Istat
Istituto Nazionale di Statistica

DIGITALIZZAZIONE, INTEROPERABILITÀ E INTELLIGENZA ARTIFICIALE. DIRITTO DELLE NUOVE TECNOLOGIE

METODI · LETTURE STATISTICHE

DIGITALIZZAZIONE, INTEROPERABILITÀ E INTELLIGENZA ARTIFICIALE. DIRITTO DELLE NUOVE TECNOLOGIE

Contenuti a cura di: Giovanna Bellitti e Massimo Fedeli.

Supporto alla cura redazionale: Lara Parisella.

Attività editoriali: Nadia Mignolli (coordinamento), Claudio Bava, Alfredina Della Branca, Marco Farinacci, Alessandro Franzò e Manuela Marrone.

Responsabile per la grafica: Sofia Barletta.

ISBN 978-88-458-2157-8

STREETLIB

INDICE

PREFAZIONE

Se l'innovazione tecnologica e la ricerca scientifica sono elementi chiave per rispondere efficacemente alle sfide globali poste dalla trasformazione digitale, anche il contesto normativo di riferimento si sta evolvendo, in termini di sicurezza e riservatezza dei dati. La Pubblica amministrazione, chiamata ad affrontare la sfida dell'interoperabilità semantica, guarda con sempre maggiore attenzione le nuove metodologie e tecnologie introdotte per il trattamento dei dati; in tale ambito, l'evoluzione e la diffusione dei sistemi di intelligenza artificiale (IA) segnano una svolta. Già nell'analisi dell'interoperabilità semantica, l'IA si era preannunciata come il futuro dei sistemi informativi pubblici, e oggi l'utilizzo di questa tecnologia è presente nelle agende di tutte le amministrazioni pubbliche.

Il dibattito sugli strumenti digitali in grado di replicare determinate facoltà intellettive umane suggerisce l'opportunità di una lettura di tali temi con la prospettiva della statistica, anche alla luce della comune matrice matematica (l'idea di calcolo che sta alla base della scienza informatica è ben espressa, del resto, dalla "digitalizzazione", ossia "il ricondurre a cifre"). La statistica, infatti, può contribuire sia a fornire metodologie che costituiscono il fondamento di taluni modelli di IA (basti pensare all'impiego del teorema bayesiano per l'IA e l'apprendimento automatico), sia a trasferire in maniera adeguata, dalle persone alle macchine, la conoscenza sulle cose del mondo.

La statistica è fondamentale per soddisfare le esigenze informative necessarie per le politiche di sviluppo economico e sociale di un Paese, per l'aumento del benessere dei cittadini che lo abitano, per garantire il confronto con la realtà di altri Paesi. Tale funzione la contraddistingue oggi più di ieri; la capacità di osservare e registrare le informazioni la rende una scienza concentrata sul trattamento dei dati in tutto il loro ciclo di vita: dalla loro creazione al loro utilizzo e, oggi, finanche al loro riutilizzo. Proprio per questo l'Istat considera suo compito prioritario contribuire ad applicazioni robuste, significative ed eticamente fondate dell'intelligenza artificiale in tutti i campi rilevanti.

Per l'IA la statistica può svolgere un ruolo centrale sia a livello teorico sia pratico, quale partner naturale per la ricerca e lo sviluppo. La scienza statistica può intervenire, ad esempio, nella progettazione (riducendo di non poco i rischi di distorsione, convalidando e selezionando le variabili), nella valutazione della qualità dei dati (producendo gli standard di *performance* dei test diagnostici e il trattamento dei valori mancanti) e dei risultati prodotti. Del resto, tale evidenza emerge anche nel Regolamento (UE) 2024/3018 che ribadisce l'opportunità di affidare agli istituti nazionali di statistica funzioni nell'ambito dei sistemi nazionali di governance dei dati "*con l'obiettivo di promuovere l'integrazione e l'interoperabilità dei dati, la descrizione dei metadati, la garanzia della qualità e la definizione di norme, la condivisione e il riutilizzo dei dati*".

A ciò si unisca il rigore metodologico attraverso il quale l'Istat coniuga le esigenze di informazione statistica con il contesto normativo di riferimento, confrontando ogni fase di processo con il Regolamento generale sulla protezione dei dati (*General Data Protection Regulation* - GDPR). In tale contesto, l'Istituto sta programmando la calibrazione dell'impiego dell'intelligenza artificiale nella gestione del trattamento dei dati sui principi e sulle puntuali indicazioni dell'*AI Act*.

Si tratta di un grande sforzo, cui l'Istat da tempo ha deciso di non sottrarsi, a partire dalla tutela dei diritti fondamentali della persona e dai risvolti etici delle nuove forme di conoscenza generativa, che non possono essere messi in discussione.

Francesco Maria Chelli
Presidente dell'Istituto Nazionale di Statistica

1. PRINCIPI E REGOLE GENERALI[1]

1.1 Cenni sull'evoluzione delle tecnologie per il trattamento dei dati

Il trattamento quantitativo e qualitativo dei dati è sempre stato affiancato dall'utilizzo di metodi che tenessero conto dell'evoluzione tecnologica: nel tempo, si è sempre più fatto ricorso alle nuove tecnologie dell'informazione e della comunicazione per supportare la produzione dell'informazione statistica che, a sua volta, ha seguito un percorso evolutivo caratterizzato dalle tecnologie stesse sin dalle origini della statistica ufficiale, come descritto più avanti (Prospetto 1.1 e Figura 1.1)[2].

L'intervento manuale dell'addetto è stato, quindi, progressivamente sostituito dall'elaborazione tramite sistemi informativi di calcolo; ciò ha comportato un cambiamento decisivo compiutosi con la sperimentazione della lettura ottica, per passare, poi, all'epoca della microprogrammazione, che, nella logica di lavorazione dei dati, condurrà all'introduzione dei personal computer. Nell'ultima fase, ancora in atto, emerge la spiccata capacità della digitalizzazione di influenzare il percorso di aggiornamento dei metodi del trattamento quantitativo e qualitativo dei dati. Tale percorso conduce all'inclusione, da parte del Regolamento (UE) 2024/3018, degli istituti nazionali di statistica tra i soggetti cui affidare un ruolo attivo nei sistemi nazionali di *data governance* (cfr. paragrafo 1.2).

Il presente ebook intende illustrare tale evoluzione, alla luce dei più recenti sviluppi in ambito normativo e tecnologico. Prosegue e si completa, dunque, il percorso di analisi intrapreso per offrire una panoramica sul trattamento dei dati da parte dell'Istat; un requisito fondamentale dei sistemi di intelligenza artificiale (IA) è rappresentato, del resto, dall'interoperabilità dei sistemi informativi, vale a dire un pilastro dello sviluppo sostenibile e collaborativo dell'IA[3].

Oggetto di interesse è, in particolare, il rapporto tra tecnologia e Pubblica amministrazione, che vede la prima al servizio della seconda, consentendo il miglioramento dei servizi resi ai cittadini grazie a forme di crescente semplificazione amministrativa rese possibili dall'innovazione[4]. Il progresso in ambito tecnologico è, dunque, uno strumento imprescindibile nel perseguimento dell'azione amministrativa: nel trattamento dei dati è caratterizzato, però, da una crescente convinzione di potere trarre dai dati esistenti anche dati futuri, concedendo ai sistemi di IA la possibilità di fornirne di predittivi.

A tale proposito, occorre tenere in adeguata considerazione il rapporto tra le finalità di un sistema IA e gli aspetti etici: le finalità guidano le scelte etiche di un sistema IA, determinando come esso viene progettato, addestrato e utilizzato. Se l'obiettivo è, quindi, migliorare la giustizia sociale, le scelte etiche dovranno essere orientate a garantire equità, trasparenza e inclusione.

Un sistema IA con scopi ben definiti, ma che ignora gli aspetti etici, può portare a conseguenze negative. Allo stesso modo, un sistema IA sviluppato con un approccio etico rigoroso, ma con finalità poco chiare, potrebbe non raggiungere gli obiettivi desiderati.

1 A cura di Giovanna Bellitti, con ricerche, analisi e contributi (paragrafi 1.3 e 1.4) di Roberto Puglisi.
2 Per un riscontro di tale evoluzione fino al 1995, cfr. Geretto 2008.
3 L'Istat è soggetto attuatore del Progetto PNRR "Catalogo nazionale dati"; sul punto, si rinvia a Bellitti e Fedeli 2023.
4 Cfr. Giannini 1979.

Le finalità di un sistema IA e gli aspetti etici sono, infatti, due facce della stessa medaglia: le prime definiscono l'obiettivo del sistema, mentre i secondi ne guidano il percorso e ne garantiscono l'utilizzo responsabile; la comprensione del rapporto tra finalità e aspetti etici è dunque fondamentale per lo sviluppo e l'utilizzo responsabile dell'IA ed è necessario un approccio integrato, che tenga conto di entrambi gli aspetti, per garantire che l'intelligenza artificiale sia utilizzata nel rispetto della dignità umana e dei valori fondamentali dell'UE.

Nell'applicare tale schema, occorre fare attenzione ai pericoli scaturenti da un simile salto evolutivo nei processi di produzione e di impiego delle informazioni, pericoli che possono riassumersi nel trasferimento dall'uomo alle macchine del compito di disegnare il futuro della società. Le perplessità, dunque, anche etiche, spingono a porre la componente umana al centro di ogni ragionamento in materia di IA. Innanzitutto, occorre individuare la costante supervisione umana quale primario requisito di ricorso ai sistemi di IA: tale supervisione può abilitarsi solo garantendo, come si vedrà, la trasparenza, utile anche alla verifica delle finalità del trattamento, la responsabilizzazione dell'addetto, la robustezza e la sicurezza informatica, che, a loro volta, sono chiavi di accesso alla tutela di diritti fondamentali, come la non discriminazione e la protezione dei dati personali. Va, in sintesi, assunta una prospettiva antropocentrica che possa garantire l'allineamento dei sistemi di IA, anche sotto un profilo etico, con i principi garantiti e tutelati nell'UE.

Prospetto 1.1 - Evoluzione dei mezzi automatici di elaborazione dati e informatica (a)

Anni	Strumenti e metodi
1926-1936	Nel 1926 il lavoro di calcolo si basa sulle elaborazioni semi-manuali, il sistema di spoglio e di calcolo dei dati è basato su 20 perforatrici a mano, quattro perforatrici a regoli e 13 addizionatrici. Però, nel decennio, si sperimenta l'uso delle "Comptometer", nuove macchine prodotte in Germania, che svolgono le funzioni di addizionatrici e calcolatrici elettriche.
1936-1945	Nel 1936 si prosegue il rinnovo dell'attrezzatura: le perforatrici a mano sono sostituite da calcolatrici, selezionatrici e perforatrici elettriche, la strumentazione meccanografica utilizza ora un nuovo sistema a 45 colonne.
1946-1955	Durante questo decennio si completa il rinnovo dell'attrezzatura meccanografica, ma soprattutto nel 1948 vengono introdotte le prime macchine IBM e Remingtonm Rand che, grazie ai nuovi sistemi a 90 colonne, moltiplicano di molto la velocità di perforazione, di verifica, di selezione, tabulazione, calcolo e stampa dei prospetti statistici.
1956-1975	Nel 1970 il sistema di schede perforate segna definitivamente il passo, pertanto si inizia la sperimentazione della lettura ottica. Nel 1971 vengono acquistati gli elaboratori di terza generazione che sono impiegati per la prima volta per il Censimento dell'agricoltura, svoltosi nello stesso anno. Nel 1975 finisce l'epoca delle perforatrici, alle quali si sostituiscono prima registratrici a nastro magnetico, poi un sistema periferico di *data entry*.
1976-1985	Il rinnovo della tecnologia si completa in direzione della microprogrammazione e, nella seconda metà degli anni Settanta, sono introdotti i microcomputer e un completo sistema di *data entry*. Tale cambiamento radicale della logica di lavorazione dei dati comporta un forte investimento in corsi di formazione e/o di aggiornamento di tutto il personale. Nel 1984 vengono introdotti i personal computer.
1986-1995	L'informatica, a metà degli anni Novanta, subisce una trasformazione radicale sulla base del così detto progetto di architettura decentrata basato su una WAN, rete diffusa su tutto il territorio nazionale, che collega anche gli Uffici regionali e supporta sia Internet sia Intranet. In rete sono collegati personal computer e server dedicati alle varie rilevazioni e all'elaborazione dei dati che vengono custoditi dal server. La gestione della rete e dei server è centralizzata, il che comporta la cura della rete e salvataggi quotidiani dei dati che sono custoditi centralmente. Inoltre, si elabora il primo prototipo di sito web e si delinea una struttura sperimentale di Intranet.
1996-2005	Si introduce un nuovo programma per la gestione coordinata dei processi amministrativi (ad esempio ragioneria, personale, gestione delle funzioni amministrative della diffusione). Ormai il sito web Istat (www.istat.it) trova una struttura articolata, che è stata sostituita da una nuova *release*, maggiormente razionale e in linea con gli standard forniti dall'Autorità preposta all'informatica nel settore del pubblico. Anche la Segreteria centrale del Sistan si attrezza con un proprio sito web, che presto modificherà la sua struttura in portale che collega gli Uffici statistici degli Enti centrali, territoriali e periferici. Perciò le indagini *web-based* aumentano la velocità di esposizione dei dati e la flessibilità di consultazione degli stessi dati, liberamente scaricabili. Si producono, con maggiore facilità, estrazioni ad hoc di dati su richiesta.
2006-2015	Con l'introduzione del codice dell'amministrazione digitale, si introduce un sistema normativo organico funzionale a un migliore perseguimento degli obiettivi di digitalizzazione della Pubblica amministrazione. In tale contesto, la posta elettronica assurge a mezzo ordinario per le comunicazioni della PA e vengono definiti il documento informatico e la firma digitale; sono tutti strumenti chiamati a sostituire l'interazione analogica nei procedimenti amministrativi e destinati a segnare il passaggio, così, al procedimento e al fascicolo informatico.
2016-2024	L'ulteriore evoluzione della digitalizzazione conduce all'introduzione della *Piattaforma Digitale Nazionale Dati*, finalizzata a favorire la conoscenza e l'utilizzo del patrimonio informativo pubblico mediante la raccolta su un'unica piattaforma degli *e-service* utili per l'interscambio digitale dei dati. Si rafforza, altresì, lo sviluppo delle risorse semantiche (modelli di dati, ontologie e vocabolari controllati) destinate a essere impiegate per agevolare la conoscenza del significato dei dati.

Fonte: Elaborazione degli autori
(a) Il periodo dal 1926 al 2005 è tratto da Geretto 2008.

Figura 1.1 - Evoluzione dei mezzi e delle tecnologie per il trattamento dei dati quantitativi e qualitativi (a)

Fonte: Elaborazione degli autori
(a) Ricostruzione a cura di Paolo Nicolai, Massimiliano Spina e Lara Parisella. Si ringrazia Serenella Ravioli, Direttrice della Direzione centrale per la comunicazione, informazione e servizi ai cittadini e agli utenti, e Alexia Sasso, responsabile dei servizi bibliotecari e valorizzazione del patrimonio storico documentale dell'Istat.

Figura 1.1 segue **- Evoluzione dei mezzi e delle tecnologie per il trattamento dei dati quantitativi e qualitativi** (a)

Fonte: Elaborazione degli autori
(a) Ricostruzione a cura di Paolo Nicolai, Massimiliano Spina e Lara Parisella. Si ringrazia Serenella Ravioli, Direttrice della Direzione centrale per la comunicazione, informazione e servizi ai cittadini e agli utenti, e Alexia Sasso, responsabile dei servizi bibliotecari e valorizzazione del patrimonio storico documentale dell'Istat.

Figura 1.1 segue **- Evoluzione dei mezzi e delle tecnologie per il trattamento dei dati quantitativi e qualitativi** (a)

Fonte: Elaborazione degli autori
(a) Ricostruzione a cura di Paolo Nicolai, Massimiliano Spina e Lara Parisella. Si ringrazia Serenella Ravioli, Direttrice della Direzione centrale per la comunicazione, informazione e servizi ai cittadini e agli utenti, e Alexia Sasso, responsabile dei servizi bibliotecari e valorizzazione del patrimonio storico documentale dell'Istat.

1.2 Aspetti evolutivi del quadro normativo di riferimento

L'obiettivo di costruire un mercato unico digitale fondato sui dati e, al contempo, con regole in grado di garantire il rispetto dei diritti fondamentali è perseguito dall'UE nell'ambito della Pubblica amministrazione e nel privato con diverse prospettive tese ad abbracciare tutti i profili aventi a che fare con la circolazione dei dati. L'intelligenza artificiale rappresenta il più recente stadio evolutivo del processo di digitalizzazione e ha innescato riflessioni che hanno coinvolto anche aspetti etici e deontologici. Il ricorso sempre più insistente al potere computazionale per affrontare e soddisfare le esigenze delle persone modifica lo stesso rapporto tra cittadino e istituzioni, incide sui modelli (a questo punto, forse) tradizionali di democrazia e può spingere a rimetterne in discussione persino l'attualità, suscitando suggestioni prospettiche ispirate a forme di democrazia computazionale.

L'evoluzione del trattamento dei dati nel corso del tempo è stata affiancata da una serie di interventi normativi che hanno introdotto una disciplina intenta a garantire lo sfruttamento delle possibilità concesse dalla digitalizzazione.

Il punto di riferimento normativo in materia di IA è il Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 (*AI Act*), che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n. 300/2008, (UE) n. 167/2013, (UE) n. 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828.

In ambito nazionale, è all'esame del legislatore una disciplina che possa declinare su scala locale i principi e l'impostazione di fondo dettata dal Regolamento (UE) 1689/2024[5]. Al contempo, l'Agid (Agenzia per l'Italia digitale) ha pubblicato una Strategia italiana per l'intelligenza artificiale 2024-2026, in cui si sottolinea, tra l'altro, l'importanza di "rendere più efficienti i propri processi amministrativi e migliorare la qualità dei servizi offerti ai cittadini attraverso l'impiego di tecnologie di intelligenza artificiale", ferma restando l'esigenza di tenere costantemente presente i rischi strategici connessi all'IA: rischio del non fare, della omogeneizzazione culturale, della "iperregolazione nazionale", del *digital divide*, della inefficacia e dei rischi per il mondo del lavoro.

La notevole produzione normativa dell'UE in materia di trattamento dei dati è legata alla preoccupazione delle stesse istituzioni europee di non ammettere alcun arretramento sul fronte dei diritti fondamentali della persona. Per assicurare i migliori presupposti per tale obiettivo e, dunque, al fine di attuare con successo il quadro di governance dei dati, il Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla governance europea dei dati (*Data Governance Act*) indica gli istituti nazionali di statistica quali soggetti con competenze specifiche da considerare per la costituzione di un comitato europeo per l'innovazione in materia di dati[6]. Tale esortazione viene, dunque, recepita nel Regolamento (UE) 2024/3018 del Parlamento europeo e del Consiglio del 27 novembre 2024 che modifica il regolamento (CE) n. 223/2009 relativo alle statistiche europee, in cui si ribadisce l'opportunità di affidare agli istituti nazionali di statistica, a livello nazionale, funzioni stabilite nei quadri nazionali di governance dei dati con l'obiettivo di promuovere l'integrazione e l'interoperabilità dei dati, la descrizione dei metadati, la garanzia della qualità e la definizione di norme, la condivisione e il riutilizzo dei dati, nonché altri compiti e funzioni di cui al regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio (art. 17 *bis*, Regolamento (CE) n. 223/2009). Il coinvolgimento degli istituti di statistica deriva dalle specifiche competenze acquisite nel trattamento dei dati e sviluppate in simbiosi con l'evoluzione delle nuove tecnologie dell'informazione, e non si arresta al livello di governance nazionale; viene prescritta, infatti, la loro consultazione e partecipazione alla progettazione iniziale, allo sviluppo successivo e alla cessazione dell'uso delle fonti di dati, delle banche dati o dei sistemi di interoperabilità amministrativi (art. 26 *bis*, Regolamento (CE) n. 223/2009).

