingenti risorse necessarie – sono previsti investimenti fino a 500 miliardi di dollari da parte di tutti i partecipanti – è stata messo pesantemente in dubbio da E. Musk ha. Pochi giorni dopo, lo stesso E. Musk ha lanciato un’offerta ostile di 97,5 miliardi di dollari per acquistare attraverso una società di *venture capital* con vari investitori una quota di controllo di OpenAI, in vista della sua prevista conversione da società no-profit a una a scopo di lucro. Questa strategia d’investimento è stata invece sostenuta esplicitamente dal Presidente D. Trump, da poco insediato⁵⁵, fornendo così un esempio di un ambito di sovrapposizione fra campo economico e campo politico, quando le poste in gioco dell’IA riguardano gli interessi economici e questi attori cercano di giocare la politica come “carta” in una partita competitiva.

Per quanto riguarda la politica in senso stretto, la posta in gioco principale è l’esercizio del potere nelle relazioni che gli attori di governo hanno sia con gli interessi economici nazionali, sia con altri attori istituzionali nelle arene politiche internazionali. Le stesse poste in gioco economiche possono infatti assumere carature ancora più decisamente politiche in vari modi. Ciò avviene quando i governi definiscono e conducono strategie su scala nazionale per lo sviluppo dell’IA, esaminate in altri paragrafi. Ai fini di queste strategie attori politici e imprese dell’IA dichiarano per lo più di voler cooperare. Tuttavia, l’effettivo grado, maggiore o minore, di cooperazione dipende dal modo in cui sia gli attori economici, sia gli attori politici interpretano di volta in volta le poste in gioco e le rispettive convenienze. Da un lato, anche nel caso dell’IA, per proteggere i propri interessi commerciali le imprese aspirano alla maggiore possibile autonomia nei confronti di regole, vincoli e indirizzi posti dalla politica, specialmente in forma di *hard law*. Per questo, come vedremo, preferiscono l’assenza di regolazione o l’auto-regolazione. Dall’altro, le autorità politiche possono o meno assecondare queste preferenze, sulla base di considerazioni che riguardano l’“interesse pubblico”, di volta in volta variamente

⁵⁴ Fonte: <https://openai.com/index/announcing-the-stargate-project/>

⁵⁵ Fonte: <https://edition.cnn.com/2025/01/21/tech/openai-oracle-softbank-trump-ai-investment/index.html>

definito, o per assecondare le domande degli elettorati. In alcuni casi, come vedremo, scelgono di “lasciar fare” il business delle grandi imprese o, comunque, di non caricarle di pesanti oneri regolativi, come accade prevalentemente negli Usa e in UK. In altri casi, che si ritrovano più facilmente in Europa, danno priorità alla salvaguardia di interessi diffusi, ad esempio per quanto riguarda le garanzie sulle informazioni personali e nei confronti della riproduzione automatica di pregiudizi sociali. Le tensioni sull’IA fra grandi imprese tecnologiche e autorità politiche possono aggiungersi ad altre contese, che hanno sinora avuto per oggetto, di volta in volta, controversie fiscali (con richieste di pagamenti elevati), regole relative alla raccolta e uso dei dati (con misure a tutela della privacy) e situazioni di monopolio nei mercati (con conseguenti misure antitrust, che colpiscono spesso le Big Tech come Google, Meta, Amazon e Apple).

Nell’UE e nei suoi Paesi membri e negli Usa, questi campi di questo gioco fra autorità politiche e potenze tecnologiche sono più facilmente osservabili sia poiché nelle liberaldemocrazie almeno una parte di queste azioni è di pubblico dominio. Più difficile, invece, è accedere a informazioni su quanto accade in sistemi più autoritari e opachi come la Cina. Durante l’amministrazione Biden (2020-2024) il governo degli Usa ha adottato misure nei confronti delle principali aziende tecnologiche soprattutto per infrazioni, abusi e violazione delle leggi per la protezione della concorrenza. Nel periodo di interregno fra le amministrazioni Biden e Trump (l’autunno 2024), la *Federal Trade Commission*, ancora presieduta da un’esponente democratica, si è affrettata a portare a termine una revisione delle potenziali violazioni delle norme antitrust, alcune delle quali contenute in accordi tra start-up di IA e le Big Tech. Questo è accaduto per Google (il governo ha chiesto a un giudice di costringere l’azienda a vendere il suo browser Chrome), Microsoft (che ha acquisito posizioni di monopolio abbinando prodotti di *cloud computing*, IA e sicurezza informatica), Amazon (accusata di esercitare un monopolio illegale nella vendita al dettaglio online) e Meta (proprietaria di Instagram, Facebook e WhatsApp).

Questi sono solo alcuni esempi americani – cui se ne potrebbero aggiungere altri, come le tante audizioni di Big tech di fronte al Congresso – di un continuo braccio di ferro fra le autorità politiche e i grandi oligopoli tecnologici che si verifica anche nell’UE e nei Paesi membri, ad esempio per quanto riguarda l’uso dei dati personali. In queste contese, le autorità politiche cercano difficili equilibri fra l’obiettivo di favorire tecnologie decisive, come abbiamo appena visto, per lo sviluppo economico e il vantaggio nazionale nella competizione strategica e quello di mantenere un controllo politico dell’IA di fronte a Big Tech che appaiono tanto potenti (Rosso, 2023) da negoziare, di fatto, alla pari con i governi. Da un lato a causa della loro ric-

chezza, dall’altro perché i mercati della comunicazione che controllano permettono loro di influenzare modelli culturali, comportamenti sociali, consensi politici e scelte elettorali. In questo modo attivano poste in gioco squisitamente politiche, non solo influenzando i processi democratici, ma anche perché partiti e leader politici hanno interesse nel dimostrare agli elettori che sono in grado di tenere sotto controllo una tecnologia di cui molti diffidano⁵⁶.

Altre poste in gioco relative all’IA squisitamente politiche sono quelle di tipo geopolitico. La geopolitica dell’IA può essere riferita sinteticamente, nello scenario globale, da un lato ai suoi usi bellici, cui gli stati sono ovviamente interessati e, dall’altro, alla più ampia competizione fra stati – ossia fra sistemi tecnologici economici e politici nazionali – per la leadership tecnologica, la sicurezza nazionale e la governance dei dati⁵⁷. Questa competizione si riverbera sulle discussioni e i negoziati nelle arene internazionali.

Le aziende tecnologiche dei sistemi di difesa nazionale stanno applicando l’IA a varie forme di guerra elettronica, con forniture da provider privati spesso provenienti da start-up che raccolgono ingenti finanziamenti e dalle Big Tech, più specificamente Amazon, Microsoft e Google, talvolta in competizione fra loro per contratti di fornitura militari⁵⁸. Per le Big Tech il mercato bellico propone quindi opportunità di profitto, ma per coglierle devono superare dilemmi etici, importanti perché anche le imprese necessitano di ricevere fiducia e legittimazione dal proprio ambiente sociale. Lo ha fatto nel 2024 OpenAI, passando dal precedente divieto di utilizzo per scopi bellici dei suoi *LLMs* ad accordi con aziende nel campo delle tecnologie militari, come l’intermediario per gli appalti dell’esercito degli Usa Carahsoft⁵⁹ e Anduril. Quest’ultima è un’azienda tecnologica che, vendendo sistemi di comando e controllo, droni, missili, navi sottomarine e apparecchiature di sorveglianza alle forze armate di Usa, Australia e UK, è prossima a raggiungere il valore di 28 miliardi di dollari e il cui fondatore ha sostenuto la candidatura di D. Trump nelle campagne del 2016, 2020 e 2024⁶⁰. L’accordo con Anduril

⁵⁶ Una ricerca di Brookings Institution (Drecksler, 2025) ha rilevato che i cittadini statunitensi e britannici tendono a essere più preoccupati che ottimisti riguardo agli impatti dell’IA e, per quanto riguarda la sua regolazione, non si fidano di un’implementazione affidata né alle sole aziende tecnologiche né ai governi da soli.

⁵⁷ Fonte: <https://www.agendadigitale.eu/mercati-digitali/geopolitica-dell'ia-perche-ue-e-italia-rischiano-di-restare-fuori-dai-giochi/>

⁵⁸ Un esempio è l’uso di software geospaziali (GIS), che sembra aver ridotto la catena decisionale relativa a droni e artiglieria da 20 minuti a meno di un minuto (Benaich, Hogarth, 2022). Alcuni modelli di IA *general purpose*, invece, hanno una capacità di uso duplice (*dual use*), civile e militare, come già accaduto per l’energia nucleare (Scharre 2023).

⁵⁹ Fonte: <https://aldianews.com/en/leadership/corporate/openai-dod>

⁶⁰ Fonte: <https://www.nytimes.com/2025/02/07/technology/anduril-funding-valuation-28-billion.html>

sarebbe di fondamentale importanza per la strategia di difesa degli Usa, che hanno una spesa militare superiore a qualsiasi altro paese, per mantenere un vantaggio strategico attraverso sistemi di IA applicati alla difesa aerea, consentendo “agli operatori militari e di intelligence di prendere decisioni più rapide e precise in situazioni di alta pressione”⁶¹. Ancora un esempio: Anthropic collabora con Palantir e AWS per fornire strumenti di IA a supporto delle operazioni di difesa statunitense (Puliafito, 2025). Questi accordi assumono una particolare rilevanza nella competizione globale tra gli Usa e la Cina per una posizione di leadership mondiale nello sviluppo dell’IA, una sfida che si gioca nell’interdipendenza tra le poste in gioco economiche e politiche.

Così come per tecnologie preesistenti – si pensi ancora agli armamenti nucleari – l’IA diviene un fondamento della potenza militare su scenari regionali e globali, con usi strategici nei sistemi di comando e controllo, nelle attività di intelligence, sorveglianza e cognizione, nella logistica, e in altre. Questo non sta modificando solo la gerarchia geopolitica basata sulla competizione fra Usa e Cina, ma anche le contese in cui sono coinvolti come competitori altri stati, che hanno minore potenza militare. In questo modo vengono introdotti nuovi rischia causa dei possibili usi da parte di potenze emergenti, suscitando preoccupazioni nella società civile, con ONG come *Human Rights Watch* che hanno lanciato campagne contro i sistemi di arma automatizzati come *Stop Killer Robots* (Baele *et al.*, 2024). Questo intreccio fra business e politica militare – presente, del resto, anche in altri Paesi – complica le cose nelle relazioni internazionali, rendendo difficile raggiungere accordi per mettere al bando gli usi bellici dell’IA e spingendo ulteriormente i governi a porre un’inflessione crescente sul perseguitamento dell’autonomia, se non della superiorità nazionale nell’innovazione tecnologica. Ad esempio, alcuni governi dell’UE auspicano che le specifiche spese statali per l’IA negli armamenti siano considerate parte degli investimenti nella difesa, non conteggiati nel calcolo del deficit e del rapporto debito/PIL.

Le poste in gioco nella competizione tecnologica ed economica si capiscono perciò anche alla luce di quelle negli scenari geopolitici e delle loro reciproche interdipendenze: nel discorso pubblico globale la sfida per la leadership nell’IA è diventata una questione che riguarda anche la sfera della geopolitica. Questa sfida riguarda la supremazia tecnologica, misurata attraverso ranking come il *Global AI Index* o l’*AI leadership ranking*, che concordano nel collocare al primo posto gli Usa e al secondo la Cina, seguiti da India, UK, Sud Corea,

⁶¹ Fonte: <https://www.anduril.com/article/anduril-partners-with-openai-to-advance-u-s-artificial-intelligence-leadership-and-protect-u-s/>

Germania, Australia, Iran, Canada e Italia⁶². La corsa alla supremazia nell'IA può ulteriormente ampliare i divari di potere economico e politico. I Paesi leader nell'IA stanno ottenendo vantaggi strategici, ridisegnando le assimmetrie di potere e sfidando le norme e le metodologie internazionali esistenti (Korkmaz, 2024). I Paesi con mercati di IA di dimensioni ridotte rischiano invece di subire un ulteriore depotenziamento, mentre quelli emergenti di medie dimensioni, come il Brasile, il Messico, l'Indonesia, la Turchia o il Sudafrica, sono bloccati nel mezzo tra l'IA Usa e quella cinese (Boaventura, 2023). Sembra quindi emergere uno specifico IA *divide* che rischia di dividere il mondo nei due lati opposti di una nuova “cortina di ferro dell'IA” (Hung, 2025).

In questa sfida le rivalità politiche intersecano gli interessi economici: è il caso di Taiwan, che la Cina vorrebbe porre sotto il proprio controllo e di cui gli Usa difendono l'indipendenza. Lo Stato di Taiwan è estremamente importante per la produzione di *chip*, in particolare per Nvidia, grazie alla presenza di *Taiwan Semiconductor Manufacturing Company*, il più avanzato produttore di semiconduttori al mondo, che domina il mercato dei semiconduttori avanzati, producendo anche per Apple, AMD e Qualcomm. Allo stesso tempo, le GPU Nvidia di produzione statunitense sono utilizzate da tutte le principali aziende tecnologiche e dalle università cinesi. Per questo il governo Usa ha adottato provvedimenti progressivamente più restrittivi per ridurre l'esportazione dei loro più recenti *chip* di IA in Cina come mezzo per limitarne l'uso in applicazioni che ritiene minacciose per la sicurezza nazionale americana (Benaich e Hogarth, 2022, 2023; Benaich e Chalmers, 2024).

Prima della seconda presidenza Trump, Usa e UE avevano tentato di “fare massa” “insieme per competere con la Cina. In primo luogo, adottando misure per ridurre i rischi geopolitici, di sicurezza ed economici della presenza tecnologica cinese nei mercati europei (es. Huawei e altre aziende di telecomunicazioni cinesi nelle reti 5G). In secondo luogo, attraverso un approccio multilaterale alla definizione di standard tecnici globali mediando tra Stati Uniti e Cina (Wang e Galligan, 2021), ma l'allergia della nuova presidenza nei confronti del *frame* regolativo europeo ha reso più difficile questa collaborazione strategica.

Il proposito di raccogliere la sfida realizzando “modelli nazionali” di IA è presente nei discorsi pubblici, in cui intervengono non solo esponenti politici, ma anche le imprese. Nel sottolineare che *chips*, dati, energia e talento sono le chiavi per vincere con l'IA – “una gara che l'America può e deve vincere” – OpenAI raccomanda alla Presidenza Trump l'attivazione di una

⁶² la Repubblica, 12 agosto 2023

https://www.repubblica.it/editoriali/2023/08/12/news/intelligenza_artificiale_geopolitica-410894900/?ref=nl-rep-f-anr

strategia fondata sull'alleanza fra potere economico e politico: “si stima che ci siano 175 miliardi di dollari in fondi globali in attesa di investimenti in progetti di IA, e se gli Usa non attireranno questi fondi, essi confluiranno in progetti sostenuti dalla Cina, rafforzando l'influenza globale del Partito Comunista Cinese”⁶³. La Cina, a sua volta, incolpa Usa e Paesi occidentali di non voler dare seguito a iniziative realmente multilaterali e inclusive per uno sviluppo concordato dell'IA. Nel frattempo, da un lato gli Usa hanno posto limiti all'esportazione di *chip* verso la Cina (2022), una scelta imitata dai governi olandese e giapponese (Scharre, 2023). Dall'altro, l'investimento della statunitense Microsoft in G42, una società di IA degli Emirati Arabi Uniti che ha legami con Huawei e altre società cinesi, è stato letto come un'iniziativa che, permettendo a Microsoft di partecipare alla *corporate governance* di questa compagnia, “potrebbe aiutare gli Usa a contrastare la crescente influenza della Cina nella regione del Golfo”. Dunque, una collaborazione fra il business tecnologico e il governo degli Usa⁶⁴, in cui entrambi possono ottenere vantaggi.

In questo capitolo abbiamo evidenziato da un lato che l'IA e i suoi usi hanno non solo enormi potenzialità e rischi, approfondendo in particolare il campo della democrazia e delle sue pratiche, dall'altro che l'IA stessa ha poste in gioco sia economiche, sia politiche. E che esse si collocano spesso all'intersezione fra un'arena economica e una politica. Nel prossimo capitolo vedremo come questo insieme di problemi proponga dei dilemmi di governance dell'IA, un puzzle che sia le autorità politiche, sia l'ambiente imprenditoriale, stanno cercando di risolvere, attivando varie arene di confronto e di decisione e sperimentando forme di regolazione fra loro molto diverse.

⁶³ OpenAI's Economic Blueprint, 13 gennaio 2025
<https://openai.com/global-affairs/openais-economic-blueprint/>

⁶⁴ Fonte: <https://www.nytimes.com/2024/04/16/technology/microsoft-g42-uae-ai.html>

2. Governare l'intelligenza artificiale: fra mercato e arene internazionali

2.1. Le forme della regolazione: autoregolazione, *soft law* e *hard law*

Come abbiamo visto nel capitolo precedente, lo sviluppo dei sistemi di IA e il bilanciamento delle potenzialità e dei rischi associati al loro uso pongono urgenti interrogativi su *chi* e *come* debba e possa governare questi processi. Questi interrogativi riguardano da un lato l'identificazione degli attori più legittimi e adeguati a dirigere le scelte di sviluppo e uso dell'IA, dall'altro la selezione degli strumenti di governo più efficaci nel garantire libertà di ricerca e innovazione e nel ridurre la produzione di nuovi e riproduzione di vecchi rischi pubblici. In altre parole, l'IA pone rilevanti questioni di governance e *metagovernance*, che riguardano il ruolo e le forme di intervento della politica alle diverse scale e settori di applicazione delle tecnologie. All'inizio del paragrafo sono però necessarie due precisazioni relative all'uso dei due concetti. Primo, in questo volume – a differenza che in altre opere degli stessi autori¹ – usiamo il concetto di governance come sinonimo dei concetti di regolazione e di direzione (*steering*): questi tre termini si riferiscono alla formulazione, promulgazione, attuazione o applicazione, di regole, vincolanti o volontarie, che sono rilevanti per la società (Steurer, 2013). L'applicazione “estesa” del concetto di governance anche come sinonimo di regolazione consente di considerare una più ampia varietà di attori e di strumenti impiegati nel governo dello sviluppo e dell'uso dell'IA. Secondo, il governo dell'IA pone questioni di *metagovernance*, intese come processi riflessivi e reattivi attraverso i quali una serie di attori legittimi mirano a com-

¹ Cfr. ad esempio *Governance: istituzioni e sistemi di azione pubblica*, in d'Albergo, Moini (a cura di), 2014, pp. 119-120.

binare, facilitare, modellare e dirigere particolari forme di governance in conformità con regole, procedure e standard specifici che incarnano la concezione egemonica di ciò che costituisce il “buon governo” (Sørensen e Tøring, 2009). Riflettere sulle questioni di *metagovernance* significa quindi considerare la combinazione e l’interazione tra le diverse forme di regolazione che convergono sul governo, o la regolazione, di un particolare aspetto (Steurer, 2013), in questo caso sul bilanciamento del *trade-off* tra benefici e rischi dell’IA. In questo paragrafo sono pertanto discussi i principali approcci regolativi dell’IA di cui dispongono attori pubblici e privati che, nel loro insieme, concorrono alla definizione delle regole attraverso le quali provano a disciplinare lo sviluppo di questo settore.

Questi strumenti possono essere osservati in relazione all’applicazione ideal-tipica della dicotomia (Sinclair, 1997), consolidatasi nel dibattito scientifico degli ultimi tre decenni, che distingue tra forme di auto-regolazione delle imprese (*self-regulation*) e regolazione pubblica basata su comando e controllo (*control and command*). A metà dell’ideale *continuum* si collocano gli strumenti di *soft law*, dispositivi di regolazione degli attori pubblici che però non prevedono obblighi o imposizioni ma agiscono nei confronti degli attori che producono e usano IA attraverso le leve della persuasione e dell’incitazione rivolte a imprimere la direzione di cambiamento desiderata. Essa è quindi da considerarsi una “falsa dicotomia”² (*ibidem*) perché la regolazione presenta ibridazioni e innovazioni sia riguardo agli attori che ne sono responsabili, degli strumenti impiegati e, evidenziamo noi, delle scale di applicazione. Dunque, il capitolo ha l’obiettivo di esplorare il dibattito scientifico sulle questioni di *metagovernance* dell’IA, con particolare riferimento: alla soluzione dell’auto-regolazione delle imprese riportandone esempi, cornici e risorse di legittimazione; ai tentativi di costruire un’arena di governo transnazionale.

Dunque, al primo estremo del *continuum* delle possibili soluzioni regolative si colloca l’auto-regolazione dei privati. La *self-regulation* delle imprese non è un fenomeno recente, essa rappresenta piuttosto, già dal Novecento, una soluzione regolativa ampiamente adottata dalle aziende in campo ambientale come strategia di riduzione dell’inquinamento e di risparmio di energia e materiali di produzione (Andrews, 1998) e come strategia per anticipare ed evitare la regolazione politica *hard*. Esistono molteplici forme e strumenti di autoregolazione che, secondo una vecchia definizione, definiscono il campo degli ordinamenti normativi di “governi” privati, imprese, comunità pro-

² Già nel 1997, con questa espressione R. Sinclair evidenziava la necessità di analizzare il range delle varie forme di regolazione che interessavano le politiche ambientali, rifiutando l’idea dell’esistenza di forme idealizzate di auto-regolazione “pura” o di “rigida” regolazione pubblica.

fessionali e reti commerciali (Huyse e Parmentier, 1990). Come ha riassunto Steurer (2013) la *self-regulation* comprende: strumenti interni alle aziende, quali codici di condotta, strategie di *stakeholder management* e criteri di responsabilità sociale applicati ai partner aziendali (es. fornitori di materie prime o acquirenti all'ingrosso); strumenti elaborati da gruppi di aziende e applicati a un particolare settore (*industry self-regulation*) come codici, standard, accordi e sistemi di certificazione. Questi ultimi, generalmente, sono elaborati da organizzazioni di standardizzazione nazionali e transnazionali³ che prevedono una valutazione periodica esterna da parte dei propri esperti (Töller, 2011). Sebbene per ragioni di spazio non sia possibile offrire una rassegna esaustiva della letteratura sull'autoregolazione delle imprese, è opportuno trarne tre elementi caratterizzanti che la rendono una soluzione appetibile nel campo della governance dell'IA: volontarietà; flessibilità di implementazione; controllo e autonomia.

Sin dagli anni Novanta, la diffusione delle soluzioni regolative volontarie delle imprese ha rappresentato il contraltare della riduzione delle capacità e volontà regolative degli stati sui processi economici globali (Jenkins, 2001). Più che essere solo un gioco compensativo del ritiro dello Stato dalla regolazione, le "ondate" del successo dell'autoregolazione delle imprese sono state favorite dagli indubbi vantaggi associati alla sua dimensione volontaria: maggiore libertà delle aziende nel definire le regole per esse stesse rilevanti, compatibilità con le proprie strategie di business, assenza degli aspetti di *enforcement* e coercizione previsti dalla regolazione pubblica. La volontarietà dell'autoregolazione delle imprese riguarda quindi la volontà non solo di definire e rispettare regole definite internamente, ma anche di aderire a codici e standard elaborati da soggetti terzi. Questa scelta può derivare dall'applicazione di una razionalità economica che ne valuti costi e benefici. Si pensi a titolo di esempio agli *AI guardrails*, elaborati nel 2024 dalla nota società di consulenza statunitense McKinsey & Company, applicati allo sviluppo e uso dell'IA generativa. La metafora dei "parapetti" adottata da McKinsey appare particolarmente efficace nel raffigurare l'utilità degli standard: «barriera lungo il bordo della strada che proteggono i veicoli dal deviare dalla loro traiettoria e dal mettersi in pericolo» (2024, p. 2)⁴. McKinsey identifica almeno tre tipi di benefici per le aziende che intendono aderire al proprio sistema di *soft-regulation*: la sicurezza dei sistemi di IA e la difesa da attacchi

³ Ne è esempio la storica International Organization for Standardization (ISO), organizzazione non-governativa attiva dal 1946, che ha sviluppato oltre venticinque mila standard che definiscono il riferimento per sistemi produttivi multisettoriali, di trasporto, pratiche di cura e financo per le azioni volte all'inclusione e il rispetto della diversità nei contesti lavorativi.

⁴ Tuttavia, come ogni barriera in autostrada: «i *guardrails* dell'IA non garantiscono che i sistemi di IA saranno completamente sicuri, equi, conformi ed etici» (*ivi*).

che possono manipolare i risultati prodotti dall'IA generativa per usi malevoli; la *compliance* regolativa rispetto alle norme esistenti ed emergenti della regolazione pubblica, una necessità di conformità che può avere anche una funzione difensiva rispetto a eventuali sanzioni e attribuzioni di responsabilità civile; la creazione di fiducia presso i clienti dei propri servizi e del pubblico generale attraverso il monitoraggio dei contenuti dei risultati generati dal sistema di IA (*ivi*). Da un lato, è possibile cogliere come l'adesione a questo particolare sistema di autoregolazione sia veicolata mediante logiche difensive di protezione delle aziende da attacchi, sanzioni e danni di immagine. Dall'altro lato, la *self-regulation* può essere essa stessa uno strumento di creazione di valore attraverso l'acquisizione di vantaggi competitivi relativi, come sostiene la società di consulenza, alla costruzione della fiducia dei clienti, al risparmio dall'ingaggio di controversie legali e all'acquisizione di talenti "attratti" dall'uso responsabile del proprio sistema di IA. Secondo la proposta di McKinsey, i *guardrails* dell'IA rappresentano, quindi, strumenti fondamentali nel creare un ambiente favorevole allo sviluppo tecnologico e all'innovazione piuttosto che elementi di vincolo e costi da sostenere che possono ostacolare o scoraggiare gli investimenti. Dunque, un approccio amico dell'innovazione tecnologica.

Il secondo elemento di legittimazione dell'autoregolazione risiede nella sua velocità, flessibilità e adattabilità alle circostanze del mercato (Andrews, 1998). In altre parole, i tempi dell'adempimento alle pratiche di autoregolazione possono essere subordinati ai periodi, alle condizioni e ai contesti più favorevoli. Il terzo elemento di vantaggio della *self-regulation* riguarda il controllo sui suoi contenuti e, quindi, l'autonomia decisionale. Esso deve essere posto in relazione con una delle caratteristiche distintive dell'autoregolazione: l'identificazione tra il regolatore e il soggetto e obiettivo della regolazione (Coglianese e Mendelson, 2010). Le imprese definiscono sistemi di regole, codici di comportamento e standard riferiti ai loro sistemi interni – manager, dipendenti, partner aziendali – con la possibilità di esercitare massima discrezionalità e autonomia nella loro attuazione: le norme sono auto-specificate, la condotta è auto-monitorata e le regole sono auto-applicate (Bartle e Vass, 2007).

Non da ultimo l'autoregolazione trae la sua legittimazione in riferimento alla presunta preminenza delle imprese riguardo al possesso e controllo delle conoscenze necessarie a qualsiasi processo regolativo. In accordo con G. Mobilio (2020) il discorso sulla *self-regulation* enfatizza la tesi secondo cui le imprese tecnologiche sono in possesso delle conoscenze tecniche, e quindi delle soluzioni, più aggiornate, e hanno la reale contezza delle poste in gioco e dello stato delle problematiche via via emergenti associate all'IA. In tal senso, il progetto *World – The real human network* (da ora *World*) è un e-

sempio di particolare interesse: *World* è stato ideato nel 2023 da Sam Altman e Alex Blania, amministratori delegati, rispettivamente, di *OpenAI* (Chat-GPT) e *Tools for Humanity*. In estrema sintesi, il progetto ha l'obiettivo di creare un sistema di identità personale univoco su scala globale e un network finanziario ad esso associato. L'adesione al sistema di identità univoco è la precondizione per le persone per ottenere e scambiare il token *Worldcoin*, criptovaluta operativa del *World*, da spendere nelle transazioni online basate su questa moneta virtuale (es. gas, prodotti di e-game). Allo stesso tempo, nelle intenzioni dei suoi proponenti il progetto ambisce alla risoluzione di uno dei problemi più rilevanti associati allo sviluppo dell'IA: la difficoltà nel distinguere tra le azioni condotte online dagli umani da quelle attribuibili a bot e sistemi di IA. Un articolo apparso nel dicembre 2024 sul blog del progetto evidenzia l'esigenza di costruire una "controparte umana dell'IA", che consenta alle persone di riconoscere i contenuti online realmente realizzati da esseri umani⁵. Questa controparte è basata su una *proof of personhood* (PoP), un meccanismo che verifica digitalmente l'umanità e l'unicità di un individuo. Essa può essere ottenuta per mezzo di un sistema di identificazione biometrica attraverso la scansione dell'iride di una persona, acquisita mediante la tecnologia ORB⁶ e sulla sua conseguente trasposizione numerica da cui deriva un codice univoco "World ID". Il progetto di Blania e Altman mostra la capacità e la volontà delle Big Tech di tradurre gli esercizi di riflessività sui rischi pubblici dell'IA in nuove opportunità di business. Soluzioni che, allo stesso tempo, concorrono alla legittimazione pubblica dei privati come gli attori più adeguati alla loro risoluzione. Questa è di per sé è un'osservazione banale ma che deve essere tenuta in considerazione in riferimento alla discussione corrente sui limiti delle forme di regolazione pubblica, un tema su cui si tornerà nel paragrafo. Da un lato, la creazione di un sistema di identificazione privato su scala globale sfida il monopolio pubblico e la scala di applicazione dei sistemi pubblici di identificazione digitale e di *watermarking* (filigrana digitale) dei contenuti online; dall'altro lato, pone interrogativi circa la privacy degli individui che decidono di aderirvi. A questo proposito, il progetto *World* fa leva sull'applicazione del consolidato principio della *privacy by design*, ovvero integrare e costruire la privacy fin dall'inizio, inserendola direttamente nelle specifiche di progettazione e nell'architettura di nuovi sistemi e processi (Cavoukian, 2011). Secondo quanto sostenuto da *World* l'acquisizione dell'iride è essa stessa garanzia di privacy perché renderebbe non necessario raccogliere altri dati personali e il

⁵ Nell'articolo si fa riferimento ai progressi del modello linguistico di GPT-4 di OpenAI che lo scorso anno è stato in grado di superare il Test di Turing.

⁶ Maggiori informazioni sulle caratteristiche e il funzionamento di ORB sono disponibili al link: <https://www.toolsforhumanity.com/orb>

codice unico, una volta elaborato, non sarebbe trattenuto dal dispositivo di rilevazione o conservato in rete. Piuttosto, esso viene inviato, dopo un processo di anonimizzazione, al legittimo proprietario⁷. Stando al dato riportato da *World* in un video postato su X⁸ il 10 gennaio 2025, tali rassicurazioni hanno convinto dieci milioni persone nel mondo che hanno aderito al progetto. Diversa è invece la postura delle istituzioni e delle agenzie pubbliche che si occupano di protezione dei dati. Il GPDP italiano è stato tra i primi a intervenire attraverso il Provvedimento del 21 marzo 2024 in cui si avvertiva la *Worldcoin Foundation* che i trattamenti di dati biometrici che dovessero essere effettuati in Italia attraverso gli ORB, avrebbero verosimilmente potuto violare il GDPR. In altri Paesi europei provvedimenti più stringenti sono stati presi dall’Agenzia spagnola per la protezione dei dati (AEPD) e dal *Bavarian State Office for Data Protection Supervision* (Baviera – Germania) che lo stesso anno hanno intimato alla fondazione, rispettivamente, di interrompere le proprie attività di raccolta di dati e di disporre una procedura di cancellazione dei dati secondo quanto prescritto dal Regolamento europeo⁹. Questo esempio consente di cogliere in modo immediato alcuni aspetti della tensione che intercorre tra le diverse forme di regolazione che interessano l’IA e i suoi rischi. Se da un lato, il progetto di *World* sostiene di fornire su scala globale una soluzione tecnologica “oltre la regolazione” ai rischi di allucinazioni, disinformazione e manipolazione dell’IA, dall’altro lato, esso sembra porsi in aperta contraddizione con i principi e le norme sulla protezione dei dati esito di lunghi processi di discussione pubblica e di decisione politica, costringendo le autorità nazionali a intervenire, talvolta in modo frammentato, per ristabilirli. In altre parole, il caso di *World* evidenzia il vantaggio in termini di velocità e scala di applicazione delle soluzioni regolative progettate dalle aziende rispetto alle forme di regolazione pubblica, sebbene quest’ultime mantengano ancora “un potere di ultima istanza” nei diversi contesti nazionali quando intervengono in modo particolare nel campo della protezione dei dati personali.

Dunque, se l’autoregolazione è una soluzione comoda per le imprese che operano in diversi settori di business, essa può imporsi con maggiore successo in ambiti difficili da regolare come quello dello sviluppo tecnologico. Tuttavia, le prospettive sociali e politiche poste dallo sviluppo esponenziale

⁷ La descrizione degli strumenti di garanzia della privacy è disponibile al link: <https://world.org/privacy>

⁸ Fonte: <https://x.com/worldcoin/status/1877721028429849018?t=eQIBlGwXm-DfzjSzdn6QkQ&s=19>

⁹ Con questa espressione si intende il diritto alla cancellazione dei dati personali secondo come stabilito dall’Art. 17 del GDPR. Maggiori informazioni sono disponibili al link: <https://gdpr-text.com/it/read/article-17/>

dei sistemi di IA, oltre che la domanda sociale di protezione dai suoi possibili rischi pubblici, definiscono un campo di questioni di interesse collettivo che non possono essere risolte nell'ambito delle strategie aziendali che rispondono, per definizione, all'interesse privato, nonostante esso possa venire definito dalle imprese stesse come più o meno socialmente orientato. Nel 2021 L. Floridi, filosofo e direttore del *Digital Ethics Center* dell'Università di Yale, lanciava il seguente appello: «l'autoregolamentazione deve essere sostituita dalla legge; prima si fa, meglio è. *Dura lex, sed lex digitalis*» (p. 622). Nel suo famoso editoriale “The End of an Era” Floridi dichiarava il fallimento storico della *self-regulation* in ambito digitale denunciando «la sua imbarazzante vacuità» (p. 621) nel risolvere questioni etiche o nel modificare i comportamenti e le strategie d'impresa. Secondo Floridi, gli scandali associati all'uso dei dati personali (es. il caso Facebook-Cambridge Analytica¹⁰) e i fallimenti nell'attuare misure come il “diritto all'oblio” online hanno segnato la fine delle illusioni circa le potenzialità regolative delle imprese nel gestire le questioni etiche della transizione digitale.

Ciò che serve quindi nel campo dell'IA è ristabilire la preminenza del ruolo del potere politico e della regolazione pubblica attraverso l'elaborazione di strumenti di *hard law*¹¹? Le norme contenute in leggi e regolamenti sono strumenti distintivi a disposizione del potere politico che, attraverso le proprie risorse di autorità, può emanare comandi o divieti vincolanti, disporre le azioni di *enforcement* e sanzionare comportamenti devianti. Come sostenuto da A. Costabile (2003) è compito del potere politico assumere decisioni di interesse collettivo aventi forza vincolante e che, per essere legittime, devono essere sostenute da forme di consenso e legittimazione. Se questo accade, il potere politico può riuscire anche nella sfida di regolare il flusso caotico delle trasformazioni e di costruire, attraverso la regolazione, un senso collettivamente condiviso (Costabile, 2002). Ciò appare un obiettivo particolarmente ambizioso nel campo dell'IA, il cui flusso di trasformazione avviene a velocità sostenuta e ha impatti su scale multiple, rendendo obsolete e inefficaci le geografie amministrative del potere politico. Questo flusso è inoltre supportato da un'oligarchia tecnologica che può spendere le proprie risorse finanziarie e cognitive in ogni processo regolativo, anche attraverso prese di posizione, talvolta paradossali, sulle diverse scelte di regolazione

¹⁰ Per una ricostruzione del caso si rimanda a:

<https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

¹¹ In accordo con K. Abbott e D. Snidal (2020) intendiamo con questa espressione: «obblighi giuridicamente vincolanti che sono precisi (o possono essere resi precisi attraverso l'aggiudicazione o l'emanazione di regolamenti dettagliati) e che delegano l'autorità per l'interpretazione e l'attuazione della legge» (p. 421).

pubblica. Si pensi all'attivismo di E. Musk, autodefinitosi sul suo X: «(da oltre 20 anni) un sostenitore della regolamentazione dell'IA, proprio come regolamentiamo qualsiasi prodotto/tecnologia che rappresenti un potenziale rischio per il pubblico»¹².

Gli strumenti di *hard law* rappresentano evidentemente dispositivi che possono concorrere a equilibrare il bilanciamento tra potenzialità e rischi pubblici dei sistemi di IA, come nel caso dei divieti “per legge” dei sistemi ad alto rischio identificati dall'EU AI Act. Tuttavia, le soluzioni di “legge dura” si scontrano con alcuni limiti che ne caratterizzano il rapporto con le (nuove) tecnologie che ne mettono in discussione limiti e applicabilità. In tal senso, H. Hoffmann-Riem (2020) ha identificato alcuni degli ostacoli principali all'efficacia della regolazione pubblica applicata all'IA. In primo luogo, essi riguardano la condizione di incertezza con cui la legge si approccia alla rapidità degli sviluppi tecnologici e all'imprevedibilità dei rischi che possono emergere. Un caso paradigmatico è stato il rilascio sul mercato, nel novembre 2022, di ChatGPT, modello linguistico di OpenAI. La sua diffusione è stata esponenziale fino a superare, al mese di febbraio 2025, la quota di oltre 400 milioni di utenti attivi settimanalmente a livello globale¹³. Fin dai primi mesi del suo rilascio sono stati evidenti i potenziali rischi, ma anche benefici, dell'uso non regolamentato di un *chatbot* conversazionale così potente e avanzato (Luckett, 2023): rischi a breve termine relativi alla diffusione di contenuti falsi o tossici; rischi a medio termine sul mercato del lavoro con la possibilità di una progressiva sostituzione di alcune mansioni; rischi a lungo termine, più teorici, di perdita di controllo umano sul sistema (Elgesem, 2023). ChatGPT ha avuto un effetto dirompente nel processo regolativo dell'AI Act e sul suo ciclo di policy, rendendo necessaria una revisione del precedente lavoro e nuove interlocuzioni pubbliche e politiche nell'ambito del trilogo (cfr. par. 3.2). Come ricostruito in d'Albergo e Corsi (2024), i servizi di *chatbot* basati sui *LLMs*¹⁴ erano stati considerati dalla Commissione Europea nella prima proposta di Regolamento del 2021 come usi a rischio nullo o basso. L'avvento di ChatGPT ha reso dapprima necessario ridiscutere le interpretazioni circa le implicazioni degli usi di tali servizi, riflessioni che hanno condotto il legislatore fino alla definizione della categoria di rischio sistematico: «un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo

¹² Fonte: <https://x.com/elonmusk/status/1828205685386936567>

¹³ Fonte: <https://www.reuters.com/technology/artificial-intelligence/openais-weekly-active-users-surpass-400-million-2025-02-20/>

¹⁴ Un modello di IA che ha la capacità di generare testo in linguaggio naturale come *output*. Molti LLM sono *foundation model* a sé stanti: ad esempio, GPT-3 sono LLM e anche esempi canonici di modelli di fondazione.

sul mercato dell’Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l’intera catena del valore» (Art. 3 punto 65). Un processo e uno sforzo regolativo definito, anche dai suoi critici (Hacker *et al.*, 2023), come “eroico” nel cercare di seguire il passo delle innovazioni dell’IA.

In secondo luogo, secondo Hoffmann-Riem (2020) i sistemi di IA sono generalmente disponibili a livello globale. Travalicando i confini nazionali, sfidano la capacità regolativa degli stati, una caratteristica che suggerisce la necessità di elaborare soluzioni regolative transnazionali o, più realisticamente, internazionali. Terzo, lo sviluppo tecnologico si caratterizza per la segretezza e la riservatezza che, per finalità di concorrenza, riguarda i processi aziendali e i loro risultati. Dunque, sono prevedibili strategie delle imprese che mirano a ostacolare le autorità pubbliche nell’acquisire le conoscenze necessarie ad una eventuale regolazione efficace. Una condizione aggravata dalla concentrazione di potere e risorse da parte di un numero ristretto di aziende, che possono far valere le proprie posizioni di forza nella negoziazione con i regolatori pubblici. Non da ultimo, qualora si decidesse di intraprendere un percorso regolativo rivolto a particolari sistemi di IA ci si potrebbe scontrare con quello che Hagemann e colleghi (2018) hanno definito «problema del coordinamento», fenomeno già osservato negli ultimi decenni nel campo delle innovazioni tecnologiche. Esso si verifica per l’intrinseca molteplicità di livelli di problemi (*ibidem*) relativi alle diverse componenti tecnologiche di un prodotto o di un’applicazione di IA. Ad esempio, le automobili a guida autonoma¹⁵ basano il proprio funzionamento sulla raccolta ed elaborazione di Big Data, sulla connessione a internet, sulla capacità di prendere decisioni autonome, caratteristiche che si aggiungono alle già complicate tecnologie meccaniche di un’automobile (si pensi a materiali impiegati, standard di sicurezza e requisiti di omologazione).

Arrivati fin qui, le domande riguardano quindi chi e con quali strumenti di coordinamento può regolare con efficacia sistemi simili? Inoltre, quali agenzie o livelli amministrativi devono e possono applicare tale regolazione? Fino a giungere a una domanda, che può apparire paradossale e provocatoria: occorre di volta in volta strutturare agenzie regolative dedicate a un particolare ambito di applicazione dell’IA?

Le scelte regolative sono il riflesso degli obiettivi che la regolazione pubblica intende perseguire. Ci si può proporre di regolare orizzontalmente lo sviluppo dei sistemi di IA, o di intervenire verticalmente, per singole applica-

¹⁵ Per una categorizzazione delle implicazioni legislative dei veicoli a guida autonoma si rimanda a Vedaschi e Noberasco (2019).

zioni e ambiti di uso. Da un lato, la regolazione pubblica può avere un carattere reattivo rispetto ai rischi pubblici dei sistemi di decisione autonoma; dall’altro lato, può ambire a rafforzare il ruolo delle istituzioni politiche nel dirigere i processi di innovazione tecnologica. Sono evidentemente due approcci complementari che però riassumono le due direzioni intraprese dagli interventi pubblici in diversi contesti istituzionali. La regolazione basata sul rischio è stato il primo approccio con cui le istituzioni hanno reagito alla dirompenza dell’IA. Ne sono esempio l’*AI Act* in Europa e l’*AI Risk Management Framework* del NIST negli Usa, strumenti di *hard law* e *soft law*, con cui i due sistemi istituzionali hanno definito, rispettivamente, regole e procedure rivolte alle aziende per evitare e prevenire i rischi pubblici. La regolazione dell’IA basata sul rischio beneficia anche di altri dispositivi normativi non direttamente riferiti all’IA. Come ricorda M. Kaminski (2023) il GDPR europeo, il cui principale obiettivo è la protezione dei dati personali, contiene disposizioni che interessano i processi decisionali automatizzati e che stanno dimostrando effetti significativi nel regolare indirettamente molte applicazioni di IA che impiegano dati personali¹⁶. Il secondo approccio di regolazione pubblica può essere riassunto dall’idea di diritto trasformativo: «(un diritto che) sia in grado di accompagnare l’intero sistema, ogni sua componente, nell’era (net-zero) digitale [...] (e che) può realizzarsi, tuttavia, solo attraverso una presa in carico dello scenario di transizione che è, innanzitutto, politica e che richiede di assumere il presente come tempo per la costruzione di una promessa di futuro digitale e del suo ordine, anche giuridico» (Boschetti e Berti, 2024, p. 18). Secondo G. Finocchiaro (2024), esperta di diritto di internet, tali ambizioni necessitano della definizione, oltre che di regole, di principi etici che prescindano dai fatti concreti che intendono regolare. Un approccio che Finocchiaro stessa ritiene fondativo delle diverse strategie di sviluppo e regolazione dell’IA in contesti politici profondamente diversi, i “giganti”, così come da noi definiti di seguito: la Cina con i *Beijing AI Principles* del 2019, l’UE con gli *Orientamenti etici per un’IA affidabile del 2019*, gli Usa con i principi fondativi dell’*Executive Order* emanato nel 2023 da J. Biden (*ibidem*).

Questi principi etici sono esempi di una soluzione intermedia di regolazione politica: gli strumenti di *soft law*. Una consolidata tradizione di studi sull’implementazione delle politiche pubbliche ha mostrato che la politica non agisce solo imponendo regole, riscuotendo tributi e finanziando spesa pubblica; piuttosto, governare significa applicare diverse tecniche alla risolu-

¹⁶ Si pensi in tal senso agli interventi del GPDP sulle politiche di sicurezza urbana intraprese da alcuni comuni italiani basate su tecnologie di riconoscimento facciale e su algoritmi per la visione delle immagini che saranno discussi nel capitolo 4.

zione dei problemi di policy (Howlett e Ramesh, 2003). Infatti, i decisori politici hanno la possibilità di scegliere tra diversi *policy instruments*¹⁷ che comprendono anche principi etici e standard tecnici basati sull'adesione volontaria da parte dei soggetti target della regolazione. In altre parole, i regolatori pubblici possono scegliere tra "bastone e carota" (Bemelmans-Vidéc *et al.*, 2011). D'accordo con G. Shaffer e M. Pollack (2009) quello che interessa in realtà ai ricercatori e agli analisti politici è comprendere come le leggi e le norme operano nel mondo, consapevoli che talvolta le soluzioni "dure" non riescono a ottenere i risultati sperati mentre le forme "morbide" possono risultare più trasformative e capaci di imprimere il cambiamento desiderato. I "nuovi strumenti di policy" (Jordan *et al.*, 2005), così definiti originariamente nel campo delle politiche ambientali, non sono più così nuovi, piuttosto rappresentano una diffusa modalità di governo dei problemi di interesse collettivo basata su una «questione di persuasione» (Steurer, 2013, p. 8). Essi replicano gli strumenti della *self-regulation* quali i codici di condotta, o gli standard tecnici, a cui si aggiungono forme di cooperazione tra attori pubblici e privati (es. impegni volontari, partnership), o strumenti basati sull'informazione (es. campagne educative, formazione) (Stelljes *et al.*, 2017).

Le soluzioni di *soft law* presentano per le imprese gli stessi vantaggi delle forme di autoregolazione ma, secondo i loro sostenitori, possono favorire il perseguitamento degli obiettivi di regolazione degli stessi governi. Primo, secondo Gutierrez e colleghi (2020) le imprese possono essere incentivate nell'adesione alle soluzioni volontaristiche proprio per evitare una regolazione dura dei sistemi di IA, in altre parole la *soft law* può essere posta come antidoto alla minaccia di una regolazione più stringente. Secondo, non porre barriere all'ingresso può favorire l'innovazione nei processi di co-regolazione tra governi, privati e organizzazioni della società civile (*ibidem*). Terzo, la *soft law* può sopprimere alla mancanza e alle difficoltà di istituzione di un sistema di regolazione transnazionale dotato di un'autorità di applicazione (Wallach *et al.*, 2022). Tuttavia, tali difficoltà si ritrovano anche a livello nazionale se si pensa che, anche per volontà politica, quasi nessuno stato nazionale si è ancora dotato di un sistema di leggi sull'IA paragonabile all'AI Act dell'UE¹⁸. Piuttosto sono frequenti *framework* e sistemi di definizione di standard rivolti a imprese e centri di ricerca che sviluppano sistemi

¹⁷ Con *policy instruments* intendiamo: «un dispositivo che è sia tecnico che sociale, che organizza specifiche relazioni sociali tra lo stato e coloro a cui è rivolto, in base alle rappresentazioni e ai significati che porta con sé» (Lascombes e Le Galès, 2007).

¹⁸ Si pensi a titolo di esempio alle difficoltà nel percorso regolativo intrapreso nel 2022 dal Canada con l'Artificial Intelligence and Data Act. Per una ricostruzione del caso si rimanda a B. Attard-Frost, 2025.

di IA o, più verticalmente, agli utilizzatori in settori specifici di policy come la sanità o l'educazione. Un esempio di particolare interesse in tal senso è la pubblicazione nel 2024 da parte del Governo australiano del documento *Voluntary AI Safety Standard*¹⁹. Sebbene non sia possibile entrare nel merito dei suoi contenuti esso identifica 10 *guardrails*²⁰, elaborati in linea con gli *Australia's AI Ethics Principles*²¹, che applicano un approccio incentrato sulla protezione delle persone dai già noti rischi dell'IA. Nel documento si evidenzia che gli standard non creano per le aziende nuovi obblighi legali, piuttosto: «(stabiliscono) una pratica coerente per le organizzazioni (definiscono) le aspettative su come potrebbe essere la legislazione futura, mentre il governo valuta le opzioni sui *guardrail* obbligatori (forniscono) alle organizzazioni le migliori pratiche di governance e etica dell'IA, offrendo loro un vantaggio competitivo» (2024, p. 3).

La scelta di strumenti morbidi è ad oggi, e già nel prossimo futuro prospettato dalla presidenza di D. Trump, la soluzione regolativa preferita dai legislatori e presidenti degli Usa. I già citati *AI Risk Management Framework* (AI RMF) e l'ordine esecutivo di J. Biden, sul quale si tornerà nel prossimo paragrafo, sono le iniziative di regolazione pubblica dell'IA più importanti intraprese dal governo federale a guida democratica tra il 2020 e il 2024. L'origine del AI RMF può essere fatta risalire all'approvazione nel 2020 del *National Artificial Intelligence Initiative Act*²² presentato dalla Rappresentante democratica E. B. Johnson. La legge definiva gli orientamenti di governo dell'IA incentrati sulle leve della produzione e condivisione di conoscenza, della formazione, dell'educazione della forza lavoro e della costruzione di partnership tra industria, governo, enti di ricerca e organizzazioni no-profit. Tra le priorità della legge rientrava anche il: «supporto allo sviluppo di standard volontari, migliori pratiche e parametri di riferimento per lo sviluppo e l'uso di sistemi di intelligenza artificiale affidabili». La legge ha incaricato di questo compito il NIST con l'obiettivo, entro due anni dall'entrata in vigore dell'atto, di elaborare un quadro di riferimento per la gestione dei rischi dell'IA. Da un lato, tale quadro avrebbe dovuto identificare

¹⁹ La sua pubblicazione è parte integrante della strategia di regolazione australiana dell'IA basata su 5 pilastri: chiarezza regolativa; supporto di buone pratiche; supporto allo sviluppo; uso pubblico; engagement nelle arene internazionali.

²⁰ Nel documento si ritrova il concetto di “parapetto” già applicato nella tassonomia di McKinsey & Company, un interessante esempio di convergenza concettuale tra forme di regolazione pubblica e *self-regulation* mediata dalle società di consulenza.

²¹ Maggiori informazioni sono disponibili al link: <https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-principles/australias-ai-ethics-principles>

²² Il testo integrale è disponibile al link: <https://www.congress.gov/bill/116th-congress/house-bill/6216/text#HB7E2F29027F642F88F70634230CFA127>

e fornire standard, metodologie, procedure e processi per valutare l'affidabilità dei sistemi di IA e mitigarne quindi i rischi. Dall'altro lato, il NIST sarebbe stato responsabile del processo di definizione e caratterizzazione dei principi trasversali dell'IA quali trasparenza, sicurezza, privacy ed etica. Tra gli aspetti più interessanti dell'AI RMF, rilasciato come previsto nel 2022, è la sua postura abilitante e riflessiva nei confronti del ruolo dei suoi utilizzatori pubblici e privati:

il suo obiettivo è quello di offrire una risorsa per migliorare la capacità delle organizzazioni di gestire i rischi dell'IA per massimizzare i benefici e ridurre al minimo i danni [...] le organizzazioni e gli altri utenti del *Framework* sono incoraggiati a valutare periodicamente se l'AI RMF ha migliorato la loro capacità di gestire i rischi dell'IA (2022, p. 16).

Una postura che si ritrova nell'approccio regolativo al concetto di rischio:

sebbene l'AI RMF possa essere utilizzato per stabilire le priorità di rischio, non prescrive la tolleranza al rischio. La tolleranza al rischio si riferisce alla prontezza o alla propensione dell'organizzazione o delle parti interessate a sopportare il rischio al fine di raggiungere i propri obiettivi. [...] La tolleranza al rischio può essere influenzata dalle politiche e dalle norme stabilite dai proprietari dei sistemi di IA, dalle organizzazioni, dalle industrie, dalle comunità o dai responsabili politici. [...] organizzazioni diverse possono avere tolleranze al rischio diverse a causa delle diverse priorità organizzative e delle risorse disponibili. Anche all'interno di una singola organizzazione possono esserci un bilanciamento delle priorità e dei compromessi (2022, p. 9).

Si tratta dunque di un approccio evidentemente alternativo alla regolazione "dura" se si pensa al complesso processo di definizione dei diversi livelli di rischio che ha impegnato i decisori politici nell'elaborazione del Regolamento europeo sull'IA. Tuttavia, un breve riferimento all'EU AI Act consente di accennare a come la regolazione "dura" abbia quasi sempre una dimensione ibrida in particolare nelle fasi di implementazione. Il Regolamento, entrato in vigore il 1° agosto 2024, prevede una fase di regime transitorio che condurrà alla reale applicazione delle sue disposizioni nel 2026. Per questa fase la Commissione Europea ha promosso una forma di coinvolgimento su base volontaria dei soggetti target della regolazione e degli attori portatori di interesse attraverso lo strumento dell'AI Pact. Tale strumento si pone due obiettivi (pilastri). Il primo riguarda la produzione e lo scambio di conoscenze su buone pratiche di sviluppo e utilizzo dell'IA, già conformi con le indicazioni dell'AI Act, da condividere tra imprese, funzionari pubblici, organizzazioni no-profit e attori della conoscenza con il coordinamento dell'AI Office. Il secondo pilastro è rivolto alle imprese e ha l'obiettivo di

accelerare l'adozione delle disposizioni presso le imprese stesse, ingaggiate attraverso “impegni volontari”²³ a preparare una strategia aziendale per la futura conformità alla Legge, identificare e segnalare i sistemi ad alto rischio e promuovere presso i propri dipendenti lo sviluppo etico e responsabile dell'IA. Nelle intenzioni della Commissione si ritrova quindi la riposta alla necessità di accompagnare la *dura lex* con elementi volontaristici che possono da un lato preparare gradualmente i soggetti interessati dalla regolazione, potenzialmente ostacolati nelle loro attività dai suoi limiti e obblighi, e dall'altro garantire una fase di *fine tuning* allo stesso *output* regolativo. Allo stesso tempo, la regolazione dura non può prescindere dal garantire la libertà nell'innovazione, in particolare in un contesto globale in cui i vincoli imposti dall'AI Act non esistono in forme analoghe e che espone quindi una stessa impresa ad ambienti regolativi diversi. Per questo il legislatore europeo ha deciso di usare lo strumento delle “*regulatory sandboxes*”²⁴ (Art. 57-61) facendolo gestire agli stati membri. L'obiettivo di questo istituto è riassunto nel *recital* 138²⁵: «l'IA è una famiglia di tecnologie in rapido sviluppo che richiede una supervisione normativa e uno spazio sicuro e controllato per la sperimentazione, garantendo al contempo un'innovazione responsabile e l'integrazione di adeguate misure di salvaguardia e di mitigazione dei rischi». Non da ultimo, questo strumento è inteso dal legislatore come una leva di apprendimento normativo *evidence-based* per eventualmente rivedere e aggiornare le disposizioni approvate.

2.2. Il potere delle imprese e la cornice dell'etica

Come abbiamo appena visto, intorno all'IA vi sono poste in gioco economiche e politiche. Le seconde vengono trattate con strategie sia di sviluppo e innovazione, sia di regolazione. In entrambe le imprese tecnologiche sono attori importanti e, per questo, una sociologia politica dell'IA non può prescindere dal considerarne il ruolo nella risoluzione delle questioni di interesse collettive connesse con lo sviluppo e l'uso dei sistemi di decisione autonoma. A. Costabile (2024) ha enfatizzato il ruolo emergente, tra gli altri

²³ Al settembre 2024 tali impegni sono stati sottoscritti da circa 190 imprese. Meta è l'unica Big Tech che, al febbraio 2025, non ha aderito all'AI Pact. L'elenco completo dei firmatari è disponibile al link: <https://digital-strategy.ec.europa.eu/en/policies/ai-pact#ecl-inpage-Signatories-of-the-AI-Pact>

²⁴ «Strumenti normativi che consentono alle imprese di testare e sperimentare prodotti, servizi o attività nuovi e innovativi sotto la supervisione di un'autorità di regolamentazione per un periodo di tempo limitato» (European Parliament, 2022: 2).

²⁵ Fonte: <https://ai-act-law.eu/recital/138/>

poteri sociali, del potere tecnologico, un'espressione con cui l'autore identifica una particolare traiettoria di autonomia del potere della scienza e della tecnica che, alleatosi con il potere economico, lo ha condotto a intervenire in maniera dirompente in ogni ambito della vita sociale, influendo sulla riconfigurazione delle relazioni sociali, sui processi del mercato del lavoro, fino al funzionamento dei sistemi democratici. Il punto di caduta di questa riflessione può essere identificato nel ruolo sempre più pervasivo delle Big Tech che operano nel campo dell'IA, aziende che negli ultimi hanno concentrato una quota sempre maggiore di potere economico. Il 2024 ha consacrato l'emergere di un'élite dell'IA di sette colossi – Alphabet, Amazon, Apple, Meta, Microsoft, Nvidia e Tesla – che da soli sono arrivati a detenere un terzo dell'indice S&P500, un valore che monitora l'andamento del mercato azionario di cinquecento aziende quotate alla Borsa di New York che assorbono l'80% degli investimenti su scala globale. Questo ristretto gruppo di imprese è al centro di quel processo di riconfigurazione del sistema di produzione e accumulazione della ricchezza che ha ispirato un ampio dibattito concettuale sulle nuove forme di capitalismo contemporaneo. Il Capitalismo delle Piattaforme (Srnicek, 2016), Il Capitalismo della Sorveglianza (Zuboff, 2019), la Società delle Piattaforme (van Dijck *et al.*, 2018), sono i titoli di alcuni dei libri che hanno riscosso un grande successo provando a cogliere le caratteristiche e le implicazioni economiche, politiche e sociali di questa riconfigurazione, guidata da un oligopolio di imprese digitali che detengono immensi patrimoni dati e forniscono piattaforme hardware e software ormai quasi imprescindibili nella vita quotidiana, nel lavoro e anche nel funzionamento delle istituzioni. Una situazione di oligopolio che ha consentito alle Big Tech di acquisire rapidamente una posizione di dominio nel campo dell'IA basata sul controllo dei Big Data, sul possesso di infrastrutture computazionali e sulla possibilità di acquisire/assorbire potenziali concorrenti e attrarre i migliori talenti professionali (Verdegem, 2024). Secondo stime riportate dal Financial Times²⁶, l'investimento complessivo di Microsoft, Alphabet, Amazon e Meta per la costruzione di nuovi data center nel 2025 sarà pari a oltre 300 mld di dollari.