5 In tale senso è stata approvata recentemente, in Consiglio dei ministri, la bozza di disegno di legge recante la disciplina nazionale in materia di intelligenza artificiale. Dal testo emerge come il legislatore, in continuità con le disposizioni previste dall'*AI Act*, evidenzi la necessità di regolare l'impatto legato ai vantaggi che le nuove tecnologie, relative all'intelligenza artificiale, rappresentano e, al contempo, agli eventuali rischi di un utilizzo improprio. In particolare, nel documento sono indicati gli ambiti di intervento, come la Strategia nazionale, le azioni di promozione, l'introduzione di un'Autorità nazionale, la tutela dei diritti di autore e le sanzioni penali.

6 Tale impostazione si pone in linea con quanto già affermato dall'UNECE a proposito del ruolo da affidare agli istituti nazionali di statistica: "The position of NSOs may be addressed in different topics: legislation, digital policy agenda, changing functions, the international dimension, and the needs of users from policy makers to citizens. NSOs should be active in bridging these gaps. Conceptual unclearness or the so-called question of definitions is one gap where NSOs play a more active role by addressing different topics in clear professional language" (UNECE 2020).

1.2.1 Regole, definizioni e metodi per il trattamento dei dati attraverso l'uso del digitale e dell'intelligenza artificiale

La presente analisi si rivolge ai soggetti pubblici e privati che, ai sensi del Regolamento (UE) 2024/1689, siano destinati ad assumere la qualità di fornitori e/o utilizzatori di sistemi di IA nel settore del trattamento dati per la produzione di dati e informazioni quantitative e qualitative.

L'*AI Act* prevede il seguente percorso di attuazione:

- il **1° agosto 2024** (20 giorni dopo la pubblicazione nella Gazzetta ufficiale del 12 luglio 2024) è entrato formalmente in vigore;
- si applica a decorrere **dal 2 agosto 2026**, con le seguenti deroghe:
  - le disposizioni generali e il divieto sui sistemi di intelligenza artificiale considerati a rischio inaccettabile si applicano a decorrere **dal 2 febbraio 2025**;
  - le disposizioni relative all'Autorità di notifica, ai modelli per finalità generali, alla governance e alle sanzioni si applicano a decorrere **dal 2 agosto 2025**;
  - le regole di classificazione per i sistemi ad alto rischio e i relativi obblighi si applicano a decorrere **dal 2 agosto 2027**.

L'approvazione dell'*AI Act* punta a "promuovere la diffusione di un'intelligenza artificiale antropocentrica e affidabile, garantendo nel contempo un livello elevato di protezione della salute, della sicurezza e dei diritti fondamentali, compresi la democrazia, lo Stato di diritto e la protezione dell'ambiente, contro gli effetti nocivi dei sistemi di IA nell'Unione nonché promuovere l'innovazione". Il Regolamento diventa, dunque, la bussola per lo sviluppo e l'impiego dei sistemi di IA: è stato infatti adottato al termine di un prolungato dibattito tra le istituzioni europee in considerazione degli interessi in gioco.

L'*AI Act* si inserisce, ovviamente, in un panorama normativo in cui sono già disegnate le tutele per i diritti fondamentali, integrandosi o, comunque, allineandosi con esse e in particolare con: il Regolamento generale sulla protezione dei dati (*General Data Protection Regulation*), il *Digital Services Act* (DSA), il *Digital Markets Act* (DMA), il *Data Governance Act* (DGA) e la proposta di Direttiva sulla responsabilità dell'intelligenza artificiale. Nel panorama normativo va considerata, inoltre, la Direttiva 2006/54/CE del Parlamento europeo e del Consiglio del 5 luglio 2006, riguardante l'attuazione del principio delle pari opportunità e della parità di trattamento tra uomini e donne in materia di occupazione e impiego; con l'*AI Act*, quindi, l'UE persegue anche la non discriminazione di genere, affermando la necessaria neutralità dei sistemi di intelligenza artificiale.

Nel perseguimento di un ambiente europeo unico nel cui ambito assicurare un impiego dell'IA idoneo a garantire il rispetto dei diritti fondamentali, l'*AI Act* differenzia la disciplina dei sistemi IA sulla base del relativo livello di rischio. Oltre a porre un divieto generale per i sistemi IA con rischio inaccettabile, sono stabiliti requisiti specifici e obblighi per gli operatori per i sistemi di IA ad alto rischio e, per i sistemi a basso rischio, regole di trasparenza armonizzate; sono previste, poi, regole armonizzate specifiche per l'immissione sul mercato di modelli di IA di uso generale e, infine, regole sul monitoraggio del mercato, sulla governance e sull'applicazione della vigilanza del mercato stesso.

Il campo di interesse della legge europea sull'IA è delimitato da una definizione di "sistema di IA" rilevante ai fini dell'applicazione del Regolamento stesso e, dunque, idoneo a mettere in funzione l'attività di bilanciamento tra innovazione e diritti: "un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che

riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali".

Sono previsti obblighi specifici per i fornitori di modelli di IA per finalità generali, compresi i modelli di IA generativa di grandi dimensioni. Per i fornitori di modelli gratuiti e *open source* è prevista, invece, un'esenzione dagli stessi obblighi, fermi restando, comunque, gli obblighi per i fornitori di modelli di IA per finalità generali che comportino rischi sistemici.

Nel bilanciamento di interessi tra misure di protezione e innovazione, si predispongono, altresì, misure a sostegno dell'innovazione, con particolare attenzione alle piccole e medie imprese.

Nel contesto nazionale il quadro normativo di riferimento risulta sempre più composito, arricchendosi di diverse indicazioni tese a costituire il *framework* nazionale sulla digitalizzazione. Sul fronte della cybersicurezza – la sicurezza e la robustezza tecnica sono due tra i capisaldi etici nell'implementazione di sistemi IA – il decreto legislativo n. 138 del 2024 recepisce nell'ordinamento interno la Direttiva UE 2022/2555 (NIS2 - *Network* and *Information Security*), che migliora la sicurezza delle reti e delle informazioni ampliando i requisiti di protezione e gestione degli incidenti informatici per aziende e settori strategici. Allo stesso tempo, è stato adottato il Regolamento per le infrastrutture digitali e i servizi *cloud* della PA dell'Agenzia nazionale per la cybersicurezza, che mira a stabilire i livelli minimi di sicurezza per le pubbliche amministrazioni, di capacità elaborativa, di risparmio energetico e di affidabilità delle infrastrutture digitali e delle infrastrutture dei servizi *cloud* per le pubbliche amministrazioni; sono definite, dunque, le caratteristiche di qualità, di sicurezza, di *performance* e scalabilità, interoperabilità, portabilità dei servizi *cloud*. Del resto, tra i principi guida per l'implementazione di sistemi di IA figura la necessità di garantire un adeguato livello di accuratezza, robustezza e cybersicurezza e di operare in modo coerente con tali aspetti durante tutto il loro ciclo di vita.

Sul fronte strategico, il Piano triennale 2024-2026 per l'informatica nella Pubblica amministrazione si interessa direttamente del ricorso a forme di intelligenza artificiale nello svolgimento dell'attività amministrativa, indicando alcuni principi che i soggetti pubblici dovrebbero prendere in considerazione e declinare nell'ambito di ciascuna rispettiva organizzazione. Trattasi degli stessi principi ormai individuati come componenti necessarie per l'uso dell'IA.

Tale eterogeneità del quadro di regole di riferimento (cui si aggiungono, per l'ambito statistico, quelle metodologiche e di settore tematico coinvolto) impone l'adozione di un approccio multidisciplinare in grado di soddisfare adeguatamente le diverse esigenze in gioco e, al fine di perseguire un impiego delle nuove tecnologie conforme alla normativa ed etico, occorre, dunque, garantire un processo che integri i vari aspetti coinvolti[7].

1.3 Classificazione dei sistemi di IA

L'*AI Act* ha un approccio basato sui rischi per classificare diversi sistemi di IA, con l'obiettivo di bilanciarli con i benefici per il mercato unico digitale. Quindi, tenuto conto che a una maggiore gravità del rischio corrispondono regole più rigorose, si deve fare riferimento alla seguente scala di rischi: 1) rischio inaccettabile; 2) rischio alto; 3) rischio limitato; 4) rischio minimo o nullo.

7 Sul punto si rinvia al Capitolo 3, Prospetto 3.1.

Pratiche di IA vietate per rischio inaccettabile. Sono vietati i sistemi di IA che determinano un rischio inaccettabile per la sicurezza, i mezzi di sussistenza e i diritti delle persone. Non è consentito, dunque, l'impiego di sistemi di IA:

- in grado di manipolare il comportamento umano tramite l'attribuzione, ad esempio, di un "punteggio sociale" (*social scoring*), per finalità pubbliche e private, classificando le persone in base al loro comportamento sociale o alle loro caratteristiche personali;
- che sfruttano le vulnerabilità delle persone o basati su tecniche subliminali senza la consapevolezza delle persone o su tecniche volutamente manipolative o ingannevoli, aventi lo scopo o l'effetto di distorcere materialmente il comportamento di una persona o di un gruppo di persone, pregiudicando in modo considerevole la loro capacità;
- predittivi della commissione di un reato unicamente sulla base della profilazione di una persona fisica o della valutazione dei tratti e delle caratteristiche della personalità;
- che creano o ampliano le banche dati di riconoscimento facciale mediante *scraping* non mirato di immagini facciali da Internet o da filmati di telecamere a circuito chiuso;
- per inferire le emozioni di una persona fisica nell'ambito del luogo di lavoro e degli istituti di istruzione;
- di categorizzazione biometrica ovvero che classificano individualmente le persone fisiche sulla base dei loro dati biometrici per trarre deduzioni o inferenze in merito a razza, opinioni politiche, appartenenza sindacale, convinzioni religiose o filosofiche, vita sessuale od orientamento sessuale;
- di identificazione biometrica remota "in tempo reale" in spazi accessibili al pubblico a fini di attività di contrasto, salvo bilanciamento con la tutela dell'incolumità e la libertà fisica delle persone.

Sistemi a rischio alto. I sistemi di IA che possono potenzialmente avere ripercussioni negative sulla sicurezza delle persone o sui loro diritti fondamentali sono da considerarsi ad alto rischio.

Prima di immettere un sistema di IA ad alto rischio sul mercato dell'UE, o di farlo entrare in servizio, i fornitori devono effettuare una valutazione della conformità per dimostrare se il sistema di IA sia idoneo ai requisiti di affidabilità relativi alla qualità dei dati, alla documentazione e alla tracciabilità, alla trasparenza, alla sorveglianza umana, all'accuratezza, alla cybersicurezza e alla robustezza. La valutazione dovrà essere ripetuta in caso di modifica sostanziale del sistema o della sua finalità.

La valutazione consiste in una descrizione dei processi, del periodo di tempo e della frequenza in cui il sistema di IA ad alto rischio è destinato a essere utilizzato, delle categorie di persone fisiche e dei gruppi che possono essere interessati dal suo uso nel contesto specifico, dei rischi specifici di danno che possono incidere sulle categorie di persone o sui gruppi di persone interessati e in una descrizione dell'attuazione delle misure di sorveglianza umana e delle misure da adottare in caso di concretizzazione dei rischi.

I sistemi di IA che costituiscono componenti di sicurezza di prodotti disciplinati dalla legislazione settoriale dell'Unione europea saranno sempre considerati ad alto rischio, se soggetti a una valutazione della conformità da parte di terzi ai sensi della legislazione settoriale stessa. I fornitori di tali sistemi dovranno, inoltre, attuare sistemi di gestione della qualità e del rischio per garantire la conformità ai nuovi requisiti e ridurre al minimo i rischi per gli utenti e per le persone interessate, anche dopo l'immissione sul mercato di un prodotto.

I sistemi di IA ad alto rischio devono essere registrati in una banca dati pubblica dell'UE.

Sono considerati ad alto rischio i sistemi di IA:

- di identificazione biometrica remota, categorizzazione biometrica e riconoscimento delle emozioni (escluse quelle a rischio inaccettabile);
- utilizzati come componenti di sicurezza nella gestione e nel funzionamento delle infrastrutture digitali critiche, del traffico stradale e della fornitura di acqua, gas, riscaldamento ed elettricità;
- finalizzati a determinare l'accesso, l'ammissione o l'assegnazione agli istituti di istruzione e formazione professionale (ad esempio, per valutare i risultati dell'apprendimento e orientare il processo di apprendimento e il monitoraggio dei comportamenti disonesti);
- relativi alla valutazione dell'occupazione, a ottimizzare la gestione dei lavoratori e l'accesso al lavoro autonomo (ad esempio, per pubblicare annunci di lavoro mirati, analizzare e filtrare le candidature e valutare i candidati);
- usati per determinare l'accesso a servizi e a prestazioni pubblici e privati essenziali (come, ad esempio, l'assistenza sanitaria);
- finalizzati alla valutazione dell'affidabilità creditizia delle persone fisiche, alla valutazione dei rischi finanziari, nonché alla determinazione dei prezzi in relazione ad assicurazioni sulla vita e ad assicurazioni sanitarie;
- utilizzati nelle attività di contrasto, di gestione della migrazione, dell'asilo e del controllo delle frontiere, di amministrazione della giustizia, nonché nello svolgimento dei processi democratici e per la valutazione e classificazione delle chiamate di emergenza.

L'elenco dei sistemi di IA ad alto rischio, che può essere modificato per allineare la normativa all'evoluzione tecnologica, è allegato all'*AI Act*.

I sistemi di IA ad alto rischio sono accompagnati da istruzioni per l'uso in un formato digitale o non digitale appropriato, che comprendono informazioni concise, complete, corrette e chiare che siano pertinenti, accessibili e comprensibili per i *deployer*. Contengono almeno le seguenti informazioni:

a. l'identità e i dati di contatto del fornitore e, ove applicabile, del suo rappresentante autorizzato;
b. le caratteristiche, le capacità e i limiti delle prestazioni del sistema di IA ad alto rischio, tra cui:
    i. la finalità prevista;
    ii. il livello di accuratezza (comprese le metriche, di robustezza e cybersicurezza) rispetto al quale il sistema di IA ad alto rischio è stato sottoposto a prova e convalidato;
    iii. qualsiasi circostanza nota o prevedibile connessa all'uso del sistema di IA ad alto rischio in conformità della sua finalità prevista o in condizioni di uso improprio ragionevolmente prevedibile che possa comportare rischi per la salute e la sicurezza o per i diritti fondamentali di cui all'articolo 9, par. 2 dell'*AI Act*;
    iv. se del caso, le capacità e caratteristiche tecniche del sistema di IA ad alto rischio connesse alla fornitura di informazioni pertinenti per spiegarne l'output;
    v. ove opportuno, le sue prestazioni per quanto riguarda le persone o i gruppi di persone specifici sui quali il sistema è destinato a essere utilizzato;
    vi. ove opportuno, le specifiche per i dati di input o qualsiasi altra informazione pertinente in termini di *set* di dati di addestramento, convalida e prova, tenendo conto della finalità prevista del sistema di IA ad alto rischio;
    vii. se del caso, informazioni che consentano ai *deployer* di interpretare l'output del sistema di IA ad alto rischio e di usarlo in modo opportuno;

c. le eventuali modifiche apportate al sistema di IA ad alto rischio e alle sue prestazioni, che sono state predeterminate dal fornitore al momento della valutazione iniziale della conformità;
d. le misure di sorveglianza umana di cui all'articolo 14 dell'*AI Act*, comprese le misure tecniche poste in essere per facilitare l'interpretazione degli output dei sistemi di IA ad alto rischio da parte dei *deployer*;
e. le risorse computazionali e di hardware necessarie, la durata prevista del sistema di IA ad alto rischio e tutte le misure di manutenzione e cura, compresa la relativa frequenza, necessarie per garantire il corretto funzionamento di tale sistema, anche per quanto riguarda gli aggiornamenti software;
f. se del caso, una descrizione dei meccanismi inclusi nel sistema di IA ad alto rischio che consente ai *deployer* di raccogliere, conservare e interpretare correttamente i *log* in conformità dell'articolo 12 dell'*AI Act*.

Sistemi a rischio limitato. I sistemi a basso rischio si riferiscono ai pericoli legati alla scarsa trasparenza nell'impiego dell'intelligenza artificiale. L'*AI Act* impone precisi doveri di trasparenza per assicurare che le persone siano avvisate adeguatamente, favorendo così la fiducia. Per esempio, quando si usano tecnologie IA come i *chatbot*, gli utenti devono sapere che stanno dialogando con un sistema automatico, permettendo di scegliere consapevolmente se proseguire o ritirarsi. In aggiunta, i testi prodotti dall'IA e diffusi per informare la collettività su temi di rilevanza generale devono essere segnalati come creati da un'intelligenza artificiale. Questo principio si applica pure ai contenuti audiovisivi che rappresentano dei *deepfake*, ovvero foto, video e audio creati grazie a software di IA che, partendo da contenuti reali (immagini e audio), riescono a modificare o ricreare, in modo estremamente realistico, le caratteristiche e i movimenti di un volto o di un corpo e a imitare fedelmente una determinata voce.

Sistemi a rischio minimo o nullo. L'*AI Act* permette l'impiego senza restrizioni di sistemi a bassissimo rischio. Questo comprende usi come i *videogame* potenziati dall'IA o i filtri per la prevenzione dello *spam*.

Modelli per finalità generali (*General Purpose AI* - GPAI)[8] con rischio sistemico. I modelli di IA per finalità generali – compresi i modelli di IA generativa di grandi dimensioni, che possono essere utilizzati per un'ampia serie di compiti – possono comportare rischi sistemici se risultano particolarmente efficaci o molto utilizzati causando, ad esempio, incidenti gravi, o se sono utilizzati impropriamente per attacchi informatici di vasta portata. Sono classificati modelli generali a rischio sistemico quelli che presentano capacità di impatto elevato valutate sulla base di strumenti tecnici e metodologie adeguati, compresi indicatori e parametri di riferimento, ovvero sulla base di una decisione della Commissione. Per tali modelli vanno considerati eventuali codici di buone pratiche adottati a livello dell'UE.

8 Art. 3, *AI Act*: "«modello di IA per finalità generali»: un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, a eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato".

Figura 1.2 - Classificazione dei sistemi di IA

Fonte: Elaborazione degli autori

1.4 Principi per lo sviluppo e l'impiego etico di sistemi di intelligenza artificiale

"La condizione digitale è popolata di macchine pensanti. Macchine che elaborano fatti e producono idee. [...] La digitalizzazione con l'algoritmizzazione dei processi, le intelligenze artificiali e la robotica [sono] una delle frontiere più sfidanti nei confronti dei processi umani nel loro complesso, processi che da sempre coinvolgono non solo la tecnica ma anche il lato più profondo e radicale dell'autocomprensione della persona" (Benanti e Maffettone 2024). Sono queste connessioni tra etica e nuove tecnologie a suscitare l'immediato interessamento dell'UE nei confronti dei sistemi di intelligenza artificiale, con l'affermazione di alcuni principi etici: intervento e sorveglianza umani, robustezza tecnica e sicurezza, privacy e governance dei dati, trasparenza, diversità, non discriminazione ed equità, benessere sociale e ambientale e responsabilità. Successivamente, gli stessi principi vengono ripresi nell'*AI Act* e declinati in funzione del livello di rischio del sistema IA.