Da un punto di vista sociologico l'azione delle imprese non può essere osservata solo in relazione ai loro comportamenti nei mercati, ma è necessario anche ricostruire come esse intervengono nel, nel rapporto con la sfera della politica, influenzando il discorso pubblico e i processi regolativi. Come abbiamo visto la *self-regulation* delle imprese è una delle possibili soluzioni della *metagovernance* dell'IA. Le pratiche in cui può concretizzarsi, tuttavia, non si esauriscono nella sola dimensione di impresa, ma si relazionano ai

²⁶ Fonte: <https://www.ft.com/content/634b7ec5-10c3-44d3-ae49-2a5b9ad566fa>

processi in corso nei diversi contesti politici. Dunque, il potere delle Big Tech dell'IA si mostra anche nel contendere alla politica il primato nel bilanciamento del *trade-off* tra potenzialità e rischi. Anche se le Big Tech non possono stabilire regole autoritative esse possono intervenire con efficacia anche nella regolazione "dura". In primo luogo, infatti, possono costruire significati condivisi sugli obiettivi della regolazione dell'IA, definendone il perché e il come. Secondo, possono sperimentare strumenti di *soft law* in processi di auto-regolazione e co-regolazione con le istituzioni e le organizzazioni della società civile. Terzo, le imprese possono adottare soluzioni trasversali agli ambiti di uso e alle diverse scale spaziali risultando potenzialmente più efficaci della regolazione attraverso leggi la cui efficacia, per definizione, è limitata nell'estensione dai confini politici nazionali, o dell'UE. Pertanto, come mostrato nei paragrafi successivi, questa azione si dispiega nelle arene internazionali, a livello sovranazionale o nell'ambito delle strategie di sviluppo nazionali dei cosiddetti "giganti" (Usa, UK e Cina). Questo paragrafo, in particolare, fornisce alcuni elementi di riflessione riguardo: le narrazioni sui rischi dell'IA e le modalità con cui le Big Tech interpretano la domanda sociale di decisioni autoritative; le cornici di significato da cui le imprese traggono la loro legittimazione; in cosa consistono le forme di *self-regulation* delle Big Tech; il rapporto fra Big Tech e le strategie di regolazione pubblica.

Il primo aspetto riguarda la costruzione delle narrazioni sui rischi pubblici dell'IA da parte delle aziende leader nel settore. Se da un lato queste iniziative consentono loro di occupare il dibattito pubblico e politico, dall'altro, sono un veicolo per legittimare presso i decisori pubblici priorità di agenda e soluzioni di policy. Il 2023 può essere preso a riferimento come l'anno degli appelli "apocalittici" sull'impatto che l'IA potrà avere sul futuro dell'umanità. In tal senso due importanti dichiarazioni (*statement*) hanno attirato l'attenzione del grande pubblico a livello globale. La prima è la lettera aperta pubblicata il 22 marzo dal *Future of Life Institute* (FLI) che auspicava una moratoria semestrale dello sviluppo dei sistemi di IA con particolare riferimento alle evoluzioni di sistemi di IA generativa simili a ChatGPT. La lettera²⁷ denunciava la diffusione di un approccio di innovazione irresponsabile: «negli ultimi mesi i laboratori di IA sono stati impegnati in una corsa fuori controllo per sviluppare e implementare menti digitali sempre più potenti che nessuno, nemmeno i loro creatori, può comprendere, prevedere o controllare in modo affidabile». Il clamore pubblico della lettera aperta è dipeso più dai suoi firmatari che dal suo breve contenuto. L'appello ha ricevuto l'adesione di Yoshua Bengio, vincitore

²⁷ Il testo integrale è disponibile al link: <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

nel 2018 del Turing Award²⁸, e di Stuart Russell, docente di informatica presso l’Università di Berkeley, una delle maggiori personalità scientifiche nello studio dei rischi dell’IA. Al terzo e quarto posto dell’elenco dei firmatari si trovavano rispettivamente Elon Musk, nelle vesti di CEO di Space X, Tesla e X, e Steve Wozniak, co-fondatore di Apple. Il secondo *statement* è stato lanciato dall’organizzazione no-profit *Center of AI Safety* (CAIS): «mitigare il rischio di estinzione causato dall’IA dovrebbe essere una priorità globale, insieme ad altri rischi su scala sociale come le pandemie e la guerra nucleare». Anche questo appello ha ricevuto il consenso di circa 350 ricercatori accademici, “altre figure di rilievo”, fondatori, amministratori delegati e ricercatori delle imprese leader del settore²⁹. Tra esse spiccano Google DeepMind – Gemini (Alphabet), OpenAI – ChatGPT, Microsoft e Meta AI, adesioni che, considerando anche il precedente appello, ricomprendono l’intero panorama delle Big Tech occidentali che operano nel campo dell’IA. Entrambi gli appelli sono stati lanciati da due importanti organizzazioni statunitensi con status formale no-profit, con sede nella Silicon Valley (San Francisco Bay Arena – California), due attori rilevanti sia nell’implementazione dei principi di *self-regulation* delle imprese³⁰ che nelle azioni di *lobbying* rivolte ai processi di regolazione pubblica dell’IA³¹. Sin dal 2020 il FLI è una delle organizzazioni più attive nell’ambito del processo regolativo europeo attraverso le sue numerose pubblicazioni rivolte ai decisori pubblici, gli incontri con la Commissione e parlamentari europei³². Un’azione che sta accompagnando l’intero processo dell’AI Act, iniziata nel 2020 con una lettera di supporto all’iniziativa della Commissione e che, come vedremo, prosegue nella fase di implementazione, con un gruppo di esperti indipendenti a supporto delle attività della Commissione. L’azione del CAIS è diretta prevalentemente verso gli Usa, dove esercita influenza attraverso le attività del *CAIS Action Fund*, organizzazione di *advocacy* che si pone l’obiettivo: «di promuovere politiche pubbliche che mantengono

²⁸ Il premio, considerato nel settore il “Premio Nobel per l’informatica”, è assegnato annualmente dall’Association for Computing Machinery ai ricercatori che più si sono distinti nel campo dell’informatica e consiste in un finanziamento di 1 mln di dollari basato sul sostegno economico di Google. Maggiori informazioni al link: <https://awards.acm.org/turing>

²⁹ L’elenco completo dei firmatari è disponibile al link: <https://www.safe.ai/work/statement-on-ai-risk#open-letter>

³⁰ Si pensi che nel 2017 il FLI ha coordinato l’elaborazione degli *Asilomar AI Principles*, una tra le prime iniziative di definizione di principi etici e valoriali da applicare trasversalmente allo sviluppo di sistemi di IA: <https://futureoflife.org/open-letter/ai-principles/>

³¹ Le due organizzazioni sono inserite nel Registro della Trasparenza dell’UE, insieme ai lobbyisti e a diverse forme di rappresentanza degli interessi che svolgono azioni di advocacy presso le istituzioni europee: https://transparency-register.europa.eu/index_en

³² L’elenco pubblico delle attività è disponibile al link: https://transparency-register.europa.eu/searchregister-or-update/organisation-detail_en?id=787064543128-10

gano la leadership degli Usa nel campo dell'IA e proteggano dalle minacce alla sicurezza nazionale legate all'IA»³³. *Advocacy* che riguarda anche aspetti nodali dello sviluppo dell'IA come la competizione con la Cina sulla produzione di *chip* e semiconduttori necessari alle industrie statunitensi. Un articolo di POLITICO del 23 febbraio 2024³⁴ ricostruisce con efficacia le relazioni tra questo tipo di organizzazioni, con particolare riferimento al CAIS e al *Center for AI Policy* (CAIP), e le grandi industrie del settore, da cui traggono buona parte dei loro finanziamenti. Secondo l'articolo, le narrazioni “apocalittiche” corredano e integrano la tradizionale attività di *lobbying* presso i decisori politici: «concentrandosi su scenari apocalittici, questi gruppi ben finanziati stanno creando barriere all'ingresso per le aziende di IA più piccole e (distolgono) l'attenzione da problemi più immediati e concreti legati alla tecnologia, come il suo potenziale di eliminare posti di lavoro o perpetuare la discriminazione». Allo stesso tempo, tali narrazioni legittimano complesse strategie di governo e sviluppo dell'IA dei cosiddetti “giganti”. Si pensi in tal senso al progetto per la creazione di una “*Superintelligence Strategy*” per evitare *escalation* catastrofiche, elaborato da D. Hendrycks (amministratore delegato di CAIS), E. Schmidt (ex CEO di Google) e A. Wang (CEO di Scale AI). La Strategia equipara i rischi dello sviluppo dell'IA alla corsa agli armamenti atomici del passato prefigurando una *escalation* tra le grandi potenze dovuta al dominio incontrastato di uno stato sugli altri, o alla perdita di controllo sulla propria IA “superintelligente” da parte di uno di essi. Uno scenario che richiederebbe un sistema di deterrenza tra gli stati simile a quello applicato all'uso delle armi atomiche e finalizzato, quindi, a bloccare condizioni emergenti di monopolio nazionale nel campo dell'IA. Tale meccanismo, denominato *Mutual Assured AI Malfunction* (MAIM), si baserebbe su attività di spionaggio, *cyberattacks* e financo su attacchi cinetici distruttivi ai data center nazionali³⁵. Una strategia che sembra offrire una conseguenzialità logica e operativa al breve *statement* del 2023.

Il secondo aspetto del ruolo delle imprese di cui ci occupiamo riguarda ulteriori modalità di costruzione di cornici di significato, i *policy frame*, da cui trarre legittimazione pubblica. Come sostenuto nel precedente paragrafo, se si intende governare lo sviluppo dell'IA nel suo complesso, più che le sue specifiche applicazioni, diviene necessario un esercizio di astrazione. L'obiettivo è la definizione di alcuni principi e valori inderogabili, quelli che gli antropologi

³³ Fonte: <https://action.safe.ai/#about>

³⁴ Fonte: <https://www.politico.com/news/2024/02/23/ai-safety-washington-lobbying-00142783>

³⁵ Maggiori informazioni sulla strategia sono disponibili al link: <https://www.nationalsecurity.ai/chapter/executive-summary>

chiamerebbero “universal culturali”³⁶, che possano essere applicati in modo trasversale ai sistemi di decisione autonoma nelle diverse fasi di progettazione, sviluppo e uso, in ambito pubblico e privato. In altre parole, può essere necessario definire collettivamente un’etica dell’IA³⁷. L’etica dell’IA è sia un vasto campo di studi multidisciplinari che si interrogano sulle implicazioni morali dei sistemi di decisione autonoma, che una finalità dalle molteplici implicazioni, su cui convergono le strategie di sviluppo e regolazione degli attori pubblici, i processi di innovazione tecnologica dei privati e l’azione di *watchdog* delle organizzazioni di società civile. Tuttavia, diversi studi hanno mostrato che, quando ci si approccia alla definizione del rapporto tra etica e IA, si rischia di assumere due posture interpretative complementari tra loro. Da un lato, il rapporto tra IA ed etica viene inteso solo in termini di analisi di problemi emergenti, ci si trova quindi a discutere di “questioni etiche” intese come cose moralmente sbagliate (Stahl, 2021). Dall’altro lato, gli esperti di etica tendono ad attribuire ai sistemi di IA proprietà morali che in realtà non hanno, interpretazioni che conducono talvolta a prevedere esiti evolutivi fantasiosi sulle azioni minacciose dei sistemi di decisione autonoma (Powers e Ganascia, 2020). Più che ragionare su un’IA etica e dei comportamenti morali dei cosiddetti “Agenti Morali Artificiali”, può essere utile partire dall’interpretazione dell’etica dell’IA come una disciplina che: «si occupa dei comportamenti morali degli esseri umani quando progettano, costruiscono, utilizzano e interagiscono con gli agenti di IA, nonché degli impatti associati dei robot sull’umanità e sulla società» (Siau e Wang, 2020, p. 76).

R. Segatori (2021) ricorda come, da un punto di vista sociologico, l’etica rappresenti un dispositivo composto da almeno quattro elementi: un insieme di principi che definiscono cosa è giusto o sbagliato; la credenza interiorizzata della necessità di seguire quei principi; l’esistenza di motivi pratici o ideali su cui si fonda l’adesione e, in ultima istanza, la presenza di un meccanismo sanzionatorio. Quest’ultimo può essere di tipo sociale o può essere istituzionalizzato in un sistema giuridico di regole e sanzioni. Questi meccanismi, combinati assieme, definiscono i binari della vita collettiva distinguendo tra le condotte ammesse e non ammesse (*ibidem*). Un approccio so-

³⁶ Essi sono elementi reperibili in ogni cultura come, ad esempio, la parentela o il rifiuto dell’incesto. Spesso indicano elementi formali, reificati dagli antropologi in oggetti autonomi presenti in ogni cultura e con caratteristiche indipendenti dai contesti culturali che li adottano. Per una sintesi del dibattito critico sul concetto di universale culturale si rimanda a: [https://www.treccani.it/enciclopedia/cultura_\(Universo-del-Corpo\)](https://www.treccani.it/enciclopedia/cultura_(Universo-del-Corpo))

³⁷ «L’etica dell’IA è una branca dell’etica applicata che studia gli interrogativi etici posti dallo sviluppo, dalla distribuzione e dall’utilizzo dell’IA. Il suo interesse principale risiede nell’individuare come l’IA possa favorire o mettere a rischio la felicità degli individui, sia in termini di qualità della vita che di autonomia umana e libertà necessarie per una società democratica» (Commissione Europea, 2019: 10).

ciologico all'IA consente quindi di analizzare discorsi e pratiche di etica dell'IA e, in particolare, delle Big Tech, che stabiliscono cosa è giusto e cosa è sbagliato, le motivazioni per cui tali principi debbano essere rispettati e quali siano gli strumenti più adeguati a garantirne la reale efficacia. Le norme etiche dell'IA offrono quindi un potente riferimento di legittimazione sia alle azioni delle imprese, che delle istituzioni e organizzazioni della società civile che intendono governare lo sviluppo del settore. Secondo quanto osservato da N. van Dijk e colleghi (2021) nel contesto europeo questa convergenza è l'esito di un lungo processo di "etificazione" (*ethification*) del settore delle tecnologie ICT, processo che si estende ora fino alle questioni e gli strumenti di governance dell'IA. La Commissione Europea ha progressivamente promosso un uso politico del concetto di etica, dapprima rafforzando il ruolo dei controlli etici nei programmi e progetti di ricerca e successivamente valorizzando la funzione consultiva delle raccomandazioni etiche degli esperti nei dispositivi di regolazione pubblica in ambito digitale (*ibidem*). In Europa, quindi, l'etica ha assunto una funzione di collante universale: «uno strumento di governance che mette in contatto tutti i diversi tipi di comunità scientifiche, industria e cittadini e aiuta a realizzare gli obiettivi politici» (van Dijk *et al.*, 2021, p. 6). Si pensi che, sin dal 2018, l'eticità è un elemento denotativo delle strategie della Commissione che intende promuovere un'IA "etica, sicura e all'avanguardia". Dunque, la riflessione etica è un elemento imprescindibile sia nella definizione di un quadro normativo del settore che nella promozione di un'innovazione responsabile, al fine di: «costruire e mantenere una cultura e una mentalità etiche attraverso il dibattito pubblico, l'istruzione e l'apprendimento pratico» (Commissione Europea, 2019, p. 10).

La produzione del "collante" etico, inteso quindi come *policy frame* condiviso, coinvolge direttamente le imprese e le Big Tech del settore che da quasi un decennio presidiano il campo dell'etica dell'IA attraverso diverse strategie e strumenti. Essi possono essere riassunti per semplicità in tre ambiti di azione esplorati sinteticamente di seguito: la definizione di principi etici dell'IA; gli strumenti di integrazione degli stessi nei processi aziendali; la loro proiezione esterna nelle arene regolative alle diverse scale, anche ai fini della elaborazione di standard pubblici. Nel 2018 Google è stata tra le prime Big Tech a pubblicare propri principi per un'IA etica: «in qualità di leader nel campo dell'IA, sentiamo la profonda responsabilità di farlo nel modo giusto. Per questo motivo, oggi annunciamo sette principi che guideranno il nostro lavoro in futuro. Non si tratta di concetti teorici, ma di standard concreti che guideranno attivamente la nostra ricerca e lo sviluppo dei prodotti e influenzano le nostre decisioni aziendali»³⁸. In estrema sintesi,

³⁸ Fonte: <https://blog.google/technology/ai/ai-principles/>

i sette principi ruotavano intorno all’obiettivo di sviluppare sistemi che potevano garantire benefici alle persone, nel rispetto della privacy, contrari alla riproduzione di forme di discriminazione e sotto il controllo delle persone, rispettando i più elevati standard di sicurezza. Nella lettera firmata da S. Pichai, CEO di Google, l’azienda definiva anche gli ambiti di sviluppo sottoposti a forme di autolimitazione, primi tra tutti i sistemi di IA che potrebbero causare danni alle persone. Una scelta di business su cui si tornerà nel paragrafo. Le intenzioni iniziali di Google riprendono e ampliano le frontiere delle già consolidate retoriche della responsabilità sociale d’impresa, basata su una forma di “*self-interest illuminato*” (Sacconi, 2004, p. 133), e remunerata dalla costruzione di fiducia e reputazione presso i propri clienti e partner. A solo un anno di distanza dalla pubblicazione dei Principi, Google affermava di aver pubblicato 79 articoli di ricerca sulle questioni etiche dell’IA e di aver sviluppato e rilasciato 12 strumenti tecnici *open source* in grado di identificare e rilevare *bias* e discriminazioni dei modelli di IA³⁹.

È questo un aspetto di particolare interesse, perché evidenzia la capacità di queste imprese di proporre soluzioni alternative alla regolazione pubblica, che depoliticizzano l’identificazione dei problemi, la loro risoluzione e il processo di valutazione circa l’efficacia dei loro stessi sistemi di regole. A partire dal 2018, anche gli altri colossi del settore hanno intrapreso percorsi di definizione della propria visione dell’etica dell’IA, attraverso l’adozione e l’aggiornamento continuo dei propri principi di un’IA etica (es. IBM nel 2021 e Microsoft nel 2022). La contesa per il primato tra politica e imprese nella definizione dell’etica dell’IA si è fatta fin da subito impari. Si tratta, in altre parole, della contesa tra regolazione *hard*, adottata principalmente dall’UE, e strumenti *soft* applicabili da chi detiene il controllo economico e tecnologico di un settore emergente, di cui è difficile stabilire con certezza i confini e financo le definizioni di base. È divenuto quindi agevole promuovere narrazioni che valorizzano un ruolo virtuoso delle imprese da contrapporre a vincoli, obblighi e divieti di legge: piuttosto che rincorrere la *compliance* a regole imposte dalla politica che potrebbero limitare l’innovazione, si lasci alle aziende la possibilità di auto monitorare le proprie scoperte tecnologiche. Non solo: la disponibilità di programmi *open source* può essere un volano per gli altri sviluppatori che necessitano di risorse più che di sanzioni. Parafrasando O. Romano (2023), l’IA non è di certo un settore esente dalla pervasiva scommessa provvidenzialista occidentale dello “scatenamento”, secondo la quale lasciando esprimere liberamente le forze produttive, ossia chi guida lo sviluppo, si sbucherà prima o poi nel migliore dei mondi. Si pensi che nel 2019 A. Jobin e i suoi colleghi del Politecnico di

³⁹ Fonte: <https://ai.google/static/documents/ai-principles-2019-progress-update.pdf>

Zurigo censivano a livello globale circa novanta documenti relativi a linee guida e principi etici dell'IA, elaborati in maggior parte da imprese (23%) e da agenzie governative (21%). Ai primi posti per occorrenze si collocavano i principi di trasparenza, giustizia ed equità, non-malignità, responsabilità e privacy; mentre, agli ultimi posti si ritrovavano i principi di sostenibilità (intesa anche come sostenibilità dei consumi energetici), dignità e solidarietà (*ibidem*)⁴⁰.

Il dibattito accademico ha discusso criticamente questo “diluvio” di principi e codici etici delle imprese, evidenziandone limiti, ambiguità e il riferimento astratto a valori socialmente desiderabili (es. Mittelstadt, 2019; Munn, 2023). Diverse ricerche empiriche hanno avvalorato tali interpretazioni attraverso l'analisi dell'implementazione dell'etica nei processi aziendali (es. Seppälä *et al.*, 2021; Ali *et al.*, 2023); studi che hanno evidenziato come essa non sia realmente integrata nelle pratiche aziendali e quanto, piuttosto, rimanga un impegno “sulla carta” affidato a singoli dipendenti lasciati a fare i conti con le questioni di potere e le inerzie delle organizzazioni. Come già osservato in campo ambientale (*greenwashing*) o nelle questioni di genere (*pinkwashing*), anche le Big Tech IA sono state accusate di perseguire una manipolazione concettuale dell'etica riassumibile in una strategia “*ethicswashing*” (Wagner, 2018) e nella sua “banalizzazione” (Bietti, 2020). L'etica dell'IA sarebbe dunque il perno di una strategia di costruzione della reputazione aziendale, incremento dei profitti e dell'integrità aziendale (Schultz *et al.*, 2024).

Allo stesso tempo però, nonostante le critiche, continua l'incessante lavoro di definizione di principi etici dell'IA anche sulla base del consolidamento di un'alleanza strategica tra accademia e Big Tech (Ochigame, 2019). È un lavoro che impegna esperti, personalità accademiche e ricercatori di diverse discipline, che rielaborano, aggiornano e sintetizzano i principi etici esistenti “sul mercato”. Sebbene la loro finalità sia la riduzione della confusione e del sovraccarico informativo, secondo i critici è un'attività che può stimolare forme di “shopping dell'etica” (van Maanen, 2022) a cui i soggetti interessati possono attingere per definire il repertorio valoriale da applicare alle proprie strategie di business. Piuttosto, è necessario porre la questione del passaggio da *what to how* (Morley *et al.*, 2020), ovvero la traduzione dei principi etici in concrete pratiche aziendali. R. Ortega-Bolanos e colleghi

⁴⁰ Nel 2019 l'OECD ha provato a mettere ordine nel panorama emergente dell'etica dell'IA delineandone un primo framework valoriale, auspicabilmente applicabile a livello intergovernativo, fondato sui principi di “crescita inclusiva, sostenibilità economica e benessere”, “diritti umani e valori democratici”, “trasparenza e spiegabilità”, “robustezza e sicurezza” e “accountability”. Il documento aggiornato al 2024 è disponibile al link <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

(2024) hanno sostenuto che questa transizione può essere raggiunta attraverso una molteplicità di strumenti ibridi, teorici e tecnici (es. standard, libri, analisi di casi studio, corsi di formazione, *checklists*, software, assistenti vocali, apps), che, combinati assieme, forniscono agli sviluppatori e alle imprese, da un lato, le necessarie nozioni teoriche sul problema dell’etica e, dall’altro, le soluzioni pratiche per valutare il proprio operato.

A tal proposito, i Comitati Etici sono una soluzione “non tecnica” per un IA affidabile già suggerita alle aziende dal Gruppo di Esperti di Alto Livello della Commissione Europea (2019). I comitati esterni, etici, indipendenti, secondo le diverse definizioni possibili, afferiscono alla dimensione organizzativa delle imprese (governance nel linguaggio degli esperti) e possono avere sia una funzione di controllo che di consulenza. La loro azione è rivolta al monitoraggio interno e alla promozione esterna dell’eticità dei propri avanzamenti tecnologici. I comitati etici dell’IA non rappresentano una particolare innovazione in ambito organizzativo quanto, piuttosto, una riproposizione in chiave di *self-regulation* di strumenti istituiti per legge, ad esempio, nella ricerca in ambito medico e farmacologico⁴¹. Nel tempo questi strumenti hanno attirato numerose critiche riguardo la loro incapacità di assolvere il proprio ruolo sociale e per l’essere talvolta percepiti come ostacoli burocratici alla ricerca scientifica (Parziale, 2023). Tuttavia, nell’ambito dell’IA lo strumento dei Comitati Etici è ancora ad uno stato nascente, oltre che a essere recentemente entrato nella strumentazione della politica locale, attraverso un processo guidato dalle grandi imprese del settore. Un articolo apparso nell’estate 2022 sull’*Harvard Business Review* rivolgeva alle imprese l’appello di dotarsi di Comitati Etici come un elemento cruciale di una strategia etica aziendale di prevenzione dei rischi dell’IA. Un appello che insisteva anche su una genuina finalità di remunerazione reputazionale: «quando a un comitato viene conferito un potere reale, si può instaurare un rapporto di grande fiducia con i dipendenti, i clienti, i consumatori e le altre parti interessate dell’azienda, come il governo, soprattutto se l’organizzazione è trasparente riguardo alle operazioni del comitato, anche se non riguardo alle sue decisioni esatte»⁴².

Uno studio delle Università di Oxford e Stanford (Schuett *et al.*, 2024) si è interrogato sulle molteplici forme che questi strumenti (*Board*) possono assumere basandosi sulle prime esperienze delle grandi compagnie del settore (es. Microsoft e Google - DeepMind). Lo studio ha evidenziato il carattere strategico dello strumento attraverso l’analisi delle responsabilità che gli

⁴¹ Si pensi che le prime riflessioni organiche sul ruolo dei Comitati Etici in ambito medico in Italia risalgono al 1992 ad opera del Comitato Nazionale per la Bioetica e la prima regolazione a livello nazionale al 1998 con l’intervento del Ministero della Salute.

⁴² Fonte: <https://hbr.org/2022/07/why-you-need-an-ai-ethics-committee>

sono attribuite. Piuttosto che esercitare funzioni di verifica e controllo dei progetti in corso, i Comitati svolgono sovente attività di consulenza strategica rivolta ai vertici aziendali attraverso l’aggiornamento sulle innovazioni dell’IA, l’esercizio di influenza sui programmi di ricerca aziendali e financo il supporto delle strategie di commercializzazione dei prodotti (*ibidem*). Sono organismi che necessitano comunque della presenza di personalità di alto livello sul lato sia delle competenze tecniche, sia del prestigio necessario a legittimare l’organismo presso i competitor, i regolatori pubblici, l’ampia platea degli stakeholder e dei clienti. Particolarmente significativo è stato il fallimento del *Advanced Technology External Advisory Council* istituito da Google nel 2019. Il Consiglio, ideato come struttura esterna all’azienda, era composto di otto personalità accademiche e del modo dell’impresa⁴³ che avrebbero dovuto aiutare Google nell’implementazione dei suoi principi etici. Tuttavia, le controversie emerse nel dibattito pubblico su alcune delle nomine del Consiglio, avvalorate anche da una petizione degli stessi dipendenti di Google, hanno costretto l’azienda a porre fine a questa esperienza dopo poche settimane dal suo annuncio. L’opportunità di alcune candidature, la composizione interna, oltre che gli obiettivi costitutivi, dei Comitati Etici delle grandi compagnie dell’IA rappresentano quindi un altro terreno di ambiguità e critica circa l’azione delle aziende nel campo dell’etica. Per questo, nel dibattito scientifico stanno emergendo proposte per garantire a questi strumenti un reale pluralismo degli interessi, con l’obiettivo di fugare i dubbi riguardo il loro operato attraverso, ad esempio, processi di selezione casuale dei membri mutuati dalla democrazia deliberativa e dalle esperienze delle assemblee dei cittadini estratti a sorte (Conti e Seele, 2023).

Il terzo oggetto di riflessione riguarda il rapporto tra imprese e politica nel campo dei processi di regolazione dell’IA. In tal senso è utile la distinzione tra etica “dura” e “morbida” (*hard and soft ethics*) suggerita da L. Floridi (2018). L’etica dura riguarda i valori, i diritti e doveri, le responsabilità, ciò che è giusto fare o no, elementi che sono considerati nei processi regolativi: un’etica dura è quindi quella che orienta la formulazione di nuove leggi o il loro cambiamento. L’etica morbida si riferisce invece ad una «*post-compliance ethics*» (p. 5): ciò che dovrebbe essere o non essere fatto al di là della regolazione. Come sostiene U. Pagallo (2020), questa concettualizzazione non si discosta di molto dall’idea di *soft law* e può ricondurre l’etica a funzione complementare dei quadri regolativi esistenti o, nel caso dell’IA, emergenti. L’etica morbida è quindi un riferimento condiviso che si ritrova sia

⁴³ L’annuncio e le biografie degli otto membri sono disponibili al link <https://blog.google/technology/ai/external-advisory-council-help-advance-responsible-development-ai/>

negli strumenti di *soft law* pubblica che nelle iniziative di autoregolazione delle imprese. Diversi sono gli esempi di pratiche basate su un'etica *post-compliance* o, meglio, senza *compliance*, con cui – di fatto – le imprese intervengono nel dibattito pubblico e politico rivendicando la preminenza del proprio ruolo e dei principi aziendali nelle arene di governance dell'IA. Nel 2019 suscitò particolare clamore la lettera rivolta al Congresso degli Usa da A. Krishna, allora amministratore delegato di IBM, in cui affermava l'interruzione della fornitura di tecnologie di riconoscimento facciale per le operazioni di polizia: «IBM si oppone fermamente e non tollererà l'uso di alcuna tecnologia [...] per la sorveglianza di massa, la profilazione razziale, le violazioni dei diritti umani e delle libertà fondamentali o per qualsiasi scopo non coerente con i nostri valori e i nostri principi di fiducia e trasparenza»⁴⁴. La lettera rappresentò un intervento inedito e deciso da parte di una grande impresa sul tema del riconoscimento facciale, uno degli usi più controversi dell'IA in ambito pubblico e privato, su cui non si è ancora raggiunto un bilanciamento tra potenzialità e rischi e su cui anche le Big Tech hanno ripetutamente mostrato prudenza.

L'idea di perseguire un'IA etica attraverso un approccio “morbido” si ritrova nell'importante iniziativa *Rome Call for AI Ethics*, intrapresa dalla Pontificia Accademia per la Vita (Città del Vaticano). La “chiamata”, lanciata nel 2020, ambisce alla costruzione di un'Algoretica, neologismo che sta a indicare un approccio di etica *by design*, per l'appunto senza il riferimento alla necessità di forme di regolazione attraverso leggi e regolamenti. Microsoft e IBM sono state tra le prime firmatarie della Call, seguite nel 2024 da Cisco System, colosso multinazionale nel campo dei dispositivi per la connessione, da numerose università pubbliche e private di diversi Paesi e dai rappresentanti di undici religioni abramitiche e dell'Estremo Oriente. Dunque, ancora una volta un'iniziativa che pone al centro i principi etici e la loro integrazione volontaria nei processi di design e sviluppo degli algoritmi di IA. È questa la strada percorsa negli Usa dall'amministrazione Biden-Harris che nel luglio 2023 ha teso la mano alle Big Tech attraverso l'emanazione di otto impegni volontari (*voluntary commitments*)⁴⁵ volti a uno sviluppo sicuro e trasparente dei sistemi di IA. Gli impegni avevano come obiettivi prioritari la sicurezza e la creazione di fiducia presso il pubblico. La sicurezza era declinata in relazione sia ai rischi pubblici, sia alla vulnerabilità dei prodotti di IA rispetto ad attacchi esterni. La sicurezza è stata oggetto di un secondo intervento dell'amministrazione Biden-Harris, più sul versante *hard*

⁴⁴ Fonte: <https://www.ibm.com/policy/facial-recognition-sunset-racial-justice-reforms/>

⁴⁵ L'elenco dei principi è disponibile al link: <https://bidenwhitehouse.archives.gov/wp-content/uploads/2023/07/Ensuring-Safe-Secure-and-Trustworthy-AI.pdf>

law, con l'*Executive Order – Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence* (30 ottobre 2023) con cui sono stati definiti alcuni obblighi di comunicazione al governo federale per le aziende in pro-cinto di sviluppare modelli di IA che superino determinate soglie di potenza di elaborazione, ponendo potenziali rischi per la sicurezza nazionale. Gli altri “*commitment*”, stavolta finalizzati alla costruzione di un clima di fiducia, chiedevano alle aziende di impegnarsi nello sviluppo di filigrane digitali per riconoscere i contenuti prodotti dall’IA, dare priorità a studi relativi ai rischi sociali e allo sviluppo di prodotti che abbiano una ricaduta sociale come le tecnologie diagnostiche e mediche impiegate nella lotta al cancro. L’iniziativa ha raccolto da subito il consenso delle “sette sorelle” dell’IA – Amazon, Anthropic, Google, Inflection, Meta, Microsoft, OpenAI – che si sono affrettate a comunicare la loro adesione agli impegni, privi però di reali strumenti di *enforcement* e già in linea con alcune delle attività di test precedentemente intraprese dalle aziende⁴⁶. A un anno di distanza diversi analisti di settore avevano evidenziato la parzialità dei successi ottenuti attraverso gli impegni volontari e l’intenzione dell’amministrazione Biden di superarli arrivando a definire una legislazione federale più stringente⁴⁷.

Come vedremo, l’elezione di D. Trump nel 2024 ha rapidamente cambiato questa impostazione regolativa degli Usa, con la revoca dell’ordine esecutivo e l’abbandono delle politiche di mitigazione dei rischi dell’amministrazione Biden, mostrando la fragilità politica e giuridica degli approcci ispirati a un’etica morbida. Questo cambiamento repentino nella strategia statunitense ben si connette con la revisione di alcune importanti scelte aziendali delle Big Tech sul tema dell’etica. Ci riferiamo, in particolare, a due eventi di inizio 2025: l’eliminazione dai principi etici di Google del divieto di sviluppare armi e dispositivi di sorveglianza; le decisioni di Meta sulla libertà di espressione sui social e sul problema della disinformazione. Il cambiamento di direzione di Google è stato velato dall’azienda attraverso il riferimento alla volontà di promuovere collaborazione tra imprese, governi e organizzazioni di diverso tipo sugli obiettivi di protezione delle persone, crescita globale e sicurezza nazionale⁴⁸. Una scelta che ha attirato numerose critiche da parte di organizzazioni transnazionali che operano nel campo dei diritti umani, che insistono perciò sulla necessità di restituire alla politica il primato della regolazione dell’IA

⁴⁶ Si pensi a titolo di esempio alle azioni “Red Teaming”, attività in cui gli sviluppatori simulano attacchi o usi malevoli da parte di soggetti esterni per prevedere le reazioni del sistema di IA.

⁴⁷ Si veda a titolo di esempio

<https://www.technologyreview.com/2024/07/22/1095193/ai-companies-promised-the-white-house-to-self-regulate-one-year-ago-whats-changed/>

⁴⁸ Fonte: <https://blog.google/technology/ai/responsible-ai-2024-report-ongoing-work/>

e di alcune delle sue applicazioni “sensibili”, denunciando l’effetto di facciata della *self-regulation* delle imprese⁴⁹. La seconda capovolta delle Big Tech, paradigmatica del nuovo corso, riguarda la decisione di Meta – Facebook di rinunciare negli Usa al sistema di verifica esterno e indipendente della veridicità dei contenuti sulle proprie piattaforme (*fact checking*)⁵⁰. Tale scelta è stata motivata da Meta attraverso la critica agli effetti perversi del precedente sistema: «negli ultimi anni abbiamo sviluppato sistemi sempre più complessi per gestire i contenuti su tutte le nostre piattaforme, in parte in risposta alla pressione sociale e politica per moderare i contenuti. Questo approccio è andato troppo lontano [...] fino al punto in cui stiamo commettendo troppi errori, frustrando i nostri utenti e troppo spesso ostacolando la libertà di espressione che ci siamo prefissati di consentire»⁵¹. Un cambiamento di approccio in sintonia con l’agenda della nuova presidenza di D. Trump, che pure era stato sospeso da Meta nel 2021 dopo i post di commento pubblicati durante l’assalto al Congresso da parte degli estremisti di destra⁵².

La “via dell’etica” sembra avere quindi una parabola discendente da un lato per il disvelamento della sua inefficacia e delle sue contraddizioni, dall’altro lato per il venire meno dello stesso sforzo delle Big Tech di promuovere soluzioni di etica morbida attraverso *self-regulation*. Questa scelta si inscrive nell’ambito di un riposizionamento a “destra” dei giganti tecnologici americani che, in alleanza con la nuova amministrazione Trump, propendono ora per l’opposizione a qualsiasi forma di regolazione politica dell’IA. È dunque necessario guardare a cosa sta accedendo nell’ambito della regolazione dura, l’altro estremo della nostra dicotomia ideal-tipica delle forme di regolazione dell’IA, e a come essa è elaborata e implementata alle diverse scale del potere politico: transnazionale (par. 2.3), sovranazionale (par. 3.2) e nazionale (paragrafi 3.1 e 3.3).

⁴⁹ Fonte: <https://www.amnesty.org/en/latest/news/2025/02/global-googles-shameful-decision-to-reverse-its-ban-on-ai-for-weapons-and-surveillance-is-a-blow-for-human-rights/>

⁵⁰ In sintesi, il precedente sistema di verifica dei contenuti si basava su verificatori terzi, indipendenti da Meta, che si occupano del tema della disinformazione online, sostituito dal metodo delle “note di comunità” con cui gli utenti delle piattaforme possono intervenire direttamente segnalando post e contenuti ritenuti fuorvianti o confusi. Un sistema che è stato già adottato da X. Maggiori informazioni sono disponibili al link: <https://transparency.meta.com/it-it/features/community-notes/>

⁵¹ Fonte: <https://about.fb.com/news/2025/01/meta-more-speech-fewer-mistakes/>

⁵² La ricostruzione dei fatti è disponibile al link:

<https://www.oversightboard.com/news/226612455899839-oversight-board-upholds-former-president-trump-s-suspension-finds-facebook-failed-to-impose-proper-penalty/>

2.3. Le arene politiche internazionali

Nella prima sezione del volume abbiamo ricostruito gli elementi costitutivi della dirompenza globale dell'IA relativi ai suoi rischi, potenzialità e poste in gioco. Tali questioni chiedono alla politica di individuare scale di governo che possano produrre quadri regolativi adeguati. L'autoregolazione delle imprese è, come abbiamo visto, spesso insufficiente, contraddittoria e finalizzata soprattutto alla reputazione aziendale nei confronti dei clienti e dei decisori politici. Come vedremo nel prossimo capitolo, il più ambizioso tentativo di regolazione sovranazionale, l'EU AI Act, riguarda una sola regione continentale e anche la regolazione sinora prodotta dagli stati è, nell'insieme, diversificata e frammentata. Dunque, anche se nei discorsi pubblici si parla delle stesse cose, le regolazioni sinora prodotte tutelano i diritti in modo diverso e rendono anche difficile per le Big Tech adeguarsi ai diversi ambienti regolativi che frequentano. Pertanto, anche eventi come il lancio dell'IA generativa di OpenAI nel novembre del 2022 hanno rilanciato la discussione scientifica, politica e imprenditoriale sulla necessità e, allo stesso tempo, le difficoltà di pensare un governo globale dell'IA. In queste discussioni sono ritenute necessarie cornici di regole e strumenti *interoperabili*, ossia capaci di interagire in modo efficace fra loro nell'interazione fra luoghi e scale diversi (Cacciari, 2023). Non sorprende, quindi, che l'IA sia oggetto di iniziative politiche all'interno di sistemi di relazioni internazionali. Ma quali sono, a quali agende hanno dato luogo, con quali risultati e come stanno funzionando? E da cosa dipendono i successi e gli insuccessi sinora registrati?

Nel caso dell'IA non esiste un'agenzia internazionale specializzata per la definizione di regole e standard comuni⁵³, piuttosto si assiste alla formazione di un embrione di un regime, o arena di governance transnazionale (d'Albergo e Moini, 2024) dell'IA di cui proviamo ad isolare e analizzare alcuni attori ed eventi chiave (cfr. fig. 1, p. 93). Questo regime si sta configurando come un'arena di arene, sedi di dialogo nelle quali governi, imprese ed esperti ricercano un consenso minimo su definizioni condivise delle potenzialità e dei rischi dell'IA e su visioni relative a norme non vincolanti che inseriscono in accordi continuamente negoziati e in costante aggiustamento. Alcune di queste arene sono espressione di sistemi di azione multilaterali preesistenti, nelle cui agende ora è entrata anche l'IA (le UN, i G7, G20, l'OECD, il Consiglio d'Europa, il *World Economic Forum*). Altre arene si sono formate espressamente in relazione all'IA (*AI Safety summit*; *Global Partnership on Artificial Intelligence*).

⁵³ Un esempio in tal senso è la International Civil Aviation Organization, che assicura a 193 Paesi standard, procedure e misure di regolamentazione comuni per la sicurezza del traffico aereo.

Ulteriori iniziative sono state lanciate da organizzazioni di ricerca e think tank come, ad esempio, il *Forum for Cooperation on AI*, una collaborazione fra la *Brookings Institution* e il *Centre for European Policy Studies*⁵⁴.

La prima arena che esploriamo è quella che le Nazioni Unite (UN) stanno provando a costruire. Una precisazione necessaria: le UN⁵⁵ non possono approvare norme riguardanti l'IA e tantomeno attuarle. Tuttavia, possiamo identificare tre direttive dell'azione delle UN con cui l'organizzazione prova a costruire il proprio ruolo di costruttore di un regime, o arena di governance transnazionale: definizione di strumenti di *soft law* per lo sviluppo e l'uso dell'IA su cui costruire consenso tra governi, imprese e attori della conoscenza; legittimazione come *convenor* globale; definizione di strumenti per la costruzione di una governance globale inclusiva. In riferimento alla direttrice *soft law* varie agenzie e organi delle UN hanno commissionato studi e formulato raccomandazioni per lo sviluppo e la codificazione dell'IA. Il primo intervento del suo *High-level Panel* di esperti, istituito nel 2017, ha riguardato la questione della frammentazione dei documenti contenenti principi etici e di governo dell'IA, per la cui soluzione è stato proposto un approccio *multi-stakeholder* (con il coinvolgimento di più attori, statali e non), raccomandazioni formalizzate nel 2020 dal Segretario generale⁵⁶. Nello stesso anno, l'Assemblea generale ha approvato risoluzioni (non vincolanti) per la limitazione delle armi autonome letali guidate da IA, richiesta avanzata anche da reti e campagne di ONG. Nel 2021 il commissario UN per i diritti umani ha chiesto uno stop allo sviluppo di strumenti di IA, se non sono tutelati diritti di vario tipo⁵⁷. Nel 2022 le UN hanno segnato un altro passo nella “via dell'etica” attraverso l'approvazione di 10 principi, fondati sull'etica e sui diritti umani, per guidare l'uso dell'IA all'interno delle entità del sistema UN. Nel 2023, dopo il lancio di Chat GPT, il Consiglio di Sicurezza ha proposto tre principi cardine su rischi e opportunità dell'IA: rispetto delle norme umanitarie, accordo sull'uso in ambito militare e per attività terroristiche. In questo periodo, molte agenzie specializzate delle UN hanno inserito nelle loro agende una regolamentazione dell'IA nei settori di cui si occupano, come l'istruzione, o la salute⁵⁸.

Il Segretario generale ha chiesto nel 2023 a un *High-level multi-stakeholder*

⁵⁴ Fonte: <https://www.brookings.edu/projects/the-forum-for-cooperation-on-artificial-intelligence/>

⁵⁵ Le Nazioni Unite sono un'organizzazione internazionale fondata nel 1945. Attualmente è composta da 193 Stati membri (<https://www.un.org/en/about-us>).

⁵⁶ Fonte: <https://documents.un.org/doc/undoc/gen/n20/102/51/pdf/n2010251.pdf>

⁵⁷ Fonti: <https://news.un.org/en/story/2021/09/1099972>; <https://www.un.org/en/ai-advisory-body>

⁵⁸ Nel complesso, ci sono 27 strumenti delle UN vincolanti in ambiti specifici che possono essere applicati all'IA e 29 non vincolanti.

Advisory Body di produrre analisi e raccomandazioni su governance internazionale dell'IA, basate sulla comprensione di rischi e sfide e del suo possibile uso per i SDGs dell'Agenda 2030⁵⁹. Il rapporto finale (*Governing AI for Humanity*, 2024), redatto dopo una consultazione con 2.000 partecipanti, propone di affrontare il deficit della governance globale dell'IA, in cui “la pratica smentisce la retorica”, attraverso meccanismi istituzionali inclusivi e leggeri, per integrare gli sforzi già esistenti e promuovere accordi agili e adattivi, capaci di tenere il passo con l'evoluzione dell'IA⁶⁰. Un lavoro da realizzare attraverso la condivisione delle conoscenze e degli standard adottati – ben 333 fra il 2018 e il 2024 – le reti di esperti, la creazione di un fondo globale per l'IA, il coordinamento delle attività da parte di un ufficio IA nel segretariato delle UN e la costituzione di un'agenzia apposita delle UN.

Dall'interno del sistema UN (2024) è stato inoltre proposto un contributo per deliberazioni coordinate delle UN sull'IA, nella forma di un Libro Bianco⁶¹, che raccomanda di usare il potere “di convocazione” (*convening power*) delle UN per creare consenso, promuovendo la ricerca, il coinvolgimento e il coordinamento delle parti interessate, la definizione di standard e norme condivisi, lo sviluppo di capacità, il monitoraggio. Gli strumenti da usare sono sia incentivi, sia barriere di protezione commisurati alle caratteristiche dei diversi sistemi e applicazioni di IA. Inoltre, deve essere tenuta in considerazione l'attivazione delle organizzazioni settoriali delle UN: 40 agenzie delle UN hanno definito 280 progetti con l'IA al servizio dei SDGs e l'UNESCO, in particolare, ha adottato la sua Raccomandazione sull'etica dell'IA (2021) e sull'uso nel campo nell'istruzione, l'uguaglianza di genere, lo sviluppo di capacità per i governi e il sistema giudiziario⁶².

La seconda direttrice dell'azione delle UN nella costruzione di un regime di governance transnazionale è rivolta alla comunità epistemica dell'IA per l'apertura di spazi di discussione. La più importante iniziativa in tal senso riguarda l'Unione Internazionale delle Telecomunicazioni (ITU), agenzia UN che promuove standard globali per le telecomunicazioni, che nel 2017 ha lanciato il programma *AI for Good*,⁶³ per indirizzare il potenziale trasformativo dell'IA verso i SDGs. *AI for Good* è una piattaforma di discussione e scambio di conoscenze ed esperienze su scala globale fra agenzie UN, ricercatori, professionisti delle ICT, per facilitare la ricerca di un *framework* globale e di

⁵⁹ La commissione era composta di 38 esperti di governi, imprese e società civile (inclusi russi, cinesi, di Paesi arabi, Israele, Vaticano, Microsoft, Sony, Google-Alphabet, OpenAI), <https://www.un.org/ai-advisory-body>.

⁶⁰ Fonte: <https://unseeb.org/topics/artificial-intelligence>

⁶¹ Fonte: <https://unseeb.org/united-nations-system-white-paper-ai-governance>

⁶² Fonte: <https://www.unesco.org/en/artificial-intelligence>

⁶³ Fonte: <https://aiforgood.itu.int>

soluzioni tecniche per i problemi legati alla dimensione etica dell’uso di IA. I suoi principali strumenti di discussione consistono nel summit annuale e nei webinar quasi quotidiani. La piattaforma aspira ad essere *convenor* autorevole a livello globale perché il “marchio” UN può indurre fiducia in questa varietà di attori. Nel 2025, in occasione del *AI Action Summit* di Parigi (febbraio 2025, cfr. oltre) le UN si sono ulteriormente offerte, anche attraverso ITU, come forum inclusivo per riunire esperti e industrie per lo sviluppo di una rete globale per un’IA interoperabile e condivisa su scala globale.

L’ultima direttrice di azione delle UN riguarda più direttamente la definizione delle precondizioni e degli elementi necessari per un governo globale dell’IA quali un dialogo politico, l’interoperabilità dei dati e il superamento dei divari digitali tra i diversi Paesi. Sono questi gli obiettivi del *Global Digital Compact*, un più “ampio quadro per la governance globale della tecnologia digitale e dell’IA (...) basata su approccio multistakeholder” adottato dalle UN nel 2024. Questa iniziativa intende ovviamente trovare un equilibrio tra massimizzazione dei benefici dell’IA e mitigazione dei suoi rischi, ancora una volta standard e codici di condotta interoperabili su sicurezza, affidabilità, sostenibilità e diritti umani⁶⁴. La costruzione di una governance globale implica anche il perseguitamento di una difficile coabitazione nelle arene internazionali di attori, stati “giganti” dell’IA, che sono in competizione tra loro. Si pensi in tal senso che, a sostegno delle iniziative UN, si schiera sovente la Cina, che lamenta l’esclusione da altri tavoli internazionali (cfr. oltre). A poca distanza temporale dal *Global Digital Compact*, la Cina ha proposto un *AI Capacity-Building Action Plan for Good and for All* per contribuire a colmare il divario globale in materia di IA, dichiarandosi pronta a collaborare con altri Paesi, soprattutto con quelli in via di sviluppo in campo di innovazione tecnologica e “per rendere sinergiche le strategie di IA e rafforzare gli scambi di politiche, condividere politiche e le pratiche tecniche in materia di test, valutazione, certificazione e regolamentazione dell’IA e affrontare congiuntamente i rischi dell’IA legati all’etica, alla sicurezza e alla protezione”⁶⁵.

La seconda e terza arena di costruzione di un regime di governance transnazionale dell’IA possono essere identificate, rispettivamente, nelle azioni dei gruppi dei G7 e dei G20. Dopo il summit del 2018 (in Canada), in cui è stata dichiarata una visione comune per il futuro dell’IA orientata a promuovere

⁶⁴ Fonti:

<https://www.un.org/digital-emerging-technologies/global-digital-compact>;
https://www.un.org/digital-emerging-technologies/sites/www.un.org.techenvoy/files/general/GDC_Rev_3_silence_procedure.pdf

⁶⁵ Fonte: https://www.mfa.gov.cn/eng/wjbzhd/202409/t20240927_11498465.html

fiducia nella nuova tecnologia, come condizione per il suo sviluppo⁶⁶, i governi dei G7⁶⁷ hanno preso l'iniziativa innescando azioni di reti di attori più ampie (Cameron *et al.*, 2025). Hanno avviato dunque il “Processo dell'IA di Hiroshima” (2023), con gli *International Guiding Principles for Advanced AI system*, un codice di condotta per un'IA “sicura, protetta e affidabile in tutto il mondo”⁶⁸ con l'adesione di 49 governi oltre i G7⁶⁹. L'accordo raggiunto si basa sui principi guida già proposti dall'OECD e la collaborazione con la GPAI (cfr. sotto). Propone inoltre un codice di condotta volontario per le organizzazioni che sviluppano, distribuiscono e utilizzano sistemi avanzati di IA. Queste dovrebbero impegnarsi a mitigare i rischi e gli abusi, identificare le vulnerabilità, condividere responsabilmente le informazioni, segnalare gli incidenti, investire nella sicurezza informatica, e conformarsi a un sistema di etichettatura per consentire agli utenti di identificare i contenuti generati dall'IA. Il mondo accademico, il no-profit e il settore pubblico utente di IA sono invitati a seguire un approccio basato sul rischio. Anche i governi dovrebbero adottare misure per identificare, valutare e mitigare i rischi, individuare le forme di uso malevolo, condividere informazioni e segnalare incidenti, sviluppare e usare standard tecnici e *best practice* internazionali.

Nel 2024⁷⁰ le autorità per la concorrenza dei G7, richiamando le linee guida di Hiroshima, hanno delineato interventi regolativi finalizzati a ridurre i rischi per le imprese negli ecosistemi economici dell'IA: promuovere condizioni favorevoli alla concorrenza, incoraggiare nuovi ingressi, innovazione e pratiche responsabili e scoraggiare comportamenti anticoncorrenziali, fraudolenti o collusivi, come l'acquisizione di posizioni dominanti e la preclusione all'accesso di *input* tecnologici chiave. Le autorità ritengono infatti che i rischi per la competizione nei mercati IA “possano riversarsi su altri aspetti chiave delle nostre società”. Si tratta da un lato di inviti fatti alle imprese, dall'altro dell'indicazione di strumenti di regolazione e attuazione, che adattano alla vigilanza antitrust dell'IA obiettivi e strumenti già adottati per altri mercati. Un'azione in linea con la Commissione UE, che ha esortato a sua

⁶⁶ Fonte: https://www.international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-06-09-artificial-intelligence-artificielle.aspx?lang=eng

⁶⁷ Il Gruppo dei Sette (G7) è stato fondato nel 1975 e comprende Canada, Francia, Germania, Italia, Giappone, UK e Usa. Partecipa anche l'UE (<https://www.g7italy.it/it/che-cose-il-g7/>)

⁶⁸ Fonti: <https://www.mofa.go.jp/files/100573471.pdf>;
<https://www.soumu.go.jp/hiroshimaiprocess/en/index.html>

⁶⁹ La lista degli aderenti è disponibile al link
<https://www.soumu.go.jp/hiroshimaiprocess/en/supporters.html>

⁷⁰ Si veda: <https://en.agcm.it/dotcmsdoc/pressrelease/G7%202024%20-%20Digital%20Competition%20Communiqué.pdf>

volta le autorità sulla concorrenza dei Paesi membri a usare le norme europee esistenti per contrastare l’acquisizione di posizioni dominanti da parte delle imprese di IA attraverso le cosiddette “acquisizioni killer”⁷¹.

La tutela della concorrenza è oggetto anche di accordi specifici, come quello sul mercato dei *foundation model* di IA generativa fra le autorità di Usa, UK e UE (2024)⁷². In un momento decisivo, in cui l’IA inizia a trasformare significativamente mercati e società, le autorità si impegnano a vigilare, ciascuna con i suoi strumenti istituzionali, sui rischi per la concorrenza derivanti dalla concentrazione del controllo di *input chiave* (*chips*, dati...), da forme di controllo finanziario o restrizione delle facoltà di scelta da parte di consumatori e imprese, anche esercitate attraverso l’IA.

La presenza nel gruppo dei G20⁷³ di Cina e Russia rende le sue attività sui temi dell’IA particolarmente interessanti. Nel 2019 i 20 ministri del commercio e dell’economia digitale hanno adottato una *Ministerial Statement on Trade and Digital Economy*, in cui si impegnano a “fornire un ambiente favorevole all’IA incentrata sull’uomo che promuova l’innovazione e gli investimenti, (...) con uso responsabile e di aiuto per gli SDGs”. Si tratta di raccomandazioni non vincolanti, basate sui principi di crescita inclusiva, sviluppo sostenibile e benessere, “valori umani ed equità, trasparenza e comprensibilità, solidità, sicurezza, protezione e responsabilità”. Una governance favorevole all’innovazione dovrebbe essere innovativa essa stessa: per ottenere standard interoperabili e cooperazione normativa fra i Paesi del G20, “sosteniamo la condivisione di buone pratiche, insieme a standard tecnici globali guidati dall’industria e dal mercato che (...) contribuiscono a garantire l’interoperabilità e promuovere fiducia”⁷⁴. I “Principi per una gestione responsabile di un’IA affidabile” adottati contestualmente rispecchiano sostanzialmente quelli già globalmente affermati (IA incentrata sull’uomo, trasparenza e comprensibilità, robustezza e sicurezza, accountability...). I governi nazionali dovrebbero investire per lo sviluppo dell’IA, creare un ambiente politico favorevole, incoraggiare l’innovazione e la competizione e cooperare fra loro in tal senso.

⁷¹ Fonte: <https://www.nortonrosefulbright.com/en-it/knowledge/publications/d49a-49d5/key-takeaways-from-the-g7-competition-summit-on-artificial-intelligence>

⁷² Fonte: https://competition-policy.ec.europa.eu/about/news/joint-statement-competition-generative-ai-foundation-models-and-ai-products-2024-07-23_en

⁷³ Il G20 è un forum internazionale che riunisce annualmente i capi di stato e di governo delle principali economie mondiali. I suoi membri rappresentano oltre l’80% del PIL mondiale, il 75% del commercio globale e il 60% della popolazione del pianeta (<http://www.g20italy.org/about-the-g20.html>). I membri sono: Argentina, Australia, Brasile, Canada, Cina, Francia, Germania, India, Indonesia, Italia, Giappone, Repubblica di Corea, Messico, Russia, Arabia Saudita, Sudafrica, Turchia, UK, Stati Uniti, Unione Africana e Unione Europea.

⁷⁴ Fonte: <https://www.mofa.go.jp/files/000486596.pdf>

L'impegno è stato ribadito nel vertice di Roma (2021), con particolare riferimento alle necessità di piccole imprese e start-up, all'importanza della cooperazione internazionale per promuovere ricerca, sviluppo e applicazione dell'IA e della libertà dei dati attraverso i confini⁷⁵. Agli impegni di sfruttare l'IA in modo responsabile e, perseguire un approccio di regolazione e di governo favorevole all'innovazione, che massimizzi i benefici e tenga conto dei rischi associati all'uso dell'IA è stata data continuità nel vertice di Nuova Delhi del 2023⁷⁶.

Le diverse opzioni di creazione di un'arena di governance transnazionale dell'IA coinvolgono anche l'Organizzazione per la Cooperazione e lo Sviluppo Economico (OECD) e la *Global Partnership on Artificial Intelligence* (GPAI). L'OECD ha un Osservatorio sull'IA⁷⁷, che fra l'altro aggiorna automaticamente gli incidenti dell'IA e il suo impatto su vari tipi di lavoro. Come mostrato nel paragrafo sulle tassonomie del rischio, l'OECD è stata la prima organizzazione internazionale ad adottare nel 2019 il "primo standard intergovernativo sull'IA", basato sul presupposto che l'affidabilità dei sistemi di IA sia un fattore chiave per la sua diffusione e adozione. Il paradigma della Raccomandazione OECD è basato sulle potenzialità e i rischi, e lo sviluppo di un'IA affidabile è affidato all'impegno di "tutte le organizzazioni e gli individui coinvolti o interessati dai sistemi di IA, direttamente o indirettamente"⁷⁸. Essa è stata elaborata da un gruppo di oltre 50 leader di imprese, sindacati, no-profit e comunità accademiche e scientifiche e sottoscritta da 38 Paesi membri, 9 non membri e dall'UE, ma non da Cina, Russia e India. La Raccomandazione chiedeva ai governi degli stati membri di predisporre ambienti di policy favorevoli allo sviluppo dell'IA attraverso ambiti di sperimentazione regolativa, di incoraggiare l'innovazione e la concorrenza, prepararsi alla trasformazione del mercato del lavoro attraverso "dialogo sociale" (con imprenditori e lavoratori), collaborare fra loro per sviluppare conoscenze e standard tecnici globali per l'IA interoperabili e affidabili. Per fare questo, dovrebbero utilizzare termini e definizioni comuni, senza i quali è difficile realizzare interoperabilità tra le giurisdizioni nazionali e i

⁷⁵ Fonte:

https://www.governo.it/sites/governo.it/files/G20ROMELEADERSDECLARATION_0.pdf

⁷⁶ Fonte: https://www.g20.in/content/dam/gtwenty/gtwenty_new/document/G20-New-Delhi-Leaders-Declaration.pdf

⁷⁷ Per conoscere le attività dell'organizzazione sul tema dell'IA si rimanda a <https://www.oecd.org/en/topics/artificial-intelligence.html>

⁷⁸ Fonte:

<https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html>;

<https://www.oecd.org/en/topics/ai-principles.html>;

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

loro diversi approcci alla gestione della tecnologia⁷⁹. Questa raccomandazione è stata alla base della posizione dei G20 nel 2019 (cfr. sopra), nel solco di una tradizione secondo cui l’OECD è stato un loro consulente strategico di fiducia sin dal 2009.

Il secondo importante intervento risale al 2022, quando l’OECD ha adottato uno strumento per classificare sistemi di IA⁸⁰, caratterizzarli e identificare le opportunità e le sfide legate al concreto uso di ciascuno di essi. Lo strumento collega le caratteristiche dei sistemi di IA con i principi dell’IA dell’OECD e consente agli utenti di mettere a fuoco rischi specifici, come *bias*, deficit di comprensibilità e robustezza. Inoltre, l’OECD ha prodotto conoscenze utili ai *policy maker* nazionali e internazionali sui *language model*⁸¹ (OECD, 2023).

Il terzo intervento è il contributo dell’OECD alla GPAI⁸², ideata nel 2018 da Francia e Canada durante le loro presidenze del G7, avviata con 17 Paesi e poi ampliata a 44, fra cui gli Usa, dopo una titubanza iniziale dovuta alla preoccupazione di introdurre troppa regolazione, UE, Brasile, India, UK e Italia. La partnership si è proposta come “un panel internazionale” di esperti e attori di governo per “portare avanti un’agenda per l’implementazione di un’IA incentrata sull’uomo, sicura e affidabile, incarnata nei principi della Raccomandazione dell’OECD sull’IA”. La GPAI si propone come una delle sedi di confronto internazionale sull’IA e si impegna a “favorire le sinergie con le iniziative sull’IA avviate in altri forum internazionali come il G20 e il G7, tra cui il processo di Hiroshima sull’IA, il Consiglio d’Europa”, oltre ai summit che ora prenderemo in considerazione⁸³.

Alle prime tre arene, promosse da organizzazioni preesistenti che hanno esteso la propria azione nel campo dell’IA, si aggiungono nuovi spazi, più informali, di discussione per il suo governo globale: i summit internazionali sull’IA: dalla sicurezza all’“azione”. Una sequenza di incontri al vertice internazionali è stata avviata nel 2023 dal premier britannico R. Sunak nella simbolica residenza di Bletchley Park, dove nel 1940 Alan Turing decifrò il codice “E-

⁷⁹ Fonte: <https://www.oecd.org/en/topics/artificial-intelligence.html#key-messages>

⁸⁰ Fonte: <https://oecd.ai/en/classification>; <https://wp.oecd.ai/app/uploads/2022/02/Classification-2-pager-1.pdf>

⁸¹ I modelli linguistici dell’IA sono una componente chiave dell’elaborazione del linguaggio naturale (NLP), ovvero la capacità dei computer di comprendere e generare il linguaggio umano. I modelli linguistici usano algoritmi che consentono di elaborare, analizzare e generare testo, o parlato in linguaggio naturale, addestrati su vaste quantità di dati provenienti da Internet, utilizzando tecniche che vanno dagli approcci basati su regole e modelli statistici e al deep learning.

⁸² Fonte: <https://gpai.ai>; <https://www.oecd.org/en/about/programmes/global-partnership-on-artificial-intelligence.html>

⁸³ Fonte: <https://pib.gov.in/PressReleasePage.aspx?PRID=2030534>

nigma” dei tedeschi. Al vertice hanno partecipato i governi di 28 Paesi⁸⁴, il segretario generale delle UN, 40 imprese leader nel campo dell’IA (personalmente E. Musk e M. Zuckerberg), gruppi del no-profit e 46 organizzazioni di ricerca (compresa la *Chinese Academy of Sciences*)⁸⁵. Sono stati discussi i rischi dell’IA e come mitigarli attraverso un’azione coordinata a livello internazionale. Il focus iniziale sulle potenzialità commerciali ha fatto seguito a precedenti collaborazioni del governo britannico con Big Tech come *OpenAI*, *DeepMind* e *Anthropic*⁸⁶. Perciò l’iniziativa è stata variamente interpretata: da un lato come il tentativo di posizionare l’UK come luogo attrattivo per gli imprenditori dell’IA⁸⁷; dall’altro, come la definizione di un’agenda globale dell’IA ispirata dalle Big Tech Usa, con un ruolo particolarmente debole dei governi del Sud globale e della società civile no-profit (Garcia, 2022)⁸⁸. Nella dichiarazione finale di Bletchley, i partecipanti si sono impegnati a collaborare per introdurre in ciascun paese politiche dell’IA basate sul rischio e a far testare i modelli della “*Frontier AI*” prima che vengano offerti al pubblico. Sono stati inoltre siglati accordi bilaterali, come quelli fra UK e Usa su standard condivisi di sicurezza dell’IA, o fra Francia e Cina “sull’importanza di favorire lo sviluppo e la sicurezza dell’IA e di promuovere un’adeguata *governance* internazionale per raggiungere questo obiettivo (...) per una IA al servizio del bene comune (...) e dello sviluppo sostenibile”⁸⁹.

A questo incontro ha fatto seguito un summit analogo a Seoul (maggio 2024) i cui partecipanti (ma non la Cina) hanno riconosciuto – con dichiarazioni prive però di efficacia vincolante – la necessità di regolazioni per governare i *trade-off* dell’IA (sicurezza, innovazione, equa condivisione di opportunità e benefici dell’IA) e concordato lo sviluppo di quadri di governance interoperabili. A Seoul 16 aziende di tecnologia dell’IA⁹⁰ hanno anche

⁸⁴ Non solo occidentali ma anche Cina, India, Indonesia, Giappone, Arabia Saudita, Corea, Singapore, Turchia, Emirati Arabi.

⁸⁵ Per l’elenco dei partecipanti si rimanda a:

<https://www.gov.uk/government/publications/ai-safety-summit-introduction/ai-safety-summit-confirmed-governments-and-organisations>

⁸⁶ Fonte: <https://www.wired.it/article/intelligenza-artificiale-vertice-quando-edove-regno-unito-stati-uniti/>

⁸⁷ Fonte: <https://www.fanpage.it/innovazione/tecnologia/laccordo-tra-i-leader-mondiali-sullintelligenza-artificiale-sembra-scritto-da-chatgpt/#>

⁸⁸ Fonte: <https://www.atlanticcouncil.org/blogs/new-atlanticist/rival-powers-agree-that-ai-poses-new-risks-can-they-work-together-to-address-them/>

⁸⁹ Fonte: <https://www.elysee.fr/emmanuel-macron/2024/05/06/declaration-conjointe-entre-la-republique-francaise-et-la-republique-populaire-de-chine-sur-lintelligence-artificielle-et-la-gouvernance-des-enjeux-globaux>

⁹⁰ Amazon, Anthropic, Cohere, Google / Google DeepMind, G42, IBM, Inflection AI, Meta, Microsoft, Mistral AI, Naver, OpenAI, Samsung Electronics, Technology Innovation Institute, xAI, Zhipu.ai

concordato “Impegni di sicurezza dell’IA di frontiera” per pubblicare le modalità in cui misureranno i rischi dei loro modelli e sulle azioni per garantire che le soglie non vengano superate.

Nel successivo *AI Action Summit* organizzato nel 2025 a Parigi dai governi francese e indiano (che ospiterà il successivo nel 2026), cui hanno partecipato esponenti politici, a partire dalle UN e da capi di governo, e dell’industria tecnologica, sono stati capovolti approccio e obiettivi: da una preoccupazione “pessimista” per la sicurezza dell’IA all’ottimismo e alla promozione di azioni per il suo sviluppo. Sono emerse però posizioni diverse su ciò che la politica può fare per favorirlo. Da un lato si rileva un consenso sull’esigenza di coordinare massicci investimenti pubblici e privati per infrastrutture, come i data center e l’energia necessaria. E molti pongono fra le condizioni necessarie per creare fiducia e liberare le energie creative ed economiche una governance internazionale dell’IA condivisa e non frammentata. Dall’altro, visioni diverse hanno impedito a Usa e UK di firmare con altri 61 Paesi, fra cui Cina, Francia, India il documento finale per una IA “aperta”, “inclusiva” ed “etica”, oltre che sostenibile sul piano energetico. Il documento, in cui significativamente si dichiara di non riferirsi agli usi militari, ripropone valori e orientamenti per l’azione non particolarmente innovativi, né delinea azioni concrete rilevanti⁹¹. Questo non è comunque bastato agli Usa per sottoscriverlo. Nel 2025, la nuova presidenza Trump ha infatti assunto una posizione radicalmente liberista, ritenendo che per creare fiducia nell’IA sia necessario “un regime regolativo internazionale che promuova lo sviluppo delle tecnologie di IA anziché limitarlo e che gli europei guardino a questa nuova frontiera con ottimismo, piuttosto che con preoccupazione”⁹². Questa posizione esercita una pressione anche sulla politica dell’IA dell’UE, come vedremo nel prossimo capitolo.

⁹¹ Fonte: <https://www.politico.eu/wp-content/uploads/2025/02/11/02-11-AI-Action-Summit-Declaration.pdf>

⁹² Fonte:

https://www.repubblica.it/economia/2025/02/11/news/von_der_leyen_intelligenza_artificial_e_piano_ue_200_miliardi-423995528/?ref=RHLF-BG-P1-S1-T1

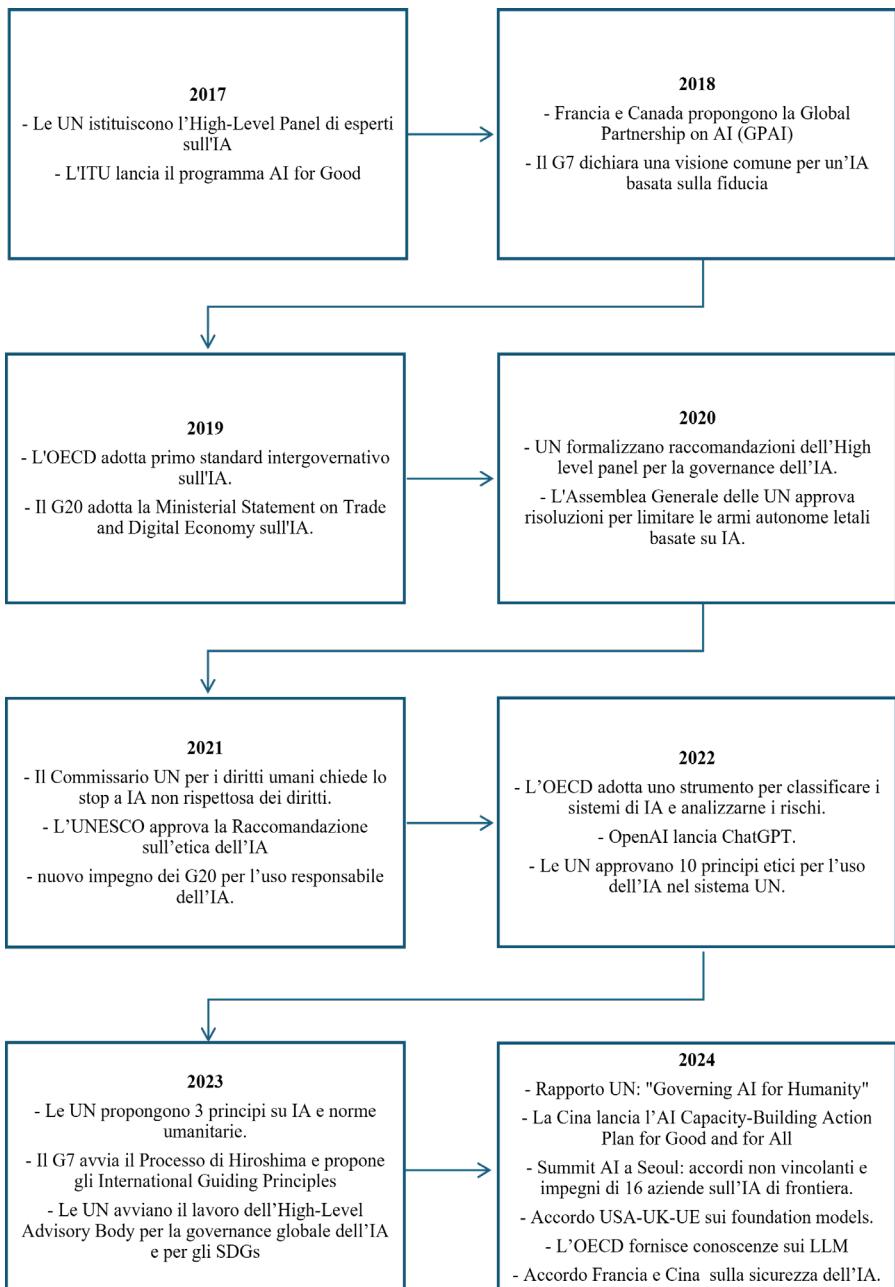


Fig. 1 - Una cronologia degli eventi principali nelle arene internazionali dell'IA, elaborazione degli autori.

Arrivati a questo punto è possibile fare una sintesi delle caratteristiche, dei limiti e dei punti di forza delle iniziative e delle reti internazionali sull'IA come quelle di OECD, G20, GPAI che sono particolarmente intrecciate fra loro. I G7 si propongono come un nucleo centrale di "imprenditori politici", mentre altri attori, statali e non, sono coinvolti in modo diversificato, dando luogo a un "plurilateralismo selettivo" (UN, 2024). In queste arene si riproducono forme di azione e pratiche consolidate: la principale è quella dei gruppi di esperti di diversa provenienza (governi, ricercatori, imprese, no-profit), che fanno consultazioni ed elaborano documenti contenenti analisi e raccomandazioni, poi adottate volontariamente dagli attori di volta in volta convenuti. Una seconda forma di azione consiste nell'offrire occasioni e piattaforme multilaterali di dialogo e scambio di esperienze, che possono funzionare se c'è una base condivisa di valori e (*frame* etico dell'IA e suo uso per finalità umanitarie, oltre che di mercato).

I discorsi su valori e interpretazioni dei *trade-off* fra potenzialità e rischi dell'IA sono complessivamente allineati intorno a tre *frame* principali: l'interpretazione dell'IA come una tecnologia rivoluzionaria, trasformativa e dirompente; la tensione fra competizione e collaborazione globale; la ricerca di approcci finalizzati a realizzare opportunità per mitigare i rischi e assicurare responsabilità (Ulinicane *et al.*, 2022). Tuttavia, il coordinamento fra le varie iniziative internazionali, nella produzione di standard e nell'attuazione degli accordi (*implementation gap*) resta debole⁹³. Su questi principi e urgenze si realizzano facilmente convergenze, ma spesso sono superficiali e generiche, e l'idea di *guidelines* etiche o *soft law* è accolta con più favore rispetto alle regolamentazioni istituzionali (Garcia, 2022). Infatti, sono stati prodotti standard tecnici e gestionali, come il *Foundational Standard ISO/IEC 22989:2022(E) - Artificial intelligence concepts and terminology* del 2022, un esempio di "para-regolazione" sinora debole, ma che potrebbe delineare i confini per sviluppi normativi successivi (Bello y Villarino, 2023). Tali standard e il lavoro delle organizzazioni che li sviluppano sono visti come risorse per la cooperazione internazionale anche in un contesto reticolare che limita la normatività all'allineamento e l'interoperabilità di norme nazionali e codici volontari (Cameron *et al.*, 2025). Tuttavia, sinora è stato impossibile a livello globale passare dalla *soft law* a un sistema di norme con meccanismi di conformità (Baele *et al.*, 2024). Nessuna delle azioni delineate per realizzare gli auspicati quadri di governance interoperabili è infatti vincolante. In altri termini, il rispetto dei criteri concordati è affidato solo alla volontà, di volta in volta, di governi o imprese. Gli stru-

⁹³ Fonte:

https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

menti di regolazione ed esecuzione identificati tendono a replicare quelli già adottati per altri ambiti (ad esempio per la concorrenza), con poteri di *enforcement* strettamente statali. Non è stato perciò sinora costituito un organismo internazionale con poteri di monitoraggio, rendicontazione, verifica e applicazione di norme e standard per la sicurezza dell'IA, pur auspicato dalle raccomandazioni degli esperti delle UN (2024) ⁹⁴.

Gli sforzi di governance globale incontrano quindi difficoltà ad andare oltre dichiarazioni di intenti e richiami a impegni congiunti, spesso generici e ripetitivi (Benaich e Chalmers, 2024) ed esiste solo un embrione di un sistema di governance globale dell'IA, attraverso le arene che abbiamo considerato. Le cause di questo regime ancora fluido sono molteplici. Un primo gruppo di fattori (*agency*) è relativo agli attori e alle loro strategie, mentre altri (*structure*) riguardano il contesto.

I fattori *agency* riguardano il modo in cui la costruzione delle poste in gioco e la valutazione del contesto ispirano strategie e comportamenti degli attori. Uno dei motivi per i quali non è stato possibile concordare interventi internazionali di tipo autoritativo per governare l'IA – cui come abbiamo visto sono associate poste in gioco sia economiche, sia politiche (di influenza e militari)⁹⁵ – è che i singoli governi cercano di perseguire strategie contemporaneamente cooperative e competitive e questo riduce i margini per proporre misure efficaci (Garcia, 2022). Gli stati confidano di poter godere dei vantaggi dell'IA e questo può incentivarli a cooperare. Però, la percezione dell'IA come risorsa strategica nello sviluppo economico e nella corsa agli armamenti mette alcuni governi uno contro l'altro ed esalta i nazionalismi (Dafoe, 2019). Inoltre, nelle arene più ampie fra quelle che abbiamo visto si incontrano anche Paesi fra loro ostili e, poiché nessuno vuole subire vincoli imposti da rivali, viene sfruttata la possibilità di porre veti formali o informali o di sottrarre il proprio consenso agli impegni collettivi. La competizione fra Usa e Cina si riflette anche nei diversi approcci che i due governi seguono nelle arene di governance dell'IA: i primi si affidano alla partecipazione delle Big Tech per perseguire una leadership nelle relazioni internazionali sull'IA, la Cina sostiene piuttosto il multilateralismo e propone un'agenda di *capacity building* per l'IA dei Paesi meno sviluppati. A questo fine, privilegia il sistema delle UN, in cui gli stati membri decidono (Cameron *et al.*, 2025). La leadership nelle arene internazionali diviene così a sua volta una posta in gioco contesa. Secondo media cinesi che possono essere considerati quasi fonti ufficiali, i Paesi occidentali “spinti da considerazioni

⁹⁴ Fonte

https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

⁹⁵ Fonte: <https://www.wired.com/story/united-nations-artificial-intelligence-report/>

ideologiche, hanno creato ‘circoli esclusivi’ che limitano una più ampia partecipazione internazionale. Ciò ha trasformato i dibattiti sulla governance dell’IA in questioni chiuse ed esclusive che minano l’efficacia delle discussioni e ostacolano la capacità di molti Paesi di impegnarsi nelle discussioni, erodendo la fiducia necessaria per raggiungere un consenso globale”⁹⁶. È evidente il riferimento ai tavoli G7, GPAI, al summit di Seul. Dal canto loro, gli Usa non hanno firmato il documento del summit di Parigi (febbraio 2025), anche per il dissenso con i valori di “sostenibilità” e “inclusione” posti alla base della ricerca di una governance globale dell’IA cui ha dato consenso anche la Cina.

Bisogna tenere a mente che il contesto in cui si svolgono queste iniziative – un insieme di fattori che condiziona le strategie degli attori – è quello della crisi della globalizzazione economica e politica, con un ambiente multipolare in cui le relazioni internazionali divengono via via più difficili e di decrescente legittimazione delle organizzazioni internazionali, cui tendono a sostituirsi iniziative unilaterali basate sulla forza, o la ricerca di accordi bilaterali fra gli stati più forti per affrontare situazioni di crisi.

Dopo decenni in cui all’ampliarsi degli scambi e delle interdipendenze economiche globali aveva corrisposto la ricerca di cooperazione politica, le risposte alle crisi economiche e finanziarie – la più importante quella iniziata nel 2007 – hanno ridotto la propensione degli stati sia a favorire l’apertura delle economie nazionali, sia ad affidare quote di sovranità a strutture di governo internazionali per fare loro condurre politiche concordate. Questi due orientamenti e le loro conseguenze emergono specialmente quando leader populisti e sovranisti di destra conquistano posizioni di governo nazionale. Particolarmente influente nell’alimentare la crisi della globalizzazione, si sta rivelando la presidenza degli Usa di D. Trump con le sue politiche di protezionismo. Le ragioni di conflitto nella geopolitica mondiale e i tentativi di formare o consolidare nuove alleanze fra Paesi come Russia, Cina, India e altre economie emergenti per contrastare i Paesi occidentali, a loro volta più divisi che nel recente passato, minano anche i potenziali regimi “pacifici” di governance dei problemi pubblici transnazionali e difficilmente l’IA potrebbe fare eccezione. Le tensioni geopolitiche ed economiche e le agende internazionali in cui prevalgono questioni di sicurezza e le preoccupazioni per gli andamenti della crescita economica si sostituiscono a quelle per la sua sostenibilità, non potevano quindi non ostacolare anche una cooperazione efficace sullo sviluppo dell’IA.

Al deficit di convergenze politiche si accompagna una debolezza del con-

⁹⁶ China shaping AI governance mechanism, China Daily 06/02/2025 (<http://www.chinadaily.com.cn/a/202502/06/WS67a3ef2ea310a2ab06eaa572.html>)

testo istituzionale. Da un lato, secondo gli esperti delle UN, “il diritto internazionale, compreso quello dei diritti umani, fornisce una base normativa condivisa per tutti gli sforzi legati all’IA, facilitando così il coordinamento e la coerenza”⁹⁷. Anche i modelli e approcci alla *compliance* internazionale già sperimentati in altri campi offrono riferimenti da imitare. Da un altro lato, però, l’innovazione radicale e velocissima dell’IA produce un gap dei processi di governance (Garcia, 2022) perché per ottenere tempi allineati a quelli dello sviluppo della tecnologia e delle imprese che lo finanziano servirebbero forme istituzionali e strumenti altrettanto innovativi (Cacciari, 2023) come *l’UN System White Paper on AI Governance*⁹⁸. Un tasso di innovazione che è al momento impedito dalla tensione fra cooperazione e competizione che caratterizza le strategie degli attori politici ed economici.

L’insieme di questi fattori mantiene il regime di governance globale dell’AI in condizioni embrionali e di debolezza. Non solo nessuna istituzione è in grado di stabilire regole efficaci in uno spazio transnazionale e nel tempo necessario per i comportamenti sui mercati dell’IA, come del resto accade per l’ambiente o il cambiamento climatico, ma non si profilano soluzioni di *metagovernance* nella direzione dell’esercizio di una sovranità condivisa e pubblica. Come vedremo nel prossimo capitolo, l’UE sembra un’eccezione, ma deve affrontare difficoltà, anche per quanto riguarda il suo posizionamento nello scenario globale.

⁹⁷ Fonte:

https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

⁹⁸ Il testo è disponibile al link: <https://unsceb.org/united-nations-system-white-paper-ai-governance>

3. Le strategie politiche

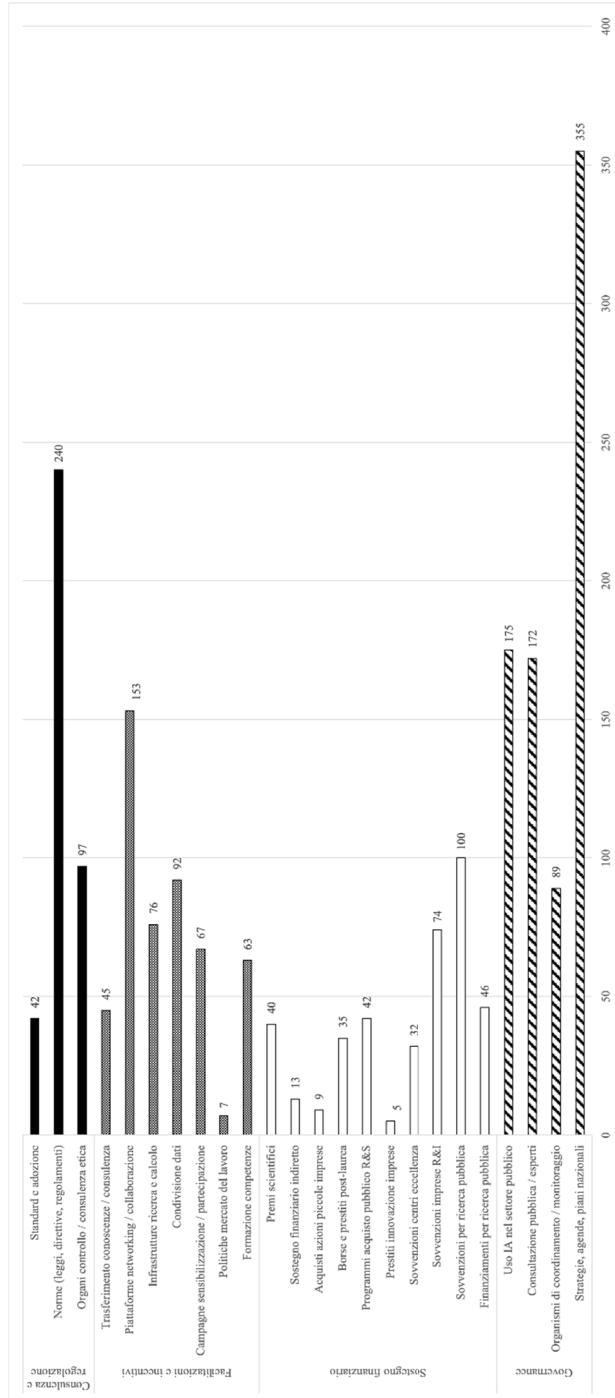
3.1. Agende e strategie politiche nazionali: i “giganti”

L’analisi delle strategie politiche dell’IA deve tenere conto della loro intrinseca multi-scalarità, causata dal fatto che il potere politico è distribuito fra diversi livelli, ciascuno con i propri poteri e anche con ambizioni di sviluppo, regolazione e uso dell’IA. Allo stesso tempo però, questa analisi non può prescindere dal riadattamento di un tradizionale approccio delle scienze sociali ispirato al cosiddetto nazionalismo metodologico (Beck, 2007). Le strategie politiche sono anche strategie di stati nazionali che prendono le loro iniziative di sviluppo e regolazione nell’ambito dei loro confini politici e amministrativi. Tuttavia, i diversi contesti nazionali si differenziano tra loro per dimensioni degli investimenti privati e pubblici nell’IA, livelli di sviluppo tecnologico raggiunti e rilevanza delle imprese del settore, oltre che per il tipo di partecipazione nelle arene internazionali. Tenendo conto di questi aspetti, nello scenario geopolitico dell’IA possiamo identificare tre “giganti” di cui ci occupiamo in questo paragrafo: gli Usa, la Cina e l’UK. Ci interroghiamo, in particolare, sugli orientamenti e le caratteristiche delle diverse strategie nazionali ricercandone analogie e differenze e identificandone i fattori di influenza.

Lo sviluppo dell’IA e le sue implicazioni sono entrati nelle agende degli stati nazionali a partire dal 2016, dando luogo a una varietà di misure e strategie, indirizzate verso le arene interne e quelle internazionali. L’Osservatorio CE-OCSE/GPAI ha registrato nel suo database¹ delle politiche sull’IA oltre 1.000 iniziative, relative a strategie e misure di oltre 60 governi, oltre che di attori internazionali. Su queste iniziative sono stati condotti vari studi.

¹ Il database, *powered by EC/OECD*, è disponibile al link: <https://oecd.ai/en/dashboards/overview>

Grafico 1 – Categorie e strumenti d’azione delle politiche nazionali sull’IA, elaborazione nostra su banca dati OECD



La maggior parte delle comparazioni ha considerato azioni precedenti l'avvento dell'IA generativa (2022) e dei *foundation model*. Tuttavia, pur in un quadro messo a fuoco via via in maggiore dettaglio, le preoccupazioni diffuse nella società e i tipi di misure politiche adottate o prospettate non sono cambiate poi molto.

Le azioni politiche sull'IA sono state classificate seguendo vari criteri. Il database dell'OECD raggruppa (febbraio 2025) misure e strumenti come mostrato nel Grafico 1.

In questo modo vengono assemblati forme e strumenti di azione, ma non emergono gli obiettivi delle strategie politiche nazionali, di cui l'OECD (2023) ha riportato la crescita nel numero (51 Paesi nel 2023).

A questo proposito, diversi studi comparativi hanno presentato i risultati di analisi prevalentemente quantitative di documenti governativi, che riguardano soprattutto le narrazioni adottate dagli attori politici circa potenzialità e rischi dell'IA, le azioni pubbliche conseguentemente necessarie e le finalità delle strategie nazionali. Da questi studi è possibile estrarre ulteriori criteri di differenziazione delle strategie, basati sulle cornici di legittimazione (*frame*), sulle priorità politiche e le modalità di implementazione.

I. Ulnicane e colleghi (2021) hanno analizzato 49 documenti del periodo 2016-2018 di governi nazionali e organizzazioni internazionali, società di consulenza, think tank e organizzazioni della società civile, nell'UE e negli Usa. Dal loro lavoro emerge una differenziazione sulla base dei *frame*, diversificati a seconda dell'importanza data alla promozione dello sviluppo tecnologico dell'IA, alla facilitazione del coinvolgimento di attori economici e sociali, o alle garanzie della sicurezza.

Da una comparazione (Radu, 2021) delle strategie di 12 Paesi (Canada, Cina, Corea del Sud, Emirati Arabi, Finlandia, Francia, Germania, Giappone, Singapore, Svezia, UK, Usa) fra il 2016 e il 2019 emerge una distinzione fra chi assegna la priorità a una guida strategica della politica industriale, chi ha l'obiettivo di non assecondare oligopoli, oppure stati che si propongono di diventare leader mondiali nel settore. Dallo studio emerge però una sostanziale convergenza tra gli stati, ad eccezione di Cina e Usa, sull'importanza della dimensione etica, che si ripropongono di definire principi che le imprese dovrebbero seguire, pur senza introdurre forme di regolazione pubblica.

A. Guendez e T. Mettler (2023) hanno rintracciato narrazioni simili e ricorrenti in 37 documenti di 33 Paesi (2017-2020). La prevalenza di ciascuna narrazione varia a seconda del paese e delle priorità che prospettano: regolare l'IA per ridurre rischi e pericoli e garantirne la compatibilità con i valori della società; facilitare lo sviluppo e l'innovazione dell'IA, garantire una leadership pubblica dei processi di ricerca e sviluppo.

Infine, G. Papshev e M. Yarime (2023) hanno studiato le funzioni che i

governi di 31 Paesi (2017-2021) intendono svolgere per implementare strategie per l'IA, distinguendo a loro volta fra tre tipi di priorità: imprimere indirizzi strategici allo sviluppo dell'IA, regolamentarla, o promuovere e facilitare l'innovazione.

Da questi diversi assemblaggi emergono sia convergenze, sia differenziazioni. Da un lato una somiglianza di fondo: governi tra loro molto diversi utilizzano narrazioni simili, affrontano gli stessi problemi e condividono la rappresentazione del *trade-off* di fondo, fra facilitare lo sviluppo di IA e mitigare i rischi. A questo isomorfismo contribuisce il fatto che spesso i governi hanno considerato mappe delle strategie adottate da altri, compresa l'UE (Ulnicane *et al.*, 2022), e ciò ha favorito processi di isomorfismo. Le opportunità di apprendimento, però, non riguardano in eguale misura tutti i Paesi e risultano più limitate per quelli che hanno preso per primi l'iniziativa (Shehan, 2023). L'OECD ha svolto un ruolo importante, fornendo raccomandazioni e consigli pratici per l'attuazione dei suoi principi del 2019 in ogni fase del ciclo dell'IA (design, implementazione, valutazione e monitoraggio). Dall'altro lato, variano l'enfasi e il peso assegnato a diverse componenti delle strategie. Dalle priorità assegnate dai governi alle politiche dell'IA emergono due polarizzazioni: da un lato, una relativa alla scala: vi sono strategie che danno priorità a ciò che avviene all'interno del paese o, al contrario, al posizionamento nelle arene globali. Una seconda polarizzazione contrappone chi adotta una prospettiva che fa leva sulla libertà dei mercati per favorire lo sviluppo dell'innovazione tecnologica, il suo valore economico e la competitività, a una che, invece, punta sulla salvaguardia di principi e tutela di diritti, mitigando i rischi attraverso varie forme di regolazione². Come vedremo, sono stati esperiti vari tentativi di combinare le diverse priorità.

Vediamo di seguito le caratteristiche dei diversi tipi di strategia, con alcuni approfondimenti relativi a quelli che si possono considerare i "giganti" nello scenario geopolitico dell'IA: gli Usa, la Cina e l'UK. Nelle agende di questi Paesi convivono, con pesi diversi, elementi riferiti a due tipi prevalenti di strategie: da un lato strategie di regolazione che possiamo definire "market oriented", che assegnano priorità alla promozione di sviluppo dell'innovazione, competitività e leadership nazionale nel campo dell'IA; dall'altro, strategie "social oriented", in cui la regolazione è finalizzata a mitigare i rischi sociali e tutelare i diritti delle persone.

² Vi sono poi specifiche strategie di utilizzazione per dell'IA per finalità pubbliche, esaminate nel capitolo 4, con riferimento alla scala locale.

3.1.1. Strategie market oriented per promuovere sviluppo dell'innovazione, competitività e leadership

Una priorità *market oriented* caratterizza soprattutto i Paesi europei dell'ex blocco sovietico, la Russia e i Paesi dell'Asia orientale (Cina, Corea del Sud e Giappone) (Papyshev e Yarime, 2023) e, nel periodo più recente, gli Usa. Quando perseguitano queste finalità, le politiche si concentrano su progetti e investimenti in infrastrutture per le necessità di ricerca e sviluppo dell'IA su scala nazionale, prevalentemente in collaborazione con il settore privato. Si investe sul capitale umano con la formazione di competenze digitali. Il ruolo dello stato di facilitatore dello sviluppo dell'IA viene spesso associato al proposito di tutelare la sicurezza e la competitività economica nazionale e legittimato attraverso di esso. Lo sviluppo del capitale umano, il sostegno al settore privato con il finanziamento di ricerca scientifica, investimenti per lo sviluppo dell'IA da parte del governo e la promozione dell'informazione sull'IA sono presenti anche in Arabia Saudita, Singapore, Nuova Zelanda, Finlandia, Lituania, Irlanda, India e Australia (Papyshev e Yarime, 2023). In questi casi, la protezione dai rischi dell'IA passa in secondo piano o, comunque, viene riposta fiducia nel potere di auto-organizzazione dell'industria.

Tuttavia, l'ambizione del primato nazionale nell'arena globale dell'IA si ritrova solo nelle strategie dei giganti. Sin dalla prima fase di elaborazione delle strategie nazionali (2017-2021, precedente l'avvento dell'IA generativa) Cina, Usa, UK (e Francia) esplicitano l'ambizione di diventare leader mondiali dell'IA dal punto di vista sia tecnico, sia politico, facendo riferimento anche a interessi militari (Radu, 2021). A questo fine costruiscono ecosistemi favorevoli alla competitività delle imprese di IA, ritenendo che lo stesso posizionamento delle intere economie nazionali in quella globale dipenda anche dal successo di queste azioni (ISPI, 2024). Quando competitività e leadership sono le priorità, le barriere sono minimizzate e le imprese vengono sostenute, si finisce però per favorire la crescita di oligarchie nel mercato. Una prassi del "chi vince prende tutto" e posizioni monopolistiche che alcuni governi cercano di contrastare, ad esempio in Norvegia, Arabia Saudita, Francia e Serbia (Guenduez e Mettler, 2023). Pur in presenza di una crescita degli scambi e della collaborazione tecnologica fra Usa e Cina, favorita anche dalle traduzioni automatiche fra le due lingue, l'intensificata competizione commerciale fra i due Paesi, esaltata dai dazi introdotti o minacciati nel 2025 sembra indirizzare lo sviluppo delle IA nazionali in modo competitivo. Sono così confermate le interpretazioni del "tecnonazionalismo", che hanno enfatizzato la competizione interstatale nelle relazioni fra Cina e Usa sull'IA (Ding, 2022).

Vediamo alcuni dettagli delle strategie dei “giganti”, con particolare riferimento agli elementi di orientamento al mercato.

La strategia degli Usa è, nel complesso, ispirata prioritariamente all’obiettivo di consolidare e proteggere una posizione nazionale di primato tecnologico e geopolitico dell’IA e di contrastare la crescita della Cina. Già la presidenza democratica di B. Obama aveva dato più spazio all’innovazione che alla regolazione dell’IA³, ma le successive presidenze di D. Trump (repubblicano e considerato un esponente del populismo di destra sovranista) e di J. Biden (democratico) le hanno impresso accentuazioni diverse.

J. Biden ha dato la priorità allo sviluppo della tecnologia e dell’innovazione statunitensi, ma non ha trascurato il suo uso sicuro e affidabile, la gestione dei rischi, il rispetto di privacy, valori democratici, diritti umani e civili, secondo principi simili a quelli dell’OECD e dell’UE. Nel 2024 ha inoltre aggiunto al perseguitamento di una governance internazionale dell’IA, l’obiettivo di proteggere l’ecosistema dell’IA degli Usa dalle minacce dell’*intelligence* straniera di gestire i rischi per la sicurezza, la protezione e l’affidabilità dell’IA⁴. Inoltre, nel 2024 ha istituito un Consiglio federale per l’IA per ottenere raccomandazioni sulla sua adozione sicura negli Usa, che riuniva esponenti di Big Tech come OpenAI, Microsoft, Google, Nvidia, rappresentanti del governo e ricercatori e, nello stesso anno, il Congresso ha costituito una task force bipartisan, che ha rilasciato le sue proposte per regolare l’IA dopo le elezioni⁵.

D. Trump ha invece messo in primo piano forme di partnership con le Big Tech finalizzate all’innovazione in funzione del predominio tecnologico, militare e politico nazionale nello scenario globale. Questo obiettivo è stato perseguito nel 2019 con l’*executive order Maintaining American Leadership in Artificial Intelligence*, che conteneva principi e misure per rafforzare le capacità della nazione nel campo dell’IA, promuovere la scoperta scientifica, la competitività economica e la sicurezza nazionale⁶. Nella seconda presi-

³ The Administration’s Report on the Future of Artificial Intelligence, Website of the Obama White House, 12 October 2016 (<https://obamawhitehouse.archives.gov/blog/2016/10/12/administrations-report-future-artificial-intelligence>).

⁴ *Memorandum on advancing the US leadership in AI. Harnessing AI to fulfill national security objectives and fostering the safety security* del 24 ottobre 2024 rimosso dal website della Casa Bianca dopo l’insediamento del nuovo Presidente.

⁵ Due anni prima era stato approvato il *Chips and Science Act* che stanziava miliardi di dollari per la ricerca e la produzione di semiconduttori negli Usa. In cambio, i beneficiari – anche non statunitensi – non dovevano aggiornare o espandere le loro attività in Cina per 10 anni (Benaich e Hogarth, 2022). Una disposizione coerente con il sostegno alla concorrenza con la Cina attraverso restrizioni all’esportazione di *chips* cruciali per l’IA e alla loro utilizzazione attraverso il Cloud (Villasenor e Turner, 2024).

⁶ Fonte: <https://trumpwhitehouse.archives.gov/ai/executive-order-ai/>

denza, D. Trump, come abbiamo visto ha adottato subito misure per liberare il mercato, rimuovendo i vincoli etici all’innovazione definiti dalla precedente amministrazione. Il modello di partnership finanziaria pubblico-privato è stato inoltre precisato e rafforzato, con lo *Stargate Project*, l’annuncio di un investimento quadriennale privato di 500 miliardi di dollari, che ha l’obiettivo di “garantire la leadership americana nell’IA, creare centinaia di migliaia di posti di lavoro americani e benefici economici per tutto il mondo”⁷. Quindi, una quota minore di spesa statale deve produrre un effetto trainante di una quota maggiore di investimenti privati, un modello che, abbiamo visto, è stato adottato anche dall’UE. E la spesa federale in IA, pur essendo frammentata, rivela una strategia sempre più orientata verso il settore della difesa (Larson *et al.*, 2024)

La Cina ha adottato una strategia pubblica di sviluppo dell’IA già nel 2015, insieme ad altre industrie ad alta tecnologia, con la campagna *Made in China 2025*⁸. Due anni dopo sono stati lanciati il *China AI Development and Safety Network* – frutto di una collaborazione fra università e istituti di ricerca pubblici e industria privata per promuovere innovazione e “campioni nazionali” nell’IA (ISPI, 2024), attraverso investimenti pubblici e privati per oltre 150 miliardi di dollari – e il *New generation AI development plan*, per la crescita economica attraverso l’IA, che prevedeva anche future leggi e regolazioni. Il primo ministro cinese ha dichiarato nel 2024 che l’IA è un motore importante per le forze produttive innovative e che ci sarà un impegno per promuovere e accelerare le innovazioni degli algoritmi e l’apertura, diffusione e utilizzo dei dati⁹. Gli obiettivi principali sono competitività e leadership internazionale anche nella sicurezza e negli usi militari¹⁰, riduzione della dipendenza dalle importazioni di tecnologia statunitense (Sheehan, 2024b), ai fini della crescita economica e del governo della società, dagli standard di vita a un maggiore controllo sul comportamento degli individui (Huw *et al.*, 2021). La politica cinese, modellata da ragioni economiche e di leadership globale dell’IA, contiene però anche un’enfasi sulla sicurezza (cfr. oltre).

La prima strategia nazionale dell’UK per l’IA è stata adottata nel 2021, ma nel 2025 il governo ha presentato al Parlamento il più ampio *AI Opportunities Action Plan*, finalizzato a fare del paese un centro mondiale dell’IA. Per non rimanere indietro rispetto ai progressi di Usa e in Cina, il programma ha l’obiettivo di “fornire una leadership globale nel cogliere in modo equo

⁷ Fonte: <https://openai.com/index/announcing-the-stargate-project/>

⁸ China making the most of a rapidly changing world, 21 11 2024, (<https://www.chinadaily.com.cn/a/202411/21/WS673e5f06a310f1265a1ceb41.html>).

⁹ Fonte: <https://www.reuters.com/technology/ai-is-an-important-engine-new-productive-forces-chinas-li-says-2024-03-13/>

¹⁰ Fonte: <https://www.cnas.org/publications/reports/understanding-chinas-ai-strategy>

ed efficace le opportunità dell'IA, come abbiamo fatto per la sicurezza dell'IA”¹¹. Tuttavia, la collocazione nello scenario globale è declinata soprattutto in chiave di autonomia nazionale: “essere la sede naturale dell'IA avanzata può fare la differenza tra il plasmare il futuro della scienza, della tecnologia e del lavoro e il vedere queste decisioni prese completamente al di fuori del nostro confine.” Obiettivo è quindi posizionare l'UK “in modo di essere un produttore di IA, non un acquirente di IA”. Per attrarre capitali, l'obiettivo è “essere il miglior partner statale per coloro che costruiscono l'IA di frontiera”. Come quello Usa, anche il piano UK si propone perciò di abbattere le barriere alla crescita dell'IA e facilitare la costruzione di infrastrutture di calcolo dei *data center*, i motori dell'era dell'AI. Sono programmati a questo fine investimenti decennali pubblici e privati in apposite *AI Growth Zones*, in ricerca, formazione di competenze e attrazione di talenti, oltre a un impegno per partenariati tra settore pubblico e privato, base di un fiorente ecosistema nazionale di IA. Da un lato, l'iniziativa governativa sembra già mobilitare capitali privati, nazionali e di Big Tech, come *Microsoft*, *Anthropic* e *OpenAI*. Dall'altro i prevedibili costi ambientali hanno provocato opposizioni dei gruppi ambientalisti¹².

3.1.2. Strategie “social oriented”: regolazione per mitigare i rischi e tutelare i diritti

Misure di regolazione pubblica per tutelare i diritti e proteggere la società dai rischi derivanti dall'IA tendono a entrare nell'agenda politica quando si fa strada la consapevolezza dell'inefficacia dell'autoregolazione delle imprese produttrici o utenti di IA secondo principi etici, o del rischio di “ethics washing”, ossia l'uso di dichiarazioni etiche da parte delle aziende per evitare regolamentazioni più stringenti (Ulnicane *et al.*, 2021). Tuttavia, i comportamenti regolativi variano notevolmente tra le strategie politiche nazionali: Paesi come UK, Nuova Zelanda, Israele, India, Giappone scommettono di poter ridurre i rischi attraverso le normative già vigenti, come la legislazione sulla privacy; i 27 Paesi membri dell'UE adottano invece una disciplina specifica comune. Il Canada ha regolamentato solo le applicazioni “ad alto rischio” (Benaich e Hogarth, 2023), mentre gli Usa hanno prima sperimentato regolazioni *soft* per poi ridurle ulteriormente (cfr. sotto). In questo scenario differenziato soprattutto le grandi imprese si trovano a navigare tra

¹¹ Fonte: <https://www.gov.uk/government/publications/ai-opportunities-action-plan/ai-opportunities-action-plan>

¹² Fonte: <https://www.ilfattoquotidiano.it/2025/01/31/intelligenza-artificiale-il-piano-fumoso-del-regno-unito-fa-gia-arrabbiare-gli-ambientalisti/>

sistemi normativi diversi mentre pianificano le loro strategie di IA a lungo termine (McKinsey, 2023).

Anche con riferimento alla regolazione, approfondiamo i casi Usa, Cina e UK, mentre lo sforzo dell'UE è oggetto del paragrafo 3.2. I primi tre casi, però, confermano che la strategia regolativa europea può esercitare un'influenza significativa, ma non determinante, sulle rimanenti norme – nazionali e internazionali – che stanno affrontando la tensione fra regolazione e innovazione in modo diverso (Feldstein, 2024).

Anche per quanto riguarda la regolazione dell'IA, sovrastata dalle preoccupazioni per l'innovazione, ma presente nell'agenda degli Usa (Schiff, 2022), le amministrazioni di J. Biden e D. Trump hanno adottato approcci diversi. Come abbiamo visto, l'amministrazione Biden ha promosso diverse iniziative per fornire risposte alle preoccupazioni sociali derivanti dall'uso trasversale dei sistemi di IA (data *bias*, sostituzione del lavoro, diritti civili). Dunque, attraverso *soft law* e autoregolazione volontaria da parte delle imprese e accordi volontari tra imprese e utenti, la strategia statunitense ha ottenuto l'impegno a collaborare di *Amazon*, *Anthropic*, *Google*, *Inflection*, *Meta*, *Microsoft* e *OpenAI*. Ha inoltre obbligato le agenzie dell'amministrazione federale a sviluppare standard di *cybersecurity* e proprie politiche di utilizzo dell'IA, ad affrontare i rischi delle infrastrutture critiche legati all'IA e a commissionare studi sugli effetti sul mercato del lavoro, chiedendo al *National Institute of Standards and Technology* di redigere linee guida, rivolte alle imprese, sulla prevenzione e gestione dei rischi (l'*AI Risk Management Framework*).

Forme blande di regolamentazione, quindi, ribadite nell'“eredità” lasciata dal Congresso alla presidenza di D. Trump nel 2024, ma ritenute in fondo poco accettabili dalle imprese dell'IA e da una parte del mondo repubblicano (Puliafito, 2025b). Un cambiamento politico dunque che ha attenuato l'orientamento sociale della strategia statunitense. Al Summit di Parigi sull'IA del febbraio 2025, infatti, il vicepresidente J.D. Vance ha chiarito per conto del Presidente che la nuova strategia regolativa degli Usa combatte “pregiudizi ideologici” e si ispira all'obiettivo di incoraggiare le politiche pro-crescita. Poiché gli Usa “sono leader nell'IA e la nostra amministrazione vuole che lo restino”, è necessario evitare che troppe regole, come quelle sul digitale (il GDPR dell'UE), rischino di uccidere sul nascere l'IA. Regole che le Big Tech vivono con sofferenza, preferendo rinunciare al mercato europeo e puntare magari sull'India, che ha poche regole. Piuttosto, possono esserci standard a cui aderire su base volontaria, fissati dagli Usa, anche in collaborazione con l'UE¹³.

La Cina è stato il primo paese (2019), ad affiancare a programmi per lo

¹³ Fonte: <https://tg24.sky.it/mondo/2025/02/11/intelligenza-artificiale-vertice-parigi>

sviluppo dell'economia dell'IA norme per il suo uso etico, anche se queste sono state viste criticamente come al servizio più di obiettivi politici retorici che di finalità sostanziali (Lucero, 2019). L'inizio di restrizioni più mirate risale all'approvazione del regolamento del 2021 sulla disseminazione di informazioni online da parte degli algoritmi di raccomandazione, uno strumento di policy riutilizzabile, che sarà impiegato dal governo cinese anche in norme successive (Sheehan, 2023, p. 12). Anche se ChatGPT non è disponibile in Cina, per disciplinare le alternative nazionali nel 2023 sono state adottate "misure provvisorie" sull'IA generativa, per garantire che "sia utilizzata in modo conforme agli obiettivi di politica nazionale, con un'attenzione particolare alla sicurezza e alla stabilità sociale, senza però comprometterne lo sviluppo" (Calderini, 2024). La sicurezza include l'uso di AI per videosorveglianza ed efficaci sistemi di riconoscimento facciale, venduti da aziende cinesi anche all'estero, ma limitati, per i privati, ad alcuni "scopi specifici strettamente necessari"¹⁴.

I modelli di IA generativa devono essere autorizzati e prodotti sotto la supervisione della *Cyberspace Administration of China* rispettando norme sul controllo e lo sviluppo di *chip*, software, strumenti, potenza di calcolo e usando fonti di dati "sicuri e affidabili"¹⁵. In particolare, gli *output* dell'IA devono evitare argomenti politicamente sensibili e aderire ai "valori socialisti" (MacCarthy, 2023). Le misure – riconfermate nel 2025 anche per la *chatbot DeepSeek*¹⁶ – riguardano anche la trasparenza che i fornitori di IA generativa devono rispettare, adottando anche misure per evitare che i loro utenti siano profilati e per rispettare la proprietà intellettuale¹⁷. Le misure regolano però solo i "servizi forniti al pubblico in generale" e non quelli per ricerca e sviluppo, per preservare l'innovazione (MacCarthy, 2023). L'implementazione è inoltre soggetta a un'ampia discrezionalità degli attuatori, finalizzata in questo caso a non ridurre con misure troppo stringenti il potenziale di innovazione¹⁸. Uno strumento è il "sistema di archiviazione algoritmica", un database online di algoritmi che hanno "proprietà di opinione pubblica o capacità di mobilitazione sociale", nel quale tutti i produttori di IA devono inserire i loro sistemi (Sheehan, 2023, p. 13). L'obiettivo dello strumento è, in sintesi, assicurare che gli algoritmi evitino di dare risposte "scorrette" – ossia non coerenti con la dottrina del Partito e il "pensiero di Xi Jinping (Cameron

¹⁴ Fonte: <https://www.wired.it/article/cina-riconoscimento-facciale-nuove-regole/>

¹⁵ Fonte: <https://edition.cnn.com/2023/07/14/tech/china-ai-regulation-intl-hnk/index.html>

¹⁶ Fonte: <https://www.theguardian.com/technology/2025/jan/28/we-tried-out-deepseek-it-works-well-until-we-asked-it-about-tiananmen-square-and-taiwan>

¹⁷ Fonte: <https://www.holisticai.com/blog/generative-ai>

¹⁸ Fonte: <https://time.com/6314790/china-ai-regulation-us/>

et al., 2025) – a domande di carattere politico degli utenti, senza però dare l’impressione di essere censurate (Benaich e Chalmers, 2024).

Preoccupazioni più specifiche per la sicurezza – intesa sia come *safety*, sia come *security* (uso malevolo) – sono entrate nell’agenda cinese nel 2024, quando il più alto grado di indirizzo politico del Partito Comunista¹⁹ ha richiesto di “istituire un sistema di supervisione e regolamentazione della sicurezza dell’IA”²⁰, alludendo indirettamente ai rischi per la sicurezza dei sistemi di IA di frontiera, simili a quelli a cui si fa riferimento nel discorso occidentale (Sheehan, 2024c).

La politica cinese per la sicurezza dell’IA è perseguita anche attraverso cooperazioni con istituti di ricerca di vari Paesi. Sin dall’inizio, infatti, la regolazione ha avuto anche finalità strategiche internazionali: da un lato fugare la percezione del ruolo della Cina come minaccia militare e di competizione economica, scientifica e tecnologica per altri Paesi, dall’altro mantenere un posto nelle arene della governance internazionale dell’IA (Sheehan, 2024c) e conquistare un vantaggio strategico attraverso una leadership internazionale sull’etica e la sicurezza dell’IA (Turner, 2019, p. 248). Come? Essere al primo posto nella regolazione e nella produzione di standard internazionali sull’IA può consentire di divenire un forte *cyber-power* attraverso il soft-power, ossia la capacità di indirizzare discorsi e narrative internazionali su questi temi (Turner, 2019, p. 235). Infatti, il governo cinese ha preso parte a più iniziative delle UN e a summit internazionali sull’IA impegnandosi per portare nelle agende sia la governance etica dell’IA sia l’accesso all’IA nei Paesi meno sviluppati, come dichiarato nel documento del *Ministry of Foreign Affairs AI Capacity-Building Action Plan for Good and for All* (2024)²¹. Questa strategia internazionale cerca inoltre di ottenere vantaggi politici dal sostenere il sistema delle UN, in cui sono gli stati membri a decidere²².

La strategia di governo dell’IA nello UK si basa su un approccio favorevole all’innovazione (Papyshev e Yarime, 2023) ed esclude regolamentazioni legislative. La strategia del 2021-2022²³ conteneva già riferimenti alla sicurezza e ai rischi a lungo termine posti da un’IA non allineata (Benaich e Hogarth, 2022). Il documento presentato al Parlamento nel 2023 con il

¹⁹ Il Terzo Plenum, riunione dei massimi dirigenti del Partito che si tiene una volta ogni cinque anni.

²⁰ Fonte: <https://www.pekingnology.com/p/full-text-resolution-of-the-central>

²¹ Fonte: https://www.mfa.gov.cn/eng/wjbzhd/202409/t20240927_11498465.html

²² China AI Development and Safety Network (<https://ai-development-and-safety-network.cn>).

²³ Fonte: <https://www.gov.uk/government/publications/national-ai-strategy/national-ai-strategy-html-version>

significativo titolo *A pro-innovation approach to AI Regulation* rassicurava circa la priorità dello sviluppo e su una normazione ispirata sì ai principi dell’OECD del 2019, ma con un approccio di *soft law* e *business-friendly* (OECD, 2023): “dopo l’uscita dall’UE siamo liberi di stabilire un approccio normativo che ci permetta di affermare lo UK come superpotenza dell’IA” e una “innovazione responsabile”²⁴. A questo scopo lo *AI Safety Institute*, una sorta di “*startup* nel governo”, deve testare sistemi avanzati di IA, ridurre il divario tra capacità dell’IA e rischi, promuovere la collaborazione tra imprese, governi e scienziati per ridurre i rischi, far progredire la ricerca e rafforzare le pratiche e le politiche di sviluppo dell’IA a livello globale²⁵. La priorità dell’innovazione sulla regolazione è trasversale e immune ai cambiamenti di colore politico del governo britannico. In altre parole, conservatori e democratici sembrano essere *on the same page* quando promuovono politiche di IA. Dopo un libro bianco del 2023 del governo conservatore “per guidare l’uso dell’IA in UK, per promuovere un’innovazione responsabile e mantenere la fiducia dell’opinione pubblica in questa tecnologia rivoluzionaria”, ispirato a sicurezza, trasparenza ed equità²⁶, si è passati all’*Action Plan* del governo laburista (2025) che ribadisce che “l’approccio pro-innovazione dello UK alla regolamentazione è una fonte di forza rispetto ad altre giurisdizioni più regolamentate” e che la protezione dei cittadini dai rischi serve a “promuovere la fiducia del pubblico nella tecnologia, in particolare considerando gli interessi dei gruppi emarginati”. Non da ultimo il Piano accorda la stessa regolazione con gli obiettivi di innovazione e afferma che “la regolamentazione, la sicurezza e la garanzia hanno anche il potere di guidare l’innovazione e la crescita economica, come dimostrato dal successo delle *sandbox* normative nel sostenere le *startup fintech* e lo sviluppo dell’industria della sicurezza informatica dello UK”. Si può considerare un’attuazione il programma *Safeguarded AI*, con il quale l’agenzia governativa ARIA finanzia (per 59 milioni di sterline) azioni per sviluppare standard di sicurezza “consentendo di sfruttare appieno i vantaggi economici e sociali dei sistemi avanzati di IA riducendo al minimo i rischi”²⁷.

Arrivati a questo punto è opportuno domandarsi da cosa dipendono gli orientamenti strategici dei “giganti”? Quali sono quindi i fattori che influenzano le loro agende politiche e strategie nazionali? Fra i fattori che deter-

²⁴ Fonte: <https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach/white-paper>

²⁵ Fonte: <https://www.aisi.gov.uk/about>

²⁶ Fonte: <https://www.gov.uk/government/news/uk-unveils-world-leading-approach-to-innovation-in-first-artificial-intelligence-white-paper-to-turbocharge-growth>

²⁷ Fonte: <https://www.aria.org.uk/opportunity-spaces/mathematics-for-safe-ai/safeguarded-ai/>

minano analogie e differenze conta in primo luogo il posizionamento di ciascun paese nei mercati globali dell'IA, il modo in cui le poste in gioco economiche e politiche esaminate nel par. 1.3 sono interpretate nei diversi sistemi politici e quanto l'IA stessa è in grado mobilitare capitali, dati e infrastrutture su scala nazionale. In un ranking globale della capacità di IA (nel 2024) gli Usa sono al primo posto, seguiti da Cina, Singapore e UK (l'Italia si colloca al 24° posto)²⁸. Nel 2022, la più alta concentrazione (32%) di *top supercomputer* si trovava in Cina, seguita da Usa (25%), Germania (7%), Giappone (6%), Francia (5%) e UK (3%) (OECD, 2023b). Alla luce del *frame* dominante – innovazione e competitività vs. regolazione – a spiegare la riluttanza della politica statunitense a regolare l'IA è quindi il timore di indebolire la propria leadership nell'economia digitale e favorire i suoi competitori. Infatti, anche se la Cina ha meno investimenti privati nell'IA rispetto agli Usa, gli investimenti stranieri sono in crescita, con l'Arabia Saudita in testa (Omaar, 2024).

Giocano però un ruolo anche fattori più strettamente relativi alla dimensione politica, come le culture politiche prevalenti, i rapporti fra élite statali e Big Tech, le caratteristiche della leadership politica. Il primo fattore di cui ci occupiamo sono quindi le culture politiche.

I contesti nazionali analizzati sono ancora caratterizzati dal prevalere di visioni e culture consolidati, che orientano i *frame* delle politiche. Nel campo delle democrazie occidentali e del cosiddetto WEIRD *framework* (*Western, Educated, Industrialized, Rich, and Democratic*) (Denford *et al.*, 2023), si intravede agevolmente una polarizzazione fra due orientamenti: quello neoliberista tipico dei Paesi anglosassoni, che anche per l'IA favorisce politiche *market friendly* per la crescita della tecnologia; quello che caratterizza complessivamente i Paesi dell'UE e l'UE stessa, più propenso a regolare i mercati, compresi quelli tecnologici e dell'IA, per mitigare i rischi sociali e tutelare i diritti, in cui la politica può assumere un ruolo di guida del sistema sociotecnico dell'IA, attraverso normative e standard etici (Guenduez e Mettler, 2023). Non a caso, l'adozione di strategie politiche per lo sviluppo dell'IA e delle tecnologie digitali attraverso regole, investimenti pubblici e incentivi a quelli privati, è stato interpretato anche come parte delle risposte ai fallimenti del neoliberismo (Tooze e Bello, 2025). In Cina, invece, i *frame* evolvono più lentamente, seguendo gli orientamenti delle élite di partito e statali senza mettere in discussione i riferimenti di fondo e al riparo da repentini cambiamenti elettorali, come quelli che caratterizzano i sistemi liberaldemocratici.