1.4.1 Trasparenza

I sistemi IA sono sviluppati e utilizzati in modo da consentire sia la tracciabilità dei dati e dei processi impiegati dal sistema IA, sia la spiegabilità dei processi tecnici seguiti dallo stesso, rendendo le persone consapevoli del fatto di comunicare o interagire con un sistema di IA e informando debitamente i *deployer* delle capacità e dei limiti dello stesso. In particolare, come indicato dal Gruppo di esperti di alto livello sull'IA (*High-level Expert Group on Artificial Intelligence* - AI HLEG)[9], i "*set* di dati e i processi che determinano la decisione del sistema di IA, compresi quelli di raccolta ed etichettatura dei dati, come pure gli algoritmi utilizzati, dovrebbero essere documentati secondo i migliori standard per consentire la tracciabilità e aumentare la trasparenza. Ciò vale anche per le decisioni prese dal sistema di IA, in quanto tale documentazione consente di capire perché un sistema di IA ha preso una decisione errata e, di conseguenza, potrebbe aiutare a prevenire errori futuri.

9 Commissione europea (Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale). 2019, p. 30. Testo redatto dal gruppo di esperti di alto livello sull'IA istituito dalla Commissione UE nel giugno 2018.

La tracciabilità facilita, quindi, la verificabilità e la spiegabilità"; d'altro canto, "affinché un sistema di IA possa essere tecnicamente spiegabile gli esseri umani devono potere capire e tenere traccia delle decisioni prese dal sistema stesso".

1.4.2 Responsabilità (accountability)

Gli esseri umani devono essere ritenuti responsabili delle decisioni prese dai sistemi di IA, garantendo che ci sia sempre qualcuno che possa rispondere delle conseguenze delle azioni della stessa. Ciò implica che chi crea e implementa i sistemi IA deve essere in grado di giustificare le scelte fatte durante la progettazione, l'addestramento e l'implementazione dell'algoritmo, nonché rispondere degli effetti prodotti dall'uso di questi sistemi, specialmente quando causano danni o discriminazioni.

I sistemi devono essere tracciabili e verificabili, garantendo la conservazione dell'opportuna documentazione, compresi i dati utilizzati per addestrare l'algoritmo, fondamentali per le indagini ex post. Infatti, la tracciabilità comporta la capacità di monitorare il comportamento del sistema IA, identificando gli algoritmi, i dati e le decisioni che hanno portato a un determinato risultato. In questo modo, è possibile intervenire quando si verificano problemi, come errori nei dati, *bias* (pregiudizi) o risultati ingiusti.

L'*accountability* si integra con gli altri principi e, in particolare, con il principio di equità. L'equità nell'IA riguarda il trattamento imparziale di tutte le persone e gruppi, evitando discriminazioni o trattamenti ingiusti nei confronti di minoranze, gruppi vulnerabili o qualsiasi altro gruppo sociale. Così, l'IA deve essere progettata in modo tale da ridurre al minimo i *bias* che potrebbero emergere dai dati utilizzati per addestrare i modelli o dalle scelte fatte dai progettisti. L'equità implica, dunque, che i sistemi di IA non producano disparità di trattamento tra individui o gruppi in base a caratteristiche protette come razza, genere, etnia, orientamento sessuale, disabilità, eccetera.

Si tratta di garantire che l'IA non perpetui o amplifichi disuguaglianze esistenti nella società, ma che contribuisca, invece, a un miglioramento delle condizioni di tutti. L'*accountability*, dunque, è funzionale all'equità nella misura in cui responsabilizza chi crea e implementa i sistemi IA a evitare che siano non equi.

1.4.3 Non discriminazione

I sistemi IA devono essere progettati, sviluppati e implementati in modo che non perpetuino o amplifichino pregiudizi e discriminazioni basati su genere, razza, origine etnica, religione, convinzioni personali, disabilità, età o orientamento sessuale. In particolare, l'obiettivo della non discriminazione di genere nell'IA è quello di creare sistemi che siano equi e giusti per tutte le persone, contribuendo così a una società più inclusiva.

Vanno considerate diverse azioni da introdurre.

Dati di addestramento equilibrati. I dati utilizzati per addestrare i modelli di IA devono essere rappresentativi e bilanciati rispetto ai generi. Dati sbilanciati possono portare a modelli che hanno *performance* peggiori o comportamenti discriminatori nei confronti di un genere specifico.

Analisi e mitigazione dei *bias*. Occorre progettare tecniche per identificare e mitigare i *bias* di genere nei dati e negli algoritmi. Questo può includere l'uso di metriche specifiche per valutare l'equità dei modelli e l'applicazione di algoritmi di *de-biasing*.

Progettazione inclusiva. Occorre coinvolgere persone di diversi generi e *background* nella progettazione e nello sviluppo dei sistemi di IA, per garantire che siano considerate diverse prospettive e che i sistemi siano progettati tenendo conto delle esigenze di tutti gli utenti.

Valutazione continua. Occorre monitorare continuamente le *performance* dei modelli di IA in termini di equità di genere, anche dopo la loro implementazione, per identificare e correggere eventuali problemi che possono emergere nel tempo.

Ovviamente, il principio di non discriminazione può essere rispettato integrandolo con altri principi dell'IA come la trasparenza, l'*accountability* e l'alfabetizzazione degli operatori.

1.4.4 Sorveglianza umana

Per quanto riguarda i sistemi ad alto rischio, i fornitori devono progettarli e svilupparli in modo da garantire la sorveglianza da parte di persone fisiche durante il periodo in cui sono in uso. Lo scopo è ridurre al minimo i rischi per la salute, la sicurezza o i diritti fondamentali, che possono emergere quando viene utilizzato un sistema di IA ad alto rischio. Le misure di sorveglianza sono commisurate ai rischi, al livello di autonomia e al contesto di utilizzo del sistema di IA ad alto rischio.

In generale, il principio di sorveglianza umana nell'uso dell'intelligenza artificiale, noto anche come *human oversight* o *human-in-the-loop*, esprime la necessità di mantenere un controllo e una supervisione attiva da parte degli esseri umani sui sistemi di IA, specialmente in contesti critici o decisionali.

Lo scopo del principio è garantire che l'IA venga utilizzata in modo responsabile, etico e sicuro, prevenendo potenziali danni o discriminazioni derivanti da decisioni automatizzate non supervisionate.

La sorveglianza umana è prevista in diverse fasi del processo:

- **Progettazione e sviluppo** con il coinvolgimento umano nella progettazione e nello sviluppo dei sistemi di IA, definendo gli obiettivi, i criteri decisionali e i limiti etici dell'IA.
- **Monitoraggio e intervento** con il coinvolgimento umano per potere monitorare l'operato dell'IA, intervenendo quando necessario per correggere errori, mitigare rischi o prendere decisioni che richiedono giudizio umano.

1.4.5 Accuratezza, robustezza e cybersicurezza

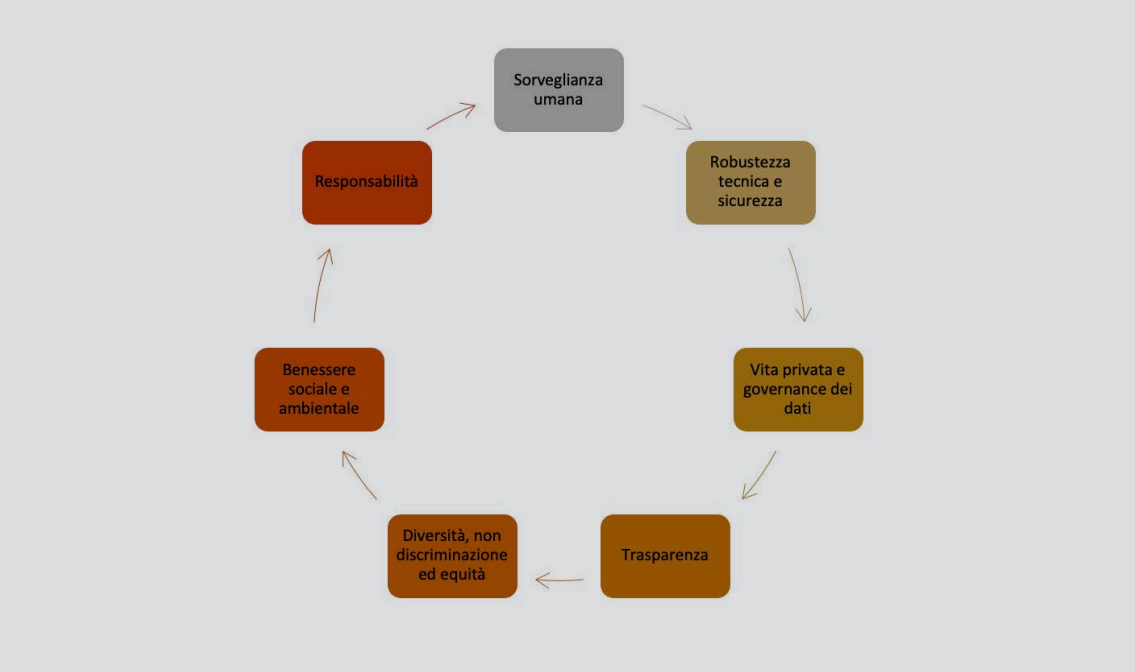
Per quanto riguarda i sistemi ad alto rischio, i fornitori devono progettarli e svilupparli in modo tale da conseguire un adeguato livello di accuratezza, robustezza e cybersicurezza e da operare in modo coerente con tali aspetti durante tutto il loro ciclo di vita; a tal fine, va fatto riferimento a eventuali parametri e metodologie di misurazione sviluppati su iniziativa della Commissione europea. Le soluzioni tecniche volte a garantire la cybersicurezza dei sistemi di IA ad alto rischio sono adeguate alle circostanze e ai rischi pertinenti e includono, ove opportuno, misure volte a prevenire, accertare, risolvere e controllare gli attacchi che cercano di manipolare il *set* di dati di addestramento (*data poisoning*, ossia "avvelenamento dei dati") o i componenti preaddestrati (*model poisoning*, ossia "avvelenamento dei modelli"), gli input progettati per fare sì che il modello di IA commetta un errore (*adversarial examples*, ossia

"esempi antagonistici", o *model evasion*, ossia "evasione dal modello"), gli attacchi alla riservatezza o i difetti del modello.

Nello specifico:

- vanno perseguiti risultati accurati e affidabili, evitando errori, distorsioni o pregiudizi indesiderati. Ciò implica l'utilizzo di dati di addestramento di alta qualità, algoritmi robusti e tecniche di validazione appropriate. I sistemi devono, dunque, essere addestrati e testati con *set* di dati sufficientemente rappresentativi per ridurre al minimo il rischio di integrare distorsioni inique nel modello e garantire che, se presenti, queste possano essere risolte mediante opportune misure di rilevazione, correzione e attenuazione;
- i sistemi di IA devono essere resistenti a eventuali guasti, attacchi o condizioni operative impreviste. Ciò include la resilienza a tentativi di manomissione, la capacità di gestire dati corrotti o mancanti e la capacità di operare in modo affidabile anche in condizioni di stress o di incertezza. I sistemi di IA ad alto rischio devono essere tecnicamente robusti per garantire che la tecnologia sia adatta allo scopo;
- vanno adottate solide misure di cybersicurezza per proteggere l'integrità dei dati, dei modelli e dei sistemi stessi da potenziali minacce informatiche, come attacchi di *hacker*, *malware* o violazioni dei dati. Ciò implica l'adozione di *best practice* di sicurezza informatica, come la crittografia, l'autenticazione robusta, la gestione degli accessi e la protezione delle comunicazioni di rete.

Figura 1.3 - Principi elaborati dall'AI HLEG



Fonte: Elaborazione degli autori sulla base del testo redatto dal Gruppo di esperti di alto livello sull'IA istituito dalla Commissione europea nel giugno 2018

1.5 Addestramento dei sistemi

Accanto agli accennati requisiti in termini di prestazione, accuratezza e robustezza, il fornitore di sistemi di IA deve garantire un adeguato addestramento del sistema con dati di elevata qualità, al fine di evitare il rischio di produrre output che non risultino in linea con i principi dell'IA, come, ad esempio, individuare le persone in modo discriminatorio o altrimenti errato o ingiusto. I dati sono, infatti, lo strumento non solo per addestrare, ma anche per convalidare e provare i sistemi di IA.

La disponibilità di dati di alta qualità assurge, dunque, a requisito essenziale per garantire le prestazioni di molti sistemi di IA, in particolare quando si utilizzano tecniche che prevedono l'addestramento di modelli.

I *set* di dati di addestramento, convalida e prova, incluse le etichette, dovrebbero essere pertinenti, sufficientemente rappresentativi e, nella misura del possibile, esenti da errori e completi nell'ottica della finalità prevista del sistema. I risultati forniti dai sistemi di IA potrebbero essere, infatti, influenzati da distorsioni intrinseche destinate ad aumentare gradualmente e quindi a perpetuare e amplificare le discriminazioni esistenti, in particolare nei confronti delle persone che appartengono a determinati gruppi vulnerabili per genere o etnia.

La qualità del dato comprende, evidentemente, anche la possibilità di garantire il rispetto del diritto alla vita privata e alla protezione dei dati personali durante l'intero ciclo di vita dei sistemi di IA. Così, nella loro progettazione, i fornitori devono includere non solo l'anonimizzazione e la cifratura, ma anche l'uso di tecnologie di addestramento senza trasmissione tra le parti di dati grezzi o strutturati o copia degli stessi.

Anche laddove il sistema di IA sia la risultante dell'integrazione di più parti, con strumenti e servizi, componenti o processi forniti da soggetti diversi, appare fondamentale garantire la catena del valore del sistema di IA. Il fornitore che integra le parti per varie finalità (addestramento dei modelli, riqualificazione dei modelli, prova e valutazione dei modelli, integrazione nel software o altri aspetti dello sviluppo dei modelli) deve essere sempre nelle condizioni di adempiere pienamente gli obblighi di legge.

LINGUAGGIO E SEMANTICA

Ai fini della qualità del dato, è utile considerare, altresì, la distinzione tra linguaggio e semantica, concetti strettamente correlati ma distinti. Da una parte, il linguaggio si riferisce alla forma con cui le informazioni vengono espresse: esiste il linguaggio naturale (ad esempio l'italiano), il linguaggio di programmazione (ad esempio *Python*, *Java*) o il linguaggio formale (ad esempio la logica matematica); dall'altra, la semantica si riferisce al significato delle informazioni: è l'interpretazione del linguaggio, il modo in cui le parole, le frasi e le strutture linguistiche vengono comprese e associate a concetti, entità e relazioni nel mondo reale o in un contesto specifico. In altre parole, la semantica si occupa del contenuto della comunicazione. Con riferimento all'IA, tale distinzione è funzionale alla comprensione del linguaggio naturale da parte delle macchine nella misura in cui: gli algoritmi di elaborazione del linguaggio naturale (*Natural Language Processing* - NLP) devono essere in grado di estrarre il significato (semantica) dalle frasi, non solo di analizzare la struttura grammaticale (linguaggio); i sistemi di generazione del testo devono essere in grado di produrre frasi che abbiano un significato coerente e appropriato (semantica), non solo frasi grammaticalmente corrette (linguaggio); i sistemi di dialogo devono comprendere il significato delle domande degli utenti (semantica) per rispondere a domande in modo pertinente, non solo per individuare parole chiave (linguaggio); per estrarre informazioni da testi, i sistemi di estrazione di informazioni devono essere in grado di identificare le entità e le relazioni descritte nel testo (semantica), non solo di analizzare la struttura del testo (linguaggio).

QUALITÀ DEI DATI E FUNZIONE STATISTICA

Il requisito della qualità del dato per l'addestramento dei sistemi di IA si concilia bene con i princìpi guida per il trattamento statistico dei dati a livello europeo raccolti nel Codice delle statistiche europee, adottato con la convinzione che la qualità sia la base del vantaggio competitivo del Sistema statistico europeo "in un contesto mondiale caratterizzato da una crescente tendenza verso informazioni istantanee di cui spesso non è attestata la qualità". L'esperienza maturata nel trattamento dati per finalità statistiche costituisce una risorsa preziosa per garantire la qualità dei dati sui quali i sistemi di IA costruiscono la propria conoscenza. È un valore aggiunto sviluppare sistemi IA con la certezza che la "qualità è un imperativo per le autorità statistiche, che individuano sistematicamente e regolarmente i punti di forza e di debolezza al fine di migliorare costantemente la qualità dei processi e dei prodotti" (principio 4 del Codice delle statistiche europee).

La qualità va perseguita assicurando strutture organizzative e strumenti adeguati, così come solide metodologie (principio 7) impiegate all'insegna dell'imparzialità e dell'obiettività (principio 6); adottando procedure per pianificare, monitorare e migliorare l'integrazione di dati provenienti da più fonti (principio 4.2) e un costante monitoraggio e valutazione della qualità; garantendo, infine, ampia trasparenza sulle metodologie applicate e sulla qualità dei prodotti (principio 15).

1.6 Intelligenza artificiale e protezione dei dati personali

L'innovazione tecnologica comporta un cambio di paradigma incentrato sul trattamento dei dati (Bellitti e Fedeli 2022). Laddove vengano in gioco dati di natura personale, se non adeguatamente regolato secondo i principi di proporzionalità e minimizzazione, il trattamento rischia di provocare violazioni capaci di determinare discriminazioni e disuguaglianze.

Il legislatore europeo ha voluto, dunque, favorire la circolazione dei dati costruendo un ecosistema reso sicuro e protetto dal GDPR, la cui necessaria applicazione va ribadita anche con riferimento ai rapporti tra impiego di forme di IA e protezione dei dati personali (Bellitti e Colasanti 2021).

Il GDPR già considera le possibili ricadute del ricorso a sistemi automatizzati per il trattamento dei dati personali, affermando il diritto di chiunque a non essere sottoposto a una decisione basata unicamente su un trattamento automatizzato dei propri dati personali e che produca effetti giuridici nei confronti dell'interessato, "quali il rifiuto automatico di una domanda di credito online o pratiche di assunzione elettronica senza interventi umani" (art. 22 e Considerando 71)[10]. Nella pratica, l'operatore del sistema IA deve individuare il giusto bilanciamento tra le possibilità offerte dalla potenza computazionale per finalità predittive e gli interessi individuali. Con particolare riferimento all'attività di trattamento dati per finalità statistiche, sono da tenere in considerazione le "Nuove regole deontologiche per trattamenti a fini statistici o di ricerca scientifica effettuati nell'ambito del Sistema statistico nazionale", in corso di aggiornamento.

In generale, riguardo ai rapporti tra l'impiego di IA e il GDPR, l'*European Data Protection Board* (EDPB) fornisce indicazioni sull'applicazione del GDPR agli *AI Models* concentrandosi sull'anonimizzazione di un modello di IA, sull'interesse legittimo come base giuridica e sugli effetti derivanti dall'uso di dati trattati illecitamente[11]. Si afferma, dunque, la necessità di verificare il grado di anonimizzazione di un modello IA attraverso una valutazione accurata in base alle casistiche; l'EDPB pone, poi, in risalto la necessità di dimostrare, in assenza di soluzioni alternative, come l'interesse legittimo possa essere una valida base giuridica per il trattamento dei dati personali e per il conseguente raggiungimento delle finalità fornendo indicazioni per l'utilizzo di un test in tre fasi; infine, si ribadisce come lo sviluppo di un modello di IA con dati personali trattati illecitamente impedisca l'impiego del modello stesso a meno che, nella fase di utilizzo, si dimostri che i dati personali non siano più oggetto di trattamento.

L'*European Data Protection Supervisor* (EDPS) ha pubblicato le prime linee guida per l'utilizzo dei sistemi di intelligenza artificiale generativa destinate alle istituzioni pubbliche e agli organi dell'UE, con l'intento di fornire un quadro dettagliato per assicurare la protezione dei dati personali in *compliance* al Regolamento (UE) 2018/1725[12].

---

10 Il riferimento è anche alle tecniche di profilazione per la previsione di aspetti riguardanti il rendimento professionale, la situazione economica, la salute, le preferenze o gli interessi personali, l'affidabilità o il comportamento, l'ubicazione o gli spostamenti dell'interessato.

11 European Data Protection Board, *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*, 17 dicembre 2024 (https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_it).

12 European Data Protection Supervisor 2024 (https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf).

Il documento fornisce un ottimo spunto di riflessione e valutazione a livello organizzativo. Nel testo vengono posti in rilievo diversi elementi, tra i quali: la definizione di IA generativa, i requisiti per l'uso di tale tecnica, la valutazione di impatto (*Data Protection Impact Assessment* - DPIA), le modalità di garanzia al principio di minimizzazione dei dati, il rispetto del principio di esattezza del dato e le modalità per informare adeguatamente gli individui. Da questi ultimi due punti si evince come sia necessario progettare tali sistemi riducendo il rischio di eventuali errori e, allo stesso tempo, assicurare l'affidabilità e la trasparenza attraverso un monitoraggio sull'algoritmo mediante la supervisione umana. Dal punto di vista organizzativo, viene evidenziato come sia necessario il coinvolgimento dell'Ufficio legale, del *Data Protection Officer* (DPO), del Servizio IT e del Responsabile locale della sicurezza informatica su eventuali sviluppi o utilizzi di sistemi di IA generativa, per verificarne la *compliance* alla normativa vigente.

Da tale ultima impostazione emerge come l'EDPS, sulla scia dell'*AI Act*, funga da precursore sul tema, fornendo ricchi spunti in termini di valutazione di impatto, organizzazione e, in particolare, di esattezza del dato personale, proprio perché l'IA generativa potrebbe, in un certo senso, attraverso l'inferenza, alterarlo e quindi generare dei rischi per la protezione dello stesso. Proprio per questo, il punto focale si concentra sulla valutazione di impatto e del rischio: sarà dunque indispensabile incrementare tale procedura, mediante fasi di test accurate per tutto il ciclo di vita del software.

Ai fini dell'attuazione dei principi di *accountability* e di *privacy by design & by default* (artt. 5 e 25 del GDPR), sin dalla fase di progettazione occorre considerare i risvolti dell'impiego di sistemi IA sulla protezione dei dati personali. Risulterà importante, dunque, la predisposizione della valutazione di impatto, prevista dall'art. 35 del GDPR, nei casi di uso di nuove tecnologie con il rischio concreto per i diritti e la libertà delle persone fisiche.

Il fornitore del sistema di IA deve, dunque, produrre adeguata documentazione sul ciclo di vita, sulle condizioni e sul funzionamento del prodotto nell'apposito contratto di licenza che fornirà all'utilizzatore, che a sua volta dovrà bilanciare gli interessi in gioco attraverso un'accurata analisi.

2. RUOLI NEL PROCESSO DI TRATTAMENTO ATTRAVERSO L'UTILIZZO DELL'INTELLIGENZA ARTIFICIALE[1]

2.1 I ruoli previsti nell'*AI Act*

Nell'ambito della disciplina generale dell'UE dei sistemi di IA, sono stabilite diverse responsabilità soggettive in funzione del ruolo svolto rispetto allo sviluppo e all'utilizzo dei sistemi di IA.

Viene, dunque, prevista la categoria generale dell'operatore di un sistema di IA, che raccoglie le seguenti figure:

- il *provider*/fornitore è il soggetto che sviluppa, produce o distribuisce sistemi di intelligenza artificiale. Il fornitore a valle è un fornitore di un sistema di IA, compreso un sistema di IA per finalità generali, che integra un modello di IA, indipendentemente dal fatto che sia fornito dallo stesso e integrato verticalmente o fornito da un'altra entità sulla base di relazioni contrattuali;
- il *deployer*/utilizzatore è il soggetto che utilizza sistemi di intelligenza artificiale per vari scopi;
- il distributore è il soggetto nella catena di approvvigionamento, diversa dal fornitore o dall'importatore, che mette a disposizione un sistema di IA sul mercato dell'Unione europea;
- l'importatore è il soggetto che immette sul mercato un sistema di IA recante il nome o il marchio di una persona fisica o giuridica stabilita in un paese terzo;
- il fabbricante del prodotto è il soggetto che immette sul mercato o mette in servizio un sistema di IA insieme al proprio prodotto e con il proprio nome o marchio;
- il rappresentante autorizzato è il soggetto che ha ricevuto e accettato un mandato scritto da un fornitore di un sistema di IA o di un modello di IA per finalità generali al fine di adempiere ed eseguire per suo conto gli obblighi e le procedure stabiliti dal Regolamento (UE) 2024/1689.

Come accennato (cfr. Capitolo 1, paragrafo 1.3), l'*AI Act*, quale principale punto di riferimento normativo in materia di intelligenza artificiale, individua requisiti e responsabilità dei vari attori nella catena del valore dell'IA. Tra le figure sopra elencate, due assumono un ruolo chiave: i fornitori e i *deployer* di sistemi di IA. La classificazione come fornitore o come *deployer* dipenderà da una valutazione di carattere fattuale, non potendosi limitare a quanto eventualmente previsto in accordi formali. Si possono fare, a tale proposito, analoghe considerazioni a quelle riguardanti l'applicazione del GDPR, dove la distinzione tra titolare e responsabile non dipende solo da quanto previsto nel contratto, ma anche dalla realtà operativa e dalla pratica quotidiana con cui vengono gestiti i dati personali; così come per la protezione dei dati personali, dunque, anche per i sistemi IA, entrambe le dimensioni (formale e fattuale) devono essere considerate per determinare il ruolo di ciascun soggetto, in modo che il ricorso all'IA rispetti i diritti fondamentali del mercato unico dell'UE. Nella pratica, il confine tra fornitore e *deployer* potrebbe, così, risultare talvolta sottile.

1 Coordinato da Roberto Puglisi, con contributi di Roberto Puglisi (paragrafi 2.1, 2.4 e 2.5) e di Paolo Nicolai (paragrafi 2.2 e 2.3 e Figure).

A seconda del ruolo svolto dal soggetto riguardo al sistema di IA emergono diversi obblighi proporzionati, in ogni caso, al relativo livello di rischio.

Per i sistemi IA ad alto rischio, il fornitore assume la responsabilità della conformità del sistema con i requisiti previsti dall'*AI Act* occupandosi di rendere conforme il sistema IA alla normativa UE, istituendo un sistema di gestione della qualità e un sistema di gestione del rischio e rendendo disponibile la relativa documentazione tecnica e la dichiarazione di conformità. Il *deployer*, viceversa, adotta idonee misure tecniche e organizzative per garantire di utilizzare tali sistemi conformemente alle istruzioni per l'uso che li accompagnano – conserva i *log* generati automaticamente dal sistema; garantisce che i dati di input siano pertinenti e sufficientemente rappresentativi; verifica l'avvenuto rispetto dell'obbligo di registrazione nella banca dati dell'UE per i sistemi ad alto rischio – e, con riferimento a taluni settori specificamente individuati dall'*AI Act*, qualora sia un organismo di diritto pubblico o un ente privato fornitore di servizi pubblici, effettua una valutazione dell'impatto sui diritti fondamentali che l'uso di tale sistema può produrre (*Fundamental Rights Impact Assessment* - FRIA). La FRIA è in stretta correlazione con la valutazione di impatto sulla protezione dei dati personali (DPIA): se alcuni diritti ricadono sotto l'area della DPIA e della FRIA, le stesse devono essere effettuate contemporaneamente.

Per i sistemi a rischio minimo (ossia non classificabile come a rischio inaccettabile o a rischio alto), i fornitori e i *deployer* devono rispettare obblighi di trasparenza variabili in funzione delle finalità del sistema IA.

Per i sistemi IA per finalità generali (GPAI)[2] senza rischio sistemico[3], i fornitori hanno obblighi relativi al rispetto del diritto di autore, alla trasparenza circa i contenuti utilizzati, alla documentazione tecnica del modello. Laddove il sistema GPAI presenti un rischio sistemico, invece, per i fornitori si aggiungono obblighi relativi alla sicurezza informatica, alla valutazione e alla mitigazione dei rischi, all'esecuzione di valutazioni del modello, all'adesione a un codice di condotta e alla notifica all'Ufficio IA dell'UE di incidenti gravi e possibili misure correttive.

Dunque, la conformità alle prescrizioni dell'*AI Act* e, soprattutto, il conseguente perseguimento dei diritti fondamentali dipende da una piena consapevolezza circa il ruolo assunto rispetto al sistema IA di volta in volta interessato. Una valutazione essenziale da compiere preliminarmente si appunta, quindi, sull'individuazione, da parte del soggetto che decide di ricorrere a sistemi di IA, della figura soggettiva di riferimento considerata dall'*AI Act* (art. 3) rispetto alla fattispecie concreta.

La definizione dei ruoli con riferimento ai sistemi IA deve tenere conto, altresì, dei codici di condotta, la cui adozione da parte dei fornitori e dei *deployer* è incoraggiata e agevolata da parte dell'Ufficio per l'IA[4]; l'intento è perseguire obiettivi chiari e indicatori chiave di prestazione volti a misurare il conseguimento degli stessi relativamente agli orientamenti etici dell'UE per un'IA affidabile, alla riduzione al minimo dell'impatto dei sistemi di IA sulla so-

2 Cfr. Capitolo 1.

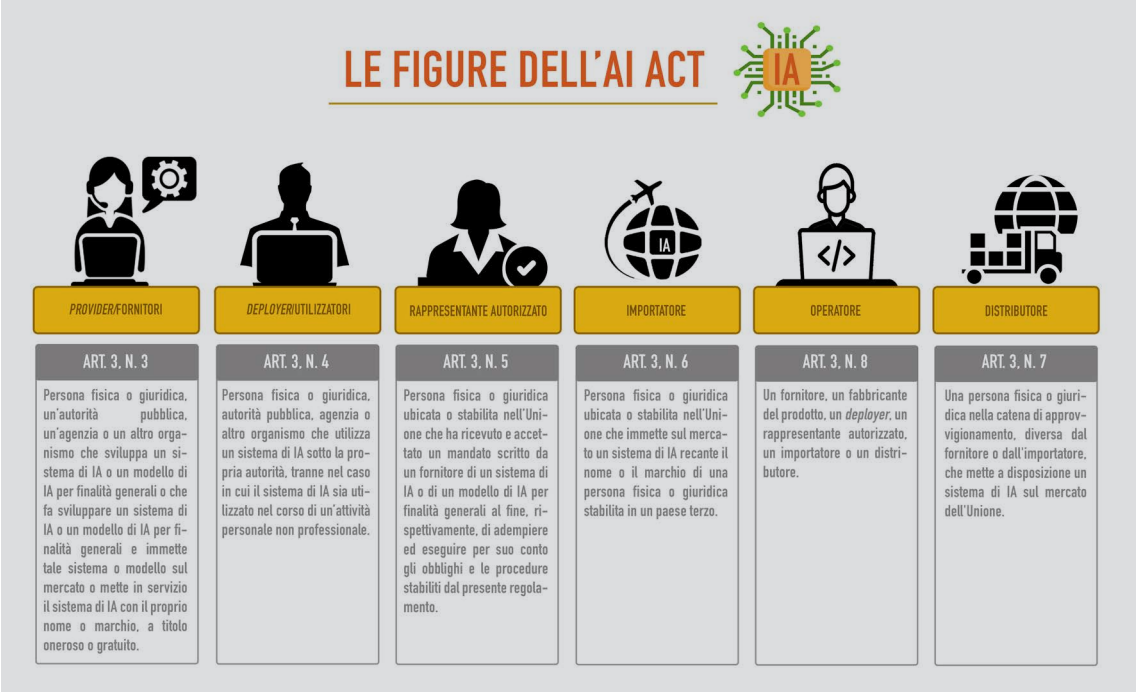
3 L'*AI Act* considera rischio sistemico "un rischio specifico per le capacità di impatto elevato dei modelli di AI per finalità generali, avente un impatto significativo sul mercato dell'Unione a causa della sua portata o di effetti negativi effettivi o ragionevolmente prevedibili sulla salute pubblica, la sicurezza, i diritti fondamentali o la società nel suo complesso, che può propagarsi su larga scala lungo l'intera catena del valore".

4 Decisione della Commissione del 24 gennaio 2024 che istituisce l'Ufficio europeo per l'intelligenza artificiale (C/2024/1459), che ha la missione di sviluppare competenze e capacità dell'Unione europea nel settore dell'IA e di contribuire all'attuazione del diritto dell'Unione in materia di IA. Si tratta della funzione della Commissione volta a contribuire all'attuazione, al monitoraggio e alla supervisione dei sistemi di IA e dei modelli di IA per finalità generali e della governance dell'IA. Così, i riferimenti all'ufficio per l'IA contenuti nel l'*AI Act* si intendono fatti alla Commissione UE.

stenibilità ambientale, alla promozione dell'alfabetizzazione in materia di IA, alla facilitazione di una progettazione inclusiva e diversificata dei sistemi di IA e alla prevenzione dell'impatto negativo dei sistemi di IA sulle persone vulnerabili.

Tale assetto previsto per l'impiego dell'intelligenza artificiale deve coniugarsi con le responsabilità riguardanti altri aspetti relativi al trattamento di dati nell'ambito dello svolgimento dell'attività istituzionale; l'obiettivo è assicurare una gestione organica dei processi dell'organizzazione in grado di affrontare tutte le implicazioni normative e relativi adempimenti (cfr. Capitolo 3) legati al trattamento di dati (IA, protezione dati personali, *cybersecurity* e digitalizzazione).

Figura 2.1 - Definizione e compiti dei soggetti preposti allo sviluppo e all'utilizzo di sistemi di IA (art. 3 dell'*AI Act*)



Fonte: Elaborazione degli autori sull'*AI Act*

2.2 Protezione dei dati personali

Nell'organizzazione del trattamento dei dati, laddove ricorrano categorie ricadenti sotto l'egida del GDPR, devono, ovviamente, essere tenuti in considerazione i ruoli individuati dalla disciplina stessa, per garantire l'effettiva protezione dei dati personali.

Tra le figure previste dal Regolamento, riveste un ruolo centrale il titolare (*data controller*) che, oltre ad avere potere decisionale, definisce le finalità e i mezzi del trattamento, avendo l'onere di implementare adeguate misure tecniche organizzative e di garantire la conformità del trattamento al GDPR. Il Regolamento stabilisce gli obblighi a cui il titolare deve attenersi, ma occorre considerare in particolare l'art. 5, comma 2 che introduce il principio di *accountability*, che pone un'accezione diversa rispetto al concetto classico di responsabilizzazione. Il titolare è incentivato a operare in autonomia attraverso un approccio innovativo e proattivo, tenendo in considerazione i principi (liceità, correttezza, trasparenza, limitazione delle finalità, minimizzazione, esattezza, limitazione della conservazione,

integrità e riservatezza) che caratterizzano l'equo trattamento dei dati personali. Sarà quindi onere dello stesso titolare dimostrare e dare conto riguardo alla corretta gestione del trattamento e all'adozione di misure tecniche idonee, mettendo al centro la tutela dei dati sin dalla progettazione.

Nel caso in cui due o più titolari individuino e definiscano di concerto le finalità e i mezzi del trattamento, ai sensi dell'art. 26 del GDPR, sono considerati contitolari (*joint controller*): tra gli elementi principali per stabilire la contitolarità risulta determinante la partecipazione congiunta alle operazioni di trattamento e l'indissolubilità dell'accordo.

Tra i ruoli principali indicati dal Regolamento, indubbiamente, occorre considerare le funzioni e i compiti svolti dal Responsabile del trattamento (*data processor*): per poter ricoprire tale ruolo il responsabile deve risultare un soggetto diverso dal titolare e trattare i dati per conto di quest'ultimo, rispettando gli obblighi previsti dall'art. 28 del GDPR, fornendo adeguate garanzie e mettendo in atto misure tecniche organizzative che garantiscano la tutela dei diritti dell'interessato. I trattamenti da effettuare dal responsabile sono regolati attraverso un contratto o da altro atto giuridico; inoltre, non possono essere trattati dati personali al di fuori delle istruzioni impartite dal titolare e il responsabile può avvalersi di un sub-responsabile solo previa autorizzazione scritta del titolare.

Il titolare o il responsabile, qualora lo ritengano opportuno, possono avvalersi, sotto la propria autorità, del soggetto autorizzato o del designato al trattamento: sono due figure che essenzialmente coincidono e sono disciplinate rispettivamente dal GDPR (art. 29) e dal Codice in materia di protezione dei dati personali (art. 2 *quaterdecies* del d.lgs. 196/2003). La disciplina europea ne individua un ruolo esclusivamente operativo, nel rispetto delle istruzioni impartite dal titolare e previa adeguata formazione, mentre quella nazionale attribuisce al designato, che opera sotto il controllo del titolare, compiti e funzioni specifici.

Tra le figure chiave, assume una posizione di rilievo il responsabile per la protezione dei dati personali (*data protection officer*). Per ricoprire tale ruolo il RPD deve possedere un'adeguata esperienza, competenza e formazione in materia di protezione dei dati personali e, per esercitare al meglio le proprie funzioni, deve soddisfare i requisiti di imparzialità e indipendenza. Questa figura è fondamentale e funzionale alle attività che caratterizzano il trattamento dei dati personali in quanto, ai sensi dell'art. 39 del Regolamento: fornisce consulenza al titolare o al responsabile del trattamento, nonché ai dipendenti che eseguono il trattamento in merito agli obblighi normativi; svolge la funzione di sorveglianza sulla corretta osservanza del GDPR, nonché delle politiche del titolare del trattamento o del responsabile del trattamento in materia di protezione dei dati personali, comprese l'attribuzione delle responsabilità, la sensibilizzazione e la formazione del personale che partecipa ai trattamenti e alle connesse attività di controllo; fornisce, se richiesto, un parere in merito alla valutazione di impatto sulla protezione dei dati, sorvegliandone lo svolgimento. Coopera, inoltre, con l'autorità di controllo, funge da punto di contatto per l'autorità stessa per questioni connesse al trattamento – tra cui la consultazione preventiva – e, se del caso, effettua consultazioni relativamente alla protezione dei dati personali.