Il secondo fattore riguarda la configurazione dei rapporti fra élite statali

²⁸ Fonte: <https://www.tortoisemedia.com/data/global-ai#rankings>

e Big Tech. Essi possono caratterizzarsi per la presenza di tensioni fra il creare condizioni favorevoli per le imprese tecnologiche, sostenendole e riducendo al minimo le potenziali barriere, gli ostacoli e le restrizioni pubbliche e il garantire anche diritti sociali come diversità, equità e trasparenza (Guenduez e Mettler, 2023). Queste tensioni si devono anche al fatto che i grandi interessi privati, che dominano il campo attraverso il controllo sulle principali piattaforme e sui dati (Ulnicane *et al.*, 2021), possono esercitare un'influenza sproporzionata sullo sviluppo dell'IA rispetto ad attori della società civile e agli stessi poteri pubblici. Anche negli Usa il rapporto fra politica e Big Tech non è stato privo di conflitti, come evidenziano le tante cause *antitrust* promosse dallo stato contro il dominio di Apple, Amazon, Meta, Microsoft e Google non nel campo dell'IA²⁹. Un altro caso, che ha implicazioni anche relative all'IA, è quello di TikTok³⁰. Quando le relazioni riguardano l'IA, però, il gioco sembra cambiare: l'elevato tasso di sviluppo tecnologico e i profitti consentono alle Big Tech una considerevole capacità di negoziare e orientare le politiche statali dell'IA. Diviene così difficile distinguere le politiche di interesse pubblico dagli interessi di dominio del mercato, una caratteristica della logica plurale in gioco nei sistemi di governance ibridi che è visibile, ad esempio nella composizione degli organismi consultivi istituzionali (Radu, 2021). Si pensi in tal senso al ruolo di E. Schmidt (consulente tecnico di Alphabet, proprietaria di Google, già amministratore delegato di Google) che, come presidente della *National Security Commission on AI* (NSCAI), ha avuto il compito di formulare raccomandazioni al Presidente e al Congresso. Come abbiamo visto, inoltre, nei negoziati fra poteri politici ed economici sull'IA conta anche l'autonomia di ciascuno degli interlocutori; se le imprese temono i potenziali vincoli pubblici, le élite politiche scontano anche la dipendenza cognitiva dai padroni dell'IA: “in mancanza di competenze tecnologiche, i legislatori (Usa) si affidano sempre più ad Anthropic, Microsoft, OpenAI, Google e altri produttori di IA per spiegarne il funzionamento e contribuire alla creazione di regole”³¹. Con la

²⁹ Ad esempio, nel 2024 con oggetto il browser Chrome (67% del mercato globale) e il monopolio della ricerca online di Google, o il sistema Android, con una quota di mercato stimata al 71% e la richiesta di obbligare Google a cedere qualsiasi partecipazione in società di IA che controllano tecnologie che potrebbero competere con i motori di ricerca. Fonte: <https://www.nytimes.com/2023/05/03/opinion/ai-lina-khan-ftc-technology.html>, NYT 3, 5, 2023.

³⁰ Nel conflitto fra gli Usa e TikTok la posta in gioco è relativa non solo alla sua posizione di oligopolio, ma anche al controllo dell'algoritmo che ne consente l'operatività. Il governo cinese ha a suo tempo comunicato che bloccherebbe una vendita a investitori Usa se incluodesse l'algoritmo che rende TikTok estremamente prezioso.

Fonte: <https://www.nytimes.com/live/2025/01/10/us/tiktok-ban-supreme-court>

³¹ Fonte: <https://www.nytimes.com/2023/12/06/technology/ai-regulation-policies.html>

seconda presidenza di D. Trump la “Tech Right” sembra configurarsi come una vera e propria coalizione fra élite politica e Big Tech entrambe a leadership fortemente personalizzata.

Quando è lo stato, invece, a scegliere la traiettoria di sviluppo della tecnologia e invitare gli attori privati a unirsi al percorso è perché esiste una sovrapposizione di fatto tra potere politico ed economico. Le élite politiche di Paesi come Russia, Cina, Corea del Sud e Giappone controllano le imprese dell’IA o hanno maggiore autonomia da loro (Papyshev e Yarime, 2023). In Cina, dove lo Stato svolge un ruolo fondamentale nell’economia, la forte integrazione fra pubblico e privato non ha frenato la capacità innovativa della ricerca e delle imprese nell’IA. Anche se l’industria privata contribuisce per oltre il 50% al gettito fiscale, il 60% al PIL e l’80% all’occupazione urbana³², nell’“economia socialista di mercato” lo Stato e il Partito mantengono un controllo sull’allineamento dei comportamenti del business privato alle politiche pubbliche. La stessa distinzione tra Stato, Partito e imprese private si è fatta più labile³³ ed è comunque vista come una relazione reciprocamente vantaggiosa. Ad esempio, non sarebbe stato possibile l’exploit di DeepSeek nel 2025, se questa startup non avesse beneficiato di buoni legami con il Partito e di un più ampio ecosistema che la Cina ha costruito per accelerare l’innovazione dell’IA (Scanlon, 2025), non solo finanziario – trattamento fiscale favorevole, sussidi governativi e sovvenzioni alla ricerca (1.000 miliardi di yuan per l’IA³⁴) – ma anche politico. Molte misure per l’IA non sono state infatti adottate attraverso editti politici unilaterali, ma in un sistema di azione più plurale di quanto ci potrebbe immaginare. Al suo interno diversi settori dello Stato e imprese sono in competizione per la leadership nella governance dell’IA (Sheehan, 2024c) in “un processo dinamico e iterativo guidato da un mix di attori interni ed esterni al Partito-Stato: burocrati di medio livello, accademici, tecnologi, giornalisti e ricercatori di politiche presso le aziende tecnologiche della piattaforma. (...) Grazie a un mix di advocacy pubblica, dibattiti intellettuali, workshop tecnici e contese burocratiche, questi attori hanno gettato le basi per le attuali e future normative cinesi sull’IA” (Sheehan, 2024, p. 2). In questo sistema di azione, per ottenere controllo e supervisione del settore privato e limitare l’autonomia dei leader delle imprese tecnologiche, Stato e Partito adottano misure di vario tipo (Fu, 2023): joint venture e investimenti a partecipazione mista, che

³² Fonte: http://english.scio.gov.cn/m/pressroom/2022-06/29/content_78295756.htm

³³ Fonte: <https://www.cna.org/our-media/indepth/2024/09/fused-together-the-chinese-communist-party-moves-inside-chinas-private-sector>

³⁴ Bank of China, gennaio 2025, https://www.boc.cn/aboutboc/bi1/202501/t2025-0123_25254674.html)

offuscano la distinzione tra pubblico e privato³⁵; cooptazione di esponenti delle imprese private nella struttura politica e inserimento di funzionari politici nelle gerarchie aziendali per il controllo politico delle decisioni di business. In questo modo si producono sia una penetrazione nel mercato da parte del partito, sia l'attribuzione di identità politiche agli imprenditori privati che, partecipando alle attività del Partito, ne rafforzano la legittimità e il controllo (Ma e He, 2018). Nel campo dell'IA questo permette di perseguire innovazione e regolamentazione. Ai "campioni dell'IA" come Baidu, Alibaba, Tencent e SenseTime vengono offerti sostegno finanziario, trattamento preferenziale e controllo statale.

Il caso del gigante Usa consente di evidenziare un terzo fattore di influenza sulle strategie nazionali: il ruolo della leadership politica. Il mutare nel tempo delle relazioni fra politica e Big Tech, le capovolte delle imprese nel campo dell'etica, ci dicono anche che la leadership politica – che solo nelle democrazie costituzionali dipende da contese elettorali e alternanza al governo – è a sua volta importante. Negli Usa, la radicalizzazione di una politica di competitività dell'IA a scapito delle tutele di diritti è da ricondurre all'attuazione delle priorità politiche del programma elettorale repubblicano, che proponeva di abrogare l'*executive order* sull'IA di J. Biden, poiché "ostacola l'innovazione dell'IA e impone idee di sinistra radicale allo sviluppo di questa tecnologia"³⁶. Proprio l'aspettativa di un'era di *laissez-faire* per l'IA³⁷ e di uno spostamento del *trade-off* da equità, responsabilità, trasparenza e inclusività verso velocità, produttività e risultati economici³⁸ ha attratto il sostegno a D. Trump dei grandi nomi della *Silicon Valley* (Benaich e Chalmers, 2024). OpenAI ha preso posizione in favore della svolta politica sull'IA, aggiungendo le proprie "proposte di *policy* per estendere la leadership globale dell'America nell'innovazione dell'IA [...] (sostenendo) "l'imprenditorialità e le libertà individuali che sono state a lungo il cuore dell'ecosistema innovativo americano"³⁹. Nella *blueprint* di OpenAI "gli altri Paesi sono visti in un'ottica binaria: avversari o alleati (...) deregolamentare, dunque, per OpenAI significa garantire la supremazia statunitense" (Puliafito, 2025c), soprattutto nei confronti della Cina.

In Cina l'influenza della leadership politica assume una funzione diversa che, tuttavia, ribadisce la rilevanza di questo fattore. Come è noto, nel siste-

³⁵ Fonte: <https://www.cna.org/our-media/indepth/2024/09/fused-together-the-chinese-communist-party-moves-inside-chinas-private-sector>

³⁶ Fonte: <https://www.euronews.com/next/2024/12/05/as-donald-trumps-second-administration-begins-to-takes-shape-what-will-it-mean-for-ai>

³⁷ Fonte: <https://futurism.com/the-byte/trump-repeal-restrictions-ai>

³⁸ Fonte: <https://outofcontrol.substack.com/p/trumps-victory-top-ten-ai-policy>

³⁹ Fonte: <https://openai.com/global-affairs/openais-economic-blueprint/>

ma politico cinese non vi sono cambiamenti di leadership dovuti a elezioni. Tuttavia, i mutamenti nelle agende della leadership politica sono importanti per le imprese, pubbliche e private, che devono conformarsi alle strategie di sviluppo tecnologico nazionale e di regolazione adottate di volta in volta dai dirigenti del Partito e dello Stato.

3.2. La politica dell’Unione europea: strategie di regolazione e innovazione

In questo paragrafo sono presentati gli aspetti principali dell’agenda politica dell’UE per il governo dell’IA. Una prima parte è dedicata alla strategia di regolazione, considerata da molti un esempio virtuoso da seguire – anche se mai imitato per davvero da altri governi. La seconda parte si occupa della strategia per lo sviluppo dell’IA e l’innovazione. Nel delineare la sua politica per l’IA, l’UE ha cercato di rendere compatibili lo sviluppo tecnologico ed economico di una IA europea e la tutela dei diritti individuali e collettivi che da questo sviluppo possono essere messi a rischio. Ha quindi cercato di abbinare una strategia *market oriented* e una *social oriented*. Due prospettive collegate fra loro dall’idea che la fiducia nell’IA possa essere un ingrediente per il suo uso e sviluppo. Un’IA affidabile che diventa così da narrazione un obiettivo e un caso specifico nell’ambito della più ampia *policy* del “mercato unico europeo” (af Malmborg, 2022). In conclusione, sono tratteggiati i fattori che sinora hanno influenzato agende, strategie e quella che, all’inizio del 2025, appare una svolta, che sta modificando la priorità politica fra le strategie di regolazione e sviluppo in favore della seconda.

L’IA ha fatto ingresso nell’agenda dell’UE nel 2018⁴⁰, quando la Commissione, dopo un invito del Consiglio ad “affrontare le tendenze emergenti, fra cui IA e tecnologie *blockchain*”⁴¹, ha prodotto una comunicazione nella quale erano indicati gli obiettivi di una politica dell’IA: incrementare la capacità industriale e tecnologica dell’UE e l’adozione dell’IA in tutti i settori economici, prepararsi ai cambiamenti socioeconomici, assicurare un quadro etico e giuridico adeguato”⁴². Poco dopo, un *Coordinated Plan on AI* invitava gli stati membri a contribuire con proprie strategie, ma in modo non frammentato, allo

⁴⁰ Una cronologia delle attività della Commissione è pubblicata al termine della pagina “Approccio europeo all’intelligenza artificiale” (<https://digital-strategy.ec.europa.eu/it/policies/european-approach-artificial-intelligence>).

⁴¹ 19.10.2017 European Council meeting – Conclusions (<https://data.consilium.europa.eu/doc/document/ST-14-2017-INIT/en/pdf>).

⁴² L’intelligenza artificiale per l’Europa, 25.4.2018 (<https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52018DC0237>).

sviluppo di un'IA europea non solo all'avanguardia, ma anche etica e sicura⁴³. Dopo una dichiarazione di cooperazione tra gli Stati membri, La Commissione ha istituito un *High Level Group on AI* di 52 esperti (presieduto dall'ex presidente di Nokia), il cui primo rapporto (2019) ha indicato principi etici e raccomandazioni per una gestione dei rischi anche favorevole alla competizione (un ecosistema di fiducia in un'IA affidabile) che hanno influenzato tutta la successiva normativa europea (Malaschini, 2024, p. 38). Da questo momento in poi, la strategia *market oriented* per lo sviluppo e l'innovazione nell'IA e quella *social oriented* hanno proceduto parallelamente, ma la seconda è stata a lungo prevalente. Iniziamo dunque ricostruendo le caratteristiche, i fattori di influenza e gli esiti della strategia *social oriented*: regolare l'IA per mitigare i rischi sociali e tutelare i diritti delle persone.

Sulla base del *Coordinated Plan on AI* e di un Libro Bianco⁴⁴, elaborato a partire dal lavoro dell'*High Level Group*, la Commissione ha lanciato una consultazione di Stati membri, società civile, industria (comprese le Big Tech Google, Amazon, IBM, Microsoft, Facebook) e mondo accademico, conclusa nel giugno 2021⁴⁵. Le principali preoccupazioni emerse riguardavano la sicurezza e possibilità che l'IA violasse i diritti fondamentali e il suo uso potesse produrre effetti discriminatori, ma vi erano opinioni discordanti circa il limitare o meno alle sole applicazioni di IA ad alto rischio il rispetto di requisiti obbligatori. E anche opinioni dubbiose circa l'inopportunità di regolare troppo, come quella del governo irlandese (Justo-Hanani, 2022). Ne è derivata una proposta di regolamento della Commissione (aprile 2021), basata sull'idea che l'IA non fosse buona o cattiva di per sé, ma utilizzabile in modi e con esiti diversi. Per questo, dovevano esser regolate le modalità di uso dell'IA, più che la tecnologia sottostante, con una disciplina differenziata secondo i diversi livelli di rischio derivanti dagli usi concreti di sistemi e modelli di IA (cfr. sotto).

Il Parlamento Europeo aveva esaminato e modificato la proposta della Commissione e il Consiglio dell'UE aveva appena proposto un testo di compromesso sugli obblighi dei fornitori e degli utenti di sistemi ad alto rischio

⁴³ Comunicazione al Parlamento, al Consiglio, al Consiglio, al Comitato economico e sociale e al Comitato delle regioni, 7.12.2018 (<https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52018DC0795>).

⁴⁴ 19.02.2020, *White paper on Artificial Intelligence - A European approach to excellence and trust* (https://commission.europa.eu/document/download/d2ec4039-c5be-423a-81ef-b9e44e79825b_en?filename=commission-white-paper-artificial-intelligence-feb2020_en.pdf).

⁴⁵ Raccogliendo 1.215 contributi: 352 di aziende od organizzazioni/associazioni imprenditoriali; 152 di istituzioni accademiche e di ricerca; 73 di autorità pubbliche; 160 di società civile no-profit, ONG, sindacati (<https://digital-strategy.ec.europa.eu/en/library/white-paper-artificial-intelligence-public-consultation-towards-european-approach-excellence-and>

quando, alla fine del 2022, è arrivata l’innovazione dirompente dell’IA generativa, con ChatGPT. Questo avvento ha messo alla prova il processo regolativo, mostrando la reattività delle istituzioni europee e smentendo le opinioni sulla lentezza della politica, una delle principali debolezze attribuite alla regolazione “dura”. Nel dicembre 2022 Il Consiglio dell’UE ha aggiunto piuttosto rapidamente una specifica proposta di regolazione dei sistemi di IA *general purpose*⁴⁶, nell’ambito di una normativa generale per il rispetto dei diritti fondamentali e dei valori dell’Unione nell’uso di sistemi di IA. Nel giugno 2023 il Parlamento ha adottato la sua posizione negoziale e sei mesi dopo Parlamento e Consiglio dell’UE hanno raggiunto un accordo provvisorio. Il 13 marzo 2024 l’AI Act è stato approvato dal Parlamento con 523 voti favorevoli, 46 contrari e 49 astensioni, con posizioni favorevoli (in maggioranza) e contrarie che hanno attraversato tutti i principali gruppi parlamentari⁴⁷. L’approvazione finale del “Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio, che stabilisce regole armonizzate sull’intelligenza artificiale (...)" è arrivata nel giugno 2024⁴⁸.

L’AI Act è quindi il risultato di un denso processo politico, in cui il rispetto del copione istituzionalizzato del *policy making* dell’UE non ha impedito la negoziazione e l’adattamento sostanzioso delle norme dopo l’avvio del procedimento. Oltre alle difficoltà tecniche, dovute anche a una tecnologia in rapido mutamento, i problemi politici più importanti sono stati affrontati nella seconda parte del 2023 durante il “trilogo”, la normale trattativa fra Parlamento, Consiglio dell’UE e Commissione. L’opposizione dei governi francese, tedesco e italiano alla disciplina più rigorosa inizialmente prospet-

⁴⁶ L’Art. 3 par. 63 dell’AI Act li definisce «un modello di IA, anche quando tale modello di IA è addestrato con una grande quantità di dati utilizzando l’autocontrollo su larga scala, che mostra una generalità significativa ed è in grado di eseguire con competenza un’ampia gamma di compiti distinti indipendentemente dal modo in cui il modello è immesso sul mercato e che può essere integrato in una varietà di sistemi o applicazioni a valle». Questi sistemi, anche definiti “modelli di base” (*foundation models*), sono caratterizzati dall’uso (pubblico e privato) come modelli pre-addestrati per sistemi di AI più specializzati. Ne sono esempio i *LLMs*, divenuti famosi per l’uso di massa di Chat GPT.

⁴⁷ Hanno votato a favore quasi tutti i parlamentari del Partito Popolare Europeo (153 favorevoli, 1 contrario, 4 astenuti), dell’Alleanza Progressista dei Socialisti e Democratici (129 favorevoli, 1 astenuto), di Renew Europe (80 favorevoli, 2 contrari, 7 astenuti), dei Verdi/Alleanza Libera Europea (58 favorevoli, 6 contrari), dei Conservatori e Riformisti Europei (38 favorevoli, 8 contrari, 13 astenuti), di Identità e Democrazia (41 favorevoli, 1 contrario, 11 astenuti), mentre Sinistra Unitaria Europea/Sinistra Verde Nordica ha avuto 4 favorevoli, 21 contrari, 4 astenuti (https://www.europarl.europa.eu/doceo/document/PV-9-2024-03-13-VOT_EN.html?item=2).

⁴⁸ Una cronologia del processo si trova all’URL <https://artificialintelligenceact.eu/developments/>. Il testo del regolamento si trova all’URL https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=OJ:L_202401689

tata per *foundation model* e *IA general purpose* e la visione complessivamente più “permissiva” nel Consiglio – ad esempio sull’uso dell’identificazione biometrica da remoto da parte delle forze dell’ordine – testimoniano di una polarizzazione: nel regolare l’IA il Parlamento è stato più propenso a tutelare diritti e prevenire rischi, mentre il Consiglio è stato più sensibile alle esigenze di ricerca, sviluppo e commercializzazione fatte valere dal mercato (Malaschini, 2024, p. 49).

Quali sono stati gli esiti? Le norme più importanti dell’AI Act⁴⁹ differenziano le prescrizioni secondo una classificazione dei rischi derivanti dall’utilizzazione. In particolare, esso distingue tra quattro livelli di rischio: inaccettabile, alto, basso e minimo. Tanto maggiori sono i rischi, quanto maggiori sono gli obblighi che il fornitore di IA deve rispettare, per non incorrere in multe: fino al 7% del fatturato globale per le pratiche vietate e il 3% per altre violazioni. Tuttavia, le sue norme e divieti non si applicano ai sistemi di IA *open source*, o usati per finalità militari, di difesa, sicurezza nazionale, o di sola ricerca scientifica.

Il Regolamento vieta le applicazioni di IA da cui può derivare un rischio inaccettabile, definito sulla base della congruenza con i valori e i diritti fondativi dell’UE. In particolare, sono vietati: sistemi di *social scoring* (valutare o classificare individui o gruppi in base al comportamento sociale o ai tratti personali); tecniche di IA subliminali, manipolative o ingannevoli; sistemi di riconoscimento facciale e identificazione biometrica in tempo reale in spazi accessibili al pubblico, anche per le forze dell’ordine (con alcune eccezioni); sistemi di categorizzazione biometrica che deducono attributi sensibili (razza, opinioni politiche, appartenenza a sindacati, convinzioni religiose o filosofiche, vita sessuale o orientamento sessuale); uso di IA predittiva da parte delle polizie per il *profiling* individuale; sistemi che sfruttano le vulnerabilità individuali e sociali, o riconoscono le emozioni nei luoghi di lavoro o in scuole o università. Di contro i sistemi di IA da cui può derivare un alto rischio sono regolamentati e dunque consentiti. Ad esempio: la profilazione di persone e trattamento automatizzato di dati per valutare aspetti della vita di una persona, come rendimento lavorativo, situazione economica, salute, preferenze, interessi, affidabilità, comportamento, ubicazione o movimento. Per mettere questi sistemi sul mercato, sviluppatori, fornitori (europei e non) e operatori pubblici o privati che li utilizzano devono garantire un sistema di gestione del rischio lungo tutto il ciclo di vita del sistema di IA e assicurare che i set di dati di formazione, convalida e test siano pertinenti, sufficientemente rappresentativi, privi di errori e completi. Oltre a dover adempiere a una serie di altri ob-

⁴⁹ Per una descrizione ufficiale (Consiglio Europeo): <https://www.consilium.europa.eu/it/policies/artificial-intelligence/>

blighi⁵⁰. I sistemi di IA a rischio limitato sono soggetti solo a *obblighi di trasparenza*: gli sviluppatori e i distributori devono garantire che gli utenti finali siano consapevoli di interagire con l'IA (ad esempio: *chatbot* e *deepfake*).

Ai modelli *general purpose* è stato invece riservato un trattamento a parte, esito di un rapido adattamento delle norme durante il loro esame, e basato sull'elaborazione del concetto di rischio sistematico: «un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore» (art. 3, 65). La definizione di rischio sistematico riflette scenari o “immaginari” di incertezza per il futuro particolarmente allarmanti. I rischi sistematici possono essere sia non intenzionali sia deliberatamente perseguiti da usi malevoli.

Fra gli usi malevoli intenzionali vi sono soprattutto reati informatici, rischi chimici, biologici, radiologici e nucleari, specialmente se per uso anche bellico, possibilità di persuasione, manipolazione e disinformazione su larga scala, con rischi per i valori democratici e i diritti umani, come l'interferenza elettorale, la perdita di fiducia nei media e l'omogeneizzazione o l'eccessiva semplificazione delle conoscenze. Immaginari di questo tipo hanno fornito a tutti gli attori coinvolti nel processo di formulazione e decisione dell'AI Act il quadro di senso, il *policy frame* fondamentale per confrontarsi con la complessità dei modelli *general purpose* (Corsi e d'Albergo, 2024) e definire le misure da adottare. Dunque, è stato necessario stabilire obblighi specifici che sono stati inseriti nel Regolamento. Tutti questi modelli sono tenuti a rispettare parametri di trasparenza più severi, che la Commissione ha in seguito dettagliato attraverso un “Codice di condotta per l'IA di uso generale”⁵¹, traducendo i principi dell'AI Act in misure e metriche attuabili. Usando questa guida, i fornitori possono dimostrare la conformità dei loro modelli alla legge durante il loro intero ciclo di vita.

Una regolazione “dura” non può prescindere dalla definizione del sistema di *enforcement* delle norme. La Commissione ha infatti già avviato l'attuazione dell'AI Act⁵², con modalità che replicano il sistema consolidato di attuazione delle leggi dell'UE. L'implementazione è gestita dall'*European AI Office*⁵³,

⁵⁰ Per dettagli da fonte ufficiale (Commissione UE): <https://digital-strategy.ec.europa.eu/it/policies/regulatory-framework-ai>; FAQ: https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683

⁵¹ Fonte: <https://digital-strategy.ec.europa.eu/en/library/first-draft-general-purpose-ai-code-practice-published-written-independent-experts>

⁵² Per informazioni sul processo di implementazione e i suoi avanzamenti: <https://digital-strategy.ec.europa.eu/it/news/first-rules-artificial-intelligence-act-are-now-applicable>

⁵³ L'AI Office è stato istituito all'interno della Commissione Europea per l'attuazione dell'AI Act e il supporto agli Stati Membri nell'implementazione nazionale del Regolamento.

dallo *AI Board*, composto da rappresentanti per ciascuno stato membro, con un *Advisory Forum* che fornisce expertise tecniche e consulenza e con uno *Scientific Panel of Independent Experts*. L'attuazione si basa su un mix di *hard* e *soft law*. Da un lato, esso prevede l'individuazione di standard, linee guida e codici di condotta, strumenti che dovrebbero facilitare il rispetto delle prescrizioni da parte delle imprese (Cameron *et al.*, 2025). In questo passaggio, critico anche per l'allineamento con altri standard internazionali, sono coinvolti attori economici (fornitori, acquirenti e utenti professionali, importatori e distributori di sistemi di IA) e no-profit, per garantire che le linee guida tengano conto sia degli interessi, sia di considerazioni etiche, anche se ciascuno ritiene che sia dato spazio sufficiente per le proprie preoccupazioni. Dall'altro lato, per disciplinare l'IA il Regolamento introduce le *regulatory sandboxes* che saranno create dalle autorità competenti per l'IA designate dagli Stati membri (artt. 57 e 58). Esse possono essere intese come una sorta di ambienti controllati, creati ad hoc, in cui le imprese possono testare sistemi innovativi di IA sotto la supervisione dell'autorità di regolamentazione senza dover affrontare immediatamente tutti gli oneri normativi, ma sapendo che stanno sviluppando le tecnologie nella direzione ritenuta giusta. La strategia regolativa dell'UE prevede un ulteriore elemento di ibridazione tra regolazione *hard* e *soft* relativo al regime transitorio che condurrà alla piena attuazione del Regolamento. Ci riferiamo all'AI Pact, avviato nel 2024 per incoraggiare e sostenere chi produce, vende e usa IA ad allinearsi volontariamente con l'AI Act anche prima della sua entrata in vigore⁵⁴: una sorta di "comunità collaborativa" di imprese, no-profit, esperti, amministratori pubblici, che condividono le esperienze e le pratiche che stanno realizzando per anticipare la conformità con la regolazione.

Questo insieme di principi, norme, strumenti e procedure sta dando luogo a un voluminoso insieme di attività da parte degli uffici e organismi istituiti, che coinvolge funzionari, esperti, esponenti del mondo imprenditoriale e della società civile no-profit. Al suo interno proseguono le tensioni e i negoziati che hanno caratterizzato le fasi di formulazione e decisione dell'AI Act, evidenziando le difficoltà di rendere coerenti, o compatibili le due strategie europee. Esaminiamo ora quella orientata al mercato, la strategia di sviluppo dell'IA, e le sue relazioni con la strategia di regolazione. Obiettivo della politica europea: investire in IA e creare un ecosistema per l'innovazione.

L'esigenza di favorire lo sviluppo di un'IA europea è stata presente fin dal

Tra i suoi compiti rientrano lo sviluppo di strumenti per la classificazione dei sistemi di IA secondo i livelli di rischio, l'elaborazione di codici di condotta rivolti agli sviluppatori, l'accertamento delle violazioni alle norme e la gestione delle comunicazioni con i fornitori delle tecnologie. Per conoscere le attività dell'AI Office si rimanda a: <https://digital-strategy.ec.europa.eu/en/policies/ai-office#ecl-inpage-tasks-of-the-ai-office>

⁵⁴ Fonte: <https://digital-strategy.ec.europa.eu/en/policies/ai-pact>

Piano Coordinato per l'IA del 2018, che proponeva un quadro strategico per le strategie nazionali e misure per favorire un mercato europeo dell'IA, come un partenariato pubblico-privato europeo ad hoc e finanziamenti per le start-up e la ricerca. Da lì ha preso forma una strategia dell'UE per l'IA nella quale è stata a lungo sostanzialmente assente l'aspirazione a un primato nella competizione mondiale per l'economia dell'IA, centrale invece nelle strategie di altri "giganti". Piuttosto l'obiettivo è stato quello di non dipendere dagli Usa, né dalla Cina (Beroche *et al.*, 2022), identificata come un "rivale sistemico che promuove modelli alternativi di governance" (Wang e Galligan, 2021).

Nel 2021 la Commissione ha precisato un "approccio europeo all'IA"⁵⁵ che, partendo dalle due facce dell'IA (opportunità e rischi), proponeva un "pacchetto", affiancando al quadro normativo per l'IA che diverrà l'AI Act una revisione del Piano coordinato, per stabilire condizioni favorevoli al suo sviluppo e adozione. A questo fine, riteneva necessario: costruire un quadro di governance per acquisire, accumulare e condividere informazioni strategiche sull'IA; sfruttare e liberare il potenziale dei dati; promuovere un'infrastruttura di calcolo di scala adeguata; rendere possibile "un'eccellenza che prospera dal laboratorio al mercato", trasferendo i risultati della ricerca alle imprese; collaborare con i portatori di interessi attraverso partenariati pubblico-privato. Erano indicati anche azioni e strumenti specifici della Commissione e degli Stati per assicurare la fiducia nei sistemi di IA, a partire dal fare del settore pubblico un apripista nell'uso di IA in vari settori di politiche – come ambiente, salute, sviluppo industriale, migrazione, contrasto alla criminalità, mobilità, agricoltura – e nelle attività delle amministrazioni pubbliche⁵⁶. Nel 2024, la Commissione ha poi prodotto una comunicazione sulla "promozione delle start-up e dell'innovazione nell'IA affidabile" per definire un quadro strategico per gli investimenti e azioni specifiche "affinché l'Unione possa sfruttare le sue risorse, in particolare la sua infrastruttura di supercalcolo leader a livello mondiale, e promuovere un ecosistema europeo innovativo di IA in cui start-up e innovatori possano collaborare strettamente con gli utenti industriali, attrarre investimenti nell'Unione e avere accesso agli ingredienti fondamentali dell'IA; dati, calcolo, algoritmi e talenti"⁵⁷.

Eppure, le opinioni sull'efficacia di questa strategia non sono state lusin-

⁵⁵ Comunicazione al Parlamento, al Consiglio, al Consiglio, al Comitato economico e sociale e al Comitato delle regioni, 21.04.2021 (<https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CELEX:52021DC0205>).

⁵⁶ AI Excellence: Build strategic leadership in high-impact sectors (<https://digital-strategy.ec.europa.eu/en/policies/build-leadership-ai>).

⁵⁷ Comunicazione al Parlamento, al Consiglio, al Consiglio, al Comitato economico e sociale e al Comitato delle regioni, 24.1.2024 (<https://digital-strategy.ec.europa.eu/it/library/communication-boosting-startups-and-innovation-trustworthy-artificial-intelligence>).

ghiere, soprattutto nell’ambiente economico, che più dovrebbe beneficiarne e che aveva accolto l’avvio della politica europea dell’IA senza opporre preferenze per la deregolazione (Justo-Hanani, 2022). Al termine di una dettagliata ricostruzione delle vicende globali delle Big Tech e delle loro implicazioni tecnologiche, economiche e politiche negli Usa, Asia orientale e Paesi arabi, A. Aresu (2024, p. 439) si chiede: “e allora, qual è il ruolo dell’Europa?”. Per concludere che “i capitali europei, privi di ambizione e di scala, hanno già perso l’occasione”. E per due motivi. In primo luogo, l’idea stessa di autonomia tecnologica europea è un concetto “reso imbarazzante dall’enorme divario finanziario, militare e tecnologico fra le due sponde dell’Atlantico”. In altri termini l’Europa non è competitiva. In secondo luogo, lo svantaggio europeo e il rischio di restare esclusi dalla competizione globale e ridurre la crescita economica sono riconducibili al peso preponderante della regolazione nella politica dell’IA: “la tesi, che gli stati europei e UE sembrano avere sposato, della ‘potenza della regolazione’ accelera inevitabilmente e costantemente il proprio declino, la propria impotenza” (Aresu 2024, p. 446).

Come sottolineato nel Rapporto di M. Draghi alla Commissione UE *The future of European Competitiveness* (2024)⁵⁸, il 70% dei *foundation model* è stato sviluppato negli Usa, solo tre “hyperscaler” statunitensi rappresentano oltre il 65% del mercato cloud globale ed europeo, il 61% dei finanziamenti globali per le start-up di IA va alle aziende statunitensi, il 17% a quelle cinesi e solo il 6% a quelle dell’UE. Al 2024, le aziende europee di AI avevano raccolto 11 miliardi di dollari rispetto ai 47 di quelle statunitensi⁵⁹. Per sbloccare il potenziale innovativo e di produttività europeo è necessario favorire l’integrazione dell’IA nell’industria europea. Un processo che dipende da vari fattori: la collaborazione tra operatori industriali e ricercatori; la sperimentazione, da incoraggiare anche attraverso l’apertura alle imprese, il coordinamento a livello UE e l’armonizzazione di “regimi di *sandbox* per l’IA” nazionali. Queste *sandbox* sperimentali consentirebbero anche di valutare gli ostacoli normativi derivanti dalla legislazione dell’UE o nazionale e fornirebbero un feedback alle autorità di regolamentazione da parte delle imprese private e dei centri di ricerca. Infatti, nel Rapporto la regolazione delle tecnologie e dell’IA in particolare viene considerata come uno degli ostacoli allo sviluppo tecnologico ed economico: “l’UE dispone attualmente di circa 100 leggi incentrate sulla tecnologia e oltre 270 autorità di regolamentazione attive nelle reti digitali in tutti gli Stati membri. Molte leggi dell’UE adottano un approccio precauzionale,

⁵⁸ Fonte: https://commission.europa.eu/document/download/97e481fd-2dc3-412d-be4cf152a8232961_en.

⁵⁹ Future of life Institute, The EU AI Act Newsletter #66: Huge Edition, 2 dicembre 2024 (<https://artificialintelligenceact.substack.com/p/the-eu-ai-act-newsletter-66-huge>).

imponendo ex ante pratiche commerciali specifiche per evitare potenziali rischi ex post. Ad esempio, l'AI Act impone requisiti normativi aggiuntivi ai modelli di IA per uso generico che superano una soglia predefinita di potenza di calcolo, soglia che alcuni modelli all'avanguardia già superano”.

Queste valutazioni sono allineate con quelle dell'ambiente economico europeo dell'IA, a partire dalle startup del settore⁶⁰, preoccupate che regolamentare troppo, e più degli Usa, possa soffocare l'innovazione⁶¹. A sua volta, la nuova amministrazione statunitense non è preoccupata per altri possibili effetti delle regolamentazioni europee, ma per le opportunità di business delle Big Tech⁶² in Europa.

Il Rapporto di M. Draghi ha largamente ispirato la revisione di quella che si può definire la “*politics of UE policies*” e il passaggio, non solo per quanto riguarda l'IA, dal primato della sostenibilità a quello della competitività, formalizzato successivamente in indirizzo politico dalla “Bussola”, richiamata sotto.

Per questo, in occasione del Summit internazionale del febbraio del 2025, la Francia ha provato a scartare di lato dal percorso sin lì seguito dall'UE, indicando i cambiamenti da introdurre. Il Presidente francese E. Macron, preoccupato anche per l'attrattività della Francia per investimenti e localizzazioni di imprese che usano IA (Justo-Hanani, 2022), ha prima annunciato investimenti francesi in IA per 109 miliardi, e poi sostenuto che “semplificheremo (...). Su scala nazionale ed europea, è molto chiaro che dobbiamo risincronizzarci con il resto del mondo”⁶³. Ha un rilievo simbolico il fatto che nella stessa sede la Presidente della Commissione abbia di fatto annunciato una svolta nella *policy* europea dell'IA, dal primato delle regole *social oriented* introdotte con l'EU AI Act appena un anno prima a un peso maggiore della strategia *market oriented*. Un *policy change* che si sostanzierebbe nella semplificazione della regolazione, e nella valorizzazione degli investimenti per infrastrutture pubbliche di elaborazione di Big Data, con l'obiettivo di mettere insieme ricercatori, imprenditori e *policy maker* nella ricerca di un brand distintivo e di un'industria europei dell'UE, deboli di fronte alle Big Tech Usa e alle tecnologie cinesi.

Di conseguenza, all'inizio del 2025 è stato lanciato lo *AI Continent Action*

⁶⁰ Fonte: <https://www.politico.eu/article/eu-tech-scene-denounces-ai-data-rules-bad-for-growth-meta-google-survey-gdpr/>

⁶¹ Fonte:

https://www.repubblica.it/economia/2025/02/10/news/trombetti_translated_intelligenza_artificiale_europa-423993249/

⁶² Fonte

https://www.repubblica.it/economia/2025/02/12/news/vertice_parigi_ia_von_der_leyen_van_ce-423997043/?ref=RHLF-BG-P8-S1-T1&ref=nlrep-a-bgr

⁶³ <https://www.nytimes.com/2025/02/10/business/ai-summit-paris.html>

*Plan*⁶⁴ con il quale l'UA mira adesso a divenire “leader mondiale nell'IA, un continente leader”⁶⁵. Il piano comprende varie azioni, basate su 5 pilastri per garantire un'autonomia tecnologia all'UE: infrastrutture, dati, applicazioni, semplificazione della regolazione digitale e competenze. Sono previste 20 “gigafabbriche”⁶⁶ dotate di oltre 100.000 *chip* di IA di ultima generazione, per produrre modelli di IA di grandi dimensioni che consentano di creare ecosistemi di IA chiave per l'UE, rendendo l'IA accessibile alle start-up, ricercatori e industria per addestrare, sviluppare e migliorare i loro modelli, rafforzando così la competitività dell'Europa nell'uso dell'IA. Il programma, da realizzare attraverso un fondo pubblico-privato di 200 miliardi, è attivato attraverso la piattaforma *InvestAI*⁶⁷, per un ammontare di investimenti più ampio. Altri obiettivi sono migliorare l'accesso ai dati, accelerare l'adozione dell'IA nei settori strategici dell'UE, rafforzare le competenze e i talenti in materia di IA. Identifichiamo altre azioni complementari per la costruzione di un'IA europea complementari al Piano. Primo, la costituzione di un “CERN per l'IA” sulla cui necessità, da una prospettiva meno *business friendly*, concordano anche reti di attori scientifici⁶⁸, ritenendo che un centro di ricerca pubblico – e non delle grandi aziende private – possa consentire ai Paesi europei di padroneggiare l'innovazione tecnologica, anziché subirne le logiche di profitto e di mercato (Puliafito, 2025d). Secondo, l'obiettivo di promuovere *compliance* e la semplificazione delle norme. Secondo alcune prime interpretazioni, il lavoro della Commissione e degli stati membri per rendere applicabile l'AI Act sembra essere già indirizzato ad ammorbidente alcuni contenuti delle norme, “come se i legislatori fossero preoccupati di aver esagerato” (Benaich e Chalmers, 2024) e si rendessero conto che per l'IA è necessario mettere in pausa la regolamentazione e virare verso un quadro normativo “più *innovation-friendly*”⁶⁹. Terzo, l'UE cerca anche di tutelare la concorrenza, per garantire che i mercati dell'IA e quelli correlati (cloud e GPU) rimangano contendibili. Lo *High-Level Group for the Digital Markets Act*, ha invitato (maggio 2024) ad applicare

⁶⁴ Comunicazione al Parlamento, al Consiglio, al Consiglio, al Comitato economico e sociale e al Comitato delle regioni. *AI Continent Action Plan* (<https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan>)

⁶⁵ Fonte: https://ec.europa.eu/commission/presscorner/detail/en/ip_25_1013

⁶⁶ Fonte: https://ec.europa.eu/commission/presscorner/detail/it/ip_24_6302

⁶⁷ Fonti: https://ec.europa.eu/commission/presscorner/detail/en/ip_25_467; https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_25_471

⁶⁸ La Confederazione dei Laboratori per la Ricerca sull'IA in Europa (CAIRNE) è un'associazione internazionale) creata dalla comunità europea dell'IA che cerca di rafforzare l'eccellenza europea nella ricerca e nell'innovazione sull'IA, con una forte attenzione all'IA incentrata sull'uomo (<https://cairne.eu/cern-for-ai/>)

⁶⁹ Fonte: <https://www.bruegel.org/first-glance/tech-regulation-european-union-should-be-bolder>

questa normativa anche alle attività nel mercato IA, come le piattaforme digitali *gatekeeper*⁷⁰. La Commissione proporrà anche una strategia per migliorare e facilitare la condivisione sicura dei dati, pubblici e privati. Sono previste quindi leggi europee sull’innovazione, sul *cloud*, sui materiali avanzati e sulle reti digitali. In particolare, la strategia *Applicare l’Intelligenza Artificiale* avrà l’obiettivo di promuovere nuovi usi industriali dell’IA in settori quali l’*automotive*, l’energia, la robotica, la farmaceutica e l’aeronautica, i servizi finanziari, nonché di migliorare i servizi pubblici, ad esempio nella sanità e nella giustizia.

Arrivati a questo punto proviamo a individuare i fattori che influenzano le agende e strategie di governo dell’IA dell’UE. Le preferenze mutevoli per i diversi tipi di strategia dipendono sia dagli attori coinvolti nell’arena politica europea dell’IA e dalle loro interazioni, sia da un contesto istituzionale e politico che mostra elementi di continuità e discontinuità.

I primi fattori riguardano gli attori coinvolti, le loro preferenze, culture e interessi. La definizione e il perseguitamento delle strategie per l’IA sono il risultato delle interazioni fra commissari e parlamentari europei, ministri e parlamentari statali, imprese, organizzazioni non governative no-profit (ONG) ed esperti⁷¹, ciascuno dei quali ha influenzato le decisioni, in grado e in modi diversi. Convergenze e tensioni fra gli attori politici su priorità (regolazione vs innovazione) e misure specifiche dipendono da un lato da preferenze differenziate – per gruppi politici nel Parlamento e secondo gli orientamenti dei governi statali nel Consiglio – dall’altro dalle modalità che i contesti istituzionali offrono per la loro composizione.

La priorità per una strategia regolativa nel periodo 2018-2024 ha fatto parlare di una politica europea che vuole assumere la guida del sistema sociotecnico dell’IA (Guenduez e Mettler, 2023). Una propensione che si spiegherebbe con un tratto distintivo dell’Europa nell’ambito del già visto WEIRD *framework* (Denford *et al.*, 2023): una cultura politica meno *business-friendly* e più propensa a regolare i mercati, compresi quelli tecnologici. Questa è dovuta, a sua volta, alla prevalenza di lunga data nella cultura politica europea, prima dell’interventismo in economia delle socialdemocrazie di varie latitudini, poi di un neoliberismo “roll-out”, o della “terza via”, orientato allo “stato regolatore” (Majone, 1997). Lo dimostrano precedenti misure regolative dell’economia, ad esempio sugli OGM, la “mucca pazza”, o la privacy. La maggioranza emersa dalle elezioni europee per la legislatura 2019-2024 ha fatto sì che

⁷⁰ Fonte: https://digital-markets-act.ec.europa.eu/high-level-group-digital-markets-act-public-statement-artificial-intelligence-2024-05-22_en

⁷¹ Dopo lo *High Level Group on AI* del 2018, nel marzo 2025 la Commissione ha istituito un gruppo di esperti scientifici indipendenti (<https://digital-strategy.ec.europa.eu/en/library/commission-adopts-implementing-act-establish-scientific-panel-under-ai-act>)

questo orientamento prevalesse in tutta la prima fase dell'agenda dell'IA (2018-2024) nella Commissione e nel Parlamento europeo.

La priorità per la regolazione nella politica dell'IA è stata favorita anche dalla presenza – prima in Europa occidentale, poi anche nei Paesi dell'ex blocco sovietico – di una rete transnazionale di organizzazioni non governative della società civile no-profit (ONG) e dei cosiddetti *watchdog groups*, che fanno *advocacy* in favore di regolazioni politiche a tutela di diritti umani, sociali e civili, ad esempio su biofuels, cibo, o ambiente. Una presenza più forte ed efficace che in Cina, dove lo sviluppo delle ONG e la presenza di quelle internazionali sono limitati da restrizioni politiche e legali (Li, 2011), o negli Usa, dove il settore filantropico è ben finanziato e legittimato, ma meno integrato nel *policy making*. Varie ONG hanno fatto *advocacy* per un'efficace regolazione dell'IA durante tutto il processo dell'AI Act. Nella fase di formulazione, hanno espresso circa il 10% dei contributi nella consultazione sul Libro Bianco, al pari degli attori della ricerca, ma al di sotto del 31% di quelli economici. Durante il processo decisionale sono intervenute, ad esempio, contro la deregolamentazione nella classificazione dell'IA ad alto rischio, richiesta invece nel 2023 dalle lobby economiche⁷². Più generalmente, hanno contrastato disposizioni che potessero indebolire la capacità dell'AI Act di proteggere i diritti e garantire la sicurezza, chiedendo anche di eliminare alcune scappatoie legali per fornitori e utilizzatori. Nella fase iniziale dell'attuazione dell'AI Act, varie ONG, come il *Center for Democracy and Technology*⁷³, hanno ritenuto che il codice di pratiche volontario relativo ai modelli *general purpose*, pur redatto in un'arena gestita da esperti e cui esse sono ammesse insieme a stakeholder privati⁷⁴, sia stato anacquato per soddisfare l'industria tecnologica. Ad esempio, i rischi di discriminazione e per i diritti umani sono stati eliminati dalla tassonomia principale⁷⁵, lasciando solo i rischi informatici, chimici, biologici, radiologici e

⁷² Più di 115 organizzazioni della società civile esortano gli eurodeputati a opporsi alle lobby tecnologiche e industriali e annullare le modifiche all'articolo 6 della legge sull'IA, ripristinando ripristinare il testo originale della Commissione (7.09.2023, <https://edri.org/our-work/civil-society-statement-eu-close-loophole-article-6-ai-act-tech-lobby>). Del novembre 2023 è la dichiarazione del Ad Lovelace Institute contro le esenzioni da obblighi normativi dei *foundation model*.

⁷³ Fonte: <https://cdt.org/insights/cdt-europe-statement-on-the-third-general-purpose-ai-code-of-practice-draft/>.

⁷⁴ Il Codice di pratiche per l'IA *general purpose* viene redatto con il coinvolgimento di diversi soggetti interessati: fornitori di modelli di IA generici, fornitori a valle, organizzazioni di settore, società civile, titolari di diritti, mondo accademico ed esperti indipendenti (<https://digital-strategy.ec.europa.eu/en/policies/ai-code-practice>).

⁷⁵ Fonte: <https://www.techpolicy.press/human-rights-are-universal-not-optional-dont-undermine-the-eu-ai-act-with-a-faulty-code-of-practice/>

nucleari, di manipolazione e di “perdita di controllo”⁷⁶. Il *Corporate Europe Observatory* (2025) sostiene che la definizione di questi standard è dominata da esponenti dell’industria tecnologica come Oracle, Microsoft, Amazon, Huawei, IBM e Google⁷⁷ e che le organizzazioni private stiano definendo a porte chiuse regole legali con vincoli ridotti⁷⁸.

Come avviene per altri settori del *policy making* europeo la partecipazione alle consultazioni⁷⁹ e le iniziative di *advocacy* e *lobbying* sono state le modalità con le quali anche le imprese hanno cercato di influenzare le decisioni. Speculare al ruolo delle ONG, infatti, è stata la pressione degli interessi economici. Infatti, sin dalle consultazioni iniziali, durante l’esame del Parlamento e del Consiglio, fino all’attuazione dell’AI Act, le imprese hanno chiesto una regolamentazione più *soft* e codici di condotta non vincolanti. A questo fine, da un lato hanno contribuito a costruire un *frame* in cui le normative stringenti sono controproducenti per l’innovazione e la competitività europea nell’arena globale dell’IA. Dall’altro lato hanno condotto azioni di *lobbying* verso Commissione, Parlamento e governi dei Paesi membri⁸⁰ per prevenire o ridurre specifici vincoli nel mercato dell’IA, riducendo i requisiti per i sistemi ad alto rischio ed escludendo dalla normativa il concetto stresso di *general purpose*⁸¹. Un’azione particolarmente mirata verso i governi, cosicché l’arena intergovernativa del Consiglio si è rivelata più permeabile di Commissione e Parlamento nei confronti di queste pressioni. La preferenza dei governi francese tedesco e italiano durante il Trilogo (cfr. sopra) per una mera “autoregolamentazione obbligatoria attraverso codici di condotta” per l’IA *general purpose*⁸² è stata infatti ricondotta al rifiuto di ostacolare le rispettive start-up tecnologiche nazionali, a partire dalla francese Mistral e dalla tedesca Aleph Alpha. L’influenza delle imprese è stata più forte di quella delle ONG e questo accomuna la politica dell’IA a quella di altri settori del *policy making* dell’UE.

Non solo le imprese europee, ma anche le Big Tech statunitensi – com-

⁷⁶ Fonte: <https://artificialintelligenceact.substack.com/p/the-eu-ai-act-newsletter-71-wealth>

⁷⁷ Fonte: <https://corporateeurope.org/en/2025/01/bias-baked>

⁷⁸ The EU AI Act Newsletter #69: Big Tech and AI Standards, 20 01 2025, <https://artificialintelligenceact.substack.com/p/the-eu-ai-act-newsletter-69-big-tech>

⁷⁹ Una consultazione è stata realizzata anche nella fase di implementazione sulle linee guida per la definizione di “sistema di IA” e sulle pratiche di IA che pongono rischi inaccettabili (novembre 2024), alla ricerca di esempi pratici da fornitori di sistemi di IA, imprese, autorità nazionali, università, istituti di ricerca e società civile (<https://digital-strategy.ec.europa.eu/en/news/commission-launches-consultation-ai-act-prohibitions-and-ai-system-definition>)

⁸⁰ Fonte: <https://artificialintelligenceact.substack.com/p/the-eu-ai-act-newsletter-41-big-tech>

⁸¹ Fonte: <https://dataethics.eu/corporate-europe-us-big-tech-lobbied-against-generative-ai-as-high-risk-in-the-ai-act/>

⁸² Fonte: <https://dig.watch/updates/germany-france-and-italy-reach-consensus-on-ai-act-with-mandatory-self-regulation>

prese OpenAI⁸³ e Meta⁸⁴ – dichiarano difficoltà a investire e a commercializzare i loro modelli di IA a causa della combinazione tra l’AI Act, i requisiti del GDPR in materia di privacy e trasferimento dei dati e il Digital Markets Act (Benaich e Chalmers, 2024). Un esempio: la notizia che *Apple Intelligence* non verrà rilasciato nell’UE per i dubbi sulla sua compatibilità con la normativa vigente (Malaschini, 2024b). Queste preoccupazioni riguardano anche il codice di pratiche volontario su divulgazione dei dati e formazione e gestione dei rischi sistematici derivanti dai modelli *general purpose*, copyright e test di modelli di terze parti proposto dagli esperti nominati dalla Commissione (2025). La *European Tech Alliance*, che associa 33 importanti aziende tecnologiche, chiede in tal senso che i codici lascino “spazio sufficiente per sviluppare e utilizzare modelli e sistemi di IA innovativi per competere con gli operatori mondiali (...): le iniziative dovrebbero essere basate sui rischi e sui problemi, adattarsi ai modelli di business e proporre soluzioni proporzionate, praticabili e pragmatiche” che funzionino per il business⁸⁵.

Infine, le imprese hanno rivolto la loro azione di influenza verso leva degli investimenti della strategia *market oriented* per l’IA. Dopo l’approvazione dell’AI Act, le imprese europee hanno congiunto le forze con la Commissione per creare “campioni europei” pubblici e pubblico-privati dell’IA (*EU AI Champions Initiative*), con la partecipazione di grandi gruppi industriali come Airbus, L’Oréal, Mercedes, Siemens e tecnologici, tra cui Spotify e Mistral AI (Vannuccini, 2025)⁸⁶. Queste azioni si comprendono meglio tenendo conto delle caratteristiche dell’economia dell’IA in Europa, in cui sono assenti grandi imprese del calibro di quelle americane e cinesi, ma dove piuttosto è presente un sistema di startup ma con una scala insufficiente per divenire globalmente competitive⁸⁷.

La politica europea dell’IA è però influenzata anche da fattori di contesto, relativi al rapporto tra assetti istituzionali e scelte politiche. Il contesto istituzionale si è manifestato in primo luogo attraverso la mutuazione di soluzioni dai sistemi normativi europei contigui (Justo-Hanani, 2022). Il sistema normativo esistente ha plasmato anche la politica europea per l’IA, soprattutto sul versante della regolazione. Elementi di tale dipendenza dal contesto

⁸³ Fonte: <https://time.com/6288245/openai-eu-lobbying-ai-act/>.

⁸⁴ Fonte: <https://www.politico.eu/article/zuckerberg-urges-trump-to-stop-eu-from-screwing-with-fining-us-tech-companies>

⁸⁵ Fonte: <https://eutechalliance.eu/ensuring-europes-competitiveness-in-ai-key-principles-and-actions-for-the-gpai-code-of-practice/>

⁸⁶ Fonte: <https://www.rivista.ai/2025/02/11/eu-ai-champions-initiative-le-grandi-aziende-europee-chiedono-regole-più-snelle-per-l'intelligenza-artistificiale/>

⁸⁷ Comunicazione della Commissione 2024, citata (<https://digital-strategy.ec.europa.eu/it/library/communication-boosting-startups-and-innovation-trustworthy-artificial-intelligence>).

regolativo preesistente sono rinvenibili nel fatto che alcune norme già vigenti, come GDPR, DSA e *Data Act*, si applicano anche all'uso di sistemi di IA. Non a caso, le pressioni provenienti dall'ambiente economico e le preoccupazioni espresse anche nel Rapporto Draghi si riferiscono non solo all'EU AI Act, ma anche a questo insieme di norme ad esso contigue.

Si può notare che questo insieme di norme è stato oggetto di critiche da prospettive e posizioni politiche opposte. Da un lato è interpretato come troppo vincolante rispetto alle necessità dell'innovazione e della competitività; dall'altro è ritenuto scarsamente vincolante. Ad esempio, il GDPR, pur accusato di “essere responsabile di una involuzione dei mercati, è obiettivamente diventato un modello normativo internazionale, un enorme volano di profitti e competitività grazie all’abilitazione per legge della libera circolazione dei dati (...)"⁸⁸. Anche il DSA, che regola piattaforme e motori di ricerca, da un lato è stato criticato dalle imprese digitali per l'eccessivo onere regolativo. Da una prospettiva opposta, è stato invece criticato per essere un approccio normativo classicamente neoliberista, che non vincola in modo significativo il potere delle imprese e, creando l'apparenza di responsabilità, legittima il “capitalismo delle piattaforme” monopolistico (Griffin, 2023).

Inserendo la regolazione dell'IA in questo apparato normativo preesistente, l'UE ha anche cercato di ricomporre due finalità della strategia *social-oriented*: da un lato, “realizzare quello che viene definito il *Brussels effect*: indurre le imprese informatiche ad adeguarsi agli standards europei se vogliono accedere a un mercato di 450 milioni di persone” (Malaschini, 2024b) ed esercitare un “potere soft” di influenza globale. Ad esempio, il GDPR ha norme adottate in California, Quebec e Brasile (Beroche *et al.*, 2022, p. 43); dall'altro, l'ambizione alla leadership nei processi di definizione di standard globali dell'IA, esportando i suoi orientamenti nelle arene internazionali e riconfermando le capacità di influenza dimostrate con il GDPR (Justo-Hanani, 2022). Proprio dall'esterno, però, provengono le tensioni politiche che hanno contribuito alla svolta del 2025, aggiungendosi a quelle specificamente europee. La seconda Commissione presieduta da U. von der Leyen ha adottato la “Bussola europea per la competitività”⁸⁹, un programma strategico che – traducendo in indirizzi politici le indicazioni del Rapporto Draghi – propone, tra l'altro, di creare un ecosistema più favorevole alla competitività e all'innovazione “nei settori tecnologici che saranno importanti nell'economia di domani, come l'IA (...). Un orientamento che può essere

⁸⁸ Fonte: <https://www.ilsole24ore.com/art/intelligenza-artificiale-problema-dell-europa-non-sono-regole-AGju37qC>

⁸⁹ Comunicazione della Commissione a Consiglio e Parlamento *Competitiveness Compass*, pubblicata il 29 gennaio 2025 (https://commission.europa.eu/document/download/10017eb1-4722-4333-add2-e0ed18105a34_en).

ricondotto al rafforzamento, nella composizione del Parlamento europeo e nei sistemi politici di alcuni stati membri (Francia, Germania, Italia, Austria, Belgio, Ungheria), delle posizioni di destra, la cui ostilità verso misure di regolazione dei mercati si estende dalle politiche ambientali a quelle tecnologiche. Infine, non si può non considerare l'interdipendenza europea con le azioni di contesti politici “altri”, riferibili alla competizione tra i “giganti” nazionali dell'IA ora esacerbata dalla nuova aggressività del governo degli Usa, che, attraverso la coalizione fra politica (presidenza Trump) e Big Tech, sfida a viso aperto le scelte regolative e di sviluppo dell'UE.

3.3. Dall'intelligenza artificiale europea alla “via italiana”

Questo paragrafo prosegue nell'analisi della multi-scalarità del governo dell'IA considerando il ruolo della politica alla scala nazionale, stavolta non in riferimento alle azioni di un “gigante”, ma di un contesto più periferico nello scenario geopolitico, l'Italia, inserito nel più ampio sistema delle politiche dell'UE. In particolare, sono esplorati i due ambiti di azione della politica dell'IA relativi allo sviluppo e alla regolazione dei sistemi di IA nelle politiche pubbliche e nella pubblica amministrazione (PA). Ci occupiamo dunque del ruolo che la politica svolge nel governo dell'IA all'interno dei confini nazionali e nella relazione con altre scale del potere politico che detengono “sovranità” rispetto al tema, nel caso italiano con l'UE e con i governi locali. L'analisi considera l'intervallo temporale 2019-2025 (marzo), il periodo intercorso tra l'approvazione della prima strategia nazionale sull'IA e l'iniziativa legislativa del Governo Meloni.

La prima questione di cui ci occupiamo riguarda il percorso di elaborazione di una strategia nazionale dell'IA. Nel paragrafo precedente abbiamo identificato nel 2018 l'inizio delle politiche di sviluppo e regolazione dell'IA dell'UE, con la comunicazione della Commissione del Piano Coordinato sull'Intelligenza Artificiale. Il Piano proponeva anche un quadro di riferimento per gli Stati membri, invitati a dotarsi di una propria strategia di sviluppo nazionale entro l'anno successivo. Tale documento avrebbe dovuto indicare i livelli di investimento pubblico, rivolti al settore privato e alla ricerca, e le relative misure di attuazione (Commissione Europea, 2018). Al momento dell'emanazione del Piano solo cinque Paesi EU, sui 28 di allora, avevano adottato una strategia di investimenti sull'IA, un gruppo ristretto di cui l'Italia non faceva ancora parte. Dunque, il processo di elaborazione di una strategia italiana ha seguito i tempi dettati dall'UE configurandosi, almeno inizialmente, come un'iniziativa di attuazione dell'agenda istituzionale europea.

Fino ad allora le questioni dell'IA erano rimaste sostanzialmente assenti nelle agende politica ed istituzionale italiane. Ricercando tra le banche dati di Camera e Senato relative alla XVII legislatura (2013-2018), anni in cui si sono succeduti tre governi di centrosinistra Letta, Renzi, Gentiloni, è possibile rilevare la marginalità nel dibattito parlamentare delle questioni connesse all'IA⁹⁰. Negli anni antecedenti al Piano della Commissione, l'ingresso nell'agenda politica italiana delle questioni dell'IA è avvenuto perlopiù in riferimento al rapporto tra IA, cambiamenti nei sistemi produttivi e trasformazioni nel mercato del lavoro. Il discorso politico sull'IA era inquadrato dalla maggioranza di governo di centro-sinistra e dalle forze di opposizione nell'ambito delle riflessioni sulla cosiddetta Quarta rivoluzione industriale, ovvero il massiccio impiego di tecnologie digitali avanzate nei sistemi produttivi. In tal senso, le prime mozioni parlamentari sull'IA della maggioranza muovevano dalla volontà di integrare il Programma Industria 4.0⁹¹ del governo Renzi (2014-2016) attraverso interventi di sostegno alla ricerca e allo sviluppo della robotica e delle applicazioni di IA. Tra il 2016 e il 2017 sono stati presentati al Parlamento i primi disegni di legge relativi a iniziative su interventi sul mercato del lavoro, che da lì a breve sarebbe stato interessato da profondi mutamenti associati allo sviluppo delle nuove tecnologie digitali e dell'IA. Erano delle proposte tangenti l'IA, prive di interventi nel merito ma ispirate dai discorsi sulle potenzialità e i rischi dei sistemi di decisione autonoma. Sempre al 2017 risalgono le prime riflessioni parlamentari sul ruolo regolativo della politica con alcune mozioni di opposizione (es. Rampaelli – Fratelli d'Italia; Civati – Sinistra Italiana) che impegnavano il Governo a promuovere interventi normativi omogenei sull'uso dell'IA. Anche in questo caso, la questione centrale verteva sull'equilibrio tra le opportunità di crescita economica e la mitigazione delle possibili ripercussioni dell'IA sulla perdita di posti di lavoro e la necessità di predisporre interventi per la ricollocazione dei lavoratori.