Figura 2.2 - Definizione e compiti dei soggetti preposti al trattamento dei dati personali

Fonte: Elaborazione degli autori sul Codice in materia di protezione dei dati personali e sul Regolamento generale sulla protezione dei dati

2.3 Cybersicurezza

L'alta frequenza di incidenti informatici ha spinto il legislatore europeo a riconsiderare la disciplina sulla cybersicurezza[5] attraverso l'introduzione di nuove disposizioni volte a rafforzare la sicurezza delle reti e dei sistemi informatici in tutti i settori. Le normative in materia stabiliscono una serie di requisiti necessari per garantire un elevato livello di sicurezza e, al contempo, impongono obblighi cogenti per i fornitori di servizi essenziali e gli enti pubblici.

Punto di riferimento è la Direttiva (UE) 2022/2555 del Parlamento europeo e del Consiglio del 14 dicembre 2022 relativa a misure per un livello comune elevato di cybersicurezza nell'Unione, recante modifica del Regolamento (UE) n. 910/2014 e della Direttiva (UE) 2018/1972 e che abroga la Direttiva (UE) 2016/1148. Tale Direttiva sul *Network and Information Security* in ambito UE (NIS2), approvata nel 2022 e in vigore dal 2024, amplia il campo di applicazione e introduce nuove misure obbligatorie di sicurezza per le aziende e le organizzazioni di vari settori essenziali, come energia, trasporti, finanza, sanità, telecomunicazioni e gestione delle infrastrutture digitali, e stabilisce misure volte a garantire un livello comune elevato di cybersicurezza nell'Unione europea, in modo da migliorare il funzionamento del mercato interno.

A tale fine, prevede:

- obblighi che impongono agli Stati membri di adottare strategie nazionali in materia di cybersicurezza e di designare o creare autorità nazionali competenti, autorità di gestione delle crisi informatiche, punti di contatto unici in materia di sicurezza e *team* di risposta agli incidenti di sicurezza informatica (*Computer Security Incident Response Team* - CSIRT);

5 Direttiva (UE) 2016/1148 del Parlamento europeo e del Consiglio del 6 luglio 2016, recante misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione.

- misure in materia di gestione dei rischi di cybersicurezza e obblighi di segnalazione per i soggetti di un tipo di cui all'allegato I o II nonché per soggetti identificati come critici ai sensi della Direttiva (UE) 2022/2557;
- norme e obblighi in materia di condivisione delle informazioni sulla cybersicurezza;
- obblighi in materia di vigilanza ed esecuzione per gli Stati membri.

Sul fronte interno, la NIS2 è stata recepita con il decreto legislativo 4 settembre 2024, n. 138, entrato in vigore il 18 ottobre 2024: tale misura di recepimento mira a rendere il Paese più resiliente agli attacchi informatici e stabilisce una struttura di gestione e coordinamento delle crisi, assegnando nuovi ruoli all'Agenzia per la cybersicurezza nazionale, che fungerà da autorità centrale e punto di contatto nazionale per la *cybersecurity*.

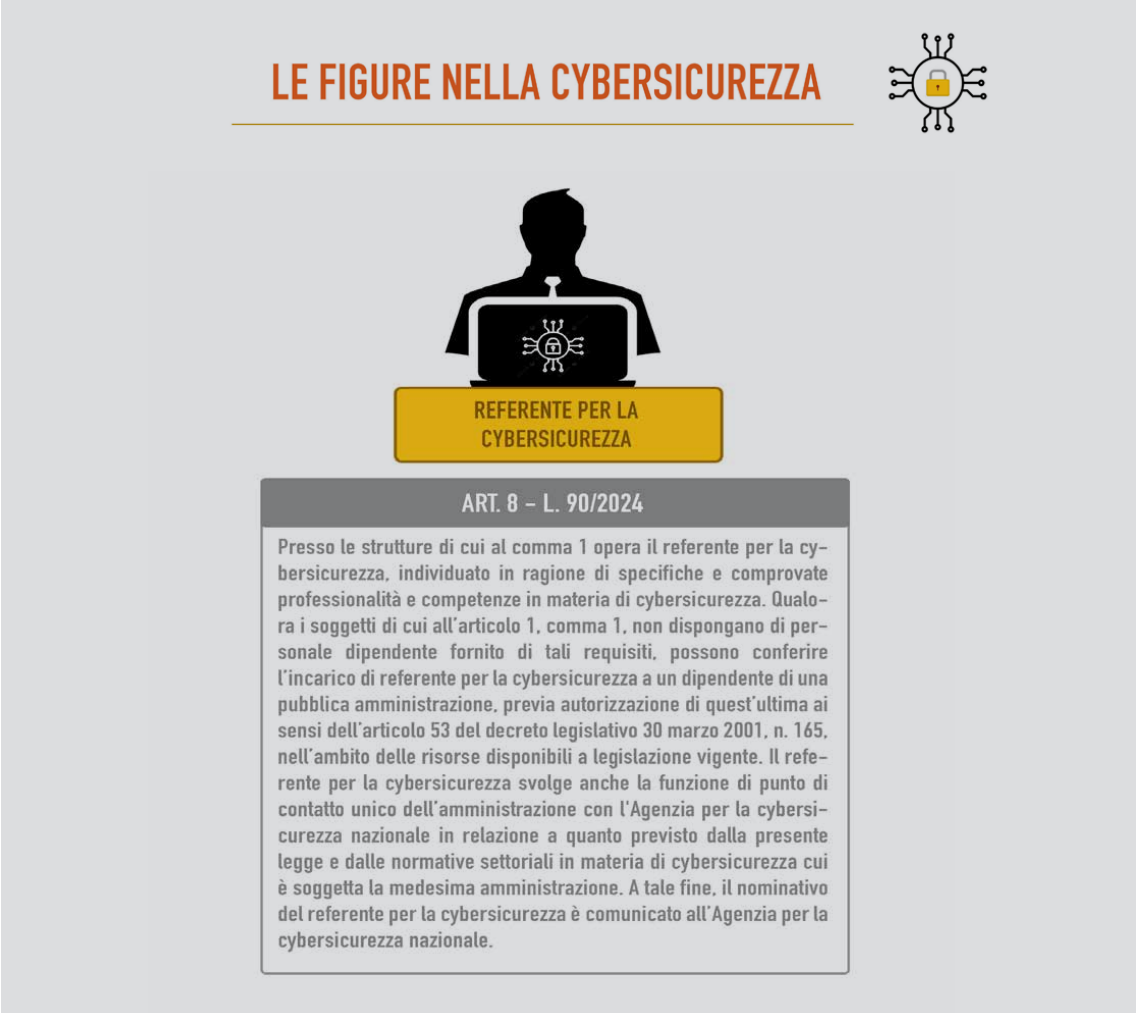
Tra gli aspetti di maggiore rilievo sulle figure principali che sono coinvolte nei processi di mitigazione del rischio vanno considerate, poi, altre indicazioni europee. Sul punto, l'Agenzia dell'Unione europea per la cybersicurezza (*European Union Agency for Cybersecurity* - Enisa), che svolge l'incarico di individuare le strategie e stabilire i punti cardine per prevenire gli attacchi informatici nell'UE, ha introdotto il quadro europeo per le competenze in materia di cybersicurezza (*European Cybersecurity Skills Framework* - ECSF). Tale documento è utile per la diffusione della cultura della cybersicurezza e offre una panoramica dettagliata sulla tipologia di profili dei professionisti[6] che operano nel settore e sui livelli di competenza essenziali per svolgere le attività professionali in tale campo. Sulla scia delle disposizioni derivanti dalla normativa europea, il legislatore nazionale è intervenuto, nell'ambito della cybersicurezza, attraverso l'adozione della legge 28 giugno 2024, n. 90 che reca "Disposizioni in materia di rafforzamento della cybersicurezza nazionale e di reati informatici", al fine di potenziare gli strumenti di prevenzione e contrasto dei reati informatici, e ne ha definito le prescrizioni e gli adempimenti per le pubbliche amministrazioni. In particolare, ha evidenziato come le PA debbano istituire una struttura per la cybersicurezza con il compito di: sviluppare politiche e procedure per la sicurezza delle informazioni; elaborare e aggiornare un piano per i rischi informatici e un piano programmatico per la sicurezza dei dati, sistemi e infrastrutture; predisporre e aggiornare un documento con la definizione dei ruoli e l'organizzazione della sicurezza delle informazioni; pianificare e attuare interventi di potenziamento delle capacità per la gestione dei rischi informatici in relazione ai piani elaborati; pianificare e attuare le misure previste nelle linee guida per la cybersicurezza emanate dall'Agenzia per la cybersicurezza nazionale; verificare che i programmi e le applicazioni informatiche e di comunicazione elettronica in uso siano adeguate alle linee guida sulla crittografia e sulla conservazione delle *password* adottate dall'ACN e dall'autorità Garante per la protezione dei dati personali e accertarne, tramite il continuo monitoraggio e valutazione delle minacce alla sicurezza, la reale vulnerabilità.

L'assetto normativo prevede, inoltre, l'individuazione di un soggetto che, in ragione di specifiche e comprovate professionalità e competenze, svolga il ruolo di referente per la cybersicurezza. La nomina di questa figura deve essere previamente comunicata all'ACN, poiché tra i compiti a essa attribuiti vi è quello del punto unico di contatto dell'amministrazione con l'Agenzia. Qualora la PA non disponesse all'interno della propria organizzazione di soggetti con i requisiti richiesti, è possibile ricorrere al conferimento di un incarico a un dipendente di altra amministrazione, previa autorizzazione ai sensi dell'articolo 53 del d.lgs. 165/2001. Infine, è possibile individuare

6 I profili dell'ECSF per i professionisti della cybersicurezza sono: *Chief information security office (Ciso), Cyber incident responder, Cyber legal, Policy and compliance officer, Cyber threat intelligence specialist, Cybersecurity architect, Cybersecurity auditor, Cybersecurity educator, Cybersecurity implementer, Cybersecurity researcher, Cybersecurity risk manager, Digital forensics investigator, Penetration tester*.

la struttura e il referente anche all'interno dell'ufficio del responsabile per la transizione al digitale (RTD), i cui compiti possono essere esercitati in forma associata come previsto dall'articolo 17, commi 1-*sexies* e 1-*septies* del Codice dell'amministrazione digitale (Cad).

Figura 2.3 - Soggetto che nella struttura svolge il ruolo di referente per la cybersicurezza



Fonte: Elaborazione degli autori sulla legge 90/2024

2.4 Trasformazione digitale delle pubbliche amministrazioni

Tutte le pubbliche amministrazioni, incluse quelle locali, sono obbligate a nominare un responsabile per la transizione al digitale: l'obiettivo è guidare e accelerare la digitalizzazione nella Pubblica amministrazione. L'istituzione di tale ruolo risponde alla necessità di rendere le amministrazioni pubbliche più efficienti, trasparenti e vicine ai cittadini, attraverso l'utilizzo strategico delle tecnologie digitali. Tale figura, prevista nel Codice dell'amministrazione digitale (Cad), deve possedere competenze specifiche in ambito tecnologico, manageriale e organizzativo, con una visione strategica delle opportunità offerte dalla digitalizzazione.

Si possono incontrare difficoltà nell'individuare figure adeguate a causa della mancanza di risorse o competenze interne; infatti, la complessità del ruolo richiede la convergenza di diverse professionalità in grado di collaborare sinergicamente alla realizzazione

degli obiettivi di trasformazione digitale dell'ente: esperti normativi per l'individuazione e l'interpretazione delle prescrizioni da rispettare; esperti informatici per lo sviluppo delle soluzioni tecniche; analisti di processo per adattare le scelte di digitalizzazione nell'ambito dell'organizzazione dell'ente. Così, appare necessario affiancare al responsabile un ufficio di supporto in grado di soddisfare tutte le esigenze relative alle diverse prospettive con cui si deve approcciare alla trasformazione digitale.

Un fattore, in ogni caso, determinante è la collaborazione con i dirigenti e l'allineamento agli obiettivi politici e istituzionali. Per superare tali ostacoli, è fondamentale incentivare la formazione specialistica e garantire il supporto economico necessario.

2.5 Profili organizzativi

L'impatto delle nuove tecnologie si misura, altresì, sui necessari adeguamenti degli assetti organizzativi da adottare, da una parte, per il migliore perseguimento degli obiettivi fissati nelle strategie UE e nazionali per la realizzazione del mercato unico digitale e, dall'altra, per una gestione delle risorse umane in linea con i diritti fondamentali della persona. Al fine di garantire sistemi di IA conformi alla disciplina generale applicabile all'interno dell'UE, occorre, dunque, prevedere dei ruoli interni all'organizzazione che risultino adeguati alle finalità di corretto bilanciamento tra innovazione e diritti.

Tenuto conto degli obblighi da rispettare e delle relative competenze richieste per l'adempimento agli stessi, nell'organizzazione dell'IA si possono distinguere tre livelli: strategico, di coordinamento e operativo.

Al livello strategico (imputabile agli organi di vertice), spetta la definizione degli obiettivi da perseguire con l'implementazione di sistemi di IA all'interno dell'organizzazione.

Al livello di coordinamento, è affidata la responsabilità dell'implementazione delle strategie definite al livello strategico provvedendo allo sviluppo di piani operativi dettagliati e al coordinamento delle risorse necessarie per raggiungere gli obiettivi strategici. Fonte di ispirazione per strutturare un coordinamento dell'implementazione delle strategie legate all'IA è la figura del CAIO (*Chief Artificial Intelligence Officer*)[7], sorta dall'esigenza di individuare un punto di riferimento cui affidare il compito di guidare l'integrazione dell'IA nei processi dell'organizzazione, garantendo che le tecnologie innovative risultino allineate con gli obiettivi strategici.

Per lo svolgimento di tale coordinamento, risulta necessario collaborare con le altre strutture chiave dell'organizzazione, al fine di ottimizzare l'uso dei dati e sviluppare soluzioni basate sull'IA che migliorino l'efficienza operativa e l'esperienza del cliente. La natura multidisciplinare delle attività coinvolte nella predisposizione delle misure in materia di IA comporta la necessità di individuare, dunque, una struttura apposita in cui confluiscano in modo organico competenze tecnico-informatiche (necessarie alla valutazione dei rischi del sistema IA e alla conseguente progettazione dello stesso sistema in forma adeguata), giuridiche (necessarie per una corretta e adeguata applicazione della disciplina normativa di settore) e tematiche (necessarie per la valutazione dei possibili ambiti applicativi e casi di uso dei sistemi di IA). La struttura di coordinamento dell'IA è, così, responsabile della governance e della *compliance* istituzionale in materia.

7 Cfr. US Office of Management and Budget (OMB) 2024.

Al livello operativo, si deve provvedere all'affidamento delle responsabilità delle attività quotidiane dell'organizzazione, tenuto conto che le figure operative eseguono le attività specifiche necessarie per fare funzionare il sistema IA istituzionale. Questo livello è fondamentale per l'efficienza e l'efficacia delle operazioni.

Per quanto riguarda, invece, l'esigenza di conciliare il ricorso alle nuove tecnologie per lo svolgimento dell'attività istituzionale con il rispetto delle prerogative dei lavoratori, deve essere il datore di lavoro a farsi promotore dell'adozione di misure adeguate. L'obiettivo è garantire che gli avanzamenti in termini di efficienza forniti dall'uso della tecnologia non determinino un arretramento sul fronte del livello di tutela dei diritti. A tale proposito, si rammenta che il datore di lavoro è responsabile della sicurezza degli strumenti tecnologici impiegati anche in modalità agile e, allo stesso tempo, deve farsi garante del diritto alla disconnessione[8]. Con specifico riferimento all'intelligenza artificiale, l'*AI Act* parte quindi dal presupposto per il quale, nel contesto dell'occupazione e della protezione dei lavoratori, non devono prodursi lesioni in materia di politica sociale, né sul diritto del lavoro nazionale, per quanto riguarda le condizioni di lavoro – comprese la salute e la sicurezza sul luogo di lavoro – e il rapporto tra datori di lavoro e lavoratori (Considerando 9). Puntuali indicazioni vengono, così, fornite ai *deployer* nel momento in cui assumono la veste di datore di lavoro: prima di mettere in servizio o utilizzare un sistema di IA ad alto rischio sul luogo di lavoro[9], devono essere informati i rappresentanti dei lavoratori e i lavoratori interessati che saranno soggetti all'uso del sistema di IA ad alto rischio.

8 Cfr. Risoluzione del Parlamento europeo del 21 gennaio 2021 recante raccomandazioni alla Commissione sul diritto alla disconnessione, in cui si rileva che "L'utilizzo maggiore delle tecnologie digitali ha trasformato i modelli tradizionali di lavoro e ha creato una cultura del «sempre connesso» e «sempre online». In tale contesto è importante garantire la tutela dei diritti fondamentali dei lavoratori, condizioni di lavoro eque, compresi il diritto a una retribuzione equa e l'attuazione del loro orario di lavoro, la salute e la sicurezza e la parità tra uomini e donne" (Considerando 9 della proposta di Direttiva allegata alla Risoluzione).

9 Sono da considerare sistemi ad alto rischio a norma dell'articolo 6, par. 2, quelli appartenenti al settore dell'occupazione, gestione dei lavoratori e accesso al lavoro autonomo:
a) i sistemi di IA destinati a essere utilizzati per l'assunzione o la selezione di persone fisiche, in particolare per pubblicare annunci di lavoro mirati, analizzare o filtrare le candidature e valutare i candidati;
b) i sistemi di IA destinati a essere utilizzati per adottare decisioni riguardanti le condizioni dei rapporti di lavoro, la promozione o cessazione dei rapporti contrattuali di lavoro, per assegnare compiti sulla base del comportamento individuale o dei tratti e delle caratteristiche personali o per monitorare e valutare le prestazioni e il comportamento delle persone nell'ambito di tali rapporti di lavoro.

3. STRUMENTI E STRUTTURA DEL PROCESSO DI PRODUZIONE STATISTICA E INTELLIGENZA ARTIFICIALE[1]

3.1 Struttura per il trattamento quantitativo e qualitativo dei dati e intelligenza artificiale

Nell'ambito del trattamento quantitativo e qualitativo dei dati sono individuabili numerosi strumenti tecnico-statistici, in costante miglioramento grazie all'utilizzo di strumenti digitali più all'avanguardia e di sistemi di intelligenza artificiale[2]. Il processo di produzione statistica (*Generic Statistical Business Process Model* - GSBPM), infatti, punta sempre più al ricorso a diverse applicazioni dell'IA, dall'impiego degli LLM (*Large Language Model*), all'uso di algoritmi di classificazione automatica e fino al ricorso a strumenti di IA generativa di testo per la disseminazione dell'informazione statistica. Rispetto a questi ultimi strumenti, si rinvia alle discipline tecniche e statistiche per gli aspetti più prettamente scientifici dei metodi e delle tecniche di produzione dell'informazione, oltre alle disposizioni giuridiche e tecniche per gli aspetti di regolamentazione. La disciplina normativa europea – Regolamento (CE) n. 223/2009 (art. 2, par. 1, lett. c) – prevede, sotto un profilo soggettivo, la professionalità degli addetti al trattamento quantitativo e qualitativo; sotto un profilo oggettivo, vanno considerati il Codice delle statistiche europee, che enuncia i principi di affidabilità e attendibilità da perseguire nell'effettuazione del trattamento del dato (cfr. paragrafo 3.2), e le Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica pubblicate ai sensi dell'art. 20, comma 4, del d.lgs. 10 agosto 2018, n. 101. A tale proposito, riguardo all'acquisizione dei dati con l'utilizzo delle nuove tecnologie (cfr. paragrafo 1.2.1), il Regolamento (UE) 2024/3018 del Parlamento europeo e del Consiglio, del 27 novembre 2024, che modifica il Regolamento (CE) n. 223/2009 relativo alle statistiche europee, ribadisce la necessità di agevolare l'accesso ai dati da parte degli istituti nazionali di statistica. Attualmente, il Regolamento (CE) n.223/2009 (art. 17 *bis*, par. 1) prevede, infatti, che gli organismi pubblici responsabili delle fonti di dati, delle banche dati o dei sistemi di interoperabilità amministrativi o di dati pertinenti e necessari per finalità statistiche europee consentano agli istituti di statistica di consultare, utilizzare e integrare gratuitamente tali dati e i metadati pertinenti, in modo tempestivo e con una frequenza e granularità sufficienti ai fini dello sviluppo, della produzione e della diffusione di statistiche europee. Tali indicazioni normative vanno lette alla luce del ruolo di indirizzo e razionalizzazione dei flussi informativi affidato all'Istat dal d.lgs. n. 322 del 1989 (artt. 1 e 15).

Di seguito si riportano, a fini esemplificativi, i principali strumenti di trattamento dei dati quantitativi e qualitativi a disposizione dell'Istituto: statistiche (da rilevazione, da fonti amministrative e da nuove fonti di dati); il sistema dei conti nazionali e il sistema dei conti regionali; l'elenco delle unità istituzionali appartenenti al settore della Pubblica amministrazione (S.13);

1 A cura di Giovanna Bellitti e Massimo Fedeli, con contributo di Serenella Ravioli (paragrafo 3.3.3); hanno collaborato all'analisi dei processi Paolo Nicolai (paragrafi 3.2 e 3.3.1) e Maria Saiz (paragrafi 3.1, 3.3.1 e 3.3.2); coordinamento di Roberto Puglisi.

2 Per un approfondimento, si rinvia alle discipline statistiche di settore e, sotto il profilo delle regole, a Bellitti e Fedeli 2022, Capitolo 1, paragrafo 1.

rapporti (Rapporto Bes, Rapporto annuale, Rapporto competitività, Rapporto sulla conoscenza, Rapporto sul mercato del lavoro, Rapporto sul territorio, Rapporto SDGs - *Sustainable Development Goals*); registri; particolari elaborazioni statistiche per conto di enti e privati, remunerate a condizioni di mercato; trattamenti realizzati nell'ambito di accordi e protocolli di ricerca.

L'impiego di sistemi IA nell'ambito del GSBPM ha l'obiettivo di migliorare l'informazione statistica in termini di maggiore tempestività, qualità e di ampliare il supporto e l'interoperabilità tra l'Istituto e le altre pubbliche amministrazioni. Come già accennato (cfr. Capitolo 1), il carattere composito del quadro normativo e regolamentare di riferimento per lo svolgimento delle attività di trattamento dati impone l'adozione di una prospettiva multidisciplinare, che possa fare cogliere e considerare adeguatamente le diverse azioni da mettere in campo per un'applicazione dell'IA in linea con i valori condivisi nell'UE. A tale fine, si propone una mappatura delle principali attività richieste in fase di avvio del processo di produzione statistica: rispetto a ciascuna di tali attività, per l'individuazione delle relative responsabilità occorre fare riferimento alle figure soggettive e alle rispettive strutture organizzative previste nelle singole discipline. Rinviando alle considerazioni già esposte sul tema (cfr. Capitolo 2), nonché a quanto verrà illustrato sui casi di uso (cfr. Capitolo 4), di seguito si richiamano i ruoli e le competenze nel processo di trattamento dei dati al fine di evidenziare i rispettivi adempimenti con riferimento ai diversi settori di attività.

In materia di protezione dei dati personali, occorre procedere alla nomina del Responsabile della protezione dei dati personali, previsto dagli artt. 37 e ss. del GDPR, e, quindi, agevolare lo svolgimento dei compiti del titolare del trattamento.

Nell'ambito della cybersicurezza, invece, si deve provvedere all'individuazione della struttura e del referente per la *cybersecurity* previsti dalla legge 28 giugno 2024, n. 90: il referente deve essere scelto in ragione di specifiche e comprovate professionalità e competenze in materia e il nominativo è comunicato all'Agenzia per la cybersicurezza nazionale. La struttura e il referente per la *cybersecurity* possono essere individuati, rispettivamente, nell'ufficio e nel responsabile per la transizione al digitale.

Nell'ambito della digitalizzazione della PA, va individuato l'ufficio del responsabile per la transizione al digitale, in modo da disporre di una struttura idonea a garantire l'adempimento agli obblighi previsti dall'art. 17 del Cad, tra i quali il coordinamento strategico dello sviluppo dei sistemi informativi, di telecomunicazione e fonia, in modo da assicurare: la coerenza con gli standard tecnici e organizzativi comuni; l'indirizzo e coordinamento dello sviluppo dei servizi, sia interni sia esterni, forniti dai sistemi informativi di telecomunicazione e fonia dell'amministrazione; l'analisi periodica della coerenza tra l'organizzazione dell'amministrazione e l'utilizzo delle tecnologie dell'informazione e della comunicazione, al fine di migliorare la soddisfazione dell'utenza e la qualità dei servizi, nonché di ridurre i tempi e i costi dell'azione amministrativa; la cooperazione alla revisione della riorganizzazione dell'amministrazione; la progettazione e il coordinamento delle iniziative rilevanti ai fini di una più efficace erogazione di servizi in rete a cittadini e imprese mediante gli strumenti della cooperazione applicativa tra pubbliche amministrazioni; la promozione delle iniziative attinenti all'attuazione delle direttive impartite dal Presidente del Consiglio dei ministri o dal Ministro delegato per l'innovazione e le tecnologie; la pianificazione e il coordinamento degli acquisti di soluzioni e sistemi informatici, telematici e di telecomunicazione al fine di garantirne la compatibilità con gli obiettivi di attuazione dell'agenda digitale e, in particolare, con quelli stabiliti nel piano triennale.

Con riguardo specifico all'IA, non esiste una figura prevista appositamente per sovrintendere all'implementazione di sistemi IA, dovendosi, piuttosto, fare riferimento agli obblighi previsti nell'*AI Act* per ciascun livello di rischio.

Tuttavia, come accennato precedentemente, si è diffusa la figura del *Chief Artificial Intelligence Officer*, cui vengono affidati compiti di coordinamento per l'utilizzo e la promozione dell'IA e di gestione dei rischi derivanti dall'utilizzo di tali sistemi. Per l'individuazione di tale figura, occorre considerare il possesso di specifici requisiti e competenze che consentano un idoneo coordinamento per l'attuazione delle strategie al fine di sfruttare al meglio le potenzialità offerte dai sistemi basati sull'intelligenza artificiale.

Rispetto alle questioni etiche e deontologiche connesse all'uso dell'IA, la loro complessità richiede un approccio sotto diverse prospettive, idonee a coinvolgere aspetti relativi allo sviluppo del sistema, al rispetto della disciplina di riferimento e alle esigenze dell'utente. È essenziale che vengano adottate pratiche responsabili che, sulla base di un confronto continuo tra sviluppatori, regolatori e *stakeholder*, riescano ad assicurare un approccio etico e deontologicamente responsabile, che garantisca un'IA in grado di portare benefici significativi alla società senza compromettere i valori fondamentali di giustizia, equità e rispetto per l'autonomia umana. Nel GSBPM IA risulta essenziale, dunque, includere una costante valutazione di conformità ai principi etici già approfonditi (cfr. Capitolo 1, paragrafo 4).

3.2 Mappatura delle attività in materia di privacy, cybersicurezza e intelligenza artificiale e delle regole di settore

Le figure soggettive dell'amministrazione e quelle precedentemente accennate e le relative strutture organizzative preposte all'adempimento dei diversi obblighi di legge devono, dunque, tenere presente, oltre alle regole del settore statistico, quelle relative alla privacy, alla cybersicurezza, alla trasformazione digitale e, più in particolare, all'intelligenza artificiale.

Dal punto di vista applicativo si possono individuare, oltre alle attività del processo di produzione statistica, alcune azioni tecniche di rilievo del processo di implementazione dell'IA ai fini di sviluppo e applicazione: la progettazione del sistema; l'utilizzo di tecniche di addestramento di modelli, con l'obiettivo di garantire che il sistema di IA funzioni come previsto e in maniera sicura e che non diventi una fonte di discriminazione vietata dal diritto dell'Unione europea[3]; l'elaborazione dei dati – per cui si rendono necessarie infrastrutture adeguate – e, infine, la disponibilità dell'output prodotto dal sistema IA.

La progettazione, tenuto conto della finalità del trattamento, può riguardare sia la revisione sia l'integrazione di un lavoro statistico, ovvero la costruzione di un sistema informativo o di un registro statistico, tenendo in considerazione le linee guida per la qualità dei processi statistici[4]. Si tratta, pertanto, di un momento fondamentale per la corretta impostazione delle attività, sia per quanto riguarda gli aspetti giuridico-organizzativi e amministrativo-gestionali della protezione dei dati personali, della cybersicurezza e dell'impiego di sistemi IA nel trattamento statistico dei dati, sia con riferimento ai profili tecnico-metodologici e di classificazione dei fenomeni sociali, economici e demografici, ai sensi dell'art. 15 lett. e) del d.lgs. 322/1989. Le classificazioni sono, infatti, uno strumento per organizzare sistematicamente le informazioni in insiemi di categorie strutturate ed esaurienti secondo criteri di somiglianza e hanno l'obiettivo di fornire una semplificazione del mondo reale, raggruppando e ordinando le informazioni in categorie predefinite secondo qualche criterio significativo (Bellitti e Fedeli 2022, p. 119 e ss.).

3 Cfr. *AI Act*, Considerando 67.

4 Cfr. Istituto nazionale di statistica - Istat. *Linee guida per la qualità*. https://www.istat.it/classificazioni-e-strumenti/strumenti-per-la-qualita/linee-guida/.

In questo contesto, il rispetto delle discipline normative concorre con l'applicazione delle regole tecniche e metodologiche che presiedono al trattamento, tenuto conto delle finalità dello stesso. A tale proposito, il perimetro è tracciato dai 16 principi stabiliti dal Codice delle statistiche europee adottato dal Comitato del sistema statistico europeo il 16 novembre 2017[5]. L'edizione del 2017 intende proprio "rispecchiare i cambiamenti e le innovazioni più recenti in materia di sviluppo, produzione e diffusione di statistiche ufficiali", adeguandosi a "la disponibilità di nuove fonti di dati, l'impiego di nuove tecnologie, la modernizzazione del quadro giuridico" e prevedendo che il trattamento debba essere effettuato secondo i seguenti principi. Nello specifico, il Codice delle statistiche europee individua principi da seguire nel trattamento e nella definizione della metodologia al fine di garantire un'informazione di qualità, in relazione alla tipologia di dati trattati (Bellitti e Fedeli 2022, p. 111, contributi di M. Scanu e M. Scannapieco).

Le metodologie definite in ambito europeo o nazionale previste dal Programma statistico nazionale (PSN) dovranno consentire l'individuazione della popolazione statistica di riferimento, riducendo al contempo l'onere statistico sui rispondenti (Bellitti e Fedeli 2022, p. 161, contributo di I. Diaco). Allo stesso tempo, si deve provvedere all'applicazione delle regole di digitalizzazione nel momento in cui si preveda che il trattamento verrà effettuato con piattaforme o altri strumenti digitali, quale il ricorso per alcune attività all'IA.

Qualora nel trattamento digitale si preveda l'utilizzo di dati personali, sarà necessario rispettare le prescrizioni previste dal Regolamento (UE) 2016/679 e dal d.lgs. 196/2003 – in particolare la *privacy by design* e *privacy by default* – e assicurare l'adozione, da parte dei soggetti autorizzati al trattamento, di tutte le specifiche misure organizzative e tecniche dirette ad assicurare la trasparenza, la riservatezza, nonché l'integrità e l'esattezza dei dati (Bellitti e Colasanti 2021). Si deve garantire, dunque, la conformità del trattamento al GDPR, provvedendo all'informativa agli interessati, al rispetto dei loro diritti, all'individuazione di una base giuridica, all'adozione di misure di sicurezza adeguate, a concludere accordi con i responsabili, a effettuare una valutazione di impatto sulla protezione dei dati personali (DPIA), a tenere un registro delle attività di trattamento, a collaborare con le Autorità di controllo. Nello specifico ambito della produzione statistica, le Regole deontologiche stabiliscono le norme per i soggetti ai quali sono affidati le fasi di rilevazione, che devono porre specifica attenzione nella selezione del personale preposto alla raccolta dei dati e nella definizione dell'organizzazione e delle modalità di rilevazione, in modo da garantire il rispetto delle stesse regole e la tutela dei diritti degli interessati.

Sotto il profilo delle tecniche di digitalizzazione del trattamento e in applicazione del principio *digital first*, occorre perseguire gli obiettivi di trasformazione digitale e interoperabilità dei sistemi informativi, favorendo il ricorso agli strumenti offerti dalle nuove tecnologie (Bellitti e Fedeli 2023). Con riferimento, invece, al tema della cybersicurezza, il trattamento deve essere progettato assicurando gli adempimenti previsti in materia e, dunque, adottando le misure necessarie per contrastare le minacce e i possibili attacchi informatici, che possono comportare un rischio elevato per la protezione dei dati e la sicurezza dei sistemi e delle reti.

5 I principi indicati dal Codice delle statistiche europee sono: Contesto istituzionale: Principio 1, Indipendenza professionale; Principio 1*bis*, Coordinamento e cooperazione; Principio 2, Mandato per la rilevazione dei dati; Principio 3, Adeguatezza delle risorse; Principio 4, Impegno in favore della qualità; Principio 5, Riservatezza statistica e protezione dati; Principio 6, Imparzialità e obiettività. Processi statistici: Principio 7, Solida metodologia; Principio 8, Procedure statistiche appropriate; Principio 9, Onere non eccessivo sui rispondenti; Principio 10, Rapporto costi/efficacia; Prodotti statistici: Principio 11, Pertinenza; Principio 12, Accuratezza e attendibilità; Principio 13, Tempestività e puntualità; Principio 14, Coerenza e comparabilità; Principio 15, Accessibilità e chiarezza.

Si dovrà, poi, considerare le linee guida operative pubblicate dall'Agenzia per la cybersicurezza nazionale sulla "tassonomia Cyber", che definiscono il "linguaggio comune per lo scambio delle informazioni relative a eventi e minacce di cybersicurezza". Sarà necessario, inoltre, tenere conto delle "linee guida per il rafforzamento della protezione delle banche dati rispetto al rischio di utilizzo improprio" e delle "linee guida per il rafforzamento della resilienza", rilasciate recentemente dalla stessa Agenzia. Il primo documento, tra le finalità e gli obiettivi principali, pone l'accento al contrasto a usi impropri delle banche dati e mira a consolidare la resilienza delle infrastrutture critiche del Paese.

In particolare, individua tra i punti focali sui quali prestare attenzione: il controllo degli accessi, la sicurezza dei sistemi e delle applicazioni, la gestione dei rischi e della sicurezza della catena di approvvigionamento, la formazione dei dipendenti e le attività di monitoraggio e *auditing* interno.

Il secondo documento è strutturato in due parti: nella prima parte individua le misure di sicurezza che i soggetti adottano per il rafforzamento della resilienza, nella seconda evidenzia le modalità di implementazione promosse e raccomandate per l'espletamento delle misure idonee o alternative a quelle riportate, qualora soddisfino i requisiti minimi previsti dalle linee guida.

Laddove il trattamento preveda, quindi, l'utilizzo di sistemi basati anche sull'intelligenza artificiale, occorre procedere, innanzitutto, alla valutazione dei livelli di rischio previsti dal Regolamento (UE) 2024/1689, per procedere, poi, all'individuazione dei ruoli e degli obblighi ivi previsti.

In particolare, nell'attività di progettazione occorre tenere conto del principio di non discriminazione e della classificazione dei sistemi di IA per livello di rischio. Dal punto di vista del genere, si deve considerare che la "necessità di assumere sistematicamente la prospettiva di genere nella realizzazione di trattamenti di dati effettuati in ambito pubblico per la produzione di informazioni quali-quantitative è stata recentemente sancita nel nostro ordinamento grazie all'approvazione della legge n. 53/2022 (Disposizioni in materia di statistiche in tema di violenza di genere)" (Bellitti e Fedeli 2022, p. 150, contributo di F. Albo).

3.3 Attività del processo di produzione

Nell'ambito del processo di trattamento quantitativo e qualitativo si possono individuare le attività fondamentali per la corretta gestione dello stesso (progettazione, gestione del trattamento, comunicazione e diffusione), articolate in azioni e adempimenti che gli addetti devono perseguire secondo un approccio multidisciplinare che tenga conto dei diversi aspetti coinvolti.

3.3.1 Progettazione del trattamento

Nel corso della progettazione del trattamento qualitativo e quantitativo dei dati occorre individuare i flussi di lavoro e i possibili interventi, attraverso una mappatura accurata dei processi correlati, per ottimizzare e rendere idonee le attività che lo accompagnano

L'elemento essenziale per tali attività è l'approccio coordinato e interdisciplinare, subordinato all'attività di *compliance* e prevenzione del rischio, per garantire un'adeguata pianificazione delle azioni da intraprendere.

La definizione di nuove strategie e metodologie è quindi legata a una programmazione complessiva ed efficace, in cui risulta imprescindibile documentare, misurare e razionalizzare i processi in fase di input e di output del lavoro. Tale sequenza di azioni, tra loro strettamente collegate, è utile per avere un quadro onnicomprensivo sulle soluzioni alternative disponibili nella fase gestionale successiva.

Una simile organizzazione, infatti, permette di studiare, elaborare e valutare nuove tecniche e metodologie, con l'obiettivo di promuovere, semplificare e consolidare le attività per conseguire un risultato efficace e innovativo.

Con riferimento alle attività di ricognizione, occorre individuare la relativa base giuridica, la finalità del trattamento e considerare, inoltre, le regole generali e di settore che stabiliscono le prescrizioni nei singoli ambiti (Bellitti e Fedeli 2022). Allo stesso tempo, è indispensabile tenere conto delle Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica effettuati nell'ambito Sistan, concernenti le condizioni essenziali per la liceità e la correttezza del trattamento, nonché i compiti esercitati dall'Istat, come previsto dall'art. 15, lettera e) del d.lgs. n. 322/1989 relativamente alle classificazioni, definizioni e nomenclature dei dati da trattare.

Vanno, quindi, individuate le diverse modalità e tecniche di raccolta e trattamento dei dati, provvedendo, tra l'altro, alla progettazione della piattaforma per l'acquisizione e la gestione dei dati in relazione alla tecnica metodologica acquisitiva e tenendo conto dei criteri di valutazione del rischio di reidentificazione degli interessati.

L'evoluzione della normativa che regola il processo di produzione statistica e l'esigenza di ridurre l'onere statistico sui rispondenti hanno richiesto l'adozione di numerose innovazioni.