È nel 2019 che si è assistito a un cambio di passo nella politica nazionale sull'IA con il Governo Conte I (maggioranza M5S-Lega). In quell'anno il Ministero per lo Sviluppo Economico (MISE) ha convocato il primo panel di esperti di IA che, in linea con le indicazioni dell'UE, ha elaborato le "Proposte per una strategia italiana per l'intelligenza artificiale". Secondo gli esperti, l'ecosistema dell'IA italiana poggiava sull'interdipendenza di tre

⁹⁰ Si pensi che gli atti di sindacato ispettivo relativi all'IA sono stati 22 nell'intera XVII legislatura, rispetto ai 127 della XVIII e ai 200 dell'attuale legislatura, ancora in corso.

⁹¹ Il "Piano Nazionale Industria 4.0" fu pubblicato nel settembre 2016 dall'allora Ministro per lo sviluppo economico (MISE) C. Calenda, un programma contenente misure fiscali e incentivi rivolti alle imprese italiane per favorirne la competitività nella "Quarta rivoluzione industriale".

elementi promettenti: un buon sistema di ricerca pubblico/privato e di trasferimento tecnologico; una posizione di leadership mondiale nell’automazione industriale; una ampia possibilità di adozione nella produzione industriale come seconda manifattura europea oltre che nella PA (MISE, 2019). Tuttavia, l’anno successivo S. da Empoli⁹², uno degli esperti convocati dal MISE, evidenziava al Senato gli elementi di debolezza del contesto italiano riassumibili nel basso livello di investimenti privati, nel bassissimo numero di startup e nel gap con gli altri stati europei nel grado di preparazione all’IA calcolato attraverso l’indice I-Com⁹³. Tali condizioni collocavano piuttosto l’Italia nella periferia della geografia dell’IA, a livello sia globale, sia dell’UE, una debolezza a cui il Governo Conte I ha provato a ovviare attraverso la ricerca di un posizionamento europeo della propria strategia. La necessità di tale collocazione era già evidente nelle raccomandazioni degli esperti che esortavano il Governo a rafforzare la presenza dell’Italia: «sui principali tavoli europei in tema di IA e tecnologie digitali, dove non sempre le nostre istituzioni sono presenti, e non sempre riescono a far sentire pienamente la loro voce e a rappresentare le necessità del sistema Paese» (MISE, 2019: 79). Gli esperti da un lato suggerivano alla politica nazionale un allineamento alla cornice internazionale dell’etica dell’IA; dall’altro, di dare alla futura strategia un connotato italiano attraverso un approccio fortemente antropocentrico – l’IA complementare e non sostitutiva dell’uomo – e finalizzato allo sviluppo sostenibile.

La prima “Strategia Nazionale per l’IA” è stata quindi approvata a un anno di distanza dalle proposte degli esperti, dopo una consultazione pubblica e una valutazione degli altri ministeri interessati. Il lungo documento, dei cui contenuti non è possibile offrire una sintesi esaustiva, proponeva una serie di obiettivi e iniziative finalizzati innanzitutto a colmare i gap italiani e a preparare le diverse componenti del suo ecosistema “debole”. Si pensi che il primo pacchetto di iniziative era rivolto alla sensibilizzazione e agevolazione delle piccole e medie imprese italiane all’uso e sviluppo di IA, un contesto imprenditoriale ben diverso da quello degli stati “giganti” della globalizzazione dell’IA. Altre iniziative insistevano su questioni e problemi strutturali della PA e del sistema della formazione secondaria e universitaria, come la frammentazione nei patrimoni di dati e il disallineamento tra competenze ed esigenze del mercato del lavoro. Ciò che è apparso da subito come un limite della prima Strategia era la mancanza di un impegno reale sugli investimenti pubblici per finanziare le iniziative previste. A fronte dell’obiettivo di 2.5 mld di investimenti

⁹² Presidente dell’Istituto per la Competitività (I-Com).

⁹³ L’indice si riferisce agli Stati nazionali e si basa su indicatori relativi a: l’adozione di tecnologie digitali da parte delle imprese; le competenze digitali dei lavoratori e dei laureati; i sistemi di sicurezza digitale; le condizioni delle infrastrutture di rete.

per il quinquennio 2021-2025 dedicati all'IA, la Strategia faceva riferimento a fondi con dotazioni limitate o multi-settoriali: il "Fondo per lo sviluppo delle tecnologie e delle applicazioni di IA, blockchain e internet of things", istituito dalla legge di bilancio del 2019, con una dotazione prevista di soli 45 mln di euro (sospeso nel 2022): il Fondo Nazionale Innovazione, gestito oggi da Cassa Depositi e Prestiti (CDP), strumento trasversale a diversi settori di innovazione e non rivolto in modo esclusivo all'IA.

La caduta del Governo Conte bis e la formazione nel febbraio del 2021 del Governo Draghi hanno provocato la fine del percorso tracciato dalla prima Strategia Nazionale. Il primo elemento di discontinuità del nuovo Governo ha riguardato infatti la scelta di spostare la delega allo sviluppo dell'IA dal MISE, a cui era stata attribuita di fatto, al Ministero per l'Innovazione Tecnologica e Digitalizzazione guidato dal ministro "tecnico" V. Colao⁹⁴. L'attribuzione della delega formale a un ministero senza portafoglio ha collocato la questione nell'ambito di competenze della Presidenza del Consiglio dei ministri (PCM) che, nel marzo 2021, ha istituito il Comitato interministeriale per la transizione digitale (CITD), con compiti di coordinamento e monitoraggio delle iniziative su banda larga, fascicolo sanitario elettronico, IA, IoT e blockchain (Art. 8 Legge n. 55, 2021). Nel frattempo, l'attività parlamentare sul tema dell'IA è rimasta ancora a bassa intensità con alcune iniziative conoscitive, come l'audizione del Ministro Colao alla Camera sull'EU AI Act del Marzo 2022, evidenza di un'azione "di rimessa" del Parlamento italiano (Malaschini e Pandolfelli, 2022), che ha lasciato maggiore protagonismo all'esecutivo.

A solo nove mesi dal suo insediamento il Governo Draghi ha elaborato, con il supporto di nove esperti⁹⁵, la seconda Strategia nazionale, stavolta denominata "Programma Strategico per l'IA 2022-2024" (PSIA). Essa condivideva con la precedente le considerazioni sulle caratteristiche dell'ecosistema italiano e sulla necessità dell'ancoraggio europeo. L'IA italiana avrebbe dovuto essere un'IA europea, gli sforzi nazionali sarebbero stati compiuti quindi: «affinché l'Europa possa raggiungere l'autonomia strategica e competere a livello internazionale» (PSIA, 2021: p. 14). Il carattere europeo del PSIA era espresso sia esplicitamente in termini di pertinenza delle sue azioni rispetto agli *input* del Piano Coordinato della Commissione sia, più implicitamente, riguardo alla dipendenza economica dalle risorse di investimento pubblico

⁹⁴ Al momento del suo insediamento Vittorio Colao vantava un lungo percorso professionale nel settore privato delle telecomunicazioni, tra cui spiccava il ruolo di amministratore delegato del gruppo Vodafone PLC nel periodo 2008-2018.

⁹⁵ Le biografie del Gruppo di Lavoro sulla Strategia Nazionale per l'IA sono disponibili al link: <https://innovazione.gov.it/notizie/articoli/nasce-il-gruppo-di-lavoro-sulla-strategia-nazionale-per-l-intelligenza-artificial/>

europee. L'elemento di vantaggio della seconda Strategia è stato proprio poter fare affidamento, almeno a livello previsionale, sulle risorse straordinarie del PNRR-NGEU per finanziare le 24 politiche previste. Si pensi che delle 16 politiche per cui sono state indicate le possibili linee di finanziamento, 14 si riferivano (e 11 in modo esclusivo) agli investimenti PNRR su Digitalizzazione, Istruzione e Ricerca. È proprio l'interesse sulle questioni della ricerca a segnare la discontinuità più rilevante con la precedente strategia, perseguita attraverso un forte orientamento al potenziamento del sistema della ricerca fondamentale e applicata, sia pubblica che privata, e al rafforzamento del capitale umano. Queste ultime sono state identificate come questioni strategiche per l'ecosistema italiano IA a cui il Programma ha dedicato due delle sue tre aree di intervento (talenti, ricerca, applicazioni). Dunque, il Governo Draghi ha inteso diffondere lo studio dell'IA applicata in tutti i curricula degli Istituti Tecnici Superiori, potenziare il programma nazionale di Dottorato, in continuità il Governo Conte bis, e promuovere iniziative per attrarre e trattenere i ricercatori attraverso il finanziamento di attività di ricerca. Lo scorso anno i dati dell'Osservatorio Artificial Intelligence (2024) del Politecnico di Milano hanno mostrato i successi del PSIA rispetto al raddoppio delle nuove borse di dottorato sull'IA (342 nel 2023 rispetto alle 156 del 2022) e alla creazione di nuove posizioni di ricercatori; mentre hanno rivelato l'inefficacia delle misure per l'attrazione dei talenti. L'investimento più rilevante del Programma Draghi è stata l'istituzione della Fondazione FAIR – *Future Artificial Intelligence* nel 2022, un partenariato esteso pubblico-privato con finanziamento PNRR, che riunisce i principali centri di ricerca pubblici nazionali sull'IA – CNR, Fondazione Bruno Kessler, Istituto Nazionale di Fisica Nucleare e Istituto Italiano di Tecnologia, 12 università pubbliche⁹⁶ e 5 grandi imprese private tra cui Leonardo. I membri di FAIR hanno assunto un ruolo centrale nella conduzione di progetti di ricerca sull'IA⁹⁷ di frontiera, tra cui Minerva, uno dei primi LLM italiani, attraverso il loro ruolo di soggetti attuatori (*hub*) ed esecutori (*spoke*).

Tuttavia, le ambizioni trasformative del Programma del Governo Draghi si sono esaurite rapidamente per una molteplicità di cause. Primo, il PSIA si è configurato come uno strumento complementare all'attuazione del PNRR nell'ambito delle politiche di digitalizzazione e innovazione rivolte a PA e sistema delle imprese, oltre che di riforma del sistema della formazione, piuttosto che rappresentare una strategia complessiva di governo dell'IA (Huy-

⁹⁶ Politecnico di Milano, Politecnico di Torino, Sapienza, Scuola Normale Superiore, Università Campus Biomedico di Roma, Università di Bologna, Università di Pisa, Università di Trento, Università di Bari, Università della Calabria, Università di Catania, Università di Napoli "Federico II".

⁹⁷ La panoramica sulle linee di ricerca è disponibile al link: <https://fondazione-fair.it/spoke/>

skes, 2022). In secondo luogo, esso ha indicato progetti puntuali con orizzonti temporali ravvicinati, così come voleva l’impianto del PNRR, senza prevedere un sistema successivo di valutazione e aggiornamento dello strumento. Terzo, l’anno successivo alla sua approvazione sono intervenuti cambiamenti rilevanti nel sistema politico nazionale e nello stesso scenario tecnologico globale che hanno aperto la strada alla necessità e volontà politica di elaborare una terza strategia nazionale a pochi anni di distanza dalla prima. Il riferimento è all’insediamento nell’ottobre 2022 del Governo di centro-destra della Presidente G. Meloni e al lancio su scala globale di ChatGPT che ha imposto la centralità dell’IA generativa e *general purpose* nelle agende politiche di sviluppo e regolazione.

Il cambiamento nella politica nazionale rappresenta un elemento di interesse per un approccio di sociologia politica che ci consente di comprendere ruolo, strumenti e mutamenti negli orientamenti della politica nazionale nel governo dell’IA. Prima di entrare nella stagione dell’IA “Meloniana” è però necessario fare un passo di lato. Un elemento distintivo del Governo Meloni, sostenuto anche da Lega e Forza Italia, è il suo orientamento sovranista che si applica a numerose *policy issue*, compresa l’IA. Il concetto di sovranismo, da non confondersi con quello di sovranità, indica la volontà politica di difendere le prerogative degli Stati nazionali sulle scelte politiche. Le istanze sovraniste, da tempo considerate costitutive del fenomeno del populismo (Segatori, 2015) e della sua versione identitaria (Morlino e Raniolo, 2022; Raniolo e Zucaro, 2023), emergono a seguito della sottrazione di potere subita dagli stati nazionali ad esempio da parte dell’UE, e poggiano sull’intenzione di recuperare il principio della sovranità politico-democratica contro gli interessi di ristrette élite di potere che mancherebbero, secondo questa visione, di una legittimazione politica (Damiani, 2022). Da alcuni anni questa posizione è divenuta dominante nel contesto europeo, con diverse esperienze di governo e leader politici sovranisti come, ad esempio, V. Orban in Ungheria o M. Salvini in Italia (Guetta, 2019), la crescita nelle elezioni politiche in Francia, Germania, Belgio e Romania e con una consolidata presenza nel Parlamento Europeo⁹⁸. Il manifesto dei Conservatori e Riformisti Europei, gruppo politico di cui fa parte il partito della Presidente Meloni, riassume in più punti la necessità di ridiscutere le competenze dell’UE in favore di assetto confederale, rispettando così la sovranità nazionale: “abbandonando un approccio “uguale per tutti” e agendo in modo più flessibile nel riconoscere le singole circostanze di ciascuna democrazia nazionale”

⁹⁸ Nel Parlamento formatosi a seguito delle elezioni del 2024, le posizioni sovraniste sono condivise da tre importanti gruppi politici: i “Patrioti per l’Europa” – 84 europarlamentari; i “Conservatori e Riformisti Europei” – 78 e il gruppo “Europa delle Nazioni Sovrane” – 25.

(ECR, 2024: p. 14). Una visione che può essere applicata, citando G. Meloni, nella costruzione di una via italiana all'IA, una via italiana allo sviluppo dell'IA, una via italiana al governo dell'IA?⁹⁹ Quali sono gli strumenti della politica nazionale in un contesto come quello italiano? Quali margini esistono per “sovranizzare” il governo dell'IA? Di seguito saranno fornite alcune parziali risposte a questi interrogativi.

I primi atti nel merito del governo sovranista sono stati: confermare la delega all'IA alla PCM, incardinata nel Dipartimento per la Trasformazione Digitale, affidato al sottosegretario A. Butti (FdI), riproporre il CITD e contestualmente abolire il Ministero per l'Innovazione Tecnologica e Digitalizzazione. Già nei primi mesi di lavoro del Dipartimento e del CITD, la scadenza ravvicinata del Programma Draghi-Colao, ha fatto emergere per il Governo la necessità/opportunità di dotarsi di una propria strategia nazionale sull'IA. A questo proposito nel novembre 2023 è stato costituito un nuovo gruppo di esperti a supporto della sua elaborazione. Il nuovo Comitato risulta diverso dal precedente per biografie e affiliazioni dei suoi membri. Il Gruppo di esperti del Governo Draghi si distingueva per la presenza omogenea di personalità accademiche della ricerca pubblica italiana. La composizione del Comitato nominato da A. Butti¹⁰⁰ è sembrata un tentativo di rendere questo strumento più eterogeneo e ibrido, secondo una strategia di individuazione degli esperti ritenuta altrove come elemento di innovazione (Vinci, 2022). Da un lato, la formazione del Comitato si è basata sulla sostituzione integrale dei precedenti esperti del Gruppo a supporto del Programma Draghi-Colao, una prevedibile forma di *spoil system*, dall'altro, sull'inserimento di personalità diversificate per affiliazione e collocazione internazionale¹⁰¹.

Il Comitato ha quindi redatto la terza strategia nazionale, pubblicata nel 2024 in coincidenza con l'approvazione definitiva dell'EU AI Act. Essa si presenta come uno strumento ambizioso con quattro aree strategiche di intervento: ricerca, PA, imprese e formazione. In tal senso, la Strategia assomiglia a quelle licenziate dal Conte I e dal Governo Draghi, con l'indicazione

⁹⁹ Fonte: <https://www.governo.it/it/articolo/lintelligenza-artificiale-litalia-il-videomessaggio-del-presidente-meloni/25193>

¹⁰⁰ Le affiliazioni degli esperti del Comitato sono disponibili al link: <https://innovazione.gov.it/notizie/articoli/intelligenza-artificiale-presentato-il-comitato-di-coordinamento/>.

¹⁰¹ Il riferimento è agli esperti G. Greco – docente e presidente dell'Associazione Italiana per l'IA, S. Micera – dell'Ecole polytechnique fédérale de Lausanne, V. Acquaviva – del Physics Department del Cuny Nyc College of Technology, M. Camisani Calzolari – divulgatore scientifico e anche ospite delle trasmissioni televisive Mediaset, e P. Benanti – teologo e consigliere di Papa Francesco sull'IA. Quest'ultimo poco dopo ha assunto anche la carica di Presidente della Commissione sull'IA per l'Informazione presso il Dipartimento per l'informazione e l'editoria della PCM dopo il disconoscimento da parte di G. Meloni della nomina di G. Amato – ex Presidente del Consiglio e della Corte costituzionale.

di una serie di azioni rivolte all’intero ecosistema IA italiano. In essa sono indicati però i “rischi della strategia” relativi al non fare, all’omogeneizzazione culturale, all’iper-regolazione nazionale, al problema del digital divide, al mercato del lavoro e all’inefficacia della Strategia. Questa concettualizzazione, assente nei precedenti documenti strategici, è un’interessante costruzione discorsiva per la sovranizzazione della politica dell’IA: «piuttosto che elaborare un quadro dei rischi legato a valutazioni di ordine generale, e in questo senso cross-nazionale, risulta [...] essenziale comprendere i rischi dell’attuazione della nostra specifica strategia nazionale, delineata tenendo conto delle specificità del nostro sistema-Paese». Pertanto, essi definiscono un set di problemi dello sviluppo dell’IA che, se non affrontati, lederebbero il “sistema Italia”. La principale preoccupazione è che l’Italia abbia un approccio all’innovazione basato sull’acquisto/importazione di tecnologie. Come sostenuto nel marzo 2024 dal sottosegretario A. Butti: «l’Italia usa e userà molto l’IA e dobbiamo fare in modo che questo non avvenga solo adottando soluzioni realizzate all’estero, che contribuiscono al Pil di altri Paesi e creano a noi problemi di “lock-in”, cioè di dipendenza»¹⁰². Alla dipendenza segue il rischio di un’omogeneizzazione culturale e una deviazione dai valori incardinati nell’ordinamento costituzionale nazionale. Entrambi i rischi fanno particolare riferimento all’avvento dell’IA generativa, che posiziona l’Italia ancora una volta nella periferia della globalizzazione dell’IA. Dunque, l’ambizione sovranista della Strategia si estende allo sviluppo di LLM, un mercato da cui negli ultimi anni imprese e centri di ricerca italiani sono stati sostanzialmente esclusi. In realtà i primi LLM italiani si giovano della valorizzazione di precedenti risorse economiche pubbliche dell’UE e della sinergia con investimenti dal settore privato. Le sperimentazioni più avanzate in tal senso – Modello Italia di iGenius¹⁰³, Minerva dell’Università Sapienza¹⁰⁴ e il LLM in italiano di Istella¹⁰⁵ – si basano sull’infrastruttura di calcolo del supercomputer Leonardo¹⁰⁶, cofinanziato Ministero dell’Università e della Ricerca e dall’UE nell’ambito dell’iniziativa EuroHPC Joint Undertaking¹⁰⁷,

¹⁰² Fonte: <https://innovazione.gov.it/notizie/interviste/butti-piu-ricercatori-e-campioni-nazionali-l-ai-parlera-italiano/>

¹⁰³ Fonte: <https://www.milanofinanza.it/news/l-intelligenza-artificiale-italiana-diventa-realta-l-hanno-creataigenius-e-cinea-il-modello-e-open-source-202406061243577004>

¹⁰⁴ Fonte: <https://www.uniroma1.it/it/notizia/la-italiana-ai-chatbot-del-futuro-presentata-lultima-versione-di-minerva-il-modello>

¹⁰⁵ Fonte: <https://istella.ai/products/large-language-models/>

¹⁰⁶ Maggiori informazioni su Leonardo sono disponibili al link: <https://leonardo-supercomputer.cineca.eu/it/home-it/>

¹⁰⁷ L’iniziativa ha l’obiettivo di rendere l’UE leader mondiale nel supercalcolo attraverso la creazione e messa in rete di infrastrutture tecnologiche europee che comprendono supercomputer, computer quantici e AI Factories.

messi a disposizione dal Cineca. Dunque, sebbene possa sembrare banale, anche i LLM italiani sono indirettamente, intrinsecamente e forse inevitabilmente europei.

In altre parole, sovranizzare l'IA significa anche perseguire un'autonomia nazionale sugli investimenti nel settore. A questo proposito il Governo sembra percorrere due strategie complementari. G. Meloni ha cercato e trovato la sponda di CDP¹⁰⁸. Negli ultimi due decenni i governi italiani hanno progressivamente trasformato CDP in una importante istituzione di mercato con missione pubblica (Bassanini, 2015) per adeguare il ruolo dello Stato nazionale come attore economico ai cambiamenti imposti dalla globalizzazione e ai vincoli di spesa pubblica (Dalla Balla, 2023). Come ricostruito da M. Giachetti Fantini (2018), queste trasformazioni hanno rilanciato il ruolo dello stato "investitore" nella politica industriale italiana¹⁰⁹, un ruolo evidente anche nell'attuazione degli investimenti del PNRR (Guarna Assanti e Di Martino, 2022). Il Governo ha scelto di riproporre il ruolo di CDP creando il primo fondo di investimenti dedicato in via esclusiva all'IA (anche nel settore della cybersicurezza). Nel 2024 è nato quindi il Fondo Artificial Intelligence di CDP con una dotazione iniziale di 500 mln, destinato a sostenere start up, centri di trasferimento tecnologico italiani e l'ambizione del LLM italiano. Da un lato, questo strumento si configura come la leva di investimento della terza strategia nazionale che deve far fronte alla progressiva riduzione della spinta pubblica del PNRR. Dall'altro lato, il Fondo è la porta d'accesso per rendere operativa la seconda strategia: sovranizzare attraverso Big Tech. Sono numerose le occasioni che hanno visto la Presidente G. Meloni cercare un'interlocuzione con i vertici delle "grandi sorelle" dell'IA. Si pensi ad esempio agli incontri tenuti da G. Meloni nel settembre 2024 a New York con i CEO di Google-Alphabet, Motorola e Open AI a margine della sua partecipazione alla 79^a Assemblea Generale delle Nazioni Unite. Un risultato di quelle interlocuzioni è stata proprio la firma di un protocollo di intesa tra CDP e Open AI nell'Ottobre successivo. In sintesi, l'accordo consentirà a OpenAI di investire nelle start up italiane attraverso il Fondo di CDP e di condurre iniziative di formazione nelle università italiane e commerciali con altre aziende utilizzatrici o sviluppatri di sistemi di IA¹¹⁰. La volontà del Governo sovranista di costruire un'alleanza con le Big Tech è

¹⁰⁸ Cdp è stata fondata nel 1850 dal Regno di Sardegna con l'obiettivo di finanziare la modernizzazione e infrastrutturazione del paese attraverso l'investimento pubblico dei risparmi delle persone raccolti dai depositi dei libretti postali.

¹⁰⁹ Per una analisi esaustiva dei mutamenti nella forma giuridica e nel ruolo di CdP si rimanda al prezioso lavoro di Mulazzani 2023.

¹¹⁰ Fonte: https://www.cdpventurecapital.it/cdp-venture-capital/it/dettaglio_comunicato.page?contentId=COM3621

evidente nel consolidamento dei rapporti con E. Musk non solo su questioni connesse all'IA. Nel suo doppio ruolo di imprenditore e componente dell'amministrazione Trump Musk è divenuto un interlocutore privilegiato dell'esecutivo italiano. Dapprima è stato ricevuto dal Ministro degli Esteri e dalla Presidente del Consiglio per discutere di IA e investimenti (giugno 2023), poi è stato ospite di Atreju (dicembre 2023), festa del partito di maggioranza FdI, per arrivare poi premiare egli stesso la Presidente Meloni con il "Global Citizen Award" dell'Atlantic Council¹¹¹. Un'alleanza che muove su due binari e non riguarda solo l'IA. Il primo binario, più strumentale, riguarda le risorse tecnologiche controllate da E. Musk come proprietario del sistema di connessione Internet satellitare Starlink. Questa tecnologia, anche secondo il sottosegretario A. Butti, rappresenterebbe una soluzione per raggiungere gli obiettivi della missione PNRR per la diffusione nel territorio nazionale della connessione Internet veloce (banda larga) dove non è possibile realizzare un'infrastrutturazione fisica del territorio. Il secondo binario riguarda la costruzione di un'ampia agenda di policy conservatrice e liberista, espressione di valori e interessi condivisi, sulle questioni quali la natalità, il libero scambio commerciale tra UE e Usa e la cybersicurezza nazionale.

La costruzione di una via italiana all'IA non prescinde dalla "seconda leva" del governo politico dell'IA: la regolazione "dura". Come sostiene N. Urbinati (2020) l'ideologia politica del sovranismo prospera in Europa proprio per i processi di riorganizzazione (ascendenti e discendenti) delle sovranità nazionali. Nel caso dell'IA la questione si fa ancora più difficile. A differenza degli stati nazionali qui considerati "giganti", l'Italia si muove nell'ambiente regolativo comune europeo dove la sovranità sulla *issue* dell'IA è stata costruita a livello comunitario con l'AI Act. L'autonomia regolativa può essere ricercata tuttavia sia in termini reattivi nell'implementazione del Regolamento, che attraverso nuova produzione normativa. Ci riferiamo in particolare al ddl n.1146 "Disposizioni e deleghe al Governo in materia di intelligenza artificiale" presentato da G. Meloni e dal Ministro della Giustizia C. Nordio, approvato dal Senato nel marzo 2025 e in fase di discussione alla Camera nel mese di maggio. Il disegno di legge (ddl) è per il Governo, a nostro avviso, il principale strumento per la costruzione della via italiana all'IA e per sfidare la sovranità europea. Si pensi che nella prima versione all'Art.1 era contenuto un provocatorio e vago riferimento, non all'AI Act, ma alla conformità "al diritto dell'UE" poi corretto nelle versioni successive in: «conformemente al regolamento (UE) 2024/1689 del Parlamento europeo

¹¹¹ L'Atlantic Council è un think tank con sede negli Stati Uniti che si occupa di questioni di politica internazionale.

e del Consiglio, del 13 giugno 2024»¹¹². Sebbene per ragioni di spazio non sia possibile analizzare l'intero testo della “legge italiana”, peraltro ancora in discussione, è possibile coglierne le due più importanti implicazioni rispetto al processo di “sovranizzazione” dell'IA.

Prima fra tutte, il ddl chiude il dibattito su quali saranno le autorità nazionali competenti per la sua attuazione (art. 70) identificate in due agenzie tecniche della PCM: l'AGID¹¹³ come autorità di notifica e l'Agenzia per cybersicurezza nazionale come autorità di vigilanza del mercato e punto di contatto unico con le istituzioni dell'UE. Il regolamento europeo concede agli stati membri un'ampia discrezionalità potendo scegliere, in relazione esigenze e caratteristiche organizzative nazionali, di istituire un'autorità *ad hoc* o assegnare la funzione ad autorità già esistenti (Villani, 2024). Su questa decisione si è aperta fin da subito una discussione tra politica, esperti e società civile basata sulla contrapposizione di due possibili soluzioni alternative: l'istituzione di una nuova autorità indipendente o l'affidamento delle funzioni di attuazione ad agenzie tecniche governative, le cosiddette agenzie amministrative. Le autorità indipendenti rappresentano una configurazione amministrativa particolarmente impiegata in Italia nella gestione di questioni che necessitano di una elevata competenza tecnica¹¹⁴. La sottrazione delle autorità dall'indirizzo politico governativo è l'elemento dirimente nella definizione della loro indipendenza. Esse sono sganciate da responsabilità nei confronti del Governo, che a sua volta non ha poteri di annullamento delle loro decisioni, e hanno il compito di presidiare beni e settori giuridici di rilievo costituzionale secondo la massima neutralità esercitando una serie di funzioni giuridico-amministrative (Marini, 2022). L'istituzione di un'autorità indipendente per l'IA è la soluzione preferita da un'alleanza informale di società civile, opposizione parlamentare e GPDP. Si pensi, ad esempio, all'iniziativa di The Good Lobby¹¹⁵, Hermes Center e Privacy Network che a margine dell'approvazione dell'AI Act hanno elaborato la propria proposta per un'autorità indipendente che possa garantire «una gestione autonoma da influenze politiche e interessi economici, operando come un organo terzo

¹¹² Fonte:

https://www.senato.it/japp/bgt/showdoc/19/DDLMESS/0/1449288/index.html?part=ddlmes s_ddlmess1-articolato_articolato1

¹¹³ Dal 2012 l'AgID è l'organismo tecnico della PCM per l'attuazione dell'Agenda Digitale Italiana ovvero la Strategia Nazionale per l'innovazione digitale della PA.

¹¹⁴ L'elenco completo delle autorità indipendenti italiani è disponibile al link: <https://www.senato.it/3222>

¹¹⁵ La prima è un'organizzazione di *lobbying* nata su iniziativa della società civile (<https://www.thegoodlobby.it/chi-siamo/>), la seconda e la terza sono associazioni che si occupano della tutela dei diritti digitali delle persone e del tema della privacy (<https://hermescenter.org/chi-siamo/>; <https://privacy-network.it/la-nostra-storia/>).

esperto e specializzato, guidato da un board che sia rappresentativo dei diversi interessi nella società»¹¹⁶. Sempre nel marzo del 2024 il GPDP sosteneva la propria posizione attraverso la Segnalazione al Parlamento e al Governo sull’Autorità per l’IA¹¹⁷. Primo, la segnalazione evidenziava come l’AI Act prescrivesse per le autorità nazionali requisiti di indipendenza, imparzialità e pregiudizi, suggerendo dunque la maggiore adeguatezza delle autorità rispetto alle agenzie amministrative. Secondo, il GPDP rivendicava la propria competenza nel settore dell’IA già dimostrata attraverso l’applicazione delle norme del GDPR che si sono mostrate utili nel risolvere le controversie su IA generativa o sull’adozione di algoritmi di riconoscimento facciale. Bisogna considerare che la competenza sull’IA è tuttora contesa da una molteplicità di autorità e agenzie che ne rivendicano almeno una parte. Assistendo al confronto¹¹⁸ tra i rappresentanti delle Autorità per le Garanzie nelle Comunicazioni, Autorità Garante della Concorrenza e del Mercato, GPDP, AGID e Agenzia per la cybersicurezza, abbiamo potuto constatare come ognuna di queste strutture amministrative ponga rivendicazioni legitimate proprio per la trasversalità intrinseca dell’IA: sviluppo e uso dell’IA implicano questioni di concorrenza nel mercato, pongono rischi per la privacy, coinvolgono la PA, richiedono standard di cybersicurezza, etc. Insomma, hanno tutti ragione! L’inevitabile frammentazione delle strutture di governo dell’IA si può ritrovare nell’iniziativa parlamentare del Partito Democratico che nel settembre 2024 ha proposto l’Istituzione dell’Autorità nazionale per l’intelligenza artificiale e le neurotecnologie (ddl 1245)¹¹⁹. Tuttavia, il testo configura un’Autorità fortemente settorializzata sul tema delle neurotecnologie, piuttosto che un’autorità deputata dell’attuazione dell’AI Act¹²⁰.

Di contro, il Governo sovranista ha scelto di attribuire funzioni a due agenzie amministrative che, rispetto alle autorità, hanno un rapporto più diretto con il Governo: sono sottoposte all’indirizzo politico e alla vigilanza di un ministro e i loro vertici organizzativi (direttori generali) sono espressione di una nomina governativa (Filippelli, 2016). Esse sono quindi lo strumento

¹¹⁶ Fonte: <https://www.thegoodlobby.it/wp-content/uploads/2024/03/Autorita-AI-Ita-2024.pdf>

¹¹⁷ Il testo integrale è disponibile al link: <https://www.garanteprivacy.it/home/docweb-/docweb-display/docweb/9996493>

¹¹⁸ Il dibattito è avvenuto nell’ambito dell’evento “Governare l’Intelligenza artificiale. Dove siamo, dove dobbiamo arrivare” organizzato dalla Camera di Commercio di Roma nel febbraio 2024.

¹¹⁹ La proposta è stata elaborata di concerto con la neonata Rete per i diritti Umani digitali di cui fanno parte le tre organizzazioni citate in precedenza.

¹²⁰ Il testo del ddl è disponibile al link: https://www.senato.it/japp/bgt/showdoc/19/DDL-PRES/0/1435720/index.html?part=ddlpres_ddlpres1

tecnico e operativo per la realizzazione degli obiettivi politici dei ministri o della PCM. Questa iniziativa è stata da subito criticata dalla Commissione Europea che, attraverso il parere circostanziato (C(2024) 7814), antecedente all'approvazione della norma italiana, ha evidenziato la necessità di rivedere la scelta in favore della soluzione delle autorità indipendenti. Accentramento governativo e controllo politico sono dunque le caratteristiche con cui provare a connotare politicamente uno degli organismi tecnici di *enforcement* dell'AI Act e connetterlo con le più ampie politiche di sviluppo strategico della via italiana all'IA. In tal senso l'Art. 19 del ddl ridefinisce la "governance" italiana centralizzando il processo di definizione della strategia nazionale, sempre in capo alla PCM, in accordo con le due agenzie e con i ministeri strategici del Made in Italy (politica industriale), della Ricerca e dell'Università e della Difesa.

La seconda implicazione di "sovranizzazione" riguarda la volontà di legiferare negli spazi lasciati liberi dall'AI Act secondo un approccio basato sui settori strategici nazionali. In particolare, la "legge italiana" si occupa di disciplinare i settori della sicurezza nazionale, sanità, lavoro, PA e giustizia. Il rapporto tra IA e PA è un vastissimo terreno di studio per scienziati sociali e giuristi che anche in Italia da anni si interrogano su potenzialità, implicazioni, rischi della cosiddetta "burocrazia algoritmica" (Musella, 2021). Tuttavia, come sostenuto nell'analisi tecnica della "legge italiana", l'IA è entrata progressivamente nei dispositivi di legge, nella PA e nei processi amministrativi attraverso una frammentarietà normativa che necessita di un intervento normativo unitario. In tal senso, il rapporto IA-PA è uno dei principali ambiti su cui il governo sovranista sta intervenendo. Il dossier di Legislatura sull'iter della "legge italiana" evidenzia che l'Art. 13 configura un uso strumentale dell'IA nella PA, segnando la predominanza del controllo umano in linea con le sentenze di giurisprudenza amministrativa (Servizi Studi di Camera e Senato, 2025). Allo stesso tempo, la legge vieterebbe qualsiasi sistema di giustizia predittiva (*ibidem*) in accordo con un principio di predominanza della responsabilità umana. L'intervento governativo più consistente nel rapporto tra PA e IA è tuttavia da ricondurre agli strumenti di indirizzo di AgID che per la prima volta ha inserito la "questione IA" nel il Piano Triennale per l'Informatica nella PA 2024-2026. La roadmap stabilita dal Piano prevede la progressiva pubblicazione di tre linee guida per rafforzare la consapevolezza nell'adozione, l'acquisto e lo sviluppo dell'IA nella PA. Nel febbraio 2025, AgID ha provveduto quindi a emanare le prime linee guida per l'adozione di IA nella pubblica amministrazione¹²¹ a cui seguiran-

¹²¹ La bozza sottoposta a consultazione pubblica è scaricabile al link:
https://www.AgID.gov.it/sites/AgID/files/2025-02/Linee_Guida_adozione_IA_nella_PA.pdf

no le linee guida per il *procurement* (acquisto da parte di amministrazioni pubbliche) delle tecnologie di IA, un ulteriore leva di “sovranizzazione” prefigurata dal ddl Meloni-Nordio. In riferimento all’acquisto dei sistemi di IA, gli Art. 5 e 6 della “legge italiana” introducono elementi di controllo nel governo dei dati e nell’acquisizione di servizi IA dall’estero. In sintesi: le piattaforme di *e-procurement* pubblico dovrebbero privilegiare l’acquisto di soluzioni che garantiscono la localizzazione e l’elaborazione dei dati strategici presso *data center* posti nel territorio nazionale. I sistemi di IA applicata a sicurezza e difesa, inoltre, dovrebbero essere installati su server presenti sul territorio nazionale. Una misura su cui convergono esigenze di sicurezza nazionale e di «competitività del sistema economico nazionale» (Presidenza del Consiglio dei ministri, 2024, p. 7). Ciò rappresenterebbe una limitazione sia alle possibilità di utilizzo dei sistemi più avanzati già adottati dalle PA, una violazione delle direttive europee sulla libera circolazione delle merci (Burelli, 2025) sia, potenzialmente, una azione ostile nei confronti delle innovazioni IA sviluppate in ambito europeo, che sappiamo essere il riferimento scalare dell’Italia sia per le strategie *social-oriented* che per quelle *market-oriented*.

In conclusione, possiamo identificare nel percorso di sovranizzazione dell’IA uno dei più interessanti aspetti del governo politico dell’IA su scala nazionale che, come visto nel caso italiano, si basa sulla ricerca di autonomia finanziaria e normativa e sulla ricerca di una sponda economica e valoriale nelle Big Tech, prerequisito per l’inserimento dell’Italia nella globalizzazione economica dell’IA. Esso ha evidenti implicazioni nel produrre tensioni tra scale regolative sovranazionali e nazionali, che si mostrano anche nell’adozione di alcune politiche prioritarie nell’agenda populista e sovranista. Si pensi in tal senso alle polemiche periodiche tra governo, opposizione e UE sull’interpretazione delle norme comunitarie sul riconoscimento facciale nelle politiche di sicurezza. Nel capitolo successivo “scaleremo” ancora una volta lo sguardo sul rapporto tra politica e IA occupandoci delle questioni regolative e del suo uso su scala locale.

4. L'intelligenza artificiale nelle politiche locali

4.1. Dalle Smart City all'intelligenza artificiale urbana: agende e strategie

Nelle agende politiche dell'IA oltre alla promozione dell'innovazione e dello sviluppo di questa tecnologia e alla regolazione delle sue modalità di utilizzazione, sono entrate svariate modalità di utilizzazione di sistemi di IA nell'azione pubblica, in vari settori e su varie scale territoriali. In questo capitolo ce ne occupiamo concentrandoci sugli usi dell'IA nelle politiche e nell'azione amministrativa su scala locale e urbana. Questa dimensione è particolarmente significativa perché ricomprende una larga parte degli usi (pubblici e privati) di sistemi di IA. Infatti, più della metà della popolazione mondiale vive in aree urbane, era circa un terzo nel 1950, e si prevede che raggiungerà circa i due terzi nel 2050¹ quando, peraltro, anche l'IA sarà una tecnologia ancora più matura e diffusa. Il processo di urbanizzazione può essere interpretato in termini di concentrazione. Nelle città si concentrano, investimenti, ricchezze e possibilità di innovazione in diversi settori di attività economica, fattori che hanno favorito la crescita economica dei centri urbani e la loro infrastrutturazione tecnologica. Non stupisce, quindi, che gli ambienti urbani offrano le infrastrutture tecnologiche avanzate – sensori, sistemi di comunicazione, centri di archiviazione e elaborazione dati e, spesso, milioni di individui produttori di dati – necessarie per farne ecosistemi privilegiati per la sperimentazione di forme di previsione digitale (Cugurullo *et al.*, 2024b) e, a volte, di decisione autonoma delle macchine. Un insieme di dispositivi, sistemi operativi e attori umani, insomma, che producono quella “città computazionale” (Luque-Ayala e Marvin, 2020) che, come vedremo, è al cuore sia del paradigma Smart City (SC) affermatosi una ventina di anni fa, sia della sua evoluzione in città dell'IA.

¹ Fonte: <https://www.un.org/development/desa/pd/content/urbanization-0>

Tuttavia, oltre allo sviluppo economico e tecnologico, l'urbanizzazione concentra anche esternalità negative, agendo come catalizzatore di degrado ambientale e disuguaglianze sociali, questioni che richiedono risposte e azioni pubbliche. La risposta divenuta *mainstream* in tutto il mondo dagli anni '90 al primo quarto del XXI secolo – ma messa in dubbio dall'avvento di governi populisti di destra negli Usa e in alcuni Paesi dell'UE – è consistita nell'adottare agende e politiche per lo sviluppo sostenibile che ricercano l'equilibrio tra sviluppo e sue esternalità negative. Le politiche per lo sviluppo urbano sostenibile sono adottate per contrastare crisi, rischi e incertezze derivanti dalla ristrutturazione delle economie urbane nell'epoca della globalizzazione, con modalità diverse fra le città in cui si è concentrata la produzione manifatturiera e quelle occidentali, in cui l'attrazione di capitali e investimenti tecnologici favoriti dal neoliberismo urbano (Rossi e Vanolo, 2015), ha fatto sì che si concentrassero la crescita della ricchezza e della povertà (polarizzazione sociale), il lavoro precario, la crescita dei costi delle abitazioni, la segregazione spaziale. A ciò si uniscono il degrado dell'ambiente e delle infrastrutture, più tipi di inquinamento, a loro volta cause e conseguenze del cambiamento climatico. Le azioni pubbliche, attraverso le quali si cerca di convertire l'urbanizzazione in una spinta per realizzare sviluppo sostenibile nelle dimensioni sociali, economiche e ambientali dello sviluppo urbano (Simon, 2016) mirano a limitare consumi energetici, emissioni di CO₂ e disuguaglianze sociali. Con sfumature diverse, le indicazioni per l'azione si ritrovano in documenti ufficiali e iniziative di vario tipo di organizzazioni internazionali (UN, UE, OECD), società globali di consulenza, reti di città, governi nazionali e think tank. Queste richiedono azioni condotte simultaneamente a diverse scale, o livelli di governo (mondiale, continentale, nazionale, locale) da attori pubblici e privati e la messa in campo di risorse di vario tipo: non solo finanziarie, di legittimazione e consenso, di autorità, ma anche risorse cognitive di ideazione e gestione di problemi complessi.

L'IA è entrata nelle agende e politiche urbane riproducendo lo stesso "schema" di concentrazione di potenzialità e esternalità negative che nel nostro approccio abbiamo definito come rischi pubblici. Già nel 2020 le UN affermavano la coesistenza a livello globale di città che stavano sfruttando i benefici dell'IoT, dell'IA e dell'automazione e di città che subivano l'emergere di nuovi problemi quali la sorveglianza di massa, la perdita occupazionale e l'esacerbarsi di esclusione digitale e sociale (UN-HABITAT, 2020). Ancora una volta troviamo il *trade-off* tra potenzialità e rischi "scalato" a livello urbano e locale. Da un lato, nel discorso pubblico sull'IA è spesso evidenziato il suo potenziale trasformativo, sia delle attività economiche e sociali nelle città, sia delle forme di azione pubblica, a partire proprio dal

mettere l'IA al servizio degli Obiettivi di Sviluppo Sostenibile (SDG)², che trovano nel rendere le città “inclusive, sicure, resilienti e sostenibili” una delle principali opportunità e sfide per il loro raggiungimento³. Dall'altro lato, così come avviene per gli usi privati, il dibattito pubblico e scientifico evidenzia le esternalità e implicazioni dei vari possibili usi dell'IA su scala urbana che riguardano molti dei *trade-off* esaminati anche nei precedenti capitoli. La valorizzazione delle potenzialità e il contestuale governo di queste esternalità è dunque la sfida principale per la politica locale dell'IA di cui ci occupiamo in questo capitolo.

In questo paragrafo sono ricostruite le caratteristiche dell’“IA urbana”, i suoi ambiti di utilizzazione e, rispettivamente, le continuità e le discontinuità rispetto al paradigma SC, il precedente ciclo di sviluppo e utilizzazione delle tecnologie digitali nelle politiche urbane. L’ultima parte si occupa di presentare le principali questioni politiche legate all’IA urbana. Per “IA urbana” si intende una varietà di sistemi digitali che “incorporano dati derivati dall’ambiente urbano, elaborati da algoritmi, il cui risultato ha applicazioni utili nel nesso socio-spatiale della città” (Popelka *et al.*, 2023). Detto in altri termini, “un insieme di algoritmi che apprendono dai set di dati urbani e vengono utilizzati per soluzioni che sono, o potrebbero essere, implementate in una città” (Beroche, 2021, p. 6). Come altri sistemi di IA, anche quelli urbani agiscono sulla base di modelli identificati da dati e regole. Oltre che attraverso l’addestramento, l’IA può anche imparare da sola (*deep learning*) analizzando i Big Data e arrivando, in alcuni casi, a trarre delle conclusioni e prendere decisioni. Si tratta quindi di sistemi materialmente situati nel contesto fisico della città e intrecciati con la vita quotidiana dei suoi abitanti, con finalità di uso differenti, personali, di imprese private e pubbliche, dalla salute alla cura di anziani e disabili, ai veicoli a guida autonoma, alla pianificazione urbanistica, solo per fare alcuni esempi. In questo volume ci concentriamo solo sui sistemi di IA urbana utilizzabili per perseguire, l’efficienza operativa di funzioni di interesse pubblico nelle città e supportare la pianificazione, attraverso il loro impiego nelle politiche pubbliche e nel funzionamento delle amministrazioni locali.

Le utilizzazioni dell’IA urbana sono varie e sono possibili diverse classificazioni. Di seguito (tab. 1) presentiamo una tipologia sintetica ed esemplificativa delle modalità di uso dell’IA urbana, basata su un criterio di tipo funzionale, relativo alle finalità con le quali l’IA viene impiegata nell’azione pubblica locale – operative, strategiche o per il governo dei processi – e sui settori di politiche e servizi pubblici interessati.

² Fonte: <https://sdgs.un.org/goals>

³ Fonte: <https://sdgs.un.org/goals/goal11>

Tab. 1 - Tipi e settori di uso dell'IA urbana, elaborazione degli autori.

Settore	Uso strategico	Uso operativo	Uso partecipativo
Pianificazione e gestione urbanistica	Modelli predittivi per lo sviluppo urbano, simulazioni di scenari	Approvazioni automatizzate dei permessi di costruzione Controlli urbanistici	Piattaforme per la pianificazione partecipativa (ad es. visualizzazione degli effetti di proposte di trasformazione urbana)
Ambiente e sviluppo sostenibile	Previsione delle tendenze della qualità dell'aria, simulazioni degli effetti delle azioni per il clima	Rilevazione in tempo reale di dati sull'inquinamento	Analisi di dati ambientali raccolti tramite <i>crowdsourcing</i>
Mobilità e trasporti	Simulazioni per la pianificazione a lungo termine della viabilità e delle reti di trasporto	Controllo adattivo dei semafori, gestione dei parcheggi	Analisi dei feedback dei cittadini sulla qualità e il miglioramento dei trasporti pubblici
Sicurezza pubblica (<i>Safety</i> e <i>Security</i>)	Previsione dei luoghi di concentrazione del crimine e dei possibili autori di comportamenti criminali. Decisioni su allocazione delle risorse di sicurezza	Sorveglianza in tempo reale attraverso IoT, sistemi di risposta alle emergenze	Rapporti sulla sicurezza forniti dagli abitanti per rilevare tendenze
Servizi amministrativi	Analisi delle tendenze di uso dei servizi per orientarne la pianificazione	<i>Chatbot</i> per la fornitura di servizi e la gestione automatizzata delle richieste	Sentiment analysis dei feedback pubblici

Un uso con finalità di tipo *strategico* colloca l'IA a supporto di decisioni politiche con effetti ampi e di lungo termine. Ci si aspettano miglioramenti macro e un funzionamento della città trattata come un gigantesco artefatto, che può essere gestito e ottimizzato con vari gradi di autonomia dal controllo umano (Cugurullo, 2020). Un uso con finalità *operative* impiega l'IA come aiuto nel *problem solving* quotidiano. Ciò può avvenire in sostituzione delle decisioni umane (senza comandi umani esplicativi di volta in volta) o a supporto di queste decisioni, attraverso la predizione degli impatti di scelte e comportamenti. Questa utilizzazione può avere più facilmente effetti a breve termine, da cui ci si aspettano miglioramenti micro, derivanti dall'automazione dei compiti di routine e gestionali e la conseguente liberazione degli

umani da attività di routine e compiti ripetitivi e riduzione dell’errore umano (Caprotti *et al.*, 2024).

La pianificazione urbanistica e il design urbano usano i modelli predittivi dell’IA e la simulazione di scenari per prendere decisioni strategiche relative allo sviluppo territoriale, su scala urbana e metropolitana, o regionale. Le scelte possono riguardare l’uso del suolo e dello spazio (per residenza, servizi, verde, infrastrutture) in strade, piazze, quartieri, fino ad intere aree urbane o metropolitane e la sua interdipendenza con mobilità e pendolarismo, raccolta dei rifiuti, installazione di grandi infrastrutture come centrali energetiche o impianti di smaltimento dei rifiuti. Un esempio che riguarda invece l’uso operativo: Austin (Texas, Usa) ha iniziato a usare un software di valutazione delle licenze edilizie basato sull’IA, con il fine di accelerare le approvazioni e consentire al governo locale di concentrarsi su valutazioni più complesse⁴. L’uso operativo riguarda anche il decoro urbano, con l’analisi di oggetti che indicano degrado o contribuiscono a rendere antiestetico lo spazio urbano, come buche, graffiti, rifiuti, tende, finestre sbarrate o rotte, facciate scolorite o fatiscenti, erbacce e segni su servizi pubblici (Vallebueno e Lee, 2023).

Nelle agende politiche di sviluppo sostenibile è sempre più centrale l’impiego di tecnologie digitali e dell’IA, di cui i *policy maker* stanno facendo un uso sempre maggiore, come già osservato negli ultimi due decenni con riferimento alle SC. Se le potenzialità dell’IA sono riconducibili ad ognuno degli SDGs (Isagah, 2024), nelle città sono particolarmente importanti le sue applicazioni nel campo delle politiche ambientali. Sebbene sia impossibile tenere traccia delle numerose sperimentazioni in corso, l’IA appare uno strumento sul quale si fa crescentemente affidamento in molte attività di previsione e monitoraggio ambientale attraverso la rilevazione ed elaborazione in tempo reale dei dati. In termini strategici, i modelli di IA possono facilitare le scelte nella di gestione del ciclo dei rifiuti e nelle politiche energetiche, ad esempio, attraverso la previsione della quantità di energia che un’area consumerà in (Jha *et al.*, 2021). Ancora, l’IA può essere impiegata nell’ambito delle comunità energetiche per ottimizzare l’autoproduzione e lo scambio dell’energia con la rete elettrica nazionale (Dodaro *et al.*, 2024). L’uso operativo di IA può invece aiutare a monitorare e prevedere la qualità dell’aria e i rischi di inquinamento. Un’azione che può avvenire anche attraverso la gestione automatizzata dei flussi di traffico, un uso che ci introduce all’impiego dell’IA nel campo della mobilità e del trasporto locale. Il suo uso strategico riguarda l’ottimizzazione dei percorsi, la pianificazione del traffico e

⁴ 16 10 24 Cities Today (<https://cities-today.com/austin-launches-ai-driven-building-permit-software/>).

la modulazione dei ritmi urbani. Gli algoritmi producono previsioni sui flussi di traffico e simulazioni elaborando dati forniti da sensori e telecamere (Colletta e Kitchin, 2017). L’uso operativo consiste invece in azioni materiali e fisiche attraverso veicoli autonomi, droni, regolazione dinamica e automatica dei semafori per ridurre la congestione, previsione dei ritardi e ottimizzazione dei percorsi del trasporto pubblico⁵. Ad esempio, in 23 città cinesi e malesi, *ET City Brain* di Alibaba è usato sia per prevedere ingorghi e incidenti stradali, sia per dare la possibilità di affrontare nel presente le cause strutturali dei futuri problemi di trasporto (Cugurullo e Xu, 2025). O, ancora, è possibile pagare i trasporti pubblici attraverso il riconoscimento facciale (Londra, Madrid)⁶.

L’uso dell’IA nel campo della sicurezza urbana è la funzione più controversa, soprattutto quando essa non è intesa solo come *safety* (circostanze rischiose non provocate da comportamenti intenzionali e risposte a emergenze come incendi, allagamenti), ma anche come *security* (pericoli o minacce derivanti da comportamenti intenzionali). Nel secondo caso, gli algoritmi possono essere impiegati in sistemi di sorveglianza e previsione dei reati, con o senza riconoscimento facciale ed elaborazione di dati biometrici. In varie città cinesi, ad esempio, il *city brain* attinge al database del dipartimento di polizia locale e alla sua rete di telecamere a circuito chiuso, accoppiato con il software di riconoscimento facciale, per prevedere dove si svolgeranno i crimini in città, chi sarà coinvolto in attività criminali e gli interventi di polizia da eseguire (Cugurullo e Xu, 2025).

Il lavoro delle amministrazioni locali può essere facilitato da *chatbot* per la comunicazione con i cittadini (e-government). Ad esempio, il sindaco di New York può raggiungere i newyorkesi tramite chiamate automatiche anche in spagnolo, yiddish e mandarino e un *chatbot* che potrebbe essere utilizzato per rispondere al 70% delle chiamate indirizzate alla linea di assistenza 311 della città per delle semplici domande⁷. Inoltre, l’elaborazione del linguaggio naturale aiuta a compilare moduli, formulare domande e approvare (o meno) automaticamente permessi e licenze. In Italia, ad esempio, questa funzione è stata sperimentata attraverso un progetto del Formez (centro servizi della presidenza del Consiglio dei ministri) che coinvolge 203 comuni italiani⁸.

Un uso dell’IA con finalità *partecipative* riguarda il supporto all’inclusione di attori non politici nei processi di governo ed è per questo trasversale

⁵ Fonte: <https://cities-today.com/event/artificial-intelligence-for-urban-mobility/>

⁶ Fonte: <https://gouai.cidob.org/#documentary>

⁷ Fonte: <https://www.nytimes.com/2023/10/20/nyregion/ai-robocalls-eric-adams.html>.

⁸ Fonte: <https://www.open.formez.it/notizie/lintelligenza-artistificiale-arriva-nei-comuni-della-sicilia-presentato-progetto-fast-formez-pa>

rispetto ai settori di azione pubblica considerati. L'IA rende possibile gestire le petizioni dei residenti, elaborare le loro opinioni⁹ e raffinare le modalità di gestione e facilitazione di pratiche partecipative e deliberative. La scala locale e urbana è considerata anche un buon ambito di sperimentazione: iniziare "in piccolo", applicando l'IA a decisioni con gruppi di *stakeholder* più limitati e impatti minori può consentire di costruire una tecnologia che assiste i processi democratici su scale più ampie (Schneier *et al.*, 2023). A loro volta, i progetti di IA urbana, le finalità e le modalità della sua utilizzazione per tutti i fini che abbiamo considerato possono essere oggetto di consultazione e partecipazione degli abitanti, con modalità che possono o meno essere assistite da IA (Popelka *et al.*, 2023).

Un altro esempio di utilizzazione intersetoriale riguarda la gestione del turismo con finalità relative sia alla mobilità, sia alla sicurezza. I dati prodotti dagli smartphone rendono infatti possibile sapere come si muovono abitanti e turisti nelle città, chi è in città, da dove viene, come è arrivato e l'IA trasforma i dati in informazioni utilizzabili dai decisori per identificare strategie e prendere decisioni finalizzate a ottimizzare le modalità di spostamento e migliorare l'esperienza dei turisti.

L'impiego più concentrato dell'IA urbana e della robotica si trova nel Nord del mondo (Golubchikov e Thornbush, 2020), ma Cina e Sud-est asiatico stanno velocemente riducendo le distanze. In generale, le tendenze per il futuro indicano un aumento delle applicazioni nei settori della sicurezza, della mobilità, della sorveglianza e della modellizzazione, con potenziali sviluppi nei campi delle Blockchain, dell'IA generativa, dei veicoli stradali e aerei a guida autonoma, delle applicazioni di IA geospaziale e delle reti neurali per il riconoscimento di immagini (Yigitcanlar *et al.*, 2024).

Se questo è il futuro atteso, possiamo capire meglio cosa sia l'IA urbana se consideriamo che essa ha anche dei precedenti e delle premesse, identificabili principalmente nei suoi rapporti con le SC. In altri termini, l'IA urbana si colloca all'intersezione fra l'IA in generale e la nozione di SC (Beroche, 2021, p. 6) e le iniziative di IA nelle città sono considerate tasselli all'interno di una varietà di processi di innovazione della gestione e dello sviluppo urbano basati sulle tecnologie digitali di cui la SC è il prototipo principale e più comprensivo. Infatti, quella di SC non solo è l'espressione più frequentemente usata nella letteratura scientifica e nel linguaggio delle *policy community* urbane in tutto il mondo per riferirsi a modelli computazionali di urbanistica e a tecnologie basate sui dati e mediate da algoritmi che introdu-

⁹ Ad esempio, su come promuovere turismo tenendo conto anche dei problemi che esso provoca alla città. Fonte: <https://cities-today.com/can-artificial-intelligence-make-tourism-smarter/>

conno innovazione nell'amministrazione locale per aiutarla a sostenere le sfide urbane (Botero Arcila, 2022), ma è attualmente spesso utilizzata per riferirsi anche all'adozione di IA da parte dei governi locali (Yigitcanlar *et al.*, 2024).

Il concetto di SC è stato introdotto nel 2008 a partire dalle attività delle imprese informatiche IBM e CISCO in risposta alle sfide poste dall'urbanizzazione, per definire un nuovo ambiente urbano, progettato per la crescita e le performance attraverso la tecnologia (Castillo *et al.*, 2022). A. Courmont e P. Le Galès (2019) hanno notato la coincidenza temporale con il lancio delle due più importanti piattaforme private dell'economia urbana (Uber e Airbnb) e l'avvio della massificazione dei dati individuali estratti dagli ambienti urbani, in primo luogo attraverso gli smartphone, per finalità commerciali.

L'“immaginario” urbano della SC aggrega rappresentazioni dei problemi urbani fondate su valori e senso comune condivisi, spazi iconici in cui le reti e i flussi diventano visibili e si rende possibile un passaggio dalla governance al controllo (Caprotti, 2019). Si tratta di rappresentazioni che sono state ricordate al ruolo svolto da ideologie e mercati, che esercitano influenza sulle istituzioni politiche (Botero Arcila, 2022). Queste rappresentazioni però non sono socialmente e culturalmente neutrali, ma costruiscono un modello di *smart urbanism* per il futuro che riflette e rafforza in larga misura i sistemi sociopolitici esistenti (Sadowski e Bendor, 2018). Basandosi su questo tipo di rappresentazioni si è provato a dare ai problemi delle città e alle soluzioni per essi proposte attraverso la tecnologia digitale il senso di un'agenda di governo, articolata in un ampio ventaglio di azioni, ma integrate e rese coerenti fra loro da una strategia. Nella letteratura sulle SC gli ambiti in cui è possibile usare in modo “smart” le tecnologie digitali corrispondono, con poche eccezioni, a quelli dell'IA urbana: governance e digitalizzazione, energia, sicurezza, pianificazione urbanistica, gestione degli edifici, gestione dei rifiuti, produzione industriale, istruzione, gestione dei servizi a rete (elettricità, acqua), trasporti (Bosco *et al.*, 2024).