Si deve provvedere alla definizione delle modalità di accesso alla piattaforma di acquisizione dati e fornire le relative istruzioni e, ai fini della tracciabilità documentale e della predisposizione della nota metodologica da pubblicare sul sito web ai sensi delle Regole deontologiche per trattamenti a fini statistici o di ricerca scientifica effettuati nell'ambito del Sistan, si provvede alla raccolta della documentazione tecnica e metodologica (Bellitti e Fedeli 2022, p. 153, contributo di C. Ceccarelli).

Qualora, nel corso della progettazione, il trattamento dei dati dovesse essere sottoposto a sperimentazioni che comprendano l'uso dell'intelligenza artificiale, sarà necessario individuare la fattispecie di riferimento prevista nel Regolamento (UE) 2024/1689, verificando se il sistema adottato rientri nelle categorie di rischio previste dal suddetto Regolamento, quindi accertare la tipologia di ruolo svolto (*deployer* o sviluppatore) e infine appurare, qualora il sistema rientri nella classificazione dei sistemi ad alto rischio, se le azioni da intraprendere possano implicare la violazione dei diritti e libertà degli interessati (*Fundamental Rights Impact Assessment*).

Al contempo, si deve garantire la trasparenza nell'uso dell'IA: qualora si faccia ricorso a metodi che prevedano l'utilizzo di algoritmi e tecniche automatizzate, sarà necessario rendere noto nell'informativa ai rispondenti il ricorso a tali metodologie e le modalità di elaborazione dei dati al fine di garantire la piena conoscibilità delle attività connesse al trattamento dei dati.

Considerando che il processo di utilizzo di sistemi di IA prende l'avvio dall'individuazione degli obiettivi che si vogliono perseguire, vanno analizzati i requisiti qualitativi dei dati e definite le finalità dell'impiego del sistema di IA, anche con l'obiettivo di valutare le modalità di sviluppo della soluzione (*make or buy*). Infine, si progettano le modalità di esecuzione delle procedure connesse alla metodologia adottata, quali, ad esempio, il *fine tuning*, il *prompting* e le tecniche di validazione dei risultati.

Aldilà della metodologia specifica adottata, nel processo di modellazione dei dati possono distinguersi alcune azioni comuni: la definizione del quesito oggetto di studio; la selezione del *set* di dati e l'analisi esplorativa delle relative variabili; la scelta della metodologia più idonea alla risoluzione del problema statistico; la fase di specificazione e stima del modello; la fase di test e validazione dei risultati; l'ottimizzazione del modello mediante le tecniche più all'avanguardia proposte in letteratura. Con riguardo al perseguimento della sicurezza e robustezza dei sistemi IA si devono prevedere le misure di protezione necessarie, attraverso l'adozione di strategie e programmi di cybersicurezza idonei a individuare e rilevare le minacce in maniera tempestiva, facendo riferimento ai livelli stabiliti dalla Direttiva NIS2 e dalla normativa di recepimento nell'ordinamento interno (d.lgs. 138/2024).

Analogamente, sarà necessario stimare il grado di rischio per il trattamento dei dati personali, sin dalla progettazione, in conformità ai principi di *privacy by default* e *by design* ai sensi dell'art. 25 del Regolamento (UE) 2016/679, attraverso la previsione e l'implementazione di misure tecniche organizzative adeguate al trattamento e alla tutela dei dati personali, con la conseguente valutazione di impatto (VIP/DPIA), ai sensi dell'art. 35 del GDPR. Si rende necessaria, dunque, un'analisi dei rischi per i diritti e le libertà degli interessati connessi al trattamento dei dati e la conseguente predisposizione di adeguate misure di sicurezza.

3.3.2 Gestione del trattamento

Nella gestione del trattamento finalizzato alla produzione di dati quantitativi e qualitativi, devono essere individuati tutti i soggetti coinvolti nel processo statistico con riferimento a ciascun ambito di competenza: produzione statistica, cybersicurezza, digitalizzazione e IA.

Nelle attività di raccolta dei dati vanno attivati i relativi processi di digitalizzazione: si procede, così, alla predisposizione e al caricamento sul sito della documentazione relativa alla rilevazione e degli atti di regolamentazione (piani di Censimento, circolari e istruzioni agli organi di rilevazione); vengono, quindi, messe in atto le misure di sicurezza in materia di protezione dei dati personali e di cybersicurezza progettate precedentemente; si provvede, dunque, all'invio ai soggetti coinvolti dell'informativa predisposta per il trattamento di dati personali.

A seguito delle attività di raccolta dei dati di base, sono previste le seguenti operazioni di trattamento statistico dei dati personali: anonimizzazione, controllo, correzione, trasmissione, registrazione, elaborazione e conservazione. Le operazioni di controllo, di correzione dei dati e le attività di monitoraggio dell'intero processo statistico sono finalizzate ad assicurare la qualità dell'informazione statistica, con particolare riguardo al principio di esattezza dei dati. Si provvede, dunque, allo svolgimento e alla verifica delle operazioni svolte dagli organi intermedi di rilevazione, dai rilevatori o dalle società, alla conseguente verifica della qualità del lavoro in corso e, infine, all'eventuale attivazione del procedimento per accertamento della violazione dell'obbligo di risposta e conseguente applicazione della sanzione.

In un'ottica di transizione digitale, si provvede alla digitalizzazione delle operazioni sopra elencate.

Riguardo all'attività di trasmissione dei dati all'Istat, che necessita di modalità che consentano di ridurre al minimo il rischio di alterazione, perdita e distruzione, anche accidentali, dei dati stessi e di accesso a questi ultimi da parte di soggetti non autorizzati –

qualora sia prevista la registrazione dei dati – l'operazione richiede l'adozione, da parte dei soggetti che la effettuano, di tutte le specifiche misure organizzative e tecniche dirette ad assicurarne la riservatezza, l'integrità e l'esattezza.

Quindi, intervengono le azioni di elaborazione statistica dei dati, che richiedono il rispetto delle pertinenti regole tecnico-metodologiche – oltre che sul piano della qualità dei processi, anche ai fini della valutazione della minimizzazione, correttezza e trasparenza del trattamento dei dati personali – e che necessitano, quindi, di essere adeguatamente documentate.

Relativamente alla conservazione dei dati, i diversi profili coinvolti attengono a cybersicurezza, privacy, utilizzo dei registri statistici mediante sistemi di IA.

Si provvede, dunque, all'adozione di misure di sicurezza dei dati trattati, alla conservazione nei *repository* e alla definizione dei tempi di conservazione dei dati personali, in conformità alle disposizioni vigenti. A tale riguardo si consideri quanto indicato nel Programma statistico nazionale: le disposizioni contenute nel Volume 2 del PSN possono, infatti, prevedere la conservazione dei dati anche oltre il periodo necessario per il raggiungimento degli scopi per i quali sono stati raccolti, per successivi trattamenti espressamente indicati, nel rispetto del principio di limitazione della conservazione affermato nel GDPR.

3.3.3 Attività di diffusione e comunicazione[6]

Nell'ambito del processo di produzione quantitativa e qualitativa dei dati statistici, anche attraverso l'utilizzo di sistemi di IA, si individuano le attività di diffusione e comunicazione, che prevedono due tipologie di azioni distinte effettuate attraverso l'utilizzo di sistemi digitali e dell'intelligenza artificiale. Le azioni si diversificano in base agli utenti determinati o indeterminati che ricevono informazioni statistiche, oppure in base alle caratteristiche dei dati forniti, dati aggregati oppure dati elementari.

La diffusione del prodotto statistico, in linea generale, costituisce il completamento naturale del lavoro statistico ed è compito che l'Istat deve svolgere nel rispetto del GDPR, dei principi statistici ed entro i limiti del segreto statistico, garantendo parità di accesso a tutti.

La comunicazione dei dati personali ha la specificità di prevedere l'individuazione del destinatario: si porta a conoscenza dei dati uno o più soggetti determinati, qualunque sia il mezzo attraverso cui essa avvenga, inclusa la messa a disposizione, la consultazione, l'interconnessione (Bellitti e Fedeli 2022, p. 57, contributo di S. Ravioli).

In Istat, oltre alle attività di diffusione e comunicazione dei dati statistici di cui al d.lgs. 322/1989, si aggiunge il perimetro della comunicazione e della promozione della statistica, intese come un derivato della comunicazione pubblica. La fonte normativa, come spesso richiamato, risiede nella legge 7 giugno 2000, n. 150, la cosiddetta "legge quadro sulla comunicazione pubblica", in attesa da anni di revisione rispetto alle importanti innovazioni digitali e di relazioni con il pubblico. Non è questa la sede per trattare l'argomento della comunicazione in tale accezione, ma per completezza di rappresentazione è necessario tanto menzionarla, quanto riconoscere che anche in termini di comunicazione e promozione sono in corso interessanti sviluppi e applicazioni della IA che necessitano ovviamente riflessioni persino di carattere etico.

L'uso corretto della tecnologia IA in tutti i diversi aspetti della diffusione, comunicazione e promozione può potenziare enormemente i risultati delle attività e aumentare la platea dei destinatari (il target di riferimento).

6 Contributo di Serenella Ravioli.

Per fare degli esempi: l'IA applicata all'organizzazione di eventi e convegni consente la promozione della proposta e il coinvolgimento dei partecipanti; l'interazione veloce con i partecipanti; il dinamismo del processo creativo grazie ai sistemi generativi di *deep learning*, che possono progettare materiali grafici e allestimenti personalizzati; la progettazione attività di intrattenimento interattive; l'eliminazione di barriere linguistiche, eccetera. Nella promozione può generare una campagna pubblicitaria, attraverso strumenti di analisi e di elaborazione che, grazie alla potenza computazionale, possono creare video, *slide*, infografiche, *newsletter*, comunicati.

Tornando alla diffusione dei dati, l'utilizzo di tecnologie basate su sistemi di intelligenza artificiale può tanto assumere un ruolo rilevante in termini di semplificazione, immediatezza e snellimento dei processi, quanto, contestualmente, implicare rischi elevati, qualora non siano sufficientemente analizzati gli elementi critici e qualora i sistemi non siano adeguatamente collaudati per le finalità preventivate. L'impiego di simili strumenti può essere significativo, a condizione che siano integrati con i flussi di lavoro e con le attività svolte dalle strutture competenti: occorre, dunque, procedere gradualmente, affinché si possa raggiungere un risultato che produca effetti positivi in termini di aumento della produttività e di ottimizzazione dei processi, per la produzione di informazioni da somministrare all'utente finale. L'utilizzo di assistenti virtuali (*chatbot*), integrati con siti web o *social network*, permette, invece, di interagire con gli utenti attraverso l'uso di un linguaggio naturale, fornendo informazioni esaustive su determinati argomenti. Come accennato, occorre considerare anche i rischi etici che possono coinvolgere le attività di comunicazione e diffusione; a tale proposito, il principio di *accountability* ci rammenta che l'automazione della generazione dei contenuti non fa venire meno la responsabilità umana di chi si affida al sistema IA. Così, il ricorso a sistemi di IA per la produzione e la diffusione comporta la responsabilizzazione di chi sviluppa e utilizza l'IA nel garantire l'affidabilità delle informazioni e nel valutare l'impatto sulla fiducia del pubblico, poiché la comunicazione basata su informazioni errate o manipolate minaccia il diritto delle persone ad avere accesso a informazioni veritiere e imparziali. Allo stesso tempo, devono essere attentamente monitorati i rischi di *bias*, nella misura in cui gli algoritmi di IA sono spesso addestrati su grandi quantità di dati, che possono riflettere pregiudizi e stereotipi presenti nella società; se non correttamente monitorati, i modelli di IA possono amplificare questi *bias*, influenzando la produzione di contenuti comunicativi.

Riguardo ai requisiti tecnici degli strumenti informatici, con maggiore forza quando si tratta della relazione con gli utenti, vanno osservati alcuni principi generali: a) accessibilità al contenuto del servizio da parte dell'utente; b) fruibilità delle informazioni offerte (quindi facilità e semplicità di uso, assicurando, tra l'altro, che le azioni da compiere per ottenere servizi e informazioni siano sempre uniformi tra loro); efficienza nell'uso (garantendo la separazione tra contenuto, presentazione e modalità di funzionamento delle interfacce, nonché la possibilità di rendere disponibile l'informazione attraverso differenti canali sensoriali); efficacia nell'uso e rispondenza alle esigenze dell'utente (assicurando che le azioni da compiere per ottenere in modo corretto servizi e informazioni siano indipendenti dal dispositivo utilizzato per l'accesso); soddisfazione nell'uso (quindi accesso al servizio e all'informazione senza ingiustificati disagi o vincoli per l'utente).

A tale fine, occorre fare riferimento alle linee guida Agid emanate ai sensi dell'art. 11 della legge 9 gennaio 2004, n.4, che stabiliscono: a) i requisiti tecnici per l'accessibilità degli strumenti informatici, ivi inclusi i siti web e le applicazioni mobili; b) le metodologie tecniche per la verifica dell'accessibilità degli strumenti informatici, ivi inclusi i siti web e

le applicazioni mobili; c) il modello della dichiarazione di accessibilità; d) la metodologia di monitoraggio e valutazione della conformità degli strumenti informatici, ivi inclusi i siti web e le applicazioni mobili; e) le circostanze in presenza delle quali si determina un onere sproporzionato, per cui i soggetti erogatori possono ragionevolmente limitare l'accessibilità di un sito web o applicazione mobile.

In tale contesto, nelle fasi di pianificazione strategica e di sperimentazione, è cruciale attenersi alle prescrizioni del Regolamento (UE) 2024/1689, verificando l'appartenenza alla categoria di rischio del sistema e il tipo di attività che, con l'ausilio dell'algoritmo, i soggetti coinvolti (*deployer*, *provider* e *developer*) sono chiamati a svolgere. Sulla base delle informazioni fornite alla macchina sarà necessario monitorare i processi di implementazione e di elaborazione degli output, attraverso l'imprescindibile sorveglianza umana e l'oculata verifica dell'accuratezza dei contenuti generati.

In conclusione, in ossequio al principio di trasparenza, è fondamentale informare l'utente circa l'eventuale utilizzo di sistemi di IA e laddove si ricorra a un *chatbot* è opportuno prevedere un'adeguata informazione sull'uso di tale strumento, predisponendo questionari per rilevare i *feedback* degli utenti, al fine di migliorare e di monitorare le prestazioni.

Prospetto 3.1 - Principali regole e attività sotto il profilo giuridico, tecnico e metodologico nel processo di trattamento dei dati quantitativi e qualitativi (a)

ATTIVITÀ	PRINCIPALI ADEMPIMENTI TECNICI, METODOLOGICI, STATISTICI, GIURIDICI DI RISERVATEZZA		AZIONI DI DIGITALIZZAZIONE E DI CYBERSICUREZZA	AZIONI PER L'USO DELL'INTELLIGENZA ARTIFICIALE
Progettazione e organizzazione di trattamenti per la realizzazione di strumenti gestionali o quantitativi e qualitativi previsti nel Programma statistico nazionale (PSN) o da norme	Applicazione delle regole tecniche, giuridiche, metodologiche e tecnologiche del settore di trattamento, nonché della definizione e classificazione delle unità dei dati anche sotto il profilo della non discriminazione e dell'interoperabilità		Raccolta digitale dei dati e degli atti, dei documenti di gestione del lavoro statistico ↓ Attivazione dei relativi processi di digitalizzazione ↓ Digitalizzazione dei questionari e identificazione del formato digitale e degli altri aspetti ↓ Digitalizzazione dei questionari e identificazione degli altri aspetti della rilevazione relativi alla metodologia e alla tipologia delle unità rilevate e della eventuale sanzionabilità della violazione dell'obbligo di risposta ↓ Progettazione della piattaforma per l'acquisizione e la gestione dati in relazione alla tecnica metodologica acquisitiva e di attivazione delle eventuali sanzioni ↓ Previsione delle misure di protezione necessarie, attraverso l'adozione di strategie e programmi di cybersicurezza idonei a individuare e rilevare le minacce in maniera tempestiva	Il processo prende l'avvio dall'individuazione degli obiettivi che si vogliono perseguire. A tal fine sono analizzati i requisiti qualitativi dei dati per la definizione dell'algoritmo e del livello di rischio ↓ Classificazione del dato *on premise* ↓ Definizione delle tecniche di elaborazione con le scelte sulle tecniche di cancellazione ↓ Qualora si preveda l'uso di sistemi di intelligenza artificiale rispettare le prescrizioni previste dal Regolamento (UE) 2024/1689 ↓ Individuazione dei ruoli per l'implementazione dell'algoritmo (Sviluppatori, Fornitori e Utilizzatori) --- Classificazione dei dati e delle finalità ↓ Valutazione *make or buy* ↓ Modalità di esecuzione del *fine tuning* ↓ Modalità di esecuzione del *prompting* (modalità di output) ↓ Modalità di validazione soluzione
	Qualora si preveda nel trattamento digitale l'utilizzo di **dati personali** ↓ Rispettare le prescrizioni previste dal Regolamento (UE) 2016/679 e dal d.lgs. 196/2003 ↓ Qualora il trattamento possa porre un pregiudizio per i diritti e la libertà degli interessati sarà necessario effettuare la Valutazione di impatto privacy (**VIP/DPIA**)	Qualora il trattamento preveda l'utilizzo di **sistemi basati anche sull' intelligenza artificiale**: ↓ Valutazione dei livelli di rischio previsti dal Regolamento (UE) 2024/1689 ↓ Qualora i sistemi adottati rientrino nella categoria di **alto rischio** e possano comportare un elevato pericolo per i diritti fondamentali, sarà necessario effettuare la *Fundamental Rights Impact Assessment* (**FRIA**)		

Fonte: Elaborazione degli autori
(a) Per un approfondimento, si rinvia alle discipline statistiche di settore e, sotto il profilo delle regole, a Bellitti e Fedeli 2022, Capitolo 1, paragrafo 1.

Prospetto 3.1 segue **- Principali regole e attività sotto il profilo giuridico, tecnico e metodologico nel processo di trattamento dei dati quantitativi e qualitativi** (a)

Gestione del trattamento per la produzione di informazioni quantitative e qualitative	Adozione delle misure in materia di protezione dei dati personali necessarie al processo di produzione: • Nomina dell'eventuale contitolare e del responsabile esterno del trattamento dei dati (ditte, ecc.). • Individuazione e nomina da parte del responsabile e designato del trattamento e degli incaricati del trattamento.	Inserimento sul sito web Istat della documentazione relativa al trattamento ↓ Adozione di misure di sicurezza e di cybersicurezza dei dati ↓ Definizione delle modalità di accesso alla piattaforma di acquisizione dati e relative istruzioni ↓ Attività di digitalizzazione delle attività del procedimento sanzionatorio ↓ Chiusura della piattaforma per acquisizione dati	Qualora si preveda l'uso di sistemi di intelligenza artificiale ↓ Rispettare le prescrizioni previste dal Regolamento UE 2024/1689 ——— Bilanciamento, se necessario,del *dataset* per renderlo rappresentativo dell'universo ↓ Esecuzione *fine tuning* ↓ Sviluppo del *prompting* (modalità di output) ↓ Esecuzione validazione soluzione
	Adozione di misure di trasmissione dei dati coerenti con la normativa di cui al Codice dell'amministrazione digitale.		
	Invio della informativa alle unità di rilevazione.		
	Predisposizione e invio e messa in disponibilità sul sito della documentazione relativa alla rilevazione e degli atti di regolamentazione (piani di censimento, circolari e istruzioni agli organi di rilevazione).		
	Svolgimento e verifica delle operazioni svolte dagli organi intermedi di rilevazione, dai rilevatori o dalle società.		
	Verifica della qualità del lavoro in corso.		
	Attività finalizzate allo svolgimento e all'eventuale attivazione del procedimento per accertamento della eventuale violazione obbligo di risposta e applicazione della sanzionabilità.		
	Verifica delle operazioni di chiusura del trattamento e della acquisizione dati.		
Trattamento statistico dei dati (controllo, correzione, trasmissione, registrazione, elaborazione)	Operazioni di controllo, correzione, trasmissione, registrazione, elaborazione dei dati svolte nel rispetto delle norme in materia di protezione dei dati personali e di segreto statistico applicabili a ogni singolo caso.		Qualora si preveda l'uso di sistemi di intelligenza artificiale ↓ Rispettare le prescrizioni previste dal Regolamento (UE) 2024/1689 ——— Esecuzione *fine tuning* ↓ Sviluppo del *prompting* (modalità di output) ↓ Esecuzione validazione soluzione

Fonte: Elaborazione degli autori
(a) Per un approfondimento, si rinvia alle discipline statistiche di settore e, sotto il profilo delle regole, a Bellitti e Fedeli 2022, Capitolo 1, paragrafo 1.