Questa estensione degli ambiti di policy investiti dal paradigma SC, tale da riguardare potenzialmente l'intera agenda di governo di una città, è stata vista come una spinta per riorganizzarla attraverso definizioni tecniche e depoliticizzate sia dei problemi, sia delle soluzioni efficaci (Vanolo, 2014). Questa rappresentazione è adattabile pressoché a ogni contesto urbano ed è coerente con il modello competitivo e imprenditoriale della città e dello sviluppo urbano tipici del neoliberismo urbano, che è possibile concretizzare, in questo caso, attraverso il ruolo delle ICT. Anche l'UE, che ritiene questa visione coerente con quella su cui è basata la sua agenda urbana (d'Albergo, 2015), ha contribuito a legittimare e diffondere il paradigma SC inserendolo

nei suoi programmi dal 2009 (Armondi, 2014) e cercando di utilizzare le SC per aumentare la competitività e il progresso tecnologico delle città europee (Beroche *et al.*, 2022)

Anche se le iniziative di IA urbana hanno suscitato l'interesse di diverse città indipendentemente dalle loro “aspirazioni di SC” (Yigitcanlar *et al.*, 2021, cit. in Popelka *et al.*, 2023), la preesistente diffusione globale delle esperienze SC ha quindi preparato il terreno. Il modello preesistente di SC ha infatti reso più facile che l'IA fosse sia accolta favorevolmente negli ambienti politici urbani, sia capace di attrarre l'attenzione degli studiosi sulle opportunità e i limiti del processo decisionale algoritmico nelle città (Yigitcanlar *et al.*, 2023).

Tuttavia, sulle relazioni fra IA urbana e SC ci sono opinioni diverse. Da un lato vengono sottolineati gli elementi in comune e di continuità, dall'altro quelli di differenziazione e discontinuità.

Vediamo i primi. In primo luogo, sia le SC, sia l'IA urbana hanno un carattere ibrido, derivante dall'essere modelli nella sfera digitale dotati però di una componente infrastrutturale materiale, fatta di sensori e altre tecnologie per raccogliere, archiviare e visualizzare dati. In secondo luogo, condividono la capacità della tecnologia digitale di mettere al servizio delle funzioni urbane conclusioni provenienti dai dati, “automatizzando un management urbano efficiente” (Herat e Mittal, 2022, p. 18).

Anche i sistemi di azione collettiva della SC sono comparabili con quelli dell'IA urbana. La SC è infatti costruita socialmente come un “sistema sociotecnico in grado di sostenere e abilitare l'innovazione” (ANCI – ForumPA, 2013), fondato sulla cooperazione fra governi locali, imprese di ICT, fornitori di dati e altri portatori di interesse della città. A sua volta, anche lo sviluppo e l'uso di IA urbana, intesa come un sistema sociotecnico, richiede la costituzione di “ecosistemi locali di innovazione”, ossia reti di attori – imprese, istituti di ricerca, investitori, governi locali e utenti – la qualità delle cui interazioni sono importanti per le prestazioni innovative dell'intero gruppo (Granstrand e Holgersson, 2020). Questa rappresentazione, già presente nelle SC, richiama il modello della “tripla elica”, ossia interazioni basate sull'interdipendenza tra università, industria e governo considerate uno dei fattori chiave nell’“immaginario” delle economie basate sulla conoscenza (Etzkowitz e Leydesdorff, 1995; Fasciani, 2025). La presenza in città di un sistema di questo tipo è una delle condizioni che può rendere possibile a una SC di adottare sistemi di IA (Rjab *et al.*, 2023), poiché sono necessari non solo una tecnologia adeguata, ma anche la qualità dell'ambiente sociale e della sua organizzazione.

Come avviene anche per altre tecnologie, questi sistemi locali possono essere connessi a scale più ampie, specialmente quando i governi locali si

riforniscono di tecnologie di IA presso fornitori che operano su scala nazionale, o globale (Oh *et al.*, 2016). Peraltro, questa componente dell'IA urbana – l'essere al tempo stesso locale e globale – ne fa anche una specifica posta in gioco nella geopolitica dell'IA. I sistemi nazionali, in particolare statunitense e cinese, competono per il controllo del cyberspazio e della sfera cognitiva umana – ad esempio attraverso il *soft-power* dell'economia delle piattaforme, della videosorveglianza e di altre tecnologie cinesi per le SC – e questi conflitti si svolgono in teatri urbani (Beroche *et al.*, 2022).

Alcuni anni fa Yigitcanlar e colleghi (2020) rilevavano che in una prima fase la letteratura esistente sulle intersezioni tra SC e IA si concentrava principalmente sulla descrizione di tecnologie e sistemi specifici di IA, con minore attenzione alle implicazioni sociali. In breve, però, la ricerca si è indirizzata a mettere in luce non solo ciò che è tecnicamente fattibile ma anche le sfide organizzative e le questioni etiche, sociali e relative alla partecipazione civica che si pongono quando le SC cercano di sfruttare il potere dell'IA per affrontare le sfide urbane (Medaglia *et al.*, 2023; Wolniak e Sedula, 2024; Aghdam *et al.*, 2024; Bosco *et al.*, 2024). In questo senso, SC e IA urbana sono accomunate dall'essere anche oggetto di valutazioni non ottimiste sulle loro conseguenze sociali. Infatti, è stato rilevato che i primi esperimenti di IA urbana non stanno mantenendo le loro promesse, non così diversamente da quanto è accaduto con le promesse sociali e ambientali non mantenute dagli esperimenti delle SC (Cugurullo *et al.*, 2024b). Infatti, se è vero che i promotori di sistemi di IA urbana sostengono che questi possono migliorare i processi metabolici di una città anche dal punto di vista ambientale e sociale, questi esiti non sono scontati. Anzi, questi obiettivi possono rivestire di una patina socialmente e culturalmente più presentabile propositi prevalentemente economici (Chen, 2024; Shapiro, 2024). L'IA urbana finirebbe così per condividere con le SC anche l'essere un modello selettivo, fondato sul consumo personalizzato abilitato dalle possibilità economiche, coerente con la città e le agende politiche urbane neoliberiste (Marciano *et al.*, 2020). Questo è più evidente in quello che è stato identificato come un modello statunitense delle SC, in cui la tecnologia – gestita principalmente da imprese private in collaborazione con i governi locali – mira a migliorare la qualità della vita, riducendo al contempo i costi operativi degli *stakeholder* urbani, nella logica della “città imprenditoriale”. Queste caratteristiche emergono più nitidamente se confrontate con quelle delle SC in Cina, declinate spesso come *Safe Cities*, secondo una cornice di legittimazione ispirata alle questioni di sicurezza urbana. Le SC cinesi sono orientate a usare le tecnologie e i dati generati dagli abitanti, soprattutto attraverso smartphones e videocamere di sorveglianza con riconoscimento facciale, per la sicurezza, l'ordine politico e l'armonia sociale. In questo caso le autorità politiche

hanno un ruolo più significativo nella gestione di sistemi e modelli di IA, anche se in collaborazione con le imprese (Beroche *et al.*, 2022).

Oltre a specifici insuccessi nell'uno o nell'altro settore di *policy*, vengono però segnalate criticità più ampie e radicali. Lo stesso carattere specificamente urbano delle tecnologie nelle SC e nell'IA urbana – ossia l'essere originate e collegate a specifiche città e sottoposte al controllo di chi le abita – è stato messo in dubbio e rappresentato come un obiettivo il cui raggiungimento è possibile, ma non scontato. La SC è stata infatti rappresentata come un non-luogo, non inclusivo, ma frammentario, parcellizzato, rivolto a pochi, un modello utopico intriso di contraddizioni, che dà per scontato la reale partecipazione/inclusione di tutti i cittadini all'utilizzo degli apparati digitali (Toti, 2020). Le tecnologie delle SC sarebbero “de-urbanizzate”, perché non realizzate per funzionare in un particolare contesto urbano e più adatte ad operare in contesti simili a sistemi chiusi, mentre le città sono sistemi aperti e incompleti (Sassen, 2021).

Del paradigma SC si è inoltre detto che traccia un confine tra ciò che è accettabile o meno, stigmatizzando città e attori che rimangono indietro nella competizione per essere *smart* o, peggio, non si adeguano alla “*smartmentality*” e sono perciò “*smartdeviant*” e (moralmente) irresponsabili (Vanolo, 2014). Come altri immaginari del neoliberismo, l'ottimalità del principio di intelligenza ha reso possibile presentare l'approccio SC in termini apolitici e di buon senso come inevitabile e, almeno razionalmente, non rifiutabile (Kitchin, 2014), sottraendolo perciò a ogni possibile controversia politica. Del resto, chi rifiuterebbe una simile panacea urbana (Hollands, 2014) scegliendo invece di vivere in una città “stupida”? Ma anche il carattere inevitabile con il quale si presenta oggi lo sviluppo e l'adozione dell'IA – urbana e non – sembra configurare un immaginario sociotecnico, ossia una rappresentazione che descrive i futuri potenziali e prescrive i futuri che dovrebbero essere perseguiti, con effetti non significativamente diversi.

Un esito di tutte queste considerazioni critiche è la proposta di usare l'IA in modo alternativo, per “completare il paradigma SC e urbanizzare la tecnologia”¹⁰. Anche questa ipotesi è un elemento in comune fra SC e IA urbana, perché richiama le precedenti propensioni a non ritenere per definizione lo stesso *smart urbanism* un paradigma imprenditoriale, post-politico e non locale (Rossi, 2016; Caprotti, 2019), ipotizzando che sia possibile piegare l'IA urbana alla trasparenza, comprensibilità e partecipazione dei cittadini, in funzione della sostenibilità. In altre parole, l'alternativa è “politicizzare” gli algoritmi attraverso la valorizzazione del ruolo della politica, in questo caso locale, nel governo e non solo nell'utilizzazione dell'IA.

¹⁰ Fonte: <https://urbanai.fr/call/>

Veniamo agli elementi di differenziazione e discontinuità fra SC e IA urbana. In un dibattito nelle scienze sociali su questi aspetti sinora piuttosto limitato, sono emersi tre principali elementi di differenziazione, relativi: ai rapporti fra conoscenze provenienti dall'elaborazione di dati e decisioni di gestione delle funzioni urbane, alla dimensione temporale e a quella relativa all'accettazione e alla legittimazione sociale delle tecnologie.

Mentre la SC è rappresentata come una macchina integrata, efficiente e guidabile attraverso il contributo della tecnologia alle scelte umane, la città dell'IA è più imprevedibile, adattabile e soggetta a logiche emergenti, con un potenziale considerevole per ricostruire o rimodellare le relazioni socio-tecniche. Nelle SC l'intelligenza stessa è di origine umana (Cugurullo *et al.*, 2023) e i sistemi tecnologici apportano capacità di calcolo di dati sul presente delle città (Caprotti *et al.*, 2024) e sono programmati per eseguire compiti e seguire percorsi o corsi di azione prestabiliti (Cugurullo *et al.*, 2024b). Nell'IA urbana la tecnologia ha invece una propria capacità di giudizio e di sviluppare soluzioni sulla base dell'apprendimento, fino a configurarsi potenzialmente come uno *stakeholder* attivo, quando gli attori politici la mettono nella posizione di prendere decisioni sul governo delle città (Cugurullo *et al.*, 2024b). Mentre nelle SC, la tecnologia è capace di acquisire dati dall'ambiente ed eseguire calcoli, l'IA urbana dai dati può estrarre concetti, produrre relazioni causali di un fenomeno e previsioni del futuro, e – in pochi casi relativi a progetti particolarmente avanzati – anche in autonomia rispetto alla guida umana, producendo potenzialmente anche impatti materiali (Caprotti *et al.*, 2024). Le tecnologie robotiche e i sistemi autonomi possono quindi usare l'IA e il *machine learning*¹¹ per prendere decisioni e adattare i processi alle circostanze anche senza l'intervento diretto dell'uomo. Per questo, l'IA urbana è rappresentata come “città autonoma” (Caprotti *et al.* 2024), o “città algoritmica”, in cui i sistemi tecnologici non necessitano di *input* umani per funzionare e possono operare in modo non supervisionato (Cugurullo *et al.*, 2023), fino ad esercitare un'autonomia decisionale. E non vi sono garanzie che le decisioni prese dall'IA siano automaticamente in linea con i valori umani, cosa più grave se le IA urbane possono svolgere ruoli che tradizionalmente sono stati di competenza degli esseri umani, come il governo dell'intera città (Cugurullo *et al.*, 2023).

Per quanto riguarda la dimensione temporale, a differenza dei processi di cambiamento implicati nelle SC, orientati a produrre conseguenze nel breve termine, l'IA urbana può produrre o profilare, come abbiamo visto, strategie

¹¹ Il *machine learning* (apprendimento automatico) è un tipo di IA, una categoria di algoritmi che creano altri algoritmi. Attraverso *machine learning*, i professionisti sviluppano l'IA attraverso modelli che possono “imparare” da modelli di dati senza la direzione umana.

relative a cambiamenti che riguardano il futuro in vari settori di politiche urbane. Questo dipende anche dal fatto che l'IA urbana e il suo software sottostante cambiano costantemente, producendo un'evoluzione continua delle interazioni uomo-tecnologia in tempo reale (Caprotti *et al.*, 2024). Infine, c'è anche una differenza sociale e culturale: il modello delle SC era conosciuto solo nella comunità ristretta dei policy maker e dei tecnici mentre dell'IA c'è maggiore consapevolezza e accettazione, spesso anche in assenza di effettiva conoscenza dei suoi meccanismi (Cugurullo *et al.*, 2024).

Un ultimo elemento di riflessione riguarda le questioni politiche legate all'IA urbana e le corrispondenti agende e strategie dei governi locali. Da un lato esse corrispondono a quelle che, come abbiamo visto, caratterizzano arene nazionali e internazionali. Da un altro lato presentano alcune specificità.

Nelle agende degli Stati nazionali e dell'UE è presente la finalità di contribuire con politiche pubbliche a creare ambienti regolativi e stimolare, anche con spesa pubblica, gli investimenti e l'innovazione, con un'importante impronta competitiva: far crescere le rispettive IA nazionali come componente strategica della solidità dei rispettivi sistemi economici e della loro competitività nella geopolitica della tecnologia. Questo tipo di strategia, invece, è sostanzialmente assente dalle agende locali, se non per il coinvolgimento delle città nella localizzazione di centri in cui si svolgono le funzioni di ricerca e sviluppo dell'IA. Sono invece presenti gli altri due tipi di questioni, cui corrispondono strategie elaborate e perseguitate anche in sede locale: da un lato l'utilizzazione pubblica degli algoritmi, dall'altro la regolazione di questi usi e, come vedremo nel prossimo paragrafo, in determinate circostanze, anche degli usi privati.

Gli usi dell'IA nelle politiche e nell'azione amministrativa su scala locale e urbana suscitano alcune questioni, relative alla loro efficacia, ai rapporti fra poteri pubblici, privati e con i portatori di conoscenze scientifiche e, infine, alle forme e alle sedi, nelle quali queste utilizzazioni vengono discusse, messe a punto e decise.

Consideriamo l'efficacia. L'IA è uno strumento di policy, particolarmente raffinato e potente ma che con gli altri strumenti condivide potenzialità di successo e rischi di fallimento. Entrambi dipendono da vari fattori. Uno di questi è il modo in cui vengono individuati e selezionati gli strumenti specifici, in questo caso modelli e sistemi di IA. Come abbiamo visto, essi possono essere progettati "su misura" per una specifica città e i suoi problemi di funzionamento e sviluppo o, invece, essere offerti secondo una logica *one-size-fits all* che abbatte i costi per i fornitori di tecnologia e per i consulenti, i quali possono adattare e riproporre modelli e sistemi a nuovi clienti, ma rischia di non proporre i sistemi migliori per ciascuna specifica circostanza.

Questo problema può essere considerato anche nella chiave del rapporto

fra poteri pubblici di governo delle città e ideazione, progettazione e fornitura di modelli e sistemi tecnologici. Come vedremo nel paragrafo 4.3., esistono strade alternative: a rifornire le amministrazioni locali possono essere fornitori privati, in particolare grandi imprese tecnologiche che agiscono su scala sovralocale, oppure centri di ricerca pubblica localizzati sul territorio o anche start-up ad essi collegate sulla base di finanziamenti pubblici. L'IA urbana può infatti amplificare il fenomeno dell'intrusione delle aziende tecnologiche private nella governance delle città già introdotto dalle SC, poiché queste imprese non solo dispongono degli specialisti che dovrebbero gestire i sistemi di IA, ma sovente controllano anche i Big Data necessari. Ciò "rafforza le asimmetrie di potere, creando una situazione in cui le imprese di IA conoscono una città e i suoi abitanti meglio del governo e dei cittadini stessi" e se "il diritto dei cittadini alla SC è in gran parte assente, e ciò è ancora vero nella post-SC mediata dalle IA" (Cugurullo *et al.*, 2024b, p. 371).

Sia le implicazioni etiche, che affronteremo nel prossimo paragrafo, sia le finalità specifiche di uso, strategiche od operative, possono essere oggetto di discussione e di decisioni in sede ristrette ed escludenti, oppure più ampie ed inclusive. La forte caratura tecnica di modelli e sistemi che devono essere oggetto di scelta pubblica può facilmente indurre a discutere e individuare obiettivi e prendere le relative decisioni in tavoli riservati ad accordi fra amministrazioni e fornitori delle tecnologie. All'opposto, non diversamente da quanto accade per altre scelte di governo locale, le finalità e le modalità di utilizzazione dell'IA urbana possono essere discusse e scelte coinvolgendo i *policy taker* delle politiche urbane di volta in volta interessate che abbiamo visto prima. L'uso dell'IA per finalità partecipative può aiutare a democratizzare le scelte relative a ciascuno di questi impieghi, oltre ad essere a sua volta, oggetto di scelte condivise. Nessuno di questi momenti – sedi tecniche, amministrative e partecipative – esclude l'altro. Tuttavia, la prevalenza di ciascuna di esse nei processi in cui vengono scelti e gestiti i sistemi di IA per il governo delle città può a sua volta dipendere da indirizzi più generali della politica locale, in particolare più *market-oriented*, oppure più *social-oriented*, per usare una distinzione già proposta nei capitoli precedenti. Aspetti di questo tipo erano già emersi nella messa a punto e nei processi di impianto del paradigma SC. Alternative simili, emergono con riguardo al contributo che anche le tecnologie dell'IA possono dare a una governance urbana orientata a valori e interessi privati, oppure collettivi e pubblici, con le numerose possibilità di combinazione e mediazione fra queste due polarità. Quanto più ci si trova in prossimità della prima, tanto maggiori sono le probabilità che si affermino forme di tecnicizzazione, depoliticizzazione del governo urbano e di concentrazione tecnocratica dei poteri. Nelle funzioni di tipo strategico, l'IA urbana può generare depoliticizzazione dell'azione di governo in gradi

e per motivi diversi: in misura maggiore se le viene richiesto di produrre visioni del futuro delle città e formulare politiche pubbliche con elevata autonomia dagli attori umani. In misura minore, ma comunque rilevante, se la governance urbana degli umani diventa “proattiva” perché si può basare su previsioni relative a un futuro “calcolato e reso visibile in termini matematici” e, quindi, su “verità indiscutibili” (Cugurullo e Xu, 2025).

All’opposto, una politica locale ispirata a interessi collettivi e sostenibilità dello sviluppo urbano si associa più verosimilmente alla localizzazione e alla democratizzazione delle scelte relative alle tecnologie e del governo dei sistemi sociotecnici urbani. Le stesse preoccupazioni che può destare lo sviluppo e l’utilizzazione di sistemi di IA capaci di autonomia decisionale possono essere ridotte se si mettono in opera procedure di democratizzazione che, a loro volta, possono essere efficaci solo in presenza di trasparenza, spiegabilità e comprensibilità di modelli e sistemi¹².

4.2. Governare i *trade-off* dell’intelligenza artificiale urbana

Come abbiamo visto, l’IA può essere integrata nelle SC e contribuire tecnicamente ad ottimizzarne le prestazioni, a condizione che siano disponibili le infrastrutture, capacità e linee guida chiare che possono provenire sia da strategie nazionali per l’IA, sia da ecosistemi collaborativi locali di attori pubblici e privati. Un’altra condizione che emerge dalla discussione scientifica e dalle esperienze locali realizzate è la presenza, o lo sviluppo di strumenti politici e legali per governare le esternalità e le implicazioni critiche dell’uso di sistemi di IA per le dimensioni etica e dei diritti, sociali e individuali (Alsabt *et al.*, 2024). Infatti, anche se le applicazioni di IA urbana possono avere un impatto maggiore o minore in termini di supporto, assistenza, previsione o sostituzione delle decisioni umane, potenzialmente nessuna è neutra o innocua (Saetra, 2021). Si delinea quindi una specifica sfida regolativa, che riguarda principalmente le utilizzazioni di sistemi di IA nelle politiche e nelle amministrazioni pubbliche locali, poiché in molti ordinamenti

¹² Per trasparenza dei modelli e sistemi di IA si intende l’accessibilità alla struttura interna del modello, una componente più tecnica. Per spiegabilità si intende che è possibile capire perché un sistema IA ha prodotto un certo *output*, ossia quali dati e quali regole o correlazioni sono stati usati, quali variabili hanno avuto più peso. Secondo le linee guida etiche del 2019 del Gruppo di esperti di alto livello nominato dalla Commissione dell’UE la spiegabilità riguardava “la capacità di spiegare sia i processi tecnici di un sistema di IA, sia le relative decisioni umane” (<https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>). Per comprensibilità si intende il grado in cui queste spiegazioni sono accessibili a utenti non tecnici (ad esempio cittadini, politici). Invece, molti modelli di IA, in particolare quelli basati su deep learning, sono come delle scatole nere, poiché molto complessi e opachi.

istituzionali i margini di intervento politico su scala locale per regolare l'uso di sistemi di IA da parte di attori privati sono assai ridotti.

Le prove più impegnative per la governance dell'IA urbana vanno oltre quelle già sperimentate della governance delle SC, soprattutto perché l'IA introduce, almeno potenzialmente, un “agente” nuovo nei processi di governo delle città. Rispetto alle precedenti tecnologie delle SC, essa ha infatti maggiore capacità di “pensare”, prendere decisioni (Cugurullo *et al.*, 2023) ed esercitare autonomia operativa. Le implicazioni etiche che ne derivano sono più rilevanti, quanto più ampi e diversificati gli ambiti di azione pubblica in cui avviene l'integrazione del digitale con le economie e la vita quotidiana delle città (Caprotti *et al.*, 2024): potenzialmente una miriade di funzioni urbane.

Non diversamente da quanto avviene su scala nazionale – o, nell'UE, sopra-nazionale – anche per i governi locali si tratta di bilanciare benefici e rischi derivanti dall'utilizzo dei sistemi di IA, gestendo dei compromessi tra i vantaggi dell'IA e le esternalità negative (Strover *et al.*, 2021), “massimizzando i risultati desiderati e gli impatti positivi per tutti e minimizzando quelli indesiderati” (Yigitcanlar *et al.*, 2021), anche se ciò può implicare la necessità di privilegiare dei valori rispetto ad altri (Brom, 2021)¹³.

I rischi pubblici dell'IA urbana sono diversi, in buona parte coincidenti con quelli che sono entrati nelle agende della politica dell'IA su scale diverse da quella locale. In primo luogo, possono riguardare l'inaffidabilità o la fragilità tecnica dei sistemi di IA stessi e derivano dalla qualità inadeguata dei dati, da imperfezioni nella definizione del problema, dall'effettivo grado di conoscenza del contesto da parte del fornitore, o dall'adattamento imperfetto del modello alle specifiche circostanze locali. Questi fattori possono provare errori (Wang *et al.*, 2020; Schachtner, 2021) se non addirittura fallimento dell'analisi predittiva e dei conseguenti processi decisionali algoritmici. Peraltro, più le previsioni dell'IA si estendono nel futuro, meno diventano statisticamente affidabili (Cugurullo *et al.* 2024). In secondo luogo, i rischi riguardano la trasparenza e spiegabilità degli algoritmi, oltre al grado di controllo umano dei cambiamenti che possono intervenire nei modelli durante il loro uso, unitamente alla relativa responsabilità professionale,

In terzo luogo, i rischi che l'IA urbana condivide con altri usi pubblici e privati riguardano possibili violazioni dei diritti dei cittadini, come la privacy e la sicurezza dei dati utilizzati per far funzionare modelli e sistemi, oltre al

¹³ Per valutare i possibili impatti etici dell'IA urbana durante il ciclo di vita delle iniziative SC, Bosco *et al.* (2024) hanno adottato una scala a quattro livelli di impatto (molto elevato, elevato, medio e basso) applicandola a una serie di modelli di IA e ricavandone un modello descrittivo offerto a esperti tecnici, policy maker e al pubblico per discutere le considerazioni etiche utilizzando il linguaggio comune e dei criteri definiti.

possibile approfondimento delle disuguaglianze sociali e delle discriminazioni nell'ambiente urbano (Ye 2023; Popelka *et al.*, 2023). Questo tipo di rischi può derivare da vari fattori, come il *digital divide*, la riproduzione di pregiudizi sociali e culturali dei progettisti (McQuillan, 2022), o di quelli inscritti nei dati sui quali i sistemi di IA sono stati addestrati o che vengono comunque usati per le elaborazioni e dalla conseguente discriminazione, ad esempio, nella fornitura di servizi pubblici (Herat e Mittal, 2022). Nel campo della sicurezza, le piattaforme delle IA urbane impiegate per prevedere i luoghi e gli orari delle attività criminali e identificare i soggetti che potrebbero essere autori o vittime di reati sono state accusate di riprodurre le ingiustizie sociali, discriminando le minoranze razziali e gli abitanti delle città socialmente svantaggiati. Ciò sarebbe dovuto all'interazione fra i sistemi di IA e la conformazione dell'apparato sociotecnico e istituzionale della polizia urbana (Shapiro, 2024).

Tuttavia, i rischi dell'applicazione dell'IA su scala urbana non riguardano solo le possibili conseguenze del suo uso ma includono questioni relative alle relazioni di potere tra pubblico e privato. Come accade nelle SC, la tecnicizzazione dei processi di governo locale potenziata dall'IA può favorire l'opacità nei processi decisionali e limitare la partecipazione e l'attenzione del pubblico (Strover *et al.*, 2021), tanto meno in forma deliberativa (Cugurullo *et al.*, 2024). In altri termini, da un lato c'è il rischio che l'uso di IA nelle città venga definito all'interno di spazi decisionali esclusivi e inaccessibili alla società civile, lasciando il controllo delle politiche pubbliche ai fornitori di tecnologia. In queste condizioni, gli stessi policy maker urbani possono non sapere come l'IA (specialmente basata su LLM) stia generando previsioni del futuro, ma la usano lo stesso, non diversamente da come gli antichi facevano con i responsi e le profezie degli oracoli (Cugurullo *et al.*, 2025).

Come vengono affrontate queste sfide? Vediamo in che modo, sinora, i diversi contesti, soprattutto istituzionali, stanno circoscrivendo il campo delle azioni regolative, quali sono state le finalità prioritarie e gli strumenti utilizzati.

Queste iniziative possono essere adottate in situazioni politiche e istituzionali diverse, che danno ai governi locali margini di autonomia differenziati. Infatti, essi possono agire applicando regolamentazioni dell'IA provenienti da livelli istituzionali superiori (statale, o dell'UE, per i 27 Paesi membri) e ciò può loro consentire dei semplici allineamenti, se perseguono una strategia di attuazione lineare, oppure può portarli ad entrare in tensione con queste discipline in termini sia politici, sia strettamente giuridici. Invece, l'assenza di norme di livello superiore, oppure l'attribuzione, o la delega di funzioni decisionali ad enti di governo sub-nazionali lascia maggiori margini di libertà per l'azione regolativa locale. Più generalmente, diversi ordina-

menti costituzionali possono concedere gradi di autonomia differenziati ai governi di scala urbana, metropolitana, regionale, o di stato federale, che questi possono usare per adottare proprie politiche dell'IA. A sua volta, l'UE introduce un livello ulteriore di potere politico rilevante ai nostri fini, poiché conduce politiche regionali e urbane e, in alcuni casi, si trova a interagire direttamente con i governi locali, in arene politiche multiscalari più complesse.

Laddove vi sono margini minori per l'autogoverno nell'uso delle tecnologie, la stessa agenda delle SC ha seguito spesso un approccio dall'alto verso il basso (Castillo *et al.*, 2022), con strategie che implementano nelle aree urbane sistemi incentrati sui dati seguendo un approccio (principi e strumenti) che ruota attorno a indirizzi politici nazionali (Brandusescu e Reia, 2022).

Ciò nonostante, varie città dell'UE adottano in autonomia vari strumenti di regolazione per un'IA urbana etica. Alcune amministrazioni hanno esplicitamente inserito il rispetto di norme del GDPR nelle procedure di gara per i sistemi di IA, oppure l'adozione di software open-source, come Barcellona, ma anche Boston negli Usa (Beroche *et al.*, 2022; Castillo *et al.*, 2022). Particolaramente interessante è lo strumento dei “registri dell'IA”, adottati ad Amsterdam¹⁴ e Helsinki¹⁵, o nei *Metropolitan Data Charters* di Nantes e Barcellona, nell'ambito di politiche dell'IA urbana che mirano a garantire la privacy. I registri sono siti web che rendono accessibile un archivio di algoritmi usati in una specifica area geografica, inclusa una descrizione limitata di come e perché vengono utilizzati. Questi registri sono alla base di uno *Algorithmic Transparency Standard* sviluppato dal *Digital Forum Lab* di Eurocities in collaborazione con nove città per rendere più facile ai governi locali “fornire informazioni chiare sugli strumenti algoritmici che utilizzano

¹⁴ Due esempi: Confronto automatico dei volti nel processo di registrazione RNI, algoritmo che può aiutare i registranti a evitare le “frodi basate sulla somiglianza”; Blurring as a Service, Servizio che consente di rendere anonime le immagini provenienti da spazi pubblici rimuovendo persone e targhe, <https://algoritmes.overheid.nl/en/organisatie/gemeente-amsterdam/gm0363>. Nel 2017 era stato lanciato il manifesto Tada, che ha condiviso i valori per una città digitale responsabile (<https://openresearch.amsterdam/en/page/110506/tada.city>).

¹⁵ Secondo l'amministrazione della città di Helsinki, il registro dell'IA è uno strumento standardizzato, consultabile e archiviabile che consente di documentare le decisioni e le ipotesi formulate nel corso del processo di sviluppo, implementazione, gestione e, infine, smantellamento di un algoritmo. In questo modo è possibile garantire trasparenza e, ove applicabile, spiegabilità ai fini del dibattito pubblico, dei revisori indipendenti e dei singoli cittadini. Per la società civile, è una finestra sui sistemi di IA utilizzati da un'organizzazione governativa. In definitiva, speriamo che diventi un catalizzatore per una partecipazione democratica significativa e una piattaforma per promuovere la fiducia reciproca (<https://ai.hel.fi/wp-content/uploads/White-Paper.pdf>).

e sui motivi per cui li utilizzano”¹⁶. Queste esplicitano il tipo di algoritmo, le finalità di uso, l’unità organizzativa che lo usa, l’area geografica e l’ambito di politiche urbane cui si riferisce, una categoria di rischio, la fonte dei dati usati e i dati di addestramento, le eventuali distorsioni e le corrispondenti misure di mitigazione, nonché dei dettagli circa la supervisione umana¹⁷. Gli strumenti utilizzati in queste esperienze sono da considerare ancora sperimentali. Tuttavia, sono stati sottoposti ad alcune critiche, in quanto considerati perlomeno parziali, e a rischio di idealizzare delle soluzioni di governance e normalizzare l’uso dell’IA nelle città. Questo perché escluderebbero dal novero di quelle rese pubbliche e spiegate le applicazioni più sensibili, consolidando le informazioni esistenti e quindi il potere e le disuguaglianze tra i governi e gli abitanti delle città. Perciò, secondo questi osservatori (Cath e Jansen, 2021), rischierebbero di ridursi a un mero “teatro etico” eludendo una domanda di fondo: “i governi, in questo caso i comuni, dovrebbero davvero implementare sistemi di IA?”.

Gli usi più controversi e oggetto di tensioni regolative riguardano spesso le politiche per la sicurezza urbana. Una nostra ricerca condotta con T. Faschiani su un caso di impiego locale di IA a fini di sicurezza a Torino ha evidenziato che in Europa e in Italia anche prima dell’entrata in vigore dell’EU AI Act non esisteva uno spazio per la gestione dei *trade-off* degli algoritmi nell’ambito di regolazioni locali. Nel caso considerato l’ambizione dell’amministrazione locale era bilanciare l’efficacia nel campo della sicurezza urbana degli algoritmi per la visione delle immagini e tutelare la privacy dei cittadini. Le norme di *hard law* in vigore attribuivano infatti al Garante per la protezione dei dati personali una capacità di intervento prevalente nei confronti delle autorità che si occupano di sicurezza ovvero la polizia nazionale e locale (d’Albergo *et al.*, 2023). In questo modo sono stati garantiti anche diritti fondamentali tutelati dalla Costituzione italiana (Bergonzini, 2024).

Sempre per quanto riguarda l’uso locale di IA a fini di sicurezza, le iniziative adottate in un contesto particolare, quello cinese, sono state oggetto di valutazioni che evidenziano la stretta relazione tra IA urbana e implementazione di politiche nazionali. Da un lato, si sottolinea che nelle aree urbane cinesi gestite da *city brain*, la tecnologia dell’IA è diventata uno strumento diffuso di governance centralizzata (ad opera del governo nazionale) delle città (Cugurullo *et al.*, 2024b). Dall’altro, studiando il ruolo svolto dagli immaginari

¹⁶ Eurocities è una rete di oltre 200 città europee in 38 Paesi. Gli standard sono stati sviluppati in collaborazione con Amsterdam, Barcellona, Bruxelles, Eindhoven, Mannheim, Rotterdam e Sofia (<https://www.algorithmregister.org>).

¹⁷ Wray S., 19 01 2023, *Standard launched to help cities open up about algorithms*, CitiesToday (<https://cities-today.com/standard-launched-to-help-cities-open-up-about-algorithms/>).

dell'IA, ossia la concezione condivisa circa un ordine sociale desiderabile da mantenere, o sostenere, attraverso le innovazioni tecnologiche dell'IA e le loro applicazioni sociali, sono state messe in evidenza forme diverse di centralizzazione. Data la loro eterogeneità spaziale e grazie all'elevata dotazione di infrastrutture tecnologiche, le città cinesi offrono terreno fertile per sperimentare e replicare gli immaginari dell'IA urbana definiti a livello nazionale. Il controllo dall'alto è un elemento centrale nella formulazione e implementazione di un paradigma di città ad alta tecnologia, intelligente, a basse emissioni di carbonio e innovativa. A questo fine, però, il governo non introduce misure autoritative. Da un lato, promuove consenso attraverso una narrazione secondo la quale un ambiente urbano permeato dall'IA garantisce la sicurezza personale e servizi pubblici mirati. Dall'altro, adotta misure incentivanti e restrittive per favorire *compliance* da parte degli attori che mettono in opera progetti di IA urbana, pubblici e privati. Le città sono così un campo sperimentale socio-tecnico per realizzare gli immaginari nazionali dell'IA (Chen, 2024).

Soprattutto negli Usa, invece, nelle arene locali vi sono margini maggiori per l'autogoverno nell'uso delle tecnologie, compresa l'IA. Quello che è stato definito *AI localism*¹⁸ – ossia le azioni intraprese da decisori locali per governare l'uso dell'IA e le sue implicazioni etiche e sociali in una città – è un fenomeno prevalentemente statunitense, emerso a causa delle lacune lasciate da quadri di governance statali¹⁹ o nazionali incompleti²⁰. Uno spazio che consentirebbe agli attori politici di comprendere meglio, da una prospettiva di prossimità, i *trade-off* dell'IA urbana e quindi colmare le lacune della legislazione federale e statale, in particolare quando si tratta di sorveglianza (Verhulst e Sloane, 2020; Verhulst, 2022). In questa prospettiva, i governi locali hanno anche l'opportunità di sondare il terreno e gettare le basi per una più ampia governance pubblica dell'IA, funzionando come laboratori sperimentali (Matibag, 2020) per il raggiungimento di risultati in termini di IA responsabili (Yigitcanlar *et al.*, 2021). Non a caso dei 113 casi di *AI localism*

¹⁸ Ad esempio nell'ambito della *New York City Artificial Intelligence Strategy* adottata nel 2021, la città si propone di “svolgere un ruolo politico importante nel promuovere approcci positivi all'IA nel settore privato locale e nel proteggere i newyorkesi nelle loro interazioni con essa”, secondo principi di responsabilità, non discriminazione, privacy e sicurezza, impegno e partecipazione della comunità (<https://www.nyc.gov/assets/cto/downloads/ai-strategy/nycaistrategy.pdf>).

¹⁹ Anche se sono segnalate iniziative di regolazione dell'IA su scala statale, come in California, dove a una legge approvata dal Senato e che richiedeva agli sviluppatori di modelli di IA di grandi dimensioni e a coloro che forniscono la potenza di calcolo necessaria per addestrare tali modelli di mettere in atto misure di salvaguardia e politiche per prevenire danni catastrofici è stato opposto il voto del Governatore. (<https://leginfo.legislature.ca.gov/faces/billStatusClient.xhtml?billid=202320240SB1047>).

²⁰ Fonte: <https://ailocalism.org/#about>.

mappati²¹ 79 si trovano nel Nord America. Inoltre, nel 2024 in California è stata costituita dal governo locale di San Jose (la “capitale della Silicon Valley”) una *GovAI Coalition* composta da otto agenzie locali, statali e federali, per promuovere un’IA responsabile e mirata nel settore pubblico. Questi casi statunitensi riguardano spesso la regolazione dell’uso di IA per finalità di sicurezza, a partire dai software di riconoscimento facciale, anche da parte delle forze dell’ordine (Coglianese e Lai, 2022). Una regolazione locale che può essere anche molto stringente come avvenuto nel caso di San Francisco dove, nel 2019, queste tecnologie sono state vietate per le preoccupazioni espresse dai gruppi per la difesa delle libertà civili circa i possibili abusi nel loro uso da parte del governo²².

Questi tentativi degli attori locali di «governare gli algoritmi, mentre governano con gli algoritmi» (Kuziemsky e Misuraca, 2020) hanno già dato luogo a diversi studi. Come nei casi relativi alla *regolazione* dell’IA su scale non locali, le analisi realizzate e in corso hanno finalità sia conoscitive, sia pratiche, orientate cioè a fornire agli attori politici e amministrativi strumenti di comprensione della complessità e di intervento (Yigitcanlar *et al.*, 2024). A questo proposito, lo sviluppo di attività autoregolative dei governi locali negli Usa non è passato inosservato nemmeno agli attori del business della consulenza. Ad esempio, *Madison AI* si offre negli Usa come un “assistente per la conoscenza per i governi locali” che attraverso l’IA può non solo rintracciare informazioni nel database di documenti di un ente, una funzione che può essere usata dai rappresentanti politici eletti nelle discussioni relative al governo della città e dai funzionari come strumento di ricerca specialistico, ma anche fornire indicazioni su come creare principi guida per l’uso di IA, prendendo spunto anche dall’EU AI Act²³.

È quindi opportuno chiedersi quali principi ispirano le finalità degli interventi regolativi locali? A questo proposito, è particolarmente utile l’iniziativa del *Global Observatory of Urban Artificial Intelligence* (GOUAI), lanciata nel 2021 dal Centro di Barcellona di affari internazionali²⁴ e dai governi locali di Barcellona, Amsterdam e Londra, con il sostegno di UN-Habitat nell’ambito della *Cities Coalition for Digital Rights*²⁵. Quest’ultima svolge il tipico ruolo delle reti di città, rendendo possibile per i membri imparare

²¹ Fonte: <https://list.ailocalism.org>

²² Fonte:

<https://www.nytimes.com/2019/05/14/us/facial-recognition-ban-san-francisco.html>

²³ Fonte: <https://madisonai.com/resources/ai-governance-blueprint/>

²⁴ Fonte: <https://www.cidob.org/en>

²⁵ Ne fanno parte 60 governi locali, fra cui 10 capitali europee:

<https://citiesfordigitalrights.org>. L’Osservatorio rende disponibile un atlante dell’IA urbana ed effettua un monitoraggio delle tendenze nella sua implementazione.

dai colleghi di altre città e dialogare attivamente con loro. In questo modo, come ha sostenuto il *London's Chief digital officer*, “non abbiamo alcun potere legislativo in questo ambito, ma abbiamo il potere di influenzare”²⁶. Il GOUAI è uno degli strumenti “operativi” della Coalizione che ha elaborato una classificazione dei sistemi di IA secondo i criteri etici considerati dai governi che li adottano per la formulazione e gestione delle loro politiche. In questo contesto i diritti umani sono intesi non come nuovi, ma come diritti esistenti e che devono essere protetti nell’uso delle tecnologie digitali. L’inventario contiene sinora 218 iniziative di 74 città che, in coerenza con i valori ispiratori della Coalizione, sono allineate con i principi di trasparenza e apertura, equità e non discriminazione, sicurezza e *cybersecurity*, protezione della *privacy*, sostenibilità e *accountability*²⁷. Le 218 iniziative sono classificate nel database secondo la presenza nei progetti dei diversi principi (grafico 2), tra i quali svettano la protezione della *privacy* e, soprattutto, quelli relativi alla trasparenza e apertura.

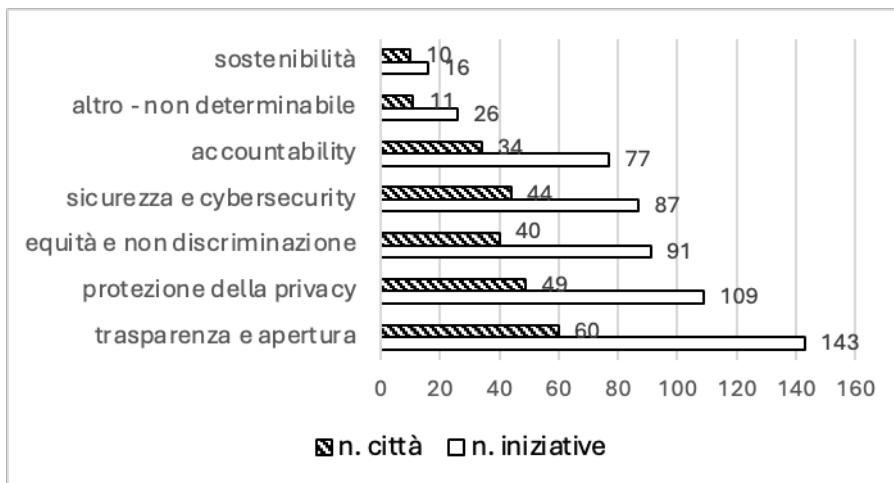


Grafico 2 - Iniziative di uso dell’IA nelle città secondo i principi etici considerati, censite dal Global Observatory of Urban Artificial Intelligence, elaborazione degli autori.

²⁶ Fonte: [https://www.computerweekly.com/ezine/Computer-Weekly/The-cities-planning-for-ethical-use-of-AI\).](https://www.computerweekly.com/ezine/Computer-Weekly/The-cities-planning-for-ethical-use-of-AI).)

²⁷. Sono censite (al 2 maggio 2025), rispettivamente, iniziative relative a uno o più dei seguenti principi: equità e non discriminazione: 91 iniziative in 40 città; trasparenza e apertura: 143 in 60 città; sicurezza e cybersecurity: 87 in 44 città, protezione della privacy: 109 iniziative in 49 città; sostenibilità: 16 in 10 città; accountability: 77 in 34 città (26 iniziative in 11 città non sono esplicitamente ispirate a uno di questi principi). Le città con il maggior numero di iniziative sono Amsterdam (22), Helsinki (19), Barcellona (12), New York e Seoul (10), San Jose e Singapore (8), Dubai (7) (<https://gouai.cidob.org/atlas/>).

Trasparenza, spiegabilità del funzionamento dei sistemi di IA e coinvolgimento del pubblico nel loro design e nella progettazione del loro uso²⁸ sono peraltro considerati accorgimenti per gestire i dilemmi relativi al come individuare i limiti alle decisioni che l'IA può prendere da sola, ad esempio, quando le si affida il compito di selezionare i destinatari di benefici pubblici. Anche i rischi per la qualità democratica del governo locale e l'individuazione di chi debba essere considerato responsabile, quando la tecnologia compie errori, sono considerati suscettibili di riduzione attraverso la possibilità di rendere pubblico e spiegare il funzionamento dei sistemi di IA.

Per quanto riguarda gli strumenti, le amministrazioni utilizzano un ampio repertorio di *soft law*, che comprende: istituzione di *sandbox normative*, redazione di manuali e adozione di standard per un uso dell'IA urbana che tenga conto di rischi e pericoli, indicazione di modalità di progettazione dei sistemi di IA e loro pubblicità, clausole inserite nel *public procurement* (appalti per l'acquisto di sistemi di IA); trasparenza e coinvolgimento del pubblico nel discuterne e deciderne l'uso, iniziative per istruire cittadini, residenti e policy maker sullo sviluppo, l'uso e i possibili impatti di sistemi di IA (Marcucci *et al.*, 2022; Verhulst, 2022b). Vediamone alcuni con maggiore dettaglio.

Le iniziative con le quali vengono introdotte delle regole sotto forma di norme contestualmente all'adozione di sistemi di IA possono essere intese come pratiche regolative “dal basso” che, mentre persegono obiettivi di politiche urbane, istituiscono anche una sorta di “*sandbox normative*” locali per l'IA (Strover *et al.*, 2021)²⁹. Ad esempio, numerose città hanno adottato forme sperimentali di regolazione per quanto riguarda l'uso di robot³⁰ come veicoli senza conducente, tecnologie impiegate nella logistica e consegna personalizzate di vari tipi di merce (a partire da Amazon), polizia su larga scala, o macchine per la sorveglianza. Negli Usa, stati e comuni hanno autorizzato sperimentazioni su singole strade e marciapiedi, cercando di prevenire danni alle persone e alle cose. Da un lato, proteggendo i pedoni negli spazi pubblici, da un altro restringendo le possibilità di intervento a distanza con robot armati, reso possibile da tecnologie sviluppate in ambito militare. A San Francisco le imprese tecnologiche, a partire dalle Big Tech della Silicon Valley, hanno chiesto di sperimentare l'uso dei robot in città, ricevendo una risposta politica orientata alla “governance collaborativa”, con l'invito delle autorità alla partecipazione di cittadini e associazioni in un *Emerging*

²⁸ Fonte: <https://gouai.cidob.org/#documentary>

²⁹ Uno strumento previsto anche dalla regolazione “dura” come abbiamo visto nel caso dell'EU AI Act.

³⁰ Per robot si intendono macchine che interagiscono con il mondo fisico tramite sensori e dispositivi attuatori, con un'autonomia sempre maggiore dal controllo umano grazie all'IA.

Technology Open Working Group per definire i problemi e le regole da introdurre. Questa consultazione ha prodotto, fra l’altro, il divieto per la polizia di usare robot armati per fini di sicurezza (While, 2024). Questi esperimenti urbani (in particolare per i veicoli a guida autonoma) sono stati considerati «azioni politiche di definizione e creazione di nicchie, modificando sia le norme sociali, sia gli spazi fisici in cui nuove pratiche sociali (possono) emergere e svilupparsi» (Savini e Bertolini, 2019, p. 832).

Anche l’uso dei dati può essere oggetto di norme o accordi per la condizione e il controllo collettivo su scala locale (Paltiel, 2023)³¹

Piuttosto diffusa è la pratica dei governi locali di adottare, individualmente o in associazione fra loro, manuali e standard per l’uso etico di IA urbana. Un esempio è l’*AI Governance Handbook* della *GovAI Coalition*, che fornisce indicazioni pratiche su come affrontare i danni causati da un incidente provocato da IA³² e, più generalmente, su come muovere i primi passi per realizzare pratiche di IA responsabili³³. Il manuale indica ad esempio in quali casi è necessario chiedere ai fornitori di IA l’attivazione di procedure di controllo. I sistemi sono divisi in livelli di rischio, comparabili con quelli dell’EU AI Act che riguardano, in particolare, sistemi di polizia predittiva, sistemi di decisione automatizzata e le modalità per effettuare tale controllo: analisi e valutazione di rischi e benefici, coinvolgimento online dei cittadini con sessioni su piattaforme social come Facebook, Instagram o X in cui esperti discutono di IA e rispondono alle domande del pubblico, o di persona in riunioni pubbliche.

Altre pratiche riguardano l’elaborazione di linee guida “locali” per l’introduzione dell’IA nella pubblica amministrazione. Ad esempio, la città di Boston ha adottato delle *Guidelines for Using Generative AI* (2023) che indicano ai dipartimenti amministrativi cosa fare e cosa non fare: è necessario dichiarare di aver utilizzato l’IA per generare dei contenuti, non condividere informazioni sensibili o private nei *prompt*, redigere i contenuti con un linguaggio semplice, rivedere i testi prodotti per assicurarsi che il linguaggio sia inclusivo e rispettoso, non affidarsi all’IA generativa per ottenere risposte accurate, assicurarsi che i risultati non siano offensivi o dannosi nei confronti delle persone, in particolare dei residenti vulnerabili che sono suscettibili di

³¹ Ad esempio, l’obbligo di rendere pubblici i dati sul consumo di energia da parte degli edifici. Fonte: <https://urbanai.fr/our-works/urban-ai-report/> (p. 42).

³² Fonte: <https://www.sanjoseca.gov/home/showpublisheddocument/118645/63877-4853661770000>

³³ Fonte: <https://www.sanjoseca.gov/home/showpublisheddocument/118629/6387747-49167870000>

subire danni, inclusi gruppi etnici e razziali, individui di genere diverso³⁴. Fuori dagli Usa, linee guida di questo tipo sono state adottate dal *London Office of Technology and Innovation* per aiutare le autorità locali a utilizzare strumenti di IA generativa come ChatGPT, Bard, Midjourney e Dall-E.

Un’ulteriore pratica consiste nello scegliere modalità di progettazione dei sistemi di IA finalizzate a garantire forme di controllo democratico sull’uso di tecnologie invasive della privacy che usano dati personali, colmando i vuoti regolativi lasciati dai livelli istituzionali superiori (Verhulst e Sloane, 2020). L’analisi preliminare dei rischi relativi alla distribuzione spaziale e/o sociale dei costi e dei benefici (Yigitcanlar *et al.*, 2021) dovrebbe essere tradotta in requisiti di progettazione volti a bilanciare i costi e i benefici dell’IA. Anche a questo fine, accorgimenti specifici possono essere adottati nel *public procurement*, inserendo nei bandi di gara per l’acquisto di sistemi di IA il rispetto di clausole relative ai diritti digitali. In questo modo è possibile per i governi locali selezionare i fornitori che diverranno componenti degli ecosistemi urbani di innovazione digitale, in modo che dimostrino nelle loro attività di essere allineati rispetto a criteri di un’IA urbana etica. È possibile, inoltre, introdurre forme di partecipazione dei cittadini nei processi decisionali attraverso i quali le autorità politiche definiscono le modalità di uso dell’IA urbana. A questo proposito, è stata evidenziata la possibilità che pratiche partecipative sull’IA facciano emergere relazioni sociotecniche alternative e controegemoniche nella vita algoritmica delle città, sfidando così il dominio del settore privato nello sviluppo e nell’implementazione della tecnologia digitale (Lynch, 2020).

Come abbiamo visto, i governi locali provano a usare l’IA e, al tempo stesso, definire e adottare norme per allineare queste utilizzazioni con gli standard etici che nel frattempo prendono forma all’interno di arene di governance di questa tecnologia costituite da entità private, dall’interazione fra pubblico e privato e, nel caso dell’UE, da arene di legislazione pubblica. In questi sforzi il tempo è un fattore complessi. L’uso di questi margini di autonomia in presenza di tecnologie dall’evoluzione rapidissima richiede infatti ai *policy maker* locali di «correre il doppio di quanto normalmente fanno», come la regina ricorda ad Alice nel paese delle meraviglie (Marcucci *et al.*, 2022, p. 27). D’altra parte, i tempi necessari per sperimentare, sviluppare e correggere innovazioni urbane basate su IA, potrebbero essere più lunghi

³⁴ Fonte: <https://www.boston.gov/sites/default/files/file/2023/05/Guidelines-for-Using-Generative-AI-2023.pdf>. Altri casi di autoregolazione dell’uso di IA e, in particolare, di IA generativa da parte di governi locali negli Usa sono le città di Santa Cruz (California), Seattle (Washington), Boston (Massachusetts), Boise (Idaho), Washington County (Oregon), Baltimora (Maryland), Miami-Dade (Florida) e il governo metropolitano di Nashville e Davidson County (Tennessee) (<https://madisonai.com/resources/ai-governance-policy-examples/>).

degli orizzonti elettorali più ravvicinati delle autorità politiche locali (Cugurullo *et al.*, 2024b).

Vi sono inoltre rischi che queste politiche locali scarsamente coordinate – soprattutto in assenza di *frame* normativi come l’EU AI Act – producano esiti regolativi incompatibili fra loro (Verhulst, 2022). Per prevenire un esito di questo tipo può essere utile creare *framework* che aiutino i governi locali a realizzare un’innovazione responsabile attraverso pratiche di IA localizzate (Yigitcanlar *et al.*, 2021). Ciò può essere fatto non solo senza ricorrere necessariamente a regolazioni sovraordinate, ma anche tenendo conto del fatto che le differenze fra i contesti locali non sono solo di carattere istituzionale. Come abbiamo visto a partire dal *frame* ormai universale del bilanciamento dei *trade-off* fra potenzialità e rischi e, in alcune prospettive, fra innovazione e regolazione, lo sviluppo dell’IA tende ad essere un potente fattore di omologazione di linguaggi e di pratiche sociali. Può emergere un “universalismo tecnologico” standardizzato, il quale rischia di non tenere conto che nelle città si trovano concezioni della governance, gerarchie, disponibilità di risorse, eredità storiche e culture politiche diverse. Forme di IA urbana etiche possono essere perseguiti, in questa prospettiva, attraverso “un mosaico di impegni locali, negoziazioni culturali e innovazioni sensibili al contesto” (Tabrizian, 2025). Una valorizzazione di modelli e pratiche sociali e politiche di co-creazione potrebbe favorire una governance locale dell’IA culturalmente inclusiva e che cerca perlomeno di sottrarla a un altro rischio, quello della “cattura regolativa” indotta dalla presenza nelle arene dell’IA urbana di imprese tecnologiche private, come sono prevalentemente quelle che sviluppano sistemi di IA (Marx, 2022), capaci di controllare risorse cognitive e quindi di potere in misura maggiore dei partner pubblici.

4.3. La politica dei gemelli digitali urbani

Come abbiamo visto nei precedenti paragrafi, l’IA è entrata a pieno titolo nella platea di strumenti digitali che da quasi tre decenni sono applicati alla conduzione di politiche urbane e alla risoluzione di questioni di interesse collettivo nelle città. In questo paragrafo ci occupiamo di Gemelli Digitali Urbani, una delle più avanzate evoluzioni tecnologiche a disposizione delle amministrazioni locali per il governo e la gestione delle città, di cui espriamo la relazione con le politiche di digitalizzazione dell’UE, le implicazioni per i processi decisionali pubblici e le sfide, per la politica, della loro implementazione in contesti reali. Comprendere concettualmente cosa sia un Gemello Digitale – *Digital Twin* (anche urbano) – è in realtà piuttosto semplice: si tratta di una replica digitale di un elemento fisico basata sulla ripro-

duzione delle sue caratteristiche reali, che consente di simulare condizioni e situazioni in ambiente virtuale per poi intervenire su quello fisico. Questa intuizione geniale, e dall'elevata complessità ingegneristica, è da attribuire a M. Grieves, il quale nel 2002 ha presentato all'Università del Michigan il primo modello di replica virtuale di un oggetto fisico come un nuovo approccio per la gestione dinamica del ciclo di vita dei prodotti (Grieves e Vickers, 2016). La storia della tecnologia Digital Twin (da ora DT) inizia in ambito industriale, militare e nell'industria aerospaziale (Negri *et al.*, 2017; Wu *et al.*, 2020) e si estende rapidamente ad una molteplicità di settori e impieghi come, ad esempio, in ambito sanitario attraverso la creazione di repliche virtuali dei pazienti per la personalizzazione della cura (Sun *et al.*, 2023; Katsoulakis *et al.*, 2024). I DT più avanzati integrano oggi l'IA (DG Connect, 2023), perché gli algoritmi sono fondamentali per implementare, gestire e potenziare le simulazioni dei DT, specialmente quando operano su una importante mole di dati (Groshev *et al.*, 2021; Kreuzer *et al.*, 2024).

Noi ci occupiamo in particolare dell'applicazione della tecnologia DT in ambito urbano, dove è impiegata per realizzare versioni virtuali di infrastrutture, edifici, quartieri, reti di trasporto dell'energia e dell'acqua e di molte altre componenti e funzioni urbane. I primi progetti di DT urbani sono quindi iniziati con la creazione di copie di edifici ed infrastrutture fisiche di cui è più semplice replicare le caratteristiche strutturali, per poi integrare progressivamente informazioni sui flussi che le attraversano, come energia e traffico (Batty, 2018). Sebbene ne esistano diverse forme e versioni, i DT urbani si basano sempre su un sistema di infrastrutturazione tecnologica (sensori, connessioni Internet, IoT) che fornisce dei dati e una loro rappresentazione geografica ad un "cervello urbano", che li elabora e consente di condurre simulazioni attraverso una serie di applicazioni relative, ad esempio, alla gestione del traffico, all'inquinamento o allo smaltimento dei rifiuti (Deng *et al.*, 2021). Dunque, esistono applicazioni di DT urbani gestiti da soggetti privati nell'ambito della progettazione, costruzione e manutenzione di edifici e strutture urbane e progetti DT attivati dalle pubbliche amministrazioni per la gestione della pianificazione urbanistica e di diversi settori di politica urbana, funzioni oggetto di supporto tecnologico anche nelle SC, come abbiamo visto. La creazione di un DT di un'intera città è una sfida assai complessa proprio per la caratteristica distintiva di questa tecnologia: l'interazione in tempo reale tra componenti fisiche e replica digitale. I DT urbani, a nostro avviso, rappresentano una tecnologia di particolare interesse poiché si collocano tra i paradigmi delle SC e dell'IA urbana che abbiamo messo a confronto nel paragrafo precedente. Da un lato, i DT urbani non prevedono l'autonomia decisionale caratterizzante l'IA urbana; dall'altro lato, i DT consentono di impiegare i dati e le tecnologie delle SC a fini previsionali e di simulazione di scenari futuri.

Singapore è stata tra le prime città a dotarsi di una propria replica virtuale con un progetto avviato circa un decennio fa, denominato “Virtual Singapore”, gestito dall’ente governativo della *Singapore Land Authority*³⁵. Il progetto è particolarmente ambizioso perché riguarda una Città-Stato di cui è stato creato un DT basandosi su un complesso patrimonio dati su infrastrutture urbane, popolazione, fattori ambientali, da mettere a disposizione per la pianificazione urbanistica e la fornitura dei servizi dei cittadini. In Europa Helsinki è tra le città pioniere in questo ambito, perché può sfruttare la sua lunga tradizione nelle politiche SC e nella sperimentazione nel campo dei modelli 3D di città (Ruohomäki *et al.*, 2018). Un’eredità sia di tipo infrastrutturale, con la possibilità di valorizzare l’infrastrutturazione tecnologica e digitale del territorio cittadino, sia di consapevolezza politica circa le implicazioni dell’uso di questi sistemi. Come abbiamo visto, nel 2020 Helsinki è stata tra le prime città, insieme con Amsterdam, a dotarsi di un “registro degli algoritmi”³⁶ per condividere con cittadini e soggetti interessati gli ambiti, le finalità e le modalità di utilizzo dell’IA da parte dell’amministrazione locale.

La tecnologia dei DT urbani si è rapidamente diffusa a livello globale ed europeo, dove finanziamenti comunitari hanno consentito di sperimentarla in diversi contesti locali³⁷. Nelle politiche dell’UE i DT incrociano i temi e le questioni dello sviluppo e dell’uso dell’IA su scala urbana. La Commissione Europea li ha identificati come uno strumento decisivo per la digitalizzazione delle comunità locali e urbane, obiettivo del più ampio Programma Europa Digitale (DIGITAL)³⁸ della Commissione von der Leyen I (2019-2024). I discorsi della Commissione sulle politiche di digitalizzazione hanno più volte enfatizzato la necessità di acquisire controllo sull’ambiente digitale europeo, inteso come una forma di sovranità digitale (Floridi, 2020; Heidebrecht, 2024; Carver, 2024). Questo obiettivo, applicato alla diffusione dei DT urbani, implica da un lato la conoscenza dello “stato dell’arte” della maturità digitale dei diversi contesti urbani e dall’altro l’obiettivo di favorire la creazione di un sistema di fornitori di tecnologia, esperti e amministratori locali interessati e disposti a ridefinire le proprie politiche di governo delle città attraverso l’uso di questa tecnologia. Un obiettivo su cui torneremo più avanti. Nel 2023 la Commissione Europea censiva 135 progetti di DT urbani

³⁵ Per una ricostruzione del progetto si rimanda a: <https://www.sla.gov.sg/articles/press-releases/2014/virtual-singapore-a-3d-city-model-platform-for-knowledge-sharing-and-community-collaboration>

³⁶ Maggiori informazioni sono disponibili al link: <https://ai.hel.fi/en/ai-register/>

³⁷ Tra i primi progetti segnaliamo “*Digital Urban European Twins for smarter decision making – DUET*”, realizzato con un finanziamento Horizon 2020 nelle città di Atene (Grecia), Pilsen (Repubblica Ceca) e nella regione belga delle Fiandre.

³⁸ Per una ricostruzione degli obiettivi e degli strumenti del Programma DIGITAL si rimanda a: <https://digital-strategy.ec.europa.eu/it/activities/digital-programme>.

e locali³⁹ in 25 stati membri, con diversi stadi di sviluppo (DG Connect, 2023). La maggior parte dei DT erano impiegati nella pianificazione urbanistica e nella costruzione di nuove infrastrutture (52%), seguiti dalle applicazioni per la mobilità urbana (36,8%) e da un 32,8% di progetti riguardanti la gestione di politiche ambientali. Un numero residuale di applicazioni riguardava la gestione dei consumi energetici cittadini, delle risorse idriche e la logistica. Non da ultimo, i DT urbani possono essere impiegati in quello che la Commissione definisce “*community engagement*”, una questione su cui torneremo più avanti. Secondo questa mappatura, tutti i progetti vedevano le amministrazioni locali come principali beneficiarie delle elaborazioni dei DT, con il coinvolgimento sia di stakeholder privati, sia della cittadinanza. La mappatura ha evidenziato una diffusione capillare della tecnologia tra gli i Paesi membri dell’UE, sebbene questo censimento sia presumibilmente già invecchiato, considerato il susseguirsi di annunci delle amministrazioni locali che intendono dotarsi di un proprio DT⁴⁰.

Finora abbiamo accennato a cosa sono i DT urbani, ai possibili ambiti di utilizzo e a come essi siano diffusi a livello globale e comunitario. Tuttavia, possiamo porci una semplice domanda: perché un’amministrazione locale dovrebbe dotarsi di questo strumento? Da un lato, i DT e gli algoritmi di IA promettono ai policy maker di poter elaborare scenari alternativi, anticipare i cambiamenti strutturali dovuti all’attuazione delle politiche da adottare e di visualizzare in modo interattivo i risultati (Marcucci *et al.*, 2020; Ferré-Bigorra *et al.*, 2022). Dall’altro lato, la disponibilità di queste informazioni può anche migliorare la qualità della discussione pubblica, delle pratiche partecipative e incrementare il grado di fiducia nelle relazioni tra gli attori coinvolti (Nochta *et al.*, 2019). La possibilità di realizzare complesse elaborazioni di dati attraverso l’IA su un modello “reale” della città e dei suoi processi economici e sociali (come, ad esempio, il sistema della logistica) rappresenta indubbiamente un enorme potenziale per i decisori pubblici. Gli scenari elaborati dai DT consentirebbero di aumentare le basi cognitive necessarie al policy-making e di superare o mitigare il noto problema della “razionalità limitata” nei processi decisionali pubblici (Simon, 1957; Wildavsky, 1964; Jones, 2002). Secondo l’approccio della razionalità limitata, le decisioni politiche sono condizionate dall’accesso a informazioni incomplete e limitate e dalla impossibilità di valutare l’intero ventaglio delle diverse alternative decisionali. Queste limitazioni dipendono anche da una serie di

³⁹ La Commissione Europea adotta la definizione di Local Digital Twins (LDTs) per comprendere i DT applicati alla scala urbana e i progetti in villaggi e aree rurali.

⁴⁰ Tra i più recenti e a noi vicini, evidenziamo l’annuncio di Roma Capitale, che nel 2026 si dovrà di un DT per gestire politiche e progetti di rigenerazione urbana, fonte: <https://www.romatoday.it/politica/roma-mipim-cannes-2025-digital-twin.html>

vincoli riconducibili ai processi politici, alla mancanza di tempo e di competenze adeguate, e alle pressioni esterne esercitate da gruppi di interesse (Nelson, 2012). Parafrasando F. Baumgartner e B. D. Jones (2015) si potrebbe passare un’intera vita a valutare le possibili soluzioni ai problemi di policy più complessi senza agire mai, una condizione evidentemente non auspicabile in un processo decisionale pubblico e che viene sovente risolta attraverso l’“euristica della soddisfazione”, che consente di scegliere soluzioni soddisfacenti ma non ottimali (Shannon *et al.*, 2019). Con l’apporto conoscitivo dei DT, almeno a livello urbano, cambierà tutto? Se ci è ormai chiara la necessità per la politica di bilanciare potenzialità e rischi nei molteplici usi dell’IA, non possiamo abbandonarci a posizioni semplicisticamente “tecnocentusiaste” sui DT urbani. La loro applicazione in contesti urbani reali si scontra infatti con alcuni fattori che possono pregiudicarne l’efficacia e il potenziale trasformativo dei processi di governo delle città. Da un lato, l’implementazione di un progetto di DT urbano comporta una serie di questioni ingegneristiche e di progettazione su cui si concentrano gli sforzi della comunità scientifica delle cosiddette “scienze dure”. Dall’altro lato, la disponibilità di questa tecnologia è da mettere in relazione con le caratteristiche politiche e sociali dei contesti di applicazione, con gli obiettivi politici delle amministrazioni politiche e con le azioni di policy preesistenti; questioni che integrano i problemi di progettazione e, da una prospettiva di scienze sociali, necessitano di attenzione empirica. Un secondo ordine di fattori che può ostacolare l’implementazione dei DT riguarda le strutture e i processi organizzativi nelle amministrazioni locali, con questioni quali: l’integrazione intersettoriale tra i dipartimenti; le culture organizzative e gli atteggiamenti e le competenze dei dipendenti (Weil *et al.*, 2023).

Dunque, è necessario esplorare le implicazioni politiche dei DT e osservarne empiricamente le modalità di utilizzazione in contesti reali. È possibile fare della “politica dei DT” un oggetto di ricerca combinando diversi approcci teorici di sociologia politica e di *political economy* urbana, per mettere sotto osservazione l’intero ciclo di policy di un DT urbano. Abbiamo recentemente realizzato uno studio di questo tipo attraverso la comparazione⁴¹ dei DT urbani di Bologna e Milano⁴² (d’Albergo e Giovanelli, 2024), fino a quel

⁴¹ La ricerca, basata su un disegno di ricerca comparativo, è stata condotta attraverso interviste agli attori pubblici e privati referenti dei due progetti e mediante l’analisi di documenti istituzionali e di materiali divulgativi.

⁴² Entrambi i progetti sono stati finanziati con risorse economiche comunitarie nell’ambito del Programma Operativo Metropolitan 2014-2020. Il DT di Milano nasce formalmente nel 2022 nell’ambito dell’iniziativa dell’amministrazione locale e di *Milano Smart City Alliance* per la creazione dell’Ecosistema Digitale Urbano, una piattaforma che integra i dati pubblici e privati sulla città. Il DT di Bologna è lo strumento operativo per le politiche di digitalizz-

momento le iniziative più avanzate quanto a stadio di implementazione ed estensione territoriale nel contesto italiano. Questo approccio ci ha consentito di rilevare due opposte modalità di implementazione di un DT urbano e mostrare come questi strumenti, posti a servizio della politica, siano da essa interpretati e configurati a seconda degli obiettivi politici specifici, del sistema di relazioni tra attori pubblici e privati e di altri fattori contestuali ai luoghi. In tal senso, la ricerca può allargare lo sguardo su come le tecnologie digitali entrano nell'agenda politica locale, si inseriscono nelle relazioni di potere della governance urbana e su come siano realmente integrate nei processi amministrativi e decisionali.

Quali sono le cornici di legittimazione impiegate dagli attori politici per giustificare l'adozione della tecnologia DT come strumento per supportare i processi decisionali pubblici? Il marketing dei principali provider privati di questa tecnologia insiste generalmente su un linguaggio condiviso che enfatizza le potenzialità tecniche dello strumento nel creare città avanzate dal punto di vista digitale, interconnesse rispetto ai patrimoni di dati e intelligenti. Ad esempio, secondo Hexagon: «un vero gemello digitale urbano è costituito da più livelli. Il primo è una replica digitale – la città digitale. L'aggiunta di dati provenienti dall'IoT crea il secondo – la città connessa. L'applicazione dell'IA e delle simulazioni a questi dati produce il terzo: la città intelligente»⁴³. Questi linguaggi sono veicolati ai decisori politici, inseriti nei discorsi istituzionali e riprodotti in quel “discorso comunicativo” che, secondo V. Schmidt (2008), collega una molteplicità di attori coinvolti nei processi di persuasione pubblica. L'utilizzo di DT, così come di altri sistemi di IA, necessita di processi di framing delle politiche (Rein e Schön, 1994) che definiscano il senso e il significato di quella decisione e che li connetta a un'agenda di policy più ampia, sempre se non li si consideri solo “freddi” strumenti tecnologici. Dunque, in ogni progetto DT è possibile rilevare e analizzare una componente discorsiva (Schmidt, 2010) sia nelle comunicazioni con i cittadini-elettori, come fattore di costruzione del consenso, sia all'interno della “macchina amministrativa” che deve relazionarsi con la nuova tecnologia. Sussistono e convivono diverse interpretazioni della tecnologia DT che possono essere ricercate empiricamente nei diversi casi studio in relazione alle politiche in cui è collocata la sua adozione. Nel discorso degli esperti e delle istituzioni europee i DT sono sovente interpretati come strumenti della transizione verde e digitale. A titolo di esempio, i DT di Helsinki, Barcellona, Bologna sono intesi come parte dell'attuazione della mis-

zazione e sostenibilità della città, parte integrante del progetto “bandiera” dell'amministrazione del Sindaco Lepore “Città della Conoscenza”.

⁴³ Fonte: <https://hexagon.com/go/sig/urban-digital-twin>

sione “NetZeroCities” della Commissione Europea per la neutralità climatica delle città⁴⁴. Il framing della loro adozione può essere quindi orientato a obiettivi politici specifici, come accade nel caso della sostenibilità urbana. Inoltre, il framing di un progetto di DT urbano può essere riferito non a obiettivi di policy specifici, bensì alle sue potenzialità nel riconfigurare le relazioni tra amministrazioni e cittadini in favore di una maggiore qualità della partecipazione pubblica su diverse questioni di interesse collettivo. Come abbiamo avuto modo di rilevare nel caso di Milano, il framing può anche essere ricondotto a finalità di crescita economica e creazione di valore aggiunto per i cittadini e le imprese, a vantaggio sia del governo pubblico, sia degli attori economici, coerentemente con un orientamento che, come abbiamo visto, caratterizza la città imprenditoriale del paradigma SC e che, fra i casi da noi analizzati è particolarmente presente a Milano.

I DT urbani hanno aperto un altro ambito del mercato delle tecnologie digitali rivolto alle amministrazioni pubbliche sempre più propense a innovare i propri strumenti di governo. La costruzione di DT urbani richiede una serie di azioni che vanno dalla rilevazione geospaziale della città all’elaborazione dei dati per la costruzione di scenari previsionali, passando per la digitalizzazione del patrimonio di dati contenuto nelle serie storiche in possesso delle amministrazioni locali. Ancora, l’implementazione di DT urbani necessita di elevate capacità tecnologiche di integrazione dei dati, standard di sicurezza e interoperabilità tra tecnologie di rilevazione e elaborazione dati⁴⁵. Appare evidente come un simile processo necessiti della definizione di un ecosistema di attori che possano concorrere al progetto apportando vari tipi di risorse, a partire dalle tecnologie e dal personale. In ogni progetto di DT è dunque possibile ricostruire la rete degli attori coinvolti per poterne cogliere due aspetti in particolare. In primo luogo, ci riferiamo alla possibilità per le amministrazioni pubbliche di rimanere coinvolte in condizioni di *lock-in*, ovvero di rimanere “bloccati” o in condizione di dipendenza rispetto al fornitore (*vendor lock-in*)⁴⁶ nell’implementazione di queste innovazioni. Tale dipendenza può comportare l’impossibilità di cambiare tipo di DT (*technological lock-in*), essere dipendenti nella sua gestione e nell’interpretazione delle elaborazioni prodotte. Questi aspetti suggeriscono la necessità

⁴⁴ Fonte: <https://living-in.eu/news/driving-climate-transition-digital-tools-and-local-digital-twins>

⁴⁵ Per la definizione di interoperabilità rimandiamo all’enciclopedia Treccani: «(la) capacità di due o più sistemi, reti, mezzi, applicazioni o componenti, di scambiare informazioni tra loro e di essere poi in grado di utilizzarle». Fonte:

[https://www.treccani.it/enciclopedia/interoperabilita_\(enciclopedia-della-scienza-e-della-tecnica\)/](https://www.treccani.it/enciclopedia/interoperabilita_(enciclopedia-della-scienza-e-della-tecnica)/)

⁴⁶ Per una ricostruzione del fenomeno della dipendenza dal fornitore di rimanda a Sjoerdstra (2016).

per i decisori pubblici di valutare criticamente ogni decisione di acquisto di una tecnologia per la costruzione del proprio DT. In secondo luogo, è importante l'orientamento del progetto, inteso in termini di dipendenza tecnologica e controllo, verso il pubblico o il privato. La configurazione più semplice e che si ritrova anche con più frequenza nei progetti di DT urbani è quella di relazioni pubblico-privato basate sull'acquisto di tecnologie e servizi sul mercato. Esiste un mercato fiorente di aziende che offrono servizi di rilevazione del territorio⁴⁷ (attraverso ad esempio droni, luce laser, fotografie panoramiche) a cui si aggiungono le imprese che forniscono le piattaforme per integrare e gestire i dati. Tali imprese sono talvolta attori imprescindibili nella costruzione di un DT urbano proprio per l'assenza di queste competenze e per l'impossibilità di produrre queste tecnologie all'interno delle amministrazioni.

Sebbene i DT urbani siano ancora ai primi passi del loro sviluppo, con molti progetti che sono solo annunciati o in fase di progettazione, sono già apprezzabili le evidenze dei processi di concentrazione di potere economico che, come abbiamo visto in precedenza, interessano questa particolare versione dell'IA urbana. Un recente articolo di M. Santaniello e C. A. Fonseca Diaz (2024) ha esplorato il mercato delle componenti base di un DT, quali sensori e droni, reti internet e sistemi cloud per l'archiviazione dei dati. Lo studio ha mostrato come i mercati digitali di queste tecnologie siano dominati in condizioni di monopolio da parte di aziende cinesi e americane. Secondo gli autori, quando un governo locale intraprende una politica di digitalizzazione su scala urbana, esso è vincolato dalle condizioni di mercato e da una concentrazione di potere che apre alla possibilità di produrre tre città: una città fisica governata da autorità locali deboli e dipendenti, uno spazio infrastrutturale dominato dalla Cina e uno spazio informativo sotto il controllo degli Stati Uniti (*ibidem*). Al dominio delle quote di mercato si associa la capacità dei giganti tecnologici di organizzare gli interessi e di influenzare le amministrazioni pubbliche nella scelta delle soluzioni tecnologiche più innovative per il governo urbano. Si pensi in tal senso al caso del *Digital Twin Consortium*, fondato da Microsoft e Dell nel 2020, che riunisce a livello globale le principali imprese fornitrice di tecnologie DT operanti in vari settori di attività, centri di ricerca e università. L'obiettivo del Consorzio è promuovere l'adozione intersetoriale della tecnologia DT, compreso il governo urbano, e la creazione di un ecosistema globale di fornitori e utilizzatori delle

⁴⁷ Si pensi a Cyclomedia, tra le principali aziende al mondo nell'acquisizione di dati dal mondo reale, che fornisce alle amministrazioni pubbliche la possibilità di rilevare immagini e dati sugli elementi urbani a fini di pianificazione urbanistica, gestione dello spazio pubblico, manutenzione del verde urbano e, soprattutto, costruzione di DT urbani. Fonte: <https://www.cyclomedia.com/it>

tecnologie⁴⁸. Tra le aziende aderenti troviamo, ad esempio, la multinazionale ESRI, proprietaria di ArcGIS una delle piattaforme geospaziali più avanzate al mondo su cui sono e saranno “costruiti” i DT di Milano e Roma. Essa è probabilmente l’azienda leader che sta guidando l’adozione a livello locale di questa tecnologia potendo dispiegare una combinazione di risorse tecnologiche e di capacità di relazionarsi con esponenti dei governi locali⁴⁹.

Tuttavia, non sarebbe adeguato interpretare il fenomeno dei DT urbani guardando esclusivamente alle relazioni di mercato e intenderli solo come un altro dei prodotti “venduti” alle amministrazioni locali. In altre parole, l’implementazione di un DT urbano si relaziona con gli ambienti politici, economici e culturali delle città, nonché con l’eredità di iniziative precedenti nel campo dell’infrastrutturazione digitale. A questo proposito ci vengono in aiuto le riflessioni della lunga tradizione del pensiero neoistituzionalista che attribuiscono un ruolo esplicativo ai contesti politici e organizzativi, intesi come “variabili” che orientano e vincolano l’azione degli attori politici (March e Olsen, 1989; Dimaggio e Powell, 1991; Brinton e Nee, 1998). Il percorso intrapreso dal Comune di Bologna rappresenta tuttora un importante caso studio per rilevare un’altra possibile configurazione dell’implementazione di un DT urbano meno orientata al “privato”. Esso è stato ideato attraverso la strutturazione di un ecosistema interamente pubblico-pubblico che comprende l’Università Alma Mater di Bologna, il Consorzio universitario CINECA, la Fondazione Bruno Kessler (FBK) e la Fondazione Innovazione Urbana (costituita da Comune di Bologna e Alma Mater). Questa configurazione risponde alla volontà politica dell’amministrazione di perseguire una peculiare forma di “politizzazione” del DT, da intendere come la ricerca di un’autonomia tecnologica dal settore privato, da realizzare sia nella gestione dei patrimoni dati che nella costruzione dell’infrastruttura tecnologica per il DT. Un’ambizione della politica locale resa possibile da almeno tre fattori di contesto. Primo, la possibilità di valorizzare un vasto patrimonio di dati pubblici per decisioni che interessano direttamente le politiche pubbliche in campo ambientale e di mobilità. Una “merce rara” per le amministrazioni locali, che spesso si scontrano con il problema della disponibilità di dati disaggregati a livello comunale, aggiornati e interoperabili.

⁴⁸ Fonte: https://www.digitaltwinconsortium.org/membership-pdf/Digital-Twin_BYLAWS.pdf

⁴⁹ In tal senso, da circa due decenni ESRI Italia organizza una propria conferenza nazionale sulle evoluzioni e applicazioni delle tecnologie geospaziali rivolta ad aziende, esperti e decisori pubblici recentemente anche sui temi connessi alle applicazioni DT nel governo urbano. Per l’elenco dei partecipanti e il programma dell’edizione 2025 si rimanda al link: <https://www.esriitalia.it/it-it/conferenza-esri-italia/conferenza-esri-italia-2025/conferenza-esri-italia-2025>

Secondo, questa autonomia è stata resa possibile dalla preesistenza di un forte ecosistema locale di ricerca pubblica con il coinvolgimento del CINECA e del suo supercomputer Leonardo, che già abbiamo visto essere impiegato nello sviluppo di LLM “italiani”, e di FBK con una lunga esperienza di ricerca nel campo dei DT. Terzo, l’integrazione del DT di Bologna in una complessa catena multiscalare di iniziative pubbliche di ricerca e sviluppo dei DT che hanno come loro ancoraggio il Tecnopolo di Bologna e le iniziative del Centro Nazionale di Ricerca in High Performance Computing, Big Data e Quantum Computing istituito con risorse PNRR⁵⁰.

Dunque, il caso di Bologna ci consente di rilevare un’altra importante componente della politica dei DT, relativa alla multiscalarità delle politiche pubbliche nello sviluppo di DT urbani interoperabili per le amministrazioni locali. A livello comunitario è in corso un complesso processo di integrazione delle tecnologie DT applicate al governo urbano, finalizzato a ricomporre l’eterogeneità del panorama dei DT europei caratterizzato da una concezione non armonizzata della tecnologia, da diversi livelli di maturità digitale e da un’assente o limitata interoperabilità tra le diverse piattaforme urbane (DG Connect, 2023). Per queste ragioni la Commissione ha intrapreso il percorso per la definizione di una “*European Local Digital Twin (LDT) toolbox*”, una cassetta degli attrezzi per realizzare DT europei che condividono tecnologie, approcci di progettazione, dati e che siano interoperabili tra loro. Per realizzare questi obiettivi sono stati messi in campo diversi strumenti di policy che possiamo ricondurre a tre macro-linee di azione. La prima riguarda la produzione di conoscenze tecniche finalizzate allo sviluppo della tecnologia, alla sua integrazione con sistemi IA e al supporto all’interoperabilità dei dati, obiettivi per i quali la Commissione ha emanato bandi⁵¹ e programmato investimenti pubblici per operatori economici e centri di ricerca europei. La seconda linea d’intervento riguarda la costruzione di un ecosistema europeo dei DT in cui riunire amministrazioni locali, esperti e attori di mercato fornitori di tecnologie. Questo obiettivo coinvolge la piattaforma Living-in EU, fondata nel 2020 dalla rete di città Eurocities insieme con Open & Agile Smart Cities & Communities ed European Network of Living Labs, due organizzazioni che si occupano di supportare le amministrazioni locali nei progetti di digitalizzazione. Living-in.EU nasce come una rete di circa duecento città europee che hanno aderito alla Dichiarazione per la costruzione di una “via europea” alla digitalizzazione di città e aree rura-

⁵⁰ In particolare, il DT di Bologna è stato indicato come il progetto di riferimento della linea di ricerca relativa alla digitalizzazione delle città intelligenti.

⁵¹ Il primo nel 2022, “Procurement of the Technical Specifications for the Local Digital Twins (LDTs) Toolbox”, e il secondo nel 2023, “Procurement for the development of the Local Digital Twins toolbox”.

li⁵². Ad essa la Commissione ha affidato il compito di costruire una rete tra amministrazioni locali, *Smart Communities Network*, per facilitarne il coinvolgimento nelle politiche di digitalizzazione e nell'adozione dei DT. La terza linea ambisce a creare un mercato europeo dei LDT attraverso il nuovo strumento dei Consorzi per un'infrastruttura digitale europea (EDIC), istituiti dalla Commissione su iniziativa degli stati membri, persone giuridiche che possono realizzare infrastrutture comuni, acquisire beni e servizi nell'ambito della realizzazione di progetti multinazionali. Nel 2024 è stato quindi istituito l'EDIC “*Networked Local Digital Twins towards the CitiVERSE*”⁵³ dedicato allo sviluppo di DT integrati con sistemi di IA e tecnologie di realtà aumentata per favorire la partecipazione dei cittadini ai processi decisionali pubblici. È su questo punto che insistono le politiche della Commissione Europea, evidenziando la necessità di favorire la partecipazione pubblica.

Questo ci conduce verso l'ultimo gruppo di questioni, relativo a cosa intendiamo per coinvolgimento dei cittadini. Possiamo distinguere tra un coinvolgimento passivo e uno attivo. Per coinvolgimento passivo intendiamo l'utilizzo dei dati di individui e cittadini che possono essere rilevati in tempo reale dai cittadini stessi attraverso sensoristica ad hoc, o condivisi da aziende private che ne sono in possesso. Sappiamo che la raccolta, gestione e conservazione dei dati necessari a un DT urbano comportano le più classiche implicazioni sulla privacy degli individui (Papyshev e Yarime, 2021; Barresi, 2023), già discusse riguardo i rischi pubblici dell'IA. Un elemento di interesse empirico è dunque la *compliance* nei confronti dei quadri regolativi esistenti che intersecano l'implementazione di un DT urbano e che connottano i diversi contesti e scale d'applicazione. Sebbene queste tecnologie possano essere trasferite da un contesto urbano all'altro, è sempre necessario interrogarsi su come gli attori pubblici e privati affrontano le questioni regolative. Dalla nostra ricerca è emerso come le attuali regole non rendano possibile collezionare i dati realmente necessari per creare la copia digitale di un sistema urbano. Come sappiamo, il contesto europeo è più vincolante di altri per imprese e amministrazioni, a causa del suo quadro normativo più rigido: GDPR, DSA, AI Act. Come sono affrontati su scala locale i vincoli regolativi vigenti? È possibile realizzare pratiche di “*AI-localism*” (Verhulst e Sloane, 2020) simili a quelle esaminate nel paragrafo precedente? Oltre alle questioni di stretta conformità alle norme sussistono altre considerazioni relative alle condizioni di ingaggio delle imprese che possono fornire i propri dati ai DT gestiti dalle amministrazioni pubbliche: per quali tipi di elaborazioni? In

⁵² Il testo è disponibile al link: https://living-in.eu/sites/default/files/files/declaration_italian.pdf

⁵³ Fonte: <https://digital-strategy.ec.europa.eu/en/factpages/citiverse>

cambio di cosa? Quali sono, e per chi, i vantaggi attesi della condivisione dei dati? Questi interrogativi sono particolarmente rilevanti quando si vogliono indagare le modalità di creazione di valore pubblico (König, 2021) da parte di un DT.

Il coinvolgimento attivo, invece, riguarda le forme e le pratiche di partecipazione politica. L'utilizzo dei DT nei processi decisionali pubblici sta aprendo la strada a una nuova ondata di pratiche partecipative mediate da innovativi strumenti digitali. Da circa tre decenni si susseguono tecnologie digitali che definiscono il campo della cosiddetta *e-participation* che viene promossa anche su scala urbana attraverso piattaforme per sondaggi online, tecnologie di mappatura partecipativa e tecnologie immersive di realtà aumentata (Medaglia, 2012; Wirtz *et al.*, 2018; Hunter *et al.*, 2021; Shin *et al.*, 2024). Il progetto europeo Urbanage è stato uno dei primi esempi di applicazione dei DT alla partecipazione, sperimentata nell'area metropolitana di Flanders-Bruxelles, e nelle città di Helsinki e Santander. L'obiettivo era sfruttare le potenzialità dei DT urbani nell'elaborare una pianificazione urbanistica attenta alle conseguenze dei cambiamenti demografici e all'invecchiamento della popolazione⁵⁴. L'analisi degli artefatti discorsivi sul rapporto tra DT e partecipazione dei cittadini evidenzia la convergenza degli attori di mercato, esperti e amministratori locali sulle sue potenzialità (d'Albergo *et al.*, 2025). Esse sono ancora una volta relative alla possibilità per i cittadini di visualizzare e interagire con scenari alternativi simulati (Adade e Timo de Vries, 2025). Sebbene queste tecnologie possano ampliare le occasioni di partecipazione in svariati ambiti di politiche urbane, favorendo la possibilità di restituire *feedback* all'amministrazione (White *et al.*, 2021), permangono anche in questo caso vecchi interrogativi sulla reale redistribuzione di potere tra cittadini e decisori pubblici e sui reali livelli della partecipazione (Arnstein, 1969), che potrebbero non essere necessariamente migliorati dalla tecnologia, per quanto avanzata. Peraltro, è possibile chiedersi come possano essere attivate utilizzazioni "dal basso" di piattaforme DT, quando esse richiedono un lungo lavoro di preparazione dei dati prima di essere inseriti nel modello e di consentire elaborazioni. Partecipazione nel campo dei DT significa, in altre parole, solo rendere possibili esperienze di visualizzazione interattiva delle possibili trasformazioni nei paesaggi o nelle funzioni urbane? Su quali questioni il coinvolgimento dei cittadini è sostanzialmente modificato dalla disponibilità di simulazioni e repliche virtuali delle componenti urbane? Quale margine di cambiamento delle decisioni politiche è possibile per i cittadini? Chi affianca queste pratiche partecipative assicurando l'expertise tecnica necessaria? La ricerca sociale sul tema non è

⁵⁴ Per i risultati del progetto si rimanda a: <https://www.urbanage.eu/results>

ancora in grado di rispondere a questi interrogativi. Serviranno dunque studi empirici che possano determinare come i DT siano e saranno effettivamente impiegati in pratiche di inclusione di cittadini e *stakeholder*, in particolare per capire se potranno alimentare ulteriore “falsa partecipazione”, in cui gli obiettivi politici sono stabiliti in anticipo e ai cittadini rimane solo la possibilità di essere meglio informati sulle possibili conseguenze di decisioni di fatto già assunte.

Bibliografia

- Abbott K.W. and Snidal D. (2000), “Hard and soft law in international governance”, *International organization*, 54, 3: 421-456.
- Adade D. and de Vries W.T. (2025), “A systematic review of digital twins’ potential for citizen participation and influence in land use agenda-setting”, *Discover Sustainability*, 6, 1: 1-21.
- Adams R. (2021), “Can artificial intelligence be decolonized?” *Interdisciplinary Science Reviews*, 46, 1-2: 176-197.
- af Malmborg F. (2022), “Narrative dynamics in European Commission AI policy—Sense-making, agency construction, and anchoring”, *Review of Policy Research*, 40: 757-780.
- Aghdam R., Morad B., Ghasemzadeh B., Itani M. and Huovila A. (2024), “Social smart city research: interconnections between participatory governance, data privacy, artificial intelligence and ethical sustainable development”, *Frontiers in Sustainable Cities*, 6, 1514040.
- Aho B. and Duffield R. (2020), “Beyond Surveillance Capitalism: Privacy, Regulation and Big Data in Europe and China”, *Economy and Society*, 49, 2: 187-212.
- Ali S.J., Christin A., Smart A. and Katila, R. (2023), Walking the Walk of AI Ethics: Organizational Challenges and the Individualization of Risk Among Ethics Entrepreneurs, Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency: 217-226. (testo disponibile al sito: <https://arxiv.org/abs/2305.09573>, 19 maggio 2025).
- Alsabt R., Adenle Y. and Alshuwaikhat H. (2024), “Exploring the Roles, Future Impacts, and Strategic Integration of Artificial Intelligence in the Optimization of Smart City”, *Sustainability*, 16, 8: 3389.
- ANCI – ForumPA (2013), *Vademecum per la città intelligente dell’Osservatorio Nazionale Smart City* (testo disponibile al sito: http://osservatoriosmartcity.it/wp-content/uploads/Vademecum_def_2_light.pdf, 19 maggio 2025).
- Andrews L. and Bucher H. (2022), “Automating Discrimination: AI Hiring Practices and Gender Inequality”, *Cardozo L. Rev.*, 44, 1: 145-202.
- Andrews R.N. (1998), “Environmental Regulation and Business’ self-regulation”, *Policy sciences*, 31, 3: 177-197.
- Aresu A. (2024), *Geopolitica dell’intelligenza artificiale*, Milano, Feltrinelli.
- Armondi S. (2014), *Equivoci e possibilità del paradigma della smart city. Quale narrazione nell’agenda urbana italiana?* in Calafati A.G., a cura di, *Città fra sviluppo e declino. Un’agenda urbana per l’Italia*, Donzelli, Roma.
- Arnstein S.R. (1969), “A ladder of citizen participation”, *Journal of the American Institute of planners*, 35, 4: 216-224.
- Arsenault A.C. and Kreps S.E. (2022), *AI and International Politics*, in Bullock J.B., Chen

Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic.

Assanti E.G. e Di Martino G. (2022), "La Cassa depositi e prestiti quale asset strategico per lo sviluppo economico nazionale", *Diritto Costituzionale*, 2: 128-148.

Attard-Frost B., De los Ríos A. and Walters, D.R. (2023), "The ethics of AI business practices: a review of 47 AI ethics guidelines", *AI and Ethics*, 3, 2: 389-406.

Australian Government (2024), *Voluntary AI Safety Standard*. (testo disponibile al sito: <https://www.industry.gov.au/publications/voluntary-ai-safety-standard>, 19 maggio 2025).

Backer L.C. (2019), "China's Social Credit System", *Current History*, 118, 809: 209-214.

Baele S.J., Bukhari I., Whyte C., Cuomo S., Jensen B., Payne K. and Garcia E.V. (2024), "AI IR: Charting International Relations in the Age of Artificial Intelligence", *International Studies Review* (2024), viae013.

Barresi A. (2023), "Urban Digital Twin e pianificazione urbana per la città sostenibile", *TECHNE. Journal of Technology for Architecture & Environment*, 25: 78-83.

Bartle I. and Vass P. (2007), "Self-regulation within the regulatory state: Towards a new regulatory paradigm?" *Public Administration*, 85, 4: 885-905.

Bassanini F. (2015), "La politica industriale dopo la crisi: il ruolo della Cassa Depositi e Prestiti", *L'industria*, 36, 3: 435-454.

Battistelli F. (2016), *La sicurezza e la sua ombra Terrorismo, panico, costruzione della minaccia*, Donzelli: Roma.

Batty M. (2018), "Digital twins", *Environment and Planning B: Urban Analytics and City Science*, 45, 5: 817-820.

Baumgartner F.R. and Jones B.D. (2015), *The politics of information: Problem definition and the course of public policy in America*, University of Chicago Press, Chicago

Beck U. (2000), *La società del rischio. Verso una seconda modernità*, Roma, Carocci.

Bello y Villarino J.M. (2023), "Global Standard-Setting for Artificial Intelligence: Para-regulating International Law for AI", *The Australian Yearbook of International Law Online*, 41, 1: 157-181.

Bemelmans-Videc M.L., Rist R.C. and Vedung E.O., eds., (2011), *Carrots, sticks, and sermons: Policy instruments and their evaluation* (Vol. 1), Routledge.

Benaich N. and Chalmers A., 2024, *State of AI Report 2024*. (testo disponibile al sito: <https://docs.google.com/presentation/d/1GmZmoWOa2O92BPrncRc-TKa15xvQGhq7g4I4hJSNIC0M/edit#slide=id.g309a25a756d085>, 19 maggio 2025).

Benaich N. and Hogarth I. (2022), *State of AI Report 2022*. (testo disponibile al sito: <https://www.stateof.ai/2022-report-launch>, 19 maggio 2025).

Benaich N. and Hogarth I., (2023), *State of AI Report 2023*. (testo disponibile al sito: <https://www.stateof.ai/2023>, 19 maggio 2025).

Benaich N., 2024, *State of AI Report 2024*. (testo disponibile al link: <https://www.stateof.ai>. 15 maggio 2025).

Benesch C., Heim B., Schelker M. and Schmid L. (2023), "Do Voting Advice Applications Change Political Behavior?" *The Journal of Politics*, 85, 2: 684-700.

Bergonzini G. (2024), "Sicurezza della città, tecnologie digitali e intelligenza artificiale: tra regole europee, garanzie costituzionali e autonomia locale", *FEDERALISMI. IT*, 25, p. 1-43.

Beroche H. (ed.) (2021), *Urban AI*. (testo disponibile al sito: <https://urbanai.fr/wp-content/uploads/2021/03/URBAN-AI-1.pdf>, 19 maggio 2025).

Beroche H., Chubinidze A. and Goelzer L. (2022), *Geopolitics of smart cities. Expression of soft Power and new order*. (testo disponibile al sito: <https://urbanai.fr/wp-content/uploads/2023/03/Geopolitics-of-Smart-Cities.pdf>, 19 maggio 2025).

Bietti E. (2020), *From ethics washing to ethics bashing: a view on tech ethics from within*

moral philosophy, in *Proceedings of the 2020 conference on fairness, accountability, and transparency* (testo disponibile al sito: <https://dl.acm.org/doi/10.1145/3351095.3372860>, 19 maggio 2025).

Black J.S. and van Esch P. (2020), “AI-enabled recruiting: What is it and how should a manager use it?” *Business Horizons*, 63, 2: 215-226.

Blanchard A. and Taddeo M. (2024), “Autonomous weapon systems and jus ad bellum” *AI & Society*, 39, 2: 705-711.

Blaith T.F., Gstrein O.J. and Zwitter A. (2022), “Artificial intelligence crime: An overview of malicious use and abuse of AI” *Ieee Access*, 10: 77110-77122.

Boaventura de Sousa Santos, (2023), AI, un approccio sociologico, Other News. (testo disponibile al sito: https://www.other-news.info/italia/wp-content/uploads/2023/03/BP_1-TA_dott_Boaventura_de_Sousa_Santos.pdf, 19 maggio 2025).

Boix C. (2022), *AI and the Economic and Informational Foundations of Democracy*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic.

Boschetti B. e Berti N. (2024), *Transizione digitale e regolazione trasformativa*, in Lalli A., a cura di, *La regolazione pubblica delle tecnologie digitali e dell'intelligenza artificiale*, Giappichelli Editore.

Bosco G., Riccardi V., Sciarrone A., D'Amore R. and Visvizi A. (2024), AI-driven innovation in smart city governance: achieving human-centric and sustainable outcomes, *Transforming Government: People, Process and Policy*, ahead-of-print. (testo disponibile al sito: <https://doi.org/10.1108/TG-04-2024-0096>, 19 maggio 2025).

Botero Arcila B. (2022), *Smart City Technologies: A Political Economy Introduction to Their Governance Challenges*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic

Bourdieu P. (1972), *Esquisse d'une théorie de la pratique*, Droz, Genève.

Bourdieu P. (1986), *The forms of capital*, in Richardson J., ed., *Handbook of Theory and Research for the Sociology od Education*, Westport, CT.

Brandusescu A. and Reia, J., eds. (2022), *Artificial intelligence in the city: Building civic engagement and public trust*. Centre for Interdisciplinary Research on Montreal, McGill University.

Brinton M.C. and Nee, V., eds. (1998), *The New Institutionalism in Sociology*, Russell Sage Foundation.

Brom D. (2021), *AI Governance in the City of Amsterdam. Scrutinising Vulnerabilities of Public Sector AI Systems*. (testo disponibile al sito: https://openresearch.amsterdam/image/2021/7/14/thesis_danielbrom_final_repository.pdf, 19 maggio 2025).

Broussard M. (2023), *More than a Glitch: Confronting Race, Gender, and Ability Bias in Tech*, MIT Press.

Brundage M., Avin S., Clark J., Toner H. et al. (2018), “The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation. *arXiv preprint*. (testo disponibile al sito: <https://arxiv.org/abs/1802.07228>, 19 maggio 2025).

Budak C., Nyhan B., Rothschild D.M., Thorson E. and Watts D.J. (2024), “Misunderstanding the harms of online misinformation”, *Nature*, 630: 45-53.

Bueno de Mesquita E., Canes-Wrone B., Hall A.B., Lum K., Martin G.J. and Velez Y.R. (2024), *Preparing for Generative AI in the 2024 Election: Recommendations and Best Practices Based on Academic Research*, University of Chicago, Harris School of Public Policy White Paper (testo disponibile al sito: https://harris.uchicago.edu/files/ai_and_elections_best_practices_no_embargo.pdf, 19 maggio 2025).

Bullock J.B., Huang H., Kim K. and Young M.M. (2022b), *The Challenge of AI*

Governance for Public Organizations, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M. and Zhang B., eds., *The Oxford Handbook of AI Governance*, Oxford University Press, Oxford, UK.

Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M. and Zhang B. (2022), *Introduction*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M. and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic.

Burelli C. (2025), “Ancora sul ‘ddl intelligenza artificiale’: il parere circostanziato della Commissione europea ai sensi dell’art. 6, par. 2, comma 2, della direttiva (UE) 2015/1535, *Quaderni AISDUE*, 1: 1-7.

Cacciari M., 2023, L’intelligenza artificiale è una tecnologia pre-potente, 24 Video, 27 novembre. (testo disponibile al sito: <https://stream24.ilsole24ore.com/video/finanza/massimo-cacciari-l-intelligenza-artificiale-e-tecnologia-pre-potente/AFosUHoB>, 19 maggio 2025).

Calderini B. (2024), *AI, meglio quella open source? Vantaggi e rischi di un’alternativa*, Agenda digitale. (testo disponibile al sito: <https://www.agendadigitale.eu/cultura-digitale/ia-open-source-lequilibrio-sottile-tra-innovazione-e-rischi-globali/#post-220935-footnote-4>, 19 maggio 2025).

Cameron F.K., Meltzer J.P., Renda A. and Wyckoff A.W. (2025), *Network Architecture for Global AI Policy*, Brookings Governance Studies. (testo disponibile al sito: <https://www.brookings.edu/articles/network-architecture-for-global-ai-policy/>, 19 maggio 2025).

Candrian C. and Scherer A. (2022), “Rise of the machines: Delegating decisions to autonomous AI”, *Computers in Human Behavior*, 134, 107308.

Caprotti F. (2019), “Spaces of Visibility in the Smart City: Flagship Urban Spaces and the Smart Urban Imaginary”, *Urban Studies*, 56, 12: 2465- 2479.

Caprotti F., Cugurullo F., Cook M., Karvonen A., Marvin S., McGuirk P. and Valdez A. (2024), “Why does urban Artificial Intelligence (AI) matter for Urban Studies? Developing Research Directions in Urban AI Research”, *Urban Geography*, 45, 5: 883-894.

Carlini N., Tramèr F., Wallace E., Jagielski M. et al., (2021), “Extracting training data from large language models”, *USENIX Security* 21: 2633-2650.

Carlson M. (2018) “Automating judgment? Algorithmic judgment, news knowledge, and journalistic professionalism”, *New media & society*, 20, 5: 1755-1772.

Carver J. (2024), “More bark than bite? European digital sovereignty discourse and changes to the European Union’s external relations policy”, *Journal of European Public Policy*, 31, 8: 2250-2286.

Castillo A., Karaiskou A. and Gandhi S. (2022), Framing AI for Sustainable Cities, Urban AI. (testo disponibile al sito: <https://medium.com/urban-ai/framing-ai-for-sustainable-cities-9723784dc97d>, 19 maggio 2025).

Cath C. and Jansen F. (2021), *Dutch Comfort: The limits of AI governance through municipal registers*. (testo disponibile al sito: <https://arxiv.org/abs/2109.02944v1>, 19 maggio 2025).

Cavaliere P. e Romeo G. (2022), “From Poisons to Antidotes: Algorithms as Democracy Boosters”, *European Journal of Risk Regulation*, 13, 3: 421-442.

Cavoukian A. (2010), “Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph.D.”, *Identity in the Information Society*, 3, 2: 247-251.

CETaS (Alan Turing Institute) (2024), *AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections*. (testo disponibile al sito: <https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections>, 19 maggio 2025).

- Chen B. (2024), *Performed imaginaries of the AI-controlled city. Conducting AI experimentation in China*, in Cugurullo F., Caprotti F., Cook M., Karvonen A., McGuirk P. and Marvin S., eds., *Artificial intelligence and the city: Urbanistic perspectives on AI*, Routledge.
- Cinelli M., De Francisci Morales G., Galeazzi A. and Starnini M. (2021), "The echo chamber effect on social media", *Psychological and cognitive sciences*, 118, 9: e2023301118.
- Coeckelberg M. (2023), "Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence", *AI and Ethics*, 3: 1341-1350.
- Coeckelbergh M. and Saetra H.S. (2023), "Climate change and the political pathways of AI: The technocracy-democracy dilemma in light of artificial intelligence and human agency", *Technology in Society*, 75: 102406.
- Coglianese C. and Lai C. (2022), *Assessing Automated Administration*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic.
- Coglianese C. and Mendelson, E. (2010), *Meta-regulation and self-regulation*, in Baldwin R., Cave C. and Lodge M., eds., *The Oxford Handbook of Regulation*.
- Coletta C. and Kitchin R. (2017), "Algorhythmic governance: Regulating the 'heartbeat' of a city using the Internet of Things", *Big Data & Society*, 4, 2: 1-16.
- Collett C. and Dillon, S. (2019), *AI and gender: Four proposals for future research*. (testo disponibile al sito: <https://www.repository.cam.ac.uk/items/2069fd59-59c9-49fa-b304-e448c07543a6>, 19 maggio 2025).
- Commissione Europea (2018), *Piano coordinato sull'intelligenza artificiale*. (testo disponibile al sito: <https://eur-lex.europa.eu/legal-content/IT/TXT/HTML/?uri=CE-LEX:52018DC0795>, 19 maggio 2025).
- Commissione Europea (2019), Orientamenti etici per un'IA affidabile. (testo disponibile al sito: <https://op.europa.eu/it/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1>, 19 maggio 2025).
- Commissione europea (2021), *Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions – Fostering a European approach to Artificial Intelligence*. (testo disponibile al sito: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021DC0205>, 19 maggio 2025).
- Conti L.G. and Seele P. (2023), "The contested role of AI ethics boards in smart societies: a step towards improvement based on board composition by sortition", *Ethics and Information Technology*, 25, 4: 51.
- Corsi R. e d'Albergo E., (2024), "La politica dell'intelligenza artificiale general purpose: immaginari socio-tecnici, democrazia e *policy frame* nel processo decisionale della regolazione europea (EU AI ACT - 2022-2024)", *Im@go. A Journal of the Social Imaginary*, 23: 109-130.
- Cortez F. (2023), "Artificial Intelligence, Climate Change and Innovative Democratic Governance" *European Journal of Risk Regulation* 14: 484-503
- Costabile A. (2002), *Il potere politico*, Carocci, Roma.
- Costabile A. (2003), "Le radici del potere politico e la globalizzazione", *Quaderni di sociologia*, 33, 121-135.
- Costabile A. (2024), "L'analisi scientifica del potere: il contributo della sociologia politica", *Società Mutamento Politica*, 15, 29: 147-161.
- Courmont A. and Le Galès P., eds. (2019), *Gouverner la ville numérique*, Puf, Paris.
- Creemers, R. (2018), "China's Social Credit System: an evolving practice of control", *SSRN*. (testo disponibile al sito: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792, 19 maggio 2025).

Critch, A. and Russell, S. (2023), "TASRA: a taxonomy and analysis of societal-scale risks from AI". (testo disponibile al sito: <https://arxiv.org/abs/2306.06924>, 19 maggio 2025).

Cugurullo F. (2020), "Urban Artificial Intelligence: From Automation to Autonomy in the Smart City", *Frontiers in Sustainable Cities*, 2: 1-14.

Cugurullo F. and Xu Y. (2025), "When AIs become oracles: generative artificial intelligence, anticipatory urban governance, and the future of cities", *Policy and Society*, 44, 1: 98-115

Cugurullo F., Barns S., Del Casino V.J *et al.* (2023), "Editorial: The governance of artificial intelligence in the "autonomous city", *Frontiers in Sustainable Cities*, 5: 1-2.

Cugurullo F., Caprotti F. and Marvin S. (2024), "The rise of AI urbanism in post-smart cities: A critical commentary on urban artificial intelligence", *Urban Studies*, 61, 6: 1168-1182.

Cugurullo F., Caprotti F., Cook M., Karvonen A., McGuirk P. and Marvin S. (2024b), Conclusions: The present of urban AI and the future of cities, in Cugurullo F., Caprotti F., Cook M., Karvonen A., McGuirk P. and Marvin S., eds., *Artificial intelligence and the city: Urbanistic perspectives on AI*, Routledge.

Cupić J., Schopmans H. and Tuncer-Ebetürk İ, (2024), "Democratization in the age of artificial intelligence: introduction to the special issue" *Democratization*, 31, 5, 899-921.

Cupić, J. and Sienknecht, M. (2024), "Regulate against the machine: how the EU mitigates AI harm to democracy", *Democratization*, 31, 5: 1067-1090.

d'Albergo E. (2015), *Azione pubblica, imprese ed egemonia in una politica neoliberista: l'Agenda urbana italiana e il paradigma Smart City*, in Moini G., a cura di, *Neoliberismo e azione pubblica. Il caso italiano*, ediesse, Roma.

d'Albergo E., Fasciani T. e Giovanelli G. (2023), "La governance dell'Intelligenza Artificiale nelle politiche locali: trade-off e potere nel caso della videosorveglianza a Torino", *Rivista Trimestrale di Scienza dell'Amministrazione*, 4: 1-26.

d'Albergo E., Giovanelli G. e Nupieri T. (2025), "Local Digital Twins and participatory processes: discourses, representation and practices", *rivista di Digital Politics*, 1: 25-50.

Dafoe A. (2019), "Global politics and governance of artificial intelligence", *Journal of International Affairs*, 72, 1: 121-126.

d'Albergo E. e Giovanelli G. (2024), "Urban Sustainability and the Local Politics of Digital Twins: the Cases of Bologna and Milan", *SocietàMutamentoPolitica*, 15, 30: 65-78.

Dalla Balla, F. (2023), "Fare Stato ad ogni effetto: l'intervento di Cassa Depositi e Prestiti nel mercato" *Munus: rivista giuridica dei servizi pubblici*: 3: 695-730.

Damiani M. (2022), "Sovranismo. Ideologia della sinistra populista?" *Quaderni di Teoria Sociale*, 1, 2: 22-22.

Danks D. and London A.J. (2017, August) Algorithmic Bias in Autonomous Systems", *Ijcai*, 17: 4691-4697.

Denford J.S., Dawson G.S. and Desouza K.C. (2023), WEIRD AI: Understanding what nations include in their artificial intelligence plans, Brookings. (testo disponibile al sito: <https://www.brookings.edu/articles/weird-ai-understanding-what-nations-include-in-their-artificial-intelligence-plans/>, 19 maggio 2025)

Deng T., Zhang K. and Shen Z.J. (2021), "A systematic review of a digital twin city: A new pattern of urban governance toward smart cities", *Journal of management science and engineering*, 6, 2: 125-134.

Dent C., (2020), "Identity, Technology and their Confluence: Governmentality in the Digital Age", *Law, Technology and Humans*, 2, 2: 81-96.

DG Connect (2023), Mapping EU-based LDT providers and users. (testo disponibile al sito: <https://op.europa.eu/en/publication-detail/-/publication/7892c228-359f-11ee-800c-01aa75ed71a1/language-en>, 19 maggio 2025).

Ding J. (2022), *Dueling Perspectives in AI and U.S.-China Relations: Technonationalism vs. Technoglobalism*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M and Zhang B., eds., *The Oxford Handbook of AI Governance* (online edn.), Oxford Academic.

Dodaro C., Lo Scudo F. e Maratea M. (2024), “L’intelligenza artificiale per il futuro sostenibile”, *Rivista di Arpae Agenzia regionale prevenzione, ambiente ed energia dell’Emilia-Romagna*, 4: 8-9.

Drecksler N., Clare S., Schigff K.K. et al. (2025), *What the public thinks about AI and the implications for governance*, Brookings Commentary. (testo disponibile al sito: <https://www.brookings.edu/articles/what-the-public-thinks-about-ai-and-the-implications-for-governance>, 19 maggio 2025).

Duberry J. (2021), *Democracies.AI. Exploring the promise of AI for civil society participation in the policy making process*, 7 maggio, *AI for Good Let’s Talk*. (audio disponibile al sito: <https://www.youtube.com/watch?v=WovImvRomBI>, 19 maggio 2025).

Duberry J. (2022), *Artificial Intelligence and Democracy. Risks and Promises of AI-Mediated Citizen-Government Relations*, Cheltenham, Edward Elgar.

Easton D. (1957), “An approach to the analysis of political systems”, *World politics*, 9 3: 383-400.

ECR (2024), *Diamo voce al buon senso*. (testo disponibile al sito: https://ecrgroup.eu/files/Italian_ECR_Image_Brochure-reduced.size.pdf, 19 maggio 2025).

Elgesem D. (2023), *The AI Act and the risks posed by generative AI models*, in Proceedings of the 5th Symposium of the Norwegian AI Society. (testo disponibile al sito: <https://ceur-ws.org/Vol-3431/paper3.pdf>, 19 maggio 2025).

Elkin-Koren N. and Perel M., (2022), *Democratic friction in speech governance by AI*, in Lindgren S., ed., *Handbook of Critical Studies of Artificial Intelligence*, Edward Elgar, Cheltenham.

Engel C., Linhardt L. and Schubert M. (2024), “Code is law: how COMPAS affects the way the judiciary handles the risk of recidivism”, *Artificial Intelligence and Law*: 1-23.

EPRS - European Parliamentary Research Service (2023), *Artificial intelligence, democracy and elections*. (testo disponibile al sito: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI\(2023\)751478_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2023/751478/EPRS_BRI(2023)751478_EN.pdf), 19 maggio 2025).

Etzkowitz H. and Leydesdorff L. (1995), “The Triple Helix: University-Industry-Government Relations: A Laboratory for Knowledge-Based Economic Development”. *EASST Review*, 14, 1: 14-19.

European Parliament (2022), *Artificial intelligence act and regulatory sandboxes*. (testo disponibile al sito: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI\(2022\)733544_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf), 19 maggio 2025).

Fasciani T. (2025), *For a sociology of local innovation ecosystems. A work in progress on NRRP and the Rome Technopole*, Sapienza Università Editrice, Roma.

Feldstein S. (2024), “Evaluating Europe’s push to enact AI regulations: how will this influence global norms?” *Democratization*, 31, 5: 1049-1066.

Ferré-Bigorra J., Casals M. and Gangolells M. (2022), “The adoption of urban digital twins”. *Cities*, 131: 103905.

Ferrer X., van Neunen T., Such JM., Coté M. and Criado N. (2021), “Bias and discrimination in AI: a cross-disciplinary perspective”, *IEEE Technology and Society Magazine*, 40, 2: 72-80.

Figà Talamanca G. and Arfini S. (2022), “Through the newsfeed glass: Rethinking filter bubbles and echo chambers”, *Philosophy & Technology*, 35, 1: 20.

Filippelli G. (2016), *Agenzie e Autorità nell’ordinamento italiano*. (testo disponibile al sito: <https://www.anquap.it/categorie03.asp?id=2127>, 19 maggio 2025).

- Finocchiaro G. (2024), *Intelligenza artificiale. Quali regole?*, Il Mulino, Bologna.
- Fischer F. (1998), “Ulrich Beck and the Politics of the Risk Society: The Environmental Threat as Institutional Crisis”, *Organization & Environment*, 11, 1: 111-115.
- Flaxman S., Goel S. and Rao J.M. (2016), “Filter bubbles, echo chambers, and online news consumption”, *Public opinion quarterly*, 80, 1: 298-320.
- Floridi L. (2018), “Soft ethics, the governance of the digital and the General Data Protection Regulation”, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 376, 2133: 20180081.
- Floridi L. (2020), “The fight for digital sovereignty: What it is, and why it matters, especially for the EU”, *Philosophy & technology*, 33: 369-378.
- Floridi L. (2021), “The end of an era: From self-regulation to hard law for the digital industry”, *Philosophy & Technology*, 34: 619-622.
- Floridi L. (2022), *Etica dell'intelligenza artificiale: Sviluppi, opportunità, sfide*, Raffaello Cortina Editore, Milano.
- Franzoni, V. (2023), *Gender differences and bias in artificial intelligence*, in Vallverdú J., ed., *Gender in AI and robotics: The gender challenges from an interdisciplinary perspective*, Springer International Publishing.
- Fu D. (2023), *Political Drivers of China's Private Sector Demise*, The Jamestown Foundation, China Brief, 23,16. (testo disponibile al sito: <https://jamestown.org/program/political-drivers-of-chinas-private-sector-demise/>, 19 maggio 2025).
- Gambín, Á.F., Yazidi, A., Vasilakos, A. et al. (2024), “Deepfakes: Current and future trends”, *Artificial Intelligence Review*, 57, 3: 64.
- Garcia E.V. (2022), *Multilateralism and artificial intelligence*, in Tinnirello M., ed., *The Global Politics of Artificial Intelligence*, CRC Press, Abingdon.
- Giachetti Fantin, M. (2018), “La straordinaria ‘mutazione’ del ruolo di Cassa Depositi e Prestiti nel passaggio dallo Stato azionista allo Stato investitore”, *Federalismi.it*, 6: 1-74.
- Giddens A. (1994), *Le conseguenze della modernità. Fiducia e rischio, sicurezza e pericolo*, Il Mulino, Bologna.
- Giovanola B. and Tiribelli S. (2023), “Beyond bias and discrimination: redefining the AI ethics principle of fairness in healthcare machine-learning algorithms” *AI & society*, 38, 2: 549-563.
- Goldman Sachs (2023), *The Potentially Large Effects of Artificial Intelligence on Economic Growth*. (testo disponibile al sito: <https://www.gspublishing.com/content/research/en-reports/2023/03/27/d64e052b-0f6e-45d7-967b-d7be35fabd16.html>, 19 maggio 2025).
- Golubchikov O. and Thornbush M. (2020), “Artificial Intelligence and Robotics in Smart City Strategies and Planned Smart Development”, *Smart Cities*, 3, 4: 1133-1144.
- Granstrand O. and Holgersson M. (2020), “Innovation ecosystems: A conceptual review and a new definition”, *Technovation*, 90-91: 102098.
- Grieves M. and Vickers J. (2017), *Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems*, in Kahlen J., Flumerfelt S. and Alves A., eds, *Transdisciplinary Perspectives on Complex Systems*, Springer, Cham.
- Griffin (2023), “EU Platform Regulation in the Age of Neo-Illiberalism”, *SSRN*: 1-26.
- Groshev M., Guimaraes C., Martín-Pérez J. and de la Oliva A. (2021), “Toward intelligent cyber-physical systems: Digital twin meets artificial intelligence”, *IEEE Communications Magazine*, 59, 8: 14-20.
- Guenduez A. and Mettler T. (2023), “Strategically constructed narratives on artificial intelligence: What stories are told in governmental artificial intelligence policies?” *Government Information Quarterly*, 40, 1: 101719.
- Guetta B. (2019), *I sovranisti*, Add editore, Torino.
- Gutierrez C. I., Marchant G. and Tournas L. (2020), “Lessons for artificial intelligence

from historical uses of soft law governance”, *Jurimetrics*, 61, 1: 1-17.

Hacker P., Engel A. and Mauer M. (2023), *Regulating ChatGPT and other large generative AI models*, in Proceedings of the 2023 ACM conference on fairness, accountability, and transparency: 1112-1123)

Hagemann R., Huddleston Skees J. and Thierer A. (2018), “Soft law for hard problems: The governance of emerging technologies in an uncertain future”, *Colorado Technology Law Journal* 17: 37-130.