Prospetto 3.1 segue **- Principali regole e attività sotto il profilo giuridico, tecnico e metodologico nel processo di trattamento dei dati quantitativi e qualitativi** (a)

Conservazione	Adozione di misure di sicurezza dei dati trattati e conservazione nei *repository*.	Acquisizione dati nei *repository*.	
	Definizione dei tempi di conservazione dei dati personali in conformità alle disposizioni vigenti che prevedono che i dati possono essere conservati anche oltre il periodo necessario per il raggiungimento degli scopi per i quali sono stati raccolti o successivamente trattati, nel rispetto del di cui all´art. 5, par.1, lett. e) del Regolamento e all´art. 6-*bis* del decreto legislativo 6 settembre 1989, n. 322 e successive modificazioni e integrazioni (art. 10 Regole deontologiche Provvedimento Garante n. 514 del 19 dicembre 2018; art. 6-*bis*, comma 6, d.lgs. n. 322/1989; riferimento anche agli artt. 9 e 10 del GDPR).		
	I dati identificativi, in ogni caso, devono essere conservati separatamente da ogni altro dato, salvo che ciò risulti impossibile in ragione delle particolari caratteristiche del trattamento o comporti un impiego di mezzi manifestamente sproporzionati rispetto al diritto tutelato.	Acquisizione dati nei *repository* e adozione misure di sicurezza e di cybersicurezza.	
Comunicazione e diffusione	Invio e comunicazione a Eurostat e/o ad altre istituzioni, quando previsto dalla normativa.	Piattaforme per la comunicazione e la diffusione.	Qualora si preveda l'uso di sistemi di intelligenza artificiale ⬇ Rispettare le prescrizioni previste dal Regolamento (UE) 2024/1689 ——— Valutazione *make or buy* ⬇ Esecuzione *fine tuning* ⬇ Definizione del *prompting* (modalità di output) ⬇ Validazione soluzione
	Comunicazione e diffusione dati nel rispetto delle disposizioni vigenti.		
	Rispetto dei criteri e delle procedure per la tutela del segreto statistico (minimizzazione del rischio di identificabilità delle unità statistiche).		

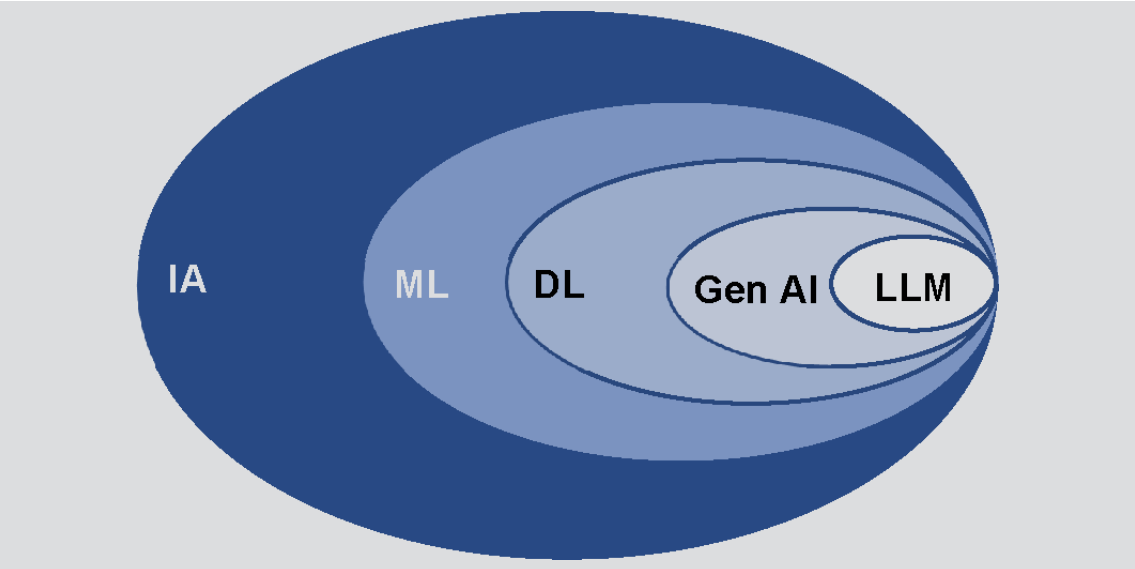
Fonte: Elaborazione degli autori
(a) Per un approfondimento, si rinvia alle discipline statistiche di settore e, sotto il profilo delle regole, a Bellitti e Fedeli 2022, Capitolo 1, paragrafo 1.

4. CASI DI USO DI APPLICAZIONE DELL' INTELLIGENZA ARTIFICIALE NEL TRATTAMENTO DEI DATI QUANTITATIVI E QUALITATIVI NELL' AMBITO GESTIONALE[1]

4.1 Introduzione e contesto di riferimento

La definizione di "sistema di intelligenza artificiale" copre diversi concetti tra loro interconnessi, come *machine learning*, *deep learning* o IA generativa, che presentano differenze importanti e sostanziali, come meglio illustrato di seguito. Riferendoci a quanto esplicitato nell'*AI Act*, pubblicato il 12 luglio 2024 sulla Gazzetta ufficiale europea, all'articolo 3, punto 1, riguardo ai sistemi di intelligenza artificiale, riportiamo, in modo sintetico, i diversi concetti a esso correlati[2].

Figura 4.1 - Sistemi di intelligenza artificiale



Fonte: UNECE 2023

- IA - Intelligenza artificiale: è un vasto campo dell'informatica riguardante la creazione di sistemi e macchine in grado di eseguire compiti specifici come l'apprendimento, il ragionamento, la comprensione del linguaggio e altro ancora che, normalmente, sono caratteristici dell'intelligenza umana;
- ML - *Machine learning*, o apprendimento automatico: è un sottoinsieme dell'IA che prevede l'uso di algoritmi e modelli statistici per consentire ai computer di migliorare le proprie capacità su un compito specifico attraverso l'apprendimento automatico dai dati, ovvero senza essere esplicitamente programmati;

1 Il Capitolo è stato curato da M. Fedeli, C. Colasanti e S. Letardi, e redatto da M. Bruno (paragrafo 4.1), S. Letardi (paragrafi 4.1 e 4.2), F. Davide e P. Torre (Prospetto 4.3 e paragrafi 4.3, 4.4, 4.5, 4.6).

2 Per questo paragrafo faremo riferimento a quanto contenuto in UNECE 2023.

- DL - *Deep learning*, o apprendimento profondo: è un sottoinsieme del ML che impiega reti neurali artificiali con molti strati interconnessi (*deep neural network*). È particolarmente adatto a compiti che coinvolgono dati complessi e non strutturati come immagini, audio e testo;
- Gen AI - *Generative AI*, Intelligenza artificiale generativa: comprende sistemi di IA in grado di generare nuovi contenuti o dati che non sono esplicitamente derivati da esempi esistenti. Ciò può includere la generazione di testo, immagini, musica e altro ancora;
- LLM - *Large Language Model*: rappresentano un'applicazione specifica della *Generative AI*, specializzata nella generazione di linguaggio naturale. In questo campo troviamo, ad esempio, modelli come LLama e GPT (*Generative Pre-trained Transformer*), che utilizzano le capacità della *Generative AI* per produrre testi con notevole qualità e coerenza.

I *Large Language Model* hanno potenzialmente diversi campi di applicazione, in cui possono coadiuvare il lavoro di ricercatori e tecnologi nel processo di produzione dell'informazione statistica, in quanto sono specializzati nella comprensione del testo, nella sintesi di grandi quantità di informazioni e nella generazione di risposte che imitano quelle umane.

Ad esempio, possono contribuire nelle seguenti fasi del processo GSBPM (*Generic Statistical Business Process Model*):

- *Design collection*: contribuire alla progettazione di questionari o alla loro traduzione in più lingue;
- *Classify and code*: classificare automaticamente i dati testuali in categorie predefinite;
- *Produce dissemination products*: generare descrizioni testuali a partire da insiemi di dati o tabelle.

Inoltre, i *Large Language Model* risultano utili ed efficienti per diverse attività trasversali e fondamentali di un'organizzazione, contribuendo a raggiungere diversi benefici e obiettivi strategici, come l'ottimizzazione delle tempistiche per la produzione di *report* e comunicati stampa, la facilitazione nell'accesso alle informazioni statistiche, il miglioramento dell'immagine dell'Istat, la personalizzazione dell'esperienza formativa per l'accrescimento delle competenze, la riduzione del *burden* statistico.

Esempi di come tali obiettivi possono essere raggiunti verranno descritti in dettaglio nel paragrafo dedicati ai casi di uso che si stanno sviluppando in Istat.

Di seguito presentiamo alcuni scenari di interesse che illustrano altre possibili applicazioni dell'intelligenza artificiale in Istat.

- Ricerca e sintesi: in tale scenario si prevede l'utilizzo di *Large Language Model* per lo sviluppo di un *chatbot* (o *Conversational* AI) in grado di dialogare con un utente, con diversi livelli di esperienza, sul contenuto di un insieme di documenti di argomento specifico. Tali documenti, a seconda della tipologia del caso di uso sviluppato, possono riguardare, ad esempio, contratti, procedure amministrative o manualistica. Utilizzando il *chatbot* l'utente può, in modo conversazionale, porre delle domande e ottenere delle risposte.
- Generazione di codice: in questo scenario, i *Large Language Model* specializzati nella generazione del codice assistono i programmatori in diverse attività relative allo sviluppo e alla manutenzione del codice come l'elaborazione, la revisione, la correzione o la documentazione del codice. Inoltre, possono coadiuvare il lavoro degli sviluppatori nella generazione di test automatici prima della messa in produzione delle applicazioni.

- *Enhanced data analysis*: questo scenario prevede che la Gen AI possa essere usata per assistere le attività dei *data analyst* e degli statistici ponendo all'applicativo di IA domande in linguaggio naturale e ricevendo risposte in forma di narrazione, di grafici o direttamente come *report* completi.
- *Learning and reskilling*: un possibile ambito di utilizzo prevede di fornire agli utenti strumenti a supporto dell'acquisizione di nuove competenze o professionalità. Mediante una interfaccia conversazionale, è possibile personalizzare l'interazione dell'utente con gli strumenti di formazione, adattandoli al proprio livello di esperienza o conoscenza dell'argomento in esame.
- Generazione di dati sintetici: in questo contesto si prevede la generazione di dati artificiali, ovvero non rilevati in campo ma aventi le stesse caratteristiche statistiche di uno specifico *dataset*. I dati sintetici generati risultano privi della presenza di dati personali e/o sensibili e possono essere utilizzati per addestrare modelli di *machine learning* con vincoli di privacy e *compliance* meno stringenti.
- Generazione di contenuti creativi: in questo caso i modelli generativi, non solo del linguaggio, ma soprattutto di immagini e video, quali DALL-E, *Stable diffusion* e *MidJourney*, vengono utilizzati per creare immagini e contenuti testuali adatti alle esigenze di comunicazione con gli utenti, sempre nel rispetto delle norme relative alla proprietà intellettuale.

4.2 Le azioni e le sperimentazioni

L'Istat ha attivato un percorso di verifica e valutazione di impatto dei sistemi basati su intelligenza artificiale mediante l'analisi delle nuove disposizioni in materia, che regolano in modo onnicomprensivo lo sviluppo, l'immissione sul mercato e l'uso dei sistemi di IA.

Tale processo si affianca alle disposizioni già esistenti in ambito tecnico-statistico e di digitalizzazione[3], che costituiscono, attraverso un approccio basato sul rischio, l'elemento essenziale per l'individuazione di potenziali vulnerabilità e aree non efficacemente protette.

Le sperimentazioni attivate in Istat riguardano, in particolare, gli ambiti relativi, da una parte, al trattamento dei dati per finalità di produzione di strumenti quantitativi e qualitativi e, dall'altra, agli aspetti che comportano l'adozione di nuove tecniche gestionali.

Da un resoconto preliminare si evidenzia come il percorso intrapreso aspiri a migliorare la qualità dei servizi interni ed esterni, attraverso l'implementazione di algoritmi che permettono di conseguire una elevata potenzialità nell'erogazione delle prestazioni.

I molteplici benefici consentono, infatti, agli operatori di semplificare e di adottare decisioni informate e accurate, con lo scopo di assicurare la qualità delle *performance*.

Se da un lato l'utilizzo o lo sviluppo di sistemi di IA può comportare, in termini di efficienza e produttività, l'ottimizzazione delle prestazioni, dall'altro concede all'utente un'esperienza migliore e una maggiore fruibilità dei contenuti.

Un esempio lampante è dato dall'integrazione di soluzioni software che prevedono procedure di calcolo di grandi quantità di dati.

3 Cfr. Bellitti e Fedeli 2022, paragrafo 1.2: "Il processo del trattamento del dato ai fini quantitativi e qualitativi e gli strumenti di produzione".

Tramite il ricorso a tali sistemi di automazione sarà semplice ottenere risultati immediati, utili all'operatore, con l'obiettivo di ridurre l'errore umano, contemplando i principi di trasparenza[4], di non esclusività della decisione algoritmica e di non discriminazione algoritmica[5].

A ogni modo, come previsto dal Regolamento, in fase di sperimentazione sarà necessario mantenere una visione antropocentrica, tenendo conto dei principi e delle regole alla base delle discipline menzionate (cfr. Capitolo 1 e Capitolo 2).

Di seguito sono analizzati i casi di uso censiti, i diversi profili tecnologici e gli obiettivi posti alla base della sperimentazione.

In particolare, nel Prospetto 4.1 sono evidenziati per ciascuna esperienza gestionale i contributi alla Strategia IA.

Prospetto 4.1 - Sperimentazioni ed esperienze gestionali

CASI DI USO	OBIETTIVI/ELEMENTI DISTINTIVI	BENEFICI
Assistente virtuale	Introduzione di un assistente virtuale per migliorare la produttività.	• Rapidità nella creazione di differenti tipologie di documenti (ad esempio documenti di testo o presentazioni). • Ricerca efficace di argomenti disponibili nei documenti.
Punto unico di contatto	Introduzione di una soluzione di IA generativa all'interno del Punto unico di contatto. In particolare, è in corso la fase di addestramento di un *chatbot* per orientare gli utenti a selezionare il servizio appropriato. Successivamente è prevista l'implementazione di una soluzione dedicata al supporto agli operatori e, laddove possibile, per le sole richieste di raccolta dati, alla soluzione automatica dei casi.	• Miglioramento della qualità delle informazioni fornite agli utenti in termini di coerenza e accuratezza. • Miglioramento della soddisfazione degli utenti grazie alla semplificazione dell'accesso al sistema. • Miglioramento della *reputation* dell'Istituto: realizzazione di un'immagine fortemente identitaria, più moderna e al passo con i tempi. • Ottimizzazione dei processi interni: miglioramento dell'efficienza nelle attività di erogazione dei servizi tramite il supporto dell'IA. • Diminuzione del tempo di risposta all'utente mediante la riduzione dei compiti manuali, frequenti, ad alta standardizzazione.
Soluzioni di IA generativa per la formazione	Miglioramento della qualità dello sviluppo delle competenze del personale Istat e ottimizzazione dei processi di formazione attraverso l'utilizzo di IA generativa.	Analisi degli ambiti, delle modalità, delle potenzialità, dei limiti dell'impiego dell'IA generativa per rendere più efficace ed efficiente l'intero processo formativo, dall'identificazione dei bisogni, alla progettazione dei corsi, alla produzione di programmi e contenuti, alle possibilità di personalizzazione dell'esperienza formativa, al monitoraggio, al *tutoring* e alla valutazione.

Fonte: Elaborazione degli autori

4 Cfr. Regolamento generale sulla protezione dei dati (GDPR) dell'Unione europea, in particolare l'articolo 22, che si occupa della "Profilazione e processi decisionali automatizzati"; *Orientamenti etici per un'IA affidabile* della Commissione europea, che sottolineano l'importanza della trasparenza nei sistemi di intelligenza artificiale.

5 Cfr. Regolamento generale sulla protezione dei dati (GDPR) dell'Unione europea, nello stesso articolo 22, che menziona il diritto degli individui di non essere soggetti a decisioni basate unicamente su processi automatizzati che hanno effetti legali significativi; Linee guida dell'OECD sull'IA che promuovono l'uso responsabile e non discriminatorio dell'intelligenza artificiale; *Orientamenti etici per un'IA affidabile* della Commissione europea, che trattano della necessità di garantire che le decisioni algoritmiche non siano discriminatorie e siano soggette a supervisione umana.

Prospetto 4.2 - Sperimentazioni ed esperienze del settore del trattamento dei dati per finalità di strumenti quantitativi e qualitativi

CASI DI USO	OBIETTIVI/ELEMENTI DISTINTIVI	BENEFICI
Motore di ricerca semantico	Predisposizione di un motore di ricerca sul nuovo sito istituzionale. L'obiettivo principale è sviluppare un sistema di estrazione delle informazioni semantiche dal sito istituzionale dell'Istat e dai comunicati in esso contenuti, per migliorare l'accessibilità e il grado di reperibilità delle informazioni al fine di indirizzare gli utenti verso le risorse (pagine o documenti) che rispondono al meglio alle loro richieste.	Possibilità di rispondere più rapidamente e accuratamente alle richieste interne ed esterne, con l'obiettivo di aumentare la qualità del servizio verso l'utenza.
Soluzione di IA per la produzione di *report* multilingue	Realizzazione di una soluzione di intelligenza artificiale generativa che supporti il processo di traduzione in più lingue dei *report* prodotti da Istat.	• Diminuzione del *time to deliver*. • Riutilizzo della soluzione per differenti ambiti e casi di uso. • Rapidità nella produzione: riduzione tempo/costi per la traduzione. • Accesso inclusivo: promozione dell'inclusività, dell'accessibilità e della fruizione delle informazioni. • Miglioramento dell'integrazione.
Introduzione di *chatbot* su sito Istat	Introduzione di un *chatbot* sul sito istituzionale per dialogare con gli utenti e restituire risposte conversazionali, al fine di rendere più accessibile l'informazione ivi contenuta.	• Miglioramento della qualità delle informazioni. • Individuazione veloce delle informazioni corrette.
Partecipazione allo sviluppo di un LLM italiano	Partecipazione a un progetto di ricerca sperimentale di metodologie di impiego di IA generativa nella produzione di informazione statistica ufficiale