Hagendorff T. (2024), “Mapping the ethics of generative ai: A comprehensive scoping review”, *Minds and Machines*, 34, 39: 1-27.

Hallin C. A. and Lipka N. (2023), *Reinventing local government through collective intelligence and artificial intelligence*, in Boucher S., Hallin C.A. and Paulson L., eds., *The Routledge Handbook of Collective Intelligence for Democracy and Governance*, Routledge

Hao K. (2022), “Artificial intelligence is creating a new colonial world order”, *MIT Technology Review*. (testo disponibile al sito: <https://www.technologyreview.com/2022/04/19/1049592/artificial-intelligence-colonialism/>, 19 maggio 2025).

Hartley K., Glen D.K., (2022), *Disrupted governance. Towards a New Policy Science*, Cambridge University Press.

Heidebrecht S. (2024), “From market liberalism to public intervention: Digital sovereignty and changing European union digital single market governance”, *Journal of Common Market Studies*, 62, 1: 205-223.

Heinelt H. (2007), *Do Policies Determine Politics?*, in Fischer F. and Miller G.J., eds., *Handbook of Public Policy Analysis*, Routledge, New York.

Helbing D., Frey B.S., Gigerenzer G. et al. (2017), “Will Democracy Survive Big Data and Artificial Intelligence?” *Scientific American*, February, 25. (testo disponibile al sito: <https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/>, 19 maggio 2025).

Herat H. and Mittal M. (2022), “Adoption of artificial intelligence in smart cities: A comprehensive review”, *International Journal of Information Management Data Insights*, 2:100076.

Hoffmann-Riem W. (2020) *Artificial intelligence as a challenge for law and regulation. Regulating artificial intelligence*, in Wischmeyer, T. and Rademacher, T., eds., *Regulating Artificial Intelligence*, Springer, Cham.

Hollands R.G. (2014), “Critical interventions into the corporate smart city”, *Cambridge Journal of Regions, Economy and Society*, 8, 1: 61-77.

Howard P.N., Woolley S. and Calo R. (2018), “Algorithms, bots, and political communication in the US 2016 election: The challenge of automated political communication for election law and administration”, *Journal of Information Technology, & Politics*, 15, 2: 81-93.

Howlett M. and Ramesh M. (2003), *Come studiare le politiche pubbliche*, Il Mulino, Bologna.

Hu L. (2023), “What is ‘race’ in algorithmic discrimination on the basis of race?”, *Journal of Moral Philosophy*, 1: 1-26.

Huang J., Shao H. and Chang K.C. (2022) “Are large pre-trained language models leaking your personal information?” *arXiv preprint*. (testo disponibile al sito: <https://arxiv.org/abs/2205.12628>, 19 maggio 2025).

Huang M.H. and Rust, R.T. (2021) “A strategic framework for artificial intelligence in marketing”, *Journal of the Academy of Marketing Science*, 49: 30-50.

Huber P. (1985), “Safety and the second best: The hazards of public risk management in the courts”, *Columbia Law Review*, 85, 2: 277-377.

Hung K. (2024), “Artificial intelligence as planetary assemblages of coloniality: The new power architecture driving a tiered global data economy”, *Big Data & Society*, 11, 4.

- Hung K. (2025), *Beyond Big Tech Geopolitics. Moving towards local and people-centred artificial intelligence*, Transnational Institute. (testo disponibile al sito: <https://www.tni.org/en/article/beyond-big-tech-geopolitics>, 19 maggio 2025).
- Hunter M., Soro A. and Brown R. (2021), “Enhancing urban conversations for smarter cities: Augmented reality as an enabler of digital civic participation”, *Interaction design and architecture*, 48: 75-99.
- Huw R., Cowls, J., Morley, J. et al. (2020), “The Chinese approach to artificial intelligence; an analysis of politics, ethics, and regulation, *AI & Society*, 36: 59-77.
- Huyse L. and Parmentier S. (1990), “Decoding codes: The dialogue between consumers and suppliers through codes of conduct in the European Community” *Journal of Consumer Policy*, 13, 3: 253-272.
- Huyskes D. (2022), “Chi governerà l’Intelligenza Artificiale Italiana?”, *Rivista il Mulino*. (testo disponibile al link: <https://www.rivistailmulino.it/a/chi-governer-l-intelligenza-artificiale-italiana>, 19 maggio 2025).
- Inglehart R. and Norris P. (2017), “Trump and the Populist Authoritarian Parties: The Silent Revolution in Reverse”, *Perspectives on politics*, 15, 2: 443-454.
- Innerarity D. (2022), *Let Siri or Alexa vote for us?*, EUI. (testo disponibile al sito: <https://www.eui.eu/news-hub?id=new-stg-chair-analyses-the-relation-between-ai-and-democracy>, 19 maggio 2025).
- Isagah T. and Ben Dhaou S. (2024), *Responsible and Inclusive Urban AI: Opportunities and Challenges for Advancing Sustainable Development Goals*, United Nations University Operating Unit on Policy-Driven Electronic Governance. (testo disponibile al link: <https://collections.unu.edu/view/UNU:10095>, 19 maggio 2025).
- ISPI (2024), *Intelligenza artificiale: chi vuole restare indietro?*. (testo disponibile al sito: <https://www.ispionline.it/it/pubblicazione/intelligenza-artificiale-chi-vuole-restare-indietro-172247>, 19 maggio 2025).
- Jenkins R. and Unies N. (2001), *April Corporate codes of conduct: Self-regulation in a global economy*, UNRISD. (testo disponibile al sito: <https://www.unrisd.org/en/library/publications/corporate-codes-of-conduct-self-regulation-in-a-global-economy>, 19 maggio 2025).
- Jha A.K., Ghimire A., Aryan M.J. et al. (2021), *A Review of AI for Urban Planning: Towards Building Sustainable Smart Cities*, in Proceedings of the Sixth International Conference on Inventive Computation Technologies (ICICT 2021).
- John P. and Cole A. (2000), “When do Institutions, Policy Sectors, and Cities Matter? Comparing Networks of Local Policy Makers in Britain and France”, *Comparative Political Studies*, 33, 2: 248-268.
- Jones B.D. (2002), “Bounded rationality and public policy: Herbert A. Simon and the decisional foundation of collective choice”, *Policy Sciences*, 35, 3: 269-284.
- Jones J. (2025), “Don’t fear artificial intelligence, question the business model: How surveillance capitalists use media to invade privacy, disrupt moral autonomy, and harm democracy” *Journal of Communication Inquiry*, 49, 1: 6-26.
- Jordan A., Wurzel R.K. and Zito A. (2005), “The rise of ‘new’ policy instruments in comparative perspective: has governance eclipsed government?” *Political studies*, 53, 3: 477-496.
- Jungherr A. (2023), “Artificial Intelligence and Democracy: A Conceptual Framework”, *Social Media + Society*, 9, 3.
- Justo-Hanani R. (2022), “The politics of Artificial Intelligence regulation and governance reform in the European Union”, *Policy Sciences*, 55: 137-159.
- Kaminski M.E. (2023), “Regulating the Risks of AI”, *Boston University Law Review*, 103: 22-21.
- Katsoulakis E., Wang Q., Wu H. et al. (2024), “Digital twins for health: a scoping review”, *NPJ digital medicine*, 7: 77.

- Kitchin R. (2014), "Making sense of smart cities: addressing present shortcomings", *Cambridge Journal of Regions, Economy and Society*, 8,1: 131-136.
- Koblentz-Stenzler L. and Klempner U. (2024), Navigating Far-Right Extremism in the Era of Artificial Intelligence, Global Network on Extremism and technology. (testo disponibile al sito: <https://gnet-research.org/2024/01/25/navigating-far-right-extremism-in-the-era-of-artificial-intelligence/>, 19 maggio 2025).
- König P.D. (2021), "Citizen-centered data governance in the smart city: From ethics to accountability", *Sustainable Cities and Society*, 75: 103308.
- König P.D. and Wenzelburger, G., (2020), "Opportunity for renewal or disruptive force? How artificial intelligence alters democratic politics", *Government Information Quarterly*, 37, 3: 101489.
- Korkmaz S. (2024), *Emerging Technologies and Power Asymmetry in International System*, in Kose U. and Demirezen M.U., eds., *Artificial intelligence. Technical and Societal Advancements*, CRC Press.
- Kouroupis K., (2023), "AI and politics: ensuring or threatening democracy?", *Tribuna Juridica*, 13, 4: 575-587.
- Kreps S. and Jakesch M. (2023), "Can AI communication tools increase legislative responsiveness and trust in democratic institutions?", *Government Information Quarterly*, 40, 3: 101829.
- Kreps S. and Kriner D., (2023), "How AI Threatens Democracy", *Journal of Democracy*, 34, 4: 122-131.
- Kreuzer T., Papapetrou P. and Zdravkovic J. (2024), "Artificial intelligence in digital twins- A systematic literature review, *Data & Knowledge Engineering*, 151:102304.
- Kuziemsky M. and Misuraca G. (2020), "AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings" *Telecommunications Policy*, 44, 6: 101976.
- Labuz M. and Nehring C. (2024), "On the way to deep fake democracy? Deep fakes in election campaigns in 2023", *European Political Science*, 23: 454-473.
- Landemore H., (2023), *Fostering More Inclusive Democracy with AI*, IMF Finance & Development Magazine. (testo disponibile al sito: <https://www.imf.org/en/Publications/fandd/Issues/2023/12/POV-Fostering-more-inclusive-democracy-with-AI-Landemore>, 19 maggio 2025).
- Larson J., Denford J.S., Dawson G.S. and Desouza K.C. (2024), *The evolution of artificial intelligence (AI) spending by the U.S. government*, Brookings Institution. (testo disponibile al sito: <https://www.brookings.edu/articles/the-evolution-of-artificial-intelligence-ai-spending-by-the-u-s-government/>, 19 maggio 2025).
- Lascoumes P. and Le Galès P. (2007), "Introduction: Understanding public policy through its instrument. From the nature of instruments to the sociology of public policy instrumentation", *Governance*, 20, 1: 1-21.
- Lee H.P., Yang Y., von Davier T.S., Forlizzi J. and Das S. (2024), May Deepfakes, *Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks*, in Proceedings of the CHI Conference on Human Factors in Computing Systems.
- Lee N.T. (2018), "Detecting racial bias in algorithms and machine learning, *Journal of Information, Communication and Ethics in Society*, 16(3), 252-260.
- Levy F. (2018), "Computers and populism: artificial intelligence, jobs, and politics in the near term", *Oxford Review of Economic Policy*, 34, 3: 393-417.
- Li Y. (2011), *Introduction: challenges and opportunities for NGOs in different parts of the world*, in Li Y., ed., *NGOs in China and Europe. Comparisons and Contrasts*, Routledge, London.
- Liang F. Das V., Kostyuk N. and Hussain M.M. (2018), "Constructing a data-driven

society: China's social credit system as a state surveillance infrastructure", *Policy & Internet*, 10, 4: 415-453.

Lindgren S. (2023), *Introducing critical studies of artificial intelligence*, in Lindgren S., ed., *Handbook of Critical Studies of Artificial Intelligence*, Edward Elgar, Cheltenham, UK.

Lucero K. (2019), "Artificial Intelligence Regulation and China's Future", *Columbia Journal of Asian Law*, 33, 1: 94-171.

Luckett J. (2023), "Regulating generative AI: A pathway to ethical and responsible implementation", *Journal of Computing Sciences in Colleges*, 39, 3: 47-65.

Luque-Ayala A. and Marvin S. (2020), *Urban Operating Systems: Producing the Computational City*, Massachussets Institute of Technology, Cambridge, Massachussets.

Ma J. and He X. (2018), "The Chinese Communist Party's integration policy towards private business and its effectiveness: An analysis of the Ninth National Survey of Chinese Private Enterprises", *Chinese Journal of Sociology*, 4, 3: 422-449.

MacCarthy M. (2023), "The US and its allies should engage with China on AI law and policy", *Brookings Institution*, October, 19. (testo disponibile al sito: <https://www.brookings.edu/articles/the-us-and-its-allies-should-engage-with-china-on-ai-law-and-policy/>, 19 maggio 2025).

Majone G. (1997), "From the Positive to the Regulatory State: Causes and Consequences of Changes in the Mode of Governance", *Journal of Public Policy*, 17: 139-167.

Malaschini A. (2024), "Il regolamento europeo sull'intelligenza artificiale, l'orientamento italiano e i diversi indirizzi di Stati Uniti e Regno Unito", *Rassegna parlamentare*, LXVI,1: 35-92.

Malaschini A. (2024b), "IA, bilanciamo le regole o è autogoa per l'Europa", *Agenda Digitale*, 13 dicembre. (testo disponibile al sito: <https://www.agendadigitale.eu/cultura-digitale/regolare-ia-un-ossimoro-prospettive-globali-per-una-governance-condivisa/>, 19 maggio 2025).

Malaschini A. e Pandolfelli M. (2022), "PARLTECH. Intelligenza Artificiale e Parlamenti: una prima riflessione", *Working Paper Series SOG-WP69/2022*. (testo disponibile al sito: <https://sog.luiss.it/sites/sog.luiss.it/files/Malaschini-Pandolfelli%20WP69%202022r.pdf>, 19 maggio 2025).

Mancera Andrade J.A. and Terán L. (2024), "From GenAI to Political Profiling Avatars: A Data-Driven Approach to Crafting Virtual Experts for Voting Advice Applications", *25th Annual International Conference on Digital Government Research (DGO 2024)*, June 11--14, 2024, Taipei, Taiwan. ACM, New York, NY. (testo disponibile al sito: <https://doi.org/10.1145/3657054.3657092>, 19 maggio 2025).

Manheim K. and Kaplan L. (2019), "Artificial intelligence: Risks to privacy and democracy", *Yale JL & Tech.*, 21, 106: 106-188.

March J.G. and Olsen J.P. (1989), *Rediscovering Institutions. The Organizational Basis of Politics*, Free Press, New York.

Marciano C., Idone Cassone V. e Onnis E. (2020), "Città in gioco. La Smart City tra neoliberismo municipale e riduzione della complessità", *Rivista Trimestrale di Scienza dell'Amministrazione*, 3: 1-28.

Marcucci E., Gatta V., Le Pira M., Hansson L. and Bråthen S. (2020), "Digital twins: A critical discussion on their potential for supporting policy-making and planning in urban logistics", *Sustainability*, 12, 24: 10623.

Marcucci S., Kalkar U. and Verhulst S. (2022), *AI localism in practice; examining how cities govern AI*. (testo disponibile al sito: <https://papers.ssrn.com/sol3/papers.cfm?abstractid=4284013>, 19 maggio 2025).

Marini F. (2022), "Le autorità amministrative indipendenti nell'architettura costituzionale", *Rivista della Corte dei Conti*, 1: 11-22.

- Martin K.D. and Zimmermann J. (2024), "Artificial Intelligence and its Implications for Data Privacy", *Current Opinion in Psychology*, 101829.
- Marx P. (2022), *Civic engagement in the face of political economic barriers*, in Brandusescu A. and Reia J., eds., *Artificial Intelligence in the City: Building Civic Engagement and Public Trust*. Centre for Interdisciplinary Research on Montreal, McGill University, Montreal, Canada.
- Matibag T. (2020), "Artificial intelligence for local governance", *Urban Lawyer*, 50, 3: 415-436.
- McKinsey & Company (2023), *As gen AI advances, regulators – and risk functions – rush to keep pace*. (testo disponibile al sito: <https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/as-gen-ai-advances-regulators-and-risk-functions-rush-to-keep-pace>, 19 maggio 2025).
- McKinsey & Company (2024), *What are AI guardrails?*. (testo disponibile al sito: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-are-ai-guardrails>, 19 maggio 2025).
- McQuillan D. (2022), *Resisting AI: an anti-fascist approach to artificial intelligence*, Bristol University Press, Bristol, UK.
- Medaglia R. (2012), "eParticipation Research: Moving Characterization Forward (2006-2011)", *Government Information Quarterly*, 29, 3: 346-360.
- Medaglia R., Gil-Garcia J.R. and Pardo T.A. (2023), "Artificial Intelligence in Government: Taking Stock and Moving Forward", *Social Science Computer Review*, 41, 1: 123-140.
- MISE (2019), *Proposte per una strategia italiana per l'intelligenza artificiale*. (testo disponibile al sito: https://www.mimit.gov.it/images/stories/documenti/Proposte_per_una_Strategia_italiana_AI.pdf, 19 maggio 2025).
- Mittelstadt B. (2019), "Principles alone cannot guarantee ethical AI", *Nature machine intelligence*, 1, 11: 501-507.
- Mobilio G. (2020), "L'intelligenza artificiale e i rischi di una 'disruption' della regolamentazione giuridica", *BioLaw Journal-Rivista di BioDiritto*, (2): 401-424.
- Morley J., Floridi L., Kinsey L. and Elhalal A. (2020), "From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices", *Science and engineering ethics*, 26, 4: 2141-2168.
- Morlino L. e Raniolo F. (2022), "La sfida populista tra rivendicazionismo e richiami identitari", *Nazioni e Regioni*, 19-20: 11-30.
- Mulazzani G. (2023), *La Cassa Depositi e Prestiti e la riforma amministrativa dell'economia*, Bologna University Press, Bologna.
- Munn L. (2023), "The uselessness of AI ethics", *AI and Ethics*, 3, 3: 869-877.
- Musella F. (2021), "Amministrazione 5.0", *Rivista di Digital Politics*, 1, 1: 95-112.
- National Institute of Standards and Technology (2022), *Ai risk management framework: Initial draft*. (testo disponibile al sito: <https://www.nist.gov/system/files/documents/2022/03/17/AI-RMF-1stdraft.pdf>, 19 maggio 2025).
- Negri E., Fumagalli L. and Macchi M. (2017), "A review of the roles of digital twin in CPS-based production systems", *Procedia manufacturing*, 11: 939-948.
- Nelson K.L. (2012), "Municipal choices during a recession: Bounded rationality and innovation", *State and Local Government Review*, 44, 1_suppl: 44S-63S.
- Nochta T., Badstuber N. and Wahby N. (2019), "On the Governance of City Digital Twins. Insights from the Cambridge case study", *Series No: CDBB_WP_0126*. (testo disponibile al sito: <https://www.repository.cam.ac.uk/items/ede5b25a-7a3e-4508-8e45-c0e2eef9ca4>, 19 maggio 2025).
- Norton Rose Fulbright (2024), *Key takeaways from the G7 Competition Summit on Artificial Intelligence*, Novembre. (testo disponibile al sito: <https://www.nortonrosefulbright.com/g7-competition-summit-artificial-intelligence>

<https://www.nortonrosefulbright.com/en-it/knowledge/publications/d49a49d5/key-takeaways-from-the-g7-competition-summit-on-artificial-intelligence>, 19 maggio 2025).

Nuseir M.T., Al Kurdi B.H., Alshurideh M.T. and Alzoubi H.M. (2021), *Gender discrimination at workplace: Do artificial intelligence (AI) and machine learning (ML) have opinions about it*, in Hassanien A.E. et al., eds., *Proceedings of the International Conference on Artificial Intelligence and Computer Vision (AICV2021)*, Springer International Publishing, Cham, Switzerland.

Ochigame R. (2019), *The invention of 'ethical AI': How Big Tech manipulates academia to avoid regulation*, in Phan T., Goldenfein J. and Kuch D., eds., *Economies of Virtue – The Circulation of 'Ethics' in AI*. Institute of Network Cultures 2022, Amsterdam.

OECD (2021), “State of implementation of the OECD AI principles. Insights from national AI policies”, *OECD Digital Economy Papers*, June, (testo disponibile al sito: https://www.oecd.org/en/publications/state-of-implementation-of-the-oecd-ai-principles_1cd40c44-en.html, 19 maggio 2025).

OECD (2023), “The state of implementation of the OECD AI principles four years on”, *OECD AI papers*, 3, October. (testo disponibile al sito: https://www.oecd.org/en/publications/the-state-of-implementation-of-the-oecd-ai-principles-four-years-on_835641c9-en.html, 19 maggio 2025).

OECD (2023b), “AI language models. Technological, socio-economic and policy considerations”, *OECD Digital Economy Papers*, April, 352. (testo disponibile al sito: https://www.oecd.org/en/publications/ai-language-models_13d38f92-en.html, 19 maggio 2025).

OECD (2023c), “A blueprint for building national compute capacity for artificial intelligence”, *OECD Digital Economy*, 350, February. (testo disponibile al sito: https://www.oecd.org/en/publications/a-blueprint-for-building-national-compute-capacity-for-artificial-intelligence_876367e3-en.html, 19 maggio 2025).

Oh D.S., Phillips F., Park S. and Lee E. (2016), “Innovation Ecosystems: A Critical Examination”, *Technovation*, 54: 1-6.

Omaar H. (2024), “How Innovative Is China in AI?”, *Information Technology & Innovation Foundation*, 26 agosto. (testo disponibile al sito: <https://itif.org/publications/2024/08/26/how-innovative-is-china-in-ai/>, 19 maggio 2025).

OMS (2023), *Regulatory considerations on artificial intelligence for health*. (testo disponibile al sito: <https://iris.who.int/bitstream/handle/10665/373421/9789240078871-eng.pdf>, 19 maggio 2025).

Ortega-Bolaños R., Bernal-Salcedo J., Germán Ortiz M., Galeano Sarmiento J., Ruz G. A. and Tabares-Soto R. (2024), “Applying the Ethics of AI: a Systematic Review of Tools for Developing and Assessing AI-based Systems”, *Artificial Intelligence Review*, 57, 110: 1-30.

Osservatorio Artificial Intelligence (2024), *Contributo dell’Osservatorio Artificial Intelligence all’Indagine Conoscitiva sull’intelligenza artificiale: opportunità e rischi per il sistema produttivo Italiano*. (testo disponibile al sito: https://www.camera.it/application/xmanager/projects/leg19/attachments/upload_file_doc_acquisiti/pdfs/000/009/894/3._Osservatorio_Artificial_Intelligence.pdf, 19 maggio 2025).

Pagallo U. (2020), “Algoritmi e conoscibilità”, *Rivista di filosofia del diritto*, 9,1: 93-106.

Paltiel G. (2023), *Re-imagining Democracy: AI’s Challenge to Political Theory*, in Lindgren S., ed., *Handbook of Critical Studies of Artificial Intelligence*, Edward Elgar, Cheltenham, UK.

Papyshev G. and Yarime M. (2023), “The State’s Role in Governing Artificial Intelligence: Development, Control, and Promotion Through National Strategies”, *Policy Design and Practice*, 6, 1: 79-102.

Papyshev G. and Yarime, M. (2021), “Exploring City Digital Twins as Policy Tools: A

Task-based Approach to Generating Synthetic Data on Urban Mobility”, *Data & Policy*, 3, e16: 1-18.

Pariser E. (2011), *The filter bubble: What the Internet is hiding from you*, Penguin Book Ltd, UK.

Parziale A. (2023), “La riforma italiana dei comitati etici nel contesto europeo: sfide, opportunità e spunti comparativi”, *Rivista italiana di medicina legale e del diritto in campo sanitario*, 1: 103-121.

Pasquale F. (2015), *The Black Box Society. The Secret Algorithms That Control Money and Information*, Harvard University Press, London, England.

Pearson J.S. (2024), “Defining Digital Authoritarianism”, *Philosophy & Technology*, 37, 73: 1-19.

Pennati E. (1963), “Scienza politica e sociologia politica”, *Il Politico*, 28,1: 113-123.

Popelka S., Narvaez Zertuche L. and Beroche H. (2023), *Urban AI Guide*(testo disponibile al sito: <https://urbanai.fr/wp-content/uploads/2023/03/Urban-AI-Guide-2023-V2.pdf>, 19 maggio 2025).

Powell W.W. and DiMaggio P.J. (1991), *The new institutionalism in organizational analysis*, University of Chicago press, Chicago, Usa.

Powers T.M. and Ganascia J.G. (2020), *The Ethics of the Ethics of AI*, in Dubber M:D., Pasquale F. and Das S., eds., *The Oxford Handbook of Ethics of AI*, Oxford University Press, Oxford, UK.

Presidenza del Consiglio dei ministri (2024), *Analisi del quadro normativo in materia di Intelligenza artificiale (D.D.L. IA e Regolamento UE su IA)*. (testo disponibile al sito: <https://www.programmagoverno.gov.it/media/je0lo4i0/focus-ia.pdf>, 19 maggio 2025).

PSIA (2021), *Programma Strategico Intelligenza Artificiale 2022-2024*. (testo disponibile al sito: <https://assets.innovazione.gov.it/1637937177-programma-strategico-iaweb-2.pdf>, 19 maggio 2025).

Puglia S. e Lamanna D. (2024), “Decrescita digitale nell’era dell’intelligenza artificiale Big Tech”, *Comune Info*, 1° settembre (testo disponibile al sito: <https://comune-info.net/decrescita-digitale-nellera-dellintelligenza-artificiale-big-tech/>, 19 maggio 2025).

Puliafito A. (2025), “Smetteremo di usare le IA per scopi militari?”, *Newsletter di Internazionale sulle intelligenze artificiali*. (testo disponibile al sito: <https://www.internazionale.it/notizie/alberto-puliafito/2025/01/13/usare-intelligenze-artificiali-scopi-militari>, 19 maggio 2025).

Puliafito A. (2025b), “La Silicon Valley incassa”, *Newsletter di Internazionale sulle intelligenze artificiali*. (testo disponibile al sito: <https://www.internazionale.it/notizie/alberto-puliafito/2025/01/27/silicon-vally-incassa-trump-altman>, 19 maggio 2025).

Puliafito A. (2025c), “La svolta nazionalista (e opportunista) della OpenAI”, *Newsletter di Internazionale sulle intelligenze artificiali*. (testo disponibile al sito: <https://www.internazionale.it/notizie/alberto-puliafito/2025/01/20/svolta-nazionalista-open-ai>, 19 maggio 2025).

Puliafito A. (2025d), “L’Europa deve creare un istituto per l’intelligenza artificiale”, *Newsletter di Internazionale sulle intelligenze artificiali*(testo disponibile al sito: <https://www.internazionale.it/notizie/alberto-puliafito/2025/02/24/europa-cern-intelligenza-artificiale>, 19 maggio 2025).

Puntoni S., Reczek R.W., Giesler M. and Botti S. (2021), “Consumers and artificial intelligence: An experiential perspective”, *Journal of Marketing*, 85, 1: 131-151.

Quattrociocchi W. (2024), “Le piattaforme digitali influenzano le interazioni sociali?”, *Pandora Rivista*. (testo disponibile al sito: <https://www.pandorarivista.it/articoli/le-piattaforme-digitali-influenzano-le-interazioni-sociali-intervista-a-walter-quattrociocchi>, 19 maggio 2025).

- Radu R. (2021), "Steering the governance of artificial intelligence: national strategies in perspective", *Policy and Society*, 40, 2: 178-193.
- Raniolo F. (2024), *La partecipazione politica*, Il Mulino, Bologna.
- Raniolo F. e Zucaro A. (2023), "Verso democrazie post-rappresentative? Sfide tecnocratiche e contraccolpi populisti in Europa", *FUTURI*, 10, 20: 63-86.
- Rein M. and Schön D. (1993), *Reframing policy discourse*, in Fischer F. and Forester J., eds., *The argumentative turn in policy analysis and planning*, Duke University Press, Durham, USA.
- Reuters Institute (2024), *Reuters Institute Digital News Report 2024*. (testo disponibile al sito: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2024-06/RISJ_DNR_-2024_Digital_v10%20lr.pdf, 19 maggio 2025).
- Rhodes S.C. (2022), "Filter bubbles, echo chambers, and fake news: How social media conditions individuals to be less critical of political misinformation", *Political Communication*, 39, 1: 1-22.
- Rjab A.B., Mellouli S. and Corbett J. (2023), "Barriers to artificial intelligence adoption in smart cities: A systematic literature review and research agenda", *Government Information Quarterly*, 40, 3: 101814.
- Robinson S.C. (2020), "Trust, transparency, and openness: How inclusion of cultural values shapes Nordic national public policy strategies for artificial intelligence (AI)", *Technology in Society*, 63: 101421.
- Rodilosso E. (2024), "Filter bubbles and the unfeeling: How AI for social media can foster extremism and polarization", *Philosophy & Technology*, 37, 71: 1-21.
- Roe M. (2019), "Who's Driving That Car: An Analysis of Regulatory and Potential Liability Frameworks for Driverless Cars", *BCL Rev.*, 60, 1: 317-347.
- Romano O. (2023), *Posizionarsi sociologicamente nel dopo-sviluppo*, in de Nardis F., Petrillo A. e Simone A., a cura di, *Sociologia di posizione. Prospettive teoriche e metodologiche*, Meltemi, Milano.
- Rossi U. (2016), "The variegated economics and the potential politics of the smart city", *Territory Politics Governance*, 4, 3: 337-353.
- Rossi U. and Vanolo A. (2015), *Urban Neoliberalism*, in Smelser N.J. and Baltes P.B., eds., *International Encyclopedia of the Social & Behavioral Sciences*, Elsevier, Oxford, UK.
- Rosso R. (2023), "La seconda conferenza sull'intelligenza artificiale: A.I. e democrazia", *Transform! Italia*. (testo disponibile al sito: <https://transform-italia.it/note-sulla-conferenza-sullintelligenza-artificiale/>, 19 maggio 2025).
- Ruohomäki T., Airaksinen E., Huuska P., Kesäniemi O., Martikka M. and Suomisto J. (2018), *Smart city platform enabling digital twin*, in *2018 International Conference on Intelligent Systems (IS)*, IEEE. (testo disponibile al sito: <https://ieeexplore.ieee.org/document/8710517>, 19 maggio 2025).
- Sacconi L. (2004), "Responsabilità sociale come governance allargata d'impresa: un'interpretazione basata sulla teoria del contratto sociale e della reputazione", *Liuc Papers* (testo disponibile al sito: <https://biblio.liuc.it/liucpap/pdf/143.pdf>, 19 maggio 2025).
- Sadowski J. and Bendor R. (2018), "Selling Smartness: Corporate Narratives and the Smart City as a Sociotechnical Imaginary", *Science, Technology, & Human Values*, 44, 3: 540-563.
- Saenz L.R. and Saura Garcia C. (2024), "China's Tech Surveillance Applications in Europe and Latin America: Analysing the Impact on Democratic Governance", *EUI, STG, Policy Brief*, 2024/17. (testo disponibile al sito: <https://hdl.handle.net/1814/77055>, 19 maggio 2025).
- Sætra H.S. (2020), "A Shallow Defence of a Technocracy of Artificial Intelligence: Examining the Political Harms of Algorithmic Governance in the Domain of Government", *Technology in Society*, 62: 101283.

- Saetra H.S. (2021), "AI in Context and the Sustainable Development Goals: Factoring in the Unsustainability of the Sociotechnical System", *Sustainability*, 13, 4: 1738
- Saheb T. (2023), "Topical review of artificial intelligence national policies: A mixed method analysis", *Technology in Society*, 74, 102316.
- Sanders N.E. and Schneier B., (2023), "How AI Could Write our Laws", *MIT Technology review*. (testo disponibile al sito: <https://www.technologyreview.com/2023/03/14/1069717/how-ai-could-write-our-laws/>, 19 maggio 2025).
- Santaniello M. and Fonseca Diaz C.A. (2024), "The Geopolitics of Digital Twins. A tale of three cities", *Rivista di Digital Politics*, 4, 3: 465-484.
- Sassen S. (2021), *Urbanized Techology*, in Beroche H., ed., *Urban AI*. (testo disponibile al sito: <https://urbanai.fr/wp-content/uploads/2021/03/URBAN-AI-1.pdf>, 19 maggio 2025).
- Savini F. and Bertolini L. (2019), "Urban experimentation as a politics of niches", *Environment and Planning A: Economy and Space*, 51, 4: 831-848.
- Scanlon R. (2025), "Beyond DeepSeek: How China's AI Ecosystem Fuels Breakthroughs", *Lawfare*. (testo disponibile al sito: <https://www.lawfaremedia.org/article/beyond-deepseek--how-china-s-ai-ecosystem-fuels-breakthroughs>, 19 maggio 2025).
- Schachtner C. (2021), "Smart Government in Local Adoption – Authorities in Strategic Change Through AI", *Smart Cities and Regional Development Journal*, 5, 3: 53-62.
- Scharre P. (2023), "AI's Gatekeepers Aren't Prepared for What's Coming", *FP*. (testo disponibile al sito: <https://foreignpolicy.com/2023/06/19/ai-regulation-development-us-china-competition-technology/>, 19 maggio 2025).
- Scherer M.U. (2015), "Regulating artificial intelligence systems: Risks, challenges, competencies, and strategies", *Harv. JL & Tech.*, 29, 353-400.
- Schiff D.S. (2023), "Looking through a policy window with tinted glasses: Setting the agenda for U.S. AI policy", *Review of Policy Research*, 40: 729-756.
- Schippers B. (2020), "Artificial Intelligence and Democratic Politics", *Political Insight*, 1, 11: 32-35.
- Schmidt V.A. (2008), "Discursive institutionalism: The explanatory power of ideas and discourse", *Annu. Rev. Polit. Sci.*, 11, 1: 303-326.
- Schmidt V.A. (2010), "Taking ideas and discourse seriously: explaining change through discursive institutionalism as the fourth 'new institutionalism'", *European political science review*, 2, 1: 1-25.
- Schmidt V.A. (2010), *Democracy and Legitimacy in the European Union Revisited: Output, Input and Throughput*, KFG Working Paper Series, No. 21, November 2010, Kolleg-Forscherguppe (KFG), *The Transformative Power of Europe*, Freie Universität Berlin, Berlin. (testo disponibile al sito: <https://core.ac.uk/reader/199429030>, 19 maggio 2025).
- Schmitt M. and Flechais I. (2024), "Digital deception: Generative artificial intelligence in social engineering and phishing", *Artificial Intelligence Review*, 57,12: 1-23.
- Schneier B. (2023), "Ten Ways AI Will Change Democracy", *Harvard Kennedy School, commentary*. (testo disponibile al sito: <https://ash.harvard.edu/articles/ten-ways-ai-will-change-democracy/>, 19 maggio 2025).
- Schneier B., Farrel H. and Sanders N.E. (2023), "How Artificial Intelligence Can Aid Democracy", *Slate*, April 21, 2023. (testo disponibile al sito: <https://slate.com/technology/2023/04/ai-public-option.html>, 19 maggio 2025).
- Schoch D., Keller F.B., Stier S. and Yang J. (2022), "Coordination Patterns Reveal Online Political Astroturfing Across the World", *Scientific Reports*, 12: 4572.
- Schröder W. (2024), Il politologo Schroeder: "Musk vuole abbattere i muri verso l'Afd. È pericoloso, la Repubblica, 21 dicembre 2024. (testo disponibile al link: <https://www.repubblica.it/esteri/2024/12/21/news/schroedermuskgermaniadestreeuropeintervista-423900513/>, 19 maggio 2025).

- Schuetz J., Reuel A.K. and Carlier A. (2024), "How to Design an AI Ethics Board", *AI and Ethics*, 5, 863-881.
- Schultz M.D., Conti L.G. and Seele, P. (2024), "Digital Ethicswashing: A Systematic Review and a Process-Perception-Outcome Framework", *AI and Ethics*, 5, 805-818.
- Segatori R. (2015), *Sintomi populisti nella crisi italiana*, in Saccà F., a cura di, *Culture politiche e mutamento nelle società complesse*, FrancoAngeli, Milano, pp. 111-31.
- Segatori R. (2021), "L'etica tra filosofia, storia e sociologia: una panoramica dei classici", *Etica pubblica: studi su legalità e partecipazione*, 1, 2021: 77-92.
- Seppälä A., Birkstedt, T. and Mäntymäki M. (2021), *From Ethical AI Principles to Governed AI*, in *ICIS 2021 Proceedings*. 10. (testo disponibile al sito: https://aisel.aisnet.org/icis2021/ai_business/ai_business/10, 19 maggio 2025).
- Servizi Studi di Camera e Senato (2025), *Disposizioni e deleghe al Governo in materia di intelligenza artificiale*. (testo disponibile al sito: <https://documenti.camera.it/leg19/dossier/pdf/ID0010a.pdf>, 19 maggio 2025).
- Shaffer G.C. and Pollack M.A. (2009), "Hard vs. Soft law: Alternatives, Complements, and Antagonists in International Governance", *Minn. L. Rev.*, 94: 706-799.
- Shannon B.N., McGee Z.A. and Jones, B.D. (2019), *Bounded Rationality and Cognitive Limits in Political Decision Making*, in *Oxford Research Encyclopedia of Politics*. (testo disponibile al sito: <https://oxfordre.com/politics/display/10.1093/acrefore/9780190228637.001.0001/acrefore-9780190228637-e-961>, 19 maggio 2025).
- Shapiro A. (2024), *Optimizing the Immeasurable. On the Techno-Ethical Limits of Predictive Policing*, in Cugurullo F., Caprotti F., Cook M., Karvonen A., McGuirk P. and Marvin S., eds., *Artificial Intelligence and the City: Urbanistic Perspectives on AI*, Routledge, London, UK.
- Sharma S. (2019), *Data Privacy and GDPR Handbook*, John Wiley & Sons, New Jersey, USA.
- Sheehan M. (2023), "China's AI Regulations and How They Get Made", *Carnegie Endowment for International Peace Working paper*. (testo disponibile al sito: <https://carnegieendowment.org/research/2023/07/chinas-ai-regulations-and-how-they-get-made?lang=en>, 19 maggio 2025).
- Sheehan M. (2024), "Tracing the Roots of China's AI Regulations", *Carnegie Endowment for International Peace Working paper*. (testo disponibile al sito: <https://carnegieendowment.org/research/2024/02/tracing-the-roots-of-chinas-ai-regulations?lang=en>, 19 maggio 2025).
- Sheehan M. (2024b), "China's Views on AI Safety Are Changing-Quickly", *Carnegie Endowment for International Peace Working paper*. (testo disponibile al sito: <https://carnegieendowment.org/research/2024/08/china-artificial-intelligence-ai-safety-regulation?lang=en>, 19 maggio 2025).
- Shin B., Floch J., Rask M., Bæk P., Edgar C., Berditchevskaia A. and Branlat M. (2024), "A systematic analysis of digital tools for citizen participation", *Government Information Quarterly*, 41,3: 101954.
- Siau K. and Wang W. (2020), "Artificial intelligence (AI) ethics: ethics of AI and ethical AI", *Journal of Database Management (JDM)*, 31,2: 74-87.
- Silverman J. (2017), "Privacy Under Surveillance Capitalism", *Social Research: An International Quarterly*, 84,1: 147-164.
- Simon D. (2016), *Rethinking Sustainable Cities, Accessible, Green and Fair*, Policy Press, Bristol, UK.
- Simon H. (1957), *A Behavioral Model of Rational Choice in Models of Man, Social and Rational*, Macmillan, New York, USA.
- Simon H.A. (2001), *Il comportamento amministrativo*, il Mulino, Bologna.

Simons J. (2023), *Algorithms for the People: Democracy in the Age of AI*, Princeton University Press, Princeton, New Jersey USA.

Sinclair D. (1997), "Self-regulation Versus Command and Control? Beyond False Dichotomies", *Law & Policy*, 19,4: 529-559.

Sjoerdstra B. (2016), *Dealing with Vendor Lock-in*. (testo disponibile al sito: https://esay.utwente.nl/70153/1/Sjoerdstra_BA_BMS.pdf, 19 maggio 2025).

Slattery P., Saeri A.K., Grundy E.A., Graham J., Noetel M., Uuk R., ...and Thompson N. (2024), "The AI Risk Repository: A Comprehensive Meta-review, Database, and Taxonomy of Risks from Artificial Intelligence", *arXiv preprint arXiv:2408.12622*. (testo disponibile al sito: <https://arxiv.org/abs/2408.12622>, 19 maggio 2025).

Snrcek N. (2016), *Platform Capitalism*, John Wiley and Sons Ltd, NJ, USA.

Sørensen E. and Torfing J. (2009), "Making governance networks effective and democratic through metagovernance", *Public administration*, 87,2: 234-258.

Stahl B.C. (2021), *Artificial intelligence for a better future: an ecosystem perspective on the ethics of AI and emerging digital technologies*, Springer Nature, Cham, Switzerland.

Stahl B.C., Schroeder D. and Rodrigues R. (2022), *Surveillance capitalism*, in Stahl B.C., Schroeder D. and Rodrigues R., eds., *Ethics of Artificial Intelligence: Case Studies and Options for Addressing Ethical Challenges*, Springer International Publishing, Cham, Switzerland.

Standford University (2024), *Artificial Intelligence Index Report*. (testo disponibile al sito: <https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAIAI-Index-Report-2024.pdf>, 19 maggio 2025).

Starke, C., and Lünich, M. (2020), "Artificial Intelligence for Political Decision-making in the European Union: Effects on Citizens' Perceptions of Input, Throughput, and Output Legitimacy", *Data & Policy*, 2: e16.

Stein A.L. (2022), "Assuming the Risks of Artificial Intelligence", *BUL Rev.*, 102: 979-1035.

Steurer R. (2013), "Disentangling Governance: a Synoptic View of Regulation by Government, Business and Civil Society", *Policy Sciences*, 46: 387-410.

Stockwell S. (2024), *AI-enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections*. (testo disponibile al sito: https://cetas.turing.ac.uk/sites/default/files/2024-09/cetas_briefing_paper_-_ai-enabled_influence_operations_-_threat_analysis_of_the_2024_uk_and_european_elections.pdf, 19 maggio 2025).

Strover S., Cao T., Esteva M. and Park S. (2021), "Smart Cities and Ethical Policies: The Challenges of Public Cameras and AI", *TPRC49: The 49th Research Conference on Communication, Information and Internet Policy* (testo disponibile al sito: <https://ssrn.com/abstract=3898220> (06/02/2023)), 19 maggio 2025.

Suleyman M. (2023), *The Coming Wave: Technology, Power, and the Twenty-First Century's Greatest Dilemma*.

Sun T., He X. and Li Z. (2023), "Digital Twin in Healthcare: Recent Updates and Challenges", *Digital Health*, 9, 20552076221149651.

Sunstein C. (2017), *#Republic: Divided Democracy in the Age of Social Media*, Princeton University Press, Princeton-NJ, USA.

Tabrizian S. (2025), "Culturally Responsive Ai Governance in Smart Cities: Cross-Cultural Ethics and Policy Pathways", *Preprint SSRN*. (testo disponibile al sito: <http://dx.doi.org/10.2139/ssrn.5238609>, 19 maggio 2025).

Tafani D. (2024), *Le Big Tech e il racconto dell'intelligenza artificiale*. (testo disponibile al sito: <https://centroriformastato.it/le-big-tech-e-il-racconto-dellintelligenza-artificiale/>, 19 maggio 2025).

Timcke S. (2024), "The need for AI to go beyond token democracy", *Transforming*

society. (testo disponibile al sito: <https://www.transformingsociety.co.uk/2024/04/18/the-need-for-ai-to-go-beyond-token-democracy/>, 19 maggio 2025).

Töller, A.E. (2011), *Voluntary Approaches to Regulation—Patterns, Causes, and Effects*, in Levi-Faur D., ed., *Handbook on the politics of regulation*, Edward Elgar Publishing, Cheltenham, UK.

Tooze A. and Bello W. (2025), *A Fractured World. Reflections on Power, Polarity and Polycrisis*, in Transnational Institute, Geopolitics of capitalism. State of power 2025. (testo disponibile al sito: <https://www.tni.org/en/publication/geopolitics-of-capitalism>, 19 maggio 2025).

Toti A. (2020), “Inclusioni ed esclusioni sociali. Utopie e distopie della smart city”, *Rivista Trimestrale di Scienza dell’Amministrazione*, 3: 1-17.

Trecanni 2017, *Echo-chamber*. (testo disponibile al link: [https://www.trecanni.it/vocabolario/echo-chamber\(Neologismi\)](https://www.trecanni.it/vocabolario/echo-chamber(Neologismi)), 19 maggio 2025).

Turner J. (2019), *Robot Rules. Regulating Artificial Intelligence*, Palgrave Macmillan, London, UK.

Ulinicane I. (2024), “Intersectionality in Artificial Intelligence: Framing Concerns and Recommendations for Action”, *Social Inclusion*, 12: 1-16.

Ulinicane I. and Erkkilä T. (2023), “Politics and policy of Artificial Intelligence”, *Review of Policy Research*, 40, 5: 607-611.

Ulinicane I., Knight W., Leach T., Stahl B.C. and Wanjiku W. (2021), “Framing Governance for a Contested Emerging Technology: Insights from AI Policy”, *Policy & Society*, 40, 2: 158-177.

UN – AI Advisory Body (2024), *Governing AI for Humanity: Final Report*. (testo disponibile al sito: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf, 19 maggio 2025).

UN (2024), *Governing AI for Humanity: Final Report*. (testo disponibile al sito: https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf, 19 maggio 2025).

UN (2024b), *United Nations System White Paper on Artificial Intelligence Governance*. (testo disponibile al sito: <https://unsceb.org/sites/default/files/2024-11/UNSystemWhitePaperAIGovernance.pdf>, 19 maggio 2025).

UN-HABITAT (2020), *World Cities Report 2020: The Value of Sustainable Urbanization*. (testo disponibile al link: <https://unhabitat.org/world-cities-report-2020-the-value-of-sustainable-urbanization>, 19 maggio 2025).

Urbinati N. (2020), “Un’analisi critica del sovranismo”, *Parolechiave*, 28, 1: 51-63.

Vallebueno A. and Lee Y.S. (2023), “Measuring Urban Quality and Change Through the Detection of Physical Attributes of Decay”, *Scientific Reports*, 13, 17316.

van Dijck J., Poell T. and de Waal M. (2018), *The Platform Society: Public Values in a Connective World*, Oxford University Press Inc, Oxford, UK.

Van Dijk N., Casiraghi S. and Gutwirth, S. (2021), “The ‘Ethification’ of ICT Governance. Artificial intelligence and Data Protection in the European Union”, *Computer Law & Security Review*, 43, 105597.

Van Maanen G. (2022), “AI Ethics, Ethics Washing, and the Need to Politicize Data Ethics”, *Digital Society*, 1, 9.

van Noordt C., Medaglia R. and Tangi L. (2023), “Policy Initiatives for Artificial Intelligence-enabled Government: An Analysis of National Strategies in Europe”, *Public Policy and Administration*, 40,2: 1-39

Van Nuenen, T., Ferrer X., Such J.M. and Coté, M. (2020), “Transparency for whom? Assessing discriminatory Artificial Intelligence”, *Computer*, 53,11: 36-44.

Vannuccini S. (2025), “Artificial intelligence and the future of European industrial

policy”, *Centro Studi sul Federalismo commentaries*, n. 318. (testo disponibile al sito: https://csfederalismo.it/images/2025/commenti/CSFCommentary318_AI-EU_Vannucini_27Feb2025.pdf, 19 maggio 2025).

Vanolo A. (2014), “Smartmentality: the smart city as disciplinary strategy”, *Urban Studies*, 71, 5: 883-898.

Vedaschi A. e Noberasco G.M. (2019), “Gli autoveicoli a guida autonoma alla prova del diritto”, *Diritto pubblico comparato ed europeo*, 21(Speciale): 769-798.

Verdegem P. (2023), *Critical AI studies meet Critical Political Economy*, in Lindgren S., ed., *Handbook of Critical Studies of Artificial Intelligence*, Edward Elgar Publishing Limited, Cheltenham, UK.

Verdegem P. (2024), “Dismantling AI Capitalism: The Commons as an Alternative to the Power Concentration of Big Tech”, *AI & society*, 39, 2: 727-737.

Verhulst S.G. (2022), *AI localism: governance of artificial intelligence at the city and local level*, in Brandusescu A. and Reia J., eds., *Artificial intelligence in the city: Building civic engagement and public trust*, Centre for Interdisciplinary Research on Montreal, McGill University, Montreal, Canada.

Verhulst S.G. (2022b), *The AI Localism Canvas: A Framework to Assess the Emergence of Governance of AI within Cities* (testo disponibile al link: <https://sverhulst.medium.com/the-ai-localism-canvas-a-framework-to-assess-the-emergence-of-governance-of-ai-within-cities-70940bd690>, 19 maggio 2025).

Verhulst S.G. and Sloane M. (2020), “Realizing the Potential of AI Localism”, Project Syndicate, 7. (testo disponibile al sito: <https://www.project-syndicate.org/commentary/local-regulation-of-artificial-intelligence-uses-by-stefaan-g-verhulst-1-and-mona-sloane-2020-02>, 19 maggio 2025).

Viehoff J. (2022), *Beyond Justice: Artificial Intelligence and the Value of Community*, in Bullock J.B., Chen Y., Himmelreich J., Hudson V.M., Korinek A., Young M.M. and Zhang B., eds., *The Oxford Handbook of AI Governance*, Oxford University Press, Oxford, UK.

Villani S. (2024), “Il sistema di vigilanza sull’applicazione dell’AI Act: ognun per sé?”, *QUADERNI AISDUE*, 2: 1-20.

Villasenor J. and Turner J. (2024), “AI Policy Directions in the New Trump Administration”, *Brookings* (testo disponibile al sito: <https://www.brookings.edu/articles/ai-policy-directions-in-the-new-trump-administration/>, 19 maggio 2025).

Vinci F. (2022), *Il rapporto tra decisori politici ed esperti in Italia durante i primi mesi della pandemia. Un’indagine esplorativa*, in Moini G. e Millefiorini A., a cura di, *Covid, azione pubblica e crisi della contemporaneità. Primato o declino della politica?*, Sapienza Università Editrice, Roma.

Viviani L. (2017), “Politica e società: le nuove sfide della sociologia politica”, *Società Mutamento Politica*, 8, 15: 7-35.

Volpi A. (2025), “DeepSeek: la tecnologia cinese mette a nudo una colossale bolla finanziaria”, *Altraeconomia*, 28 01 2025. (testo disponibile al link: <https://altreconomia.it/deepseek-la-tecnologia-cinese-mette-a-nudo-una-colossale-bolla-finanziaria>, 19 maggio 2025).

Wagner B. (2018), *Ethics as an escape from regulation. From “ethics-washing” to ethics-shopping?*, in Bayamlioglu E., Baraliuc I., Janssens L.A.W and Hildebrandt M., eds., *BEING PROFILED: COGITAS ERGO SUM. 10 Years of Profiling the European Citizen*, Amsterdam University Press, Amsterdam.

Wallach W., Reuel A. and Kaspersen A. (2022), “Soft law functions in the international governance of AI”, *Arizona State University Asulaw*. (testo disponibile al link: https://lsu.asulaw.org/wp-content/uploads/sites/7/2023/12/Wallach-et-al_Soft-Law-Functions-in-the-International-Governance-of-AI.pdf, 19 maggio 2025).

Wang J. and Galligan D. (2021), "European Policies Toward Chinese Tech Companies: European Interests, Risk Analysis and Policy Recommendations", *Oxford Global Society*. (testo disponibile al link: <https://oxgs.org/wp-content/uploads/2022/01/OXGS-POLICY-REPORT-European-policy-towards-Chinese-tech-companies.pdf>, 19 maggio 2025).

Wang K., Dong K., Wu J. and Wu J. (2025), "Patterns of artificial intelligence policies in China: a nationwide perspective", *Library Hi Tech*, 43, 1: 295-325.

Wang Y., Pan Y., Yan M., Su Z. and Luan T.H. (2023), "A survey on ChatGPT: AI-generated Contents, Challenges, and Solutions", *IEEE Open Journal of the Computer Society*, testo disponibile al link: <https://arxiv.org/pdf/2305.18339.pdf>.

Wang Y., Zhang N. and Zhao X. (2020), "Understanding the Determinants in the Different Government AI Adoption Stages: Evidence of Local Government Chatbots in China", *Social Science Computer Review*, 40, 2.

Warren S. and Brandeis L. (1890), "The Right to Privacy", *Harv. L. Rev.*, 4,5: 193-220.

Washington A.L. (2018), "How to argue with an algorithm: Lessons from the COMPAS-ProPublica debate", *Colo. Tech. LJ*, 17, 1: 131-160.

Weil C., Bibri S.E., Longchamp R., Golay, F. and Alahi, A. (2023), "Urban digital twin challenges: A systematic Review and Perspectives for Sustainable Smart Cities", *Sustainable Cities and Society*, 99: 104862.

Westerlund M. (2019), "The emergence of deepfake technology: A review", *Technology innovation management review*, 9,11: 40-53.

While A. (2024), *Regulating and Making Space for the Expanded Field of Urban Robotics*, in Cugurullo F., Caprotti F., Cook M., Karvonen A., McGuirk P. and Marvin S., eds., *Artificial Intelligence and the City: Urbanistic Perspectives on AI*, Routledge, London, UK.

White G., Zink A., Codecá L. and Clarke S. (2021), "A digital twin smart city for citizen feedback", *Cities*, 110, 103064.

Wihbey J.P. (2024), "AI and Epistemic Risk for Democracy: A Coming Crisis of Public Knowledge?", *Ethics Institute Working Paper*. (testo disponibile al link: <https://ssrn.com/abstract=405026>, 19 maggio 2025).

Wildavsky A.B. (1964), *The politics of the budgetary process*, Little, Brown and Company, Boston and Toronto.

Wirtz B.W., Daiser, P. and Binkowska B. (2018), "E-participation: A strategic framework", *International Journal of Public Administration*, 41, 1: 1-12.

Wolniak R. and Secula K. (2024), "Artificial Intelligence in Smart Cities—Applications, Barriers, and Future Directions: A Review", *Smart Cities*, 7, 3: 1346-1389.

Wu J., Yang Y., Cheng X.U.N., Zuo H. and Cheng Z. (2020), "The development of digital twin technology review", *2020 Chinese Automation Congress (CAC) IEEE*. (testo disponibile al link: <https://ieeexplore.ieee.org/document/9327756>, 19 maggio 2025).

Ye X., Newman G., Lee C., Van Zandt S. and Jourdan D. (2023), "Toward Urban Artificial Intelligence for Developing Justice-Oriented Smart Cities", *Journal of Planning Education and Research*, 43, 1: 6-7.

Yigitcanlar T., Chorcado J.M., Mehmood R., Yi Man Li R., Mossberger K. and Desouza K. (2021), "Responsible urban innovation with local government artificial intelligence (AI): A conceptual framework and research agenda", *Journal of Open Innovation: Technology, Market, and Complexity*, 7, 71.

Yigitcanlar T., Desouza K., Butler L. and Roozkhosh F. (2020), "Contributions and Risks of Artificial Intelligence (AI) in Building Smarter Cities: Insights from a Systematic Review of the Literature", *Energies*, 13, 6: 1473.

Yigitcanlar T., Duzgun A. and Degirmenci K. (2023), "Artificial intelligence in local governments: perceptions of city managers on prospects, constraints and choices", *AI & SOCIETY*, 38: 1135-1150.

Yigitcanlar T., Senadheera S., Marasinghe R., Bibri S.E., Sanchez T., Cugurullo F. and Sieber R. (2024), "Artificial intelligence and the local government: A five-decade scientometric analysis on the evolution, state-of-the-art, and emerging trends", *Cities*, 152: 105151.

Young E., Wajcman J. and Sprejer L. (2023), "Mind the gender gap: Inequalities in the emergent professions of artificial intelligence (AI) and data science", *New Technology, Work and Employment*, 38, 3: 391-414.

Zeng J. (2022), *Artificial intelligence with Chinese characteristics: national strategy, security and authoritarian governance*, Palgrave Macmillan Singapore, Singapore.

Zuboff S. (2019), *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*, PublicAffairs, New York.

Zuiderveen Borgesius F. (2018), "Discrimination, artificial intelligence, and algorithmic decision-making", *Council of Europe, Directorate General of Democracy*. (testo disponibile al link: <https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73>, 19 maggio 2025).

Vi aspettiamo su:

www.francoangeli.it

per scaricare (gratuitamente) i cataloghi delle nostre pubblicazioni

DIVISI PER ARGOMENTI E CENTINAIA DI VOCI: PER FACILITARE
LE VOSTRE RICERCHE.



Management, finanza,
marketing, operations, HR
Psicologia e psicoterapia:
teorie e tecniche
Didattica, scienze
della formazione
Economia,
economia aziendale
Sociologia
Antropologia
Comunicazione e media
Medicina, sanità



Architettura, design,
territorio
Informatica, ingegneria
Scienze
Filosofia, letteratura,
linguistica, storia
Politica, diritto
Psicologia, benessere,
autoaiuto
Efficacia personale
Politiche
e servizi sociali

FrancoAngeli

La passione per le conoscenze

Copyright © 2025 by FrancoAngeli s.r.l., Milano, Italy. ISBN 9788891731630

FrancoAngeli

a strong international commitment

Our rich catalogue of publications includes hundreds of English-language monographs, as well as many journals that are published, partially or in whole, in English.

The **FrancoAngeli**, **FrancoAngeli Journals** and **FrancoAngeli Series** websites now offer a completely dual language interface, in Italian and English.

Since 2006, we have been making our content available in digital format, as one of the first partners and contributors to the **Torrossa** platform for the distribution of digital content to Italian and foreign academic institutions. **Torrossa** is a pan-European platform which currently provides access to nearly 400,000 e-books and more than 1,000 e-journals in many languages from academic publishers in Italy and Spain, and, more recently, French, German, Swiss, Belgian, Dutch, and English publishers. It regularly serves more than 3,000 libraries worldwide.

Ensuring international visibility and discoverability for our authors is of crucial importance to us.

FrancoAngeli



Politiche dell'intelligenza artificiale

Arene, strategie, poteri

L'intelligenza artificiale (IA) – fenomeno tecnologico che si rinnova e trasforma rapidamente e dagli impatti sociali radicali e dirompenti – si è imposta in molte agende politiche. Mentre tutti concordano sulla necessità di garantire equilibrio fra le sue innumerevoli potenzialità e i suoi rischi per gli individui, i sistemi sociotecnici e gli ordinamenti democratici, le misure adottate sono fra loro diverse. Quali sono, e da che cosa dipendono convergenze e differenziazioni? Da una prospettiva di sociologia della politica, il libro ricostruisce le influenze reciproche fra fattori economici, culturali e istituzionali e il ruolo svolto dalla varietà di forze sociali, attori e sistemi politici nel complesso campo delle “politiche dell'intelligenza artificiale”.

Per rispondere alla domanda di fondo – “chi fra governi, imprese, esperti, società civile e cittadini, in che misura e come esercita potere nelle politiche dell'IA?” – il volume presenta i risultati di un'analisi di strategie politiche orientate verso varie priorità: favorire lo sviluppo dell'innovazione tecnologica, regolare il *trade-off* tra potenzialità e rischi dei modelli e dei sistemi di IA o adottarli per finalità pubbliche. Le politiche dell'IA sono osservate da una prospettiva multiscalare, che considera specificità e interdipendenze fra le arene internazionali, sovranazionali, nazionali e locali in cui si tenta di governare l'IA. Nei capitoli sono ricostruiti la varietà degli attori coinvolti e dei loro obiettivi, le poste in gioco, i *frame*, le decisioni e i principali strumenti di *policy* che stanno orientando lo sviluppo e l'adozione dell'IA, anteponendo la priorità della tutela dei diritti umani e sociali o la deregolamentazione del suo uso, oppure cercando di conciliarli.

Ernesto d'Albergo è professore ordinario presso Sapienza, Università di Roma, dove insegna Sociologia politica e Politiche per lo sviluppo urbano sostenibile. Tra le sue recenti pubblicazioni *Sociologia della politica contemporanea* (a cura di, con G. Moini), Carocci 2024.

Giorgio Giovannelli è assegnista di ricerca presso il Dipartimento di Scienze Sociali ed Economiche di Sapienza, Università di Roma. Tra le sue recenti pubblicazioni “Urban sustainability and the local politics of digital twins: the cases of Bologna and Milan” (con E. d'Albergo), *Società Mutamento Politica*, 2024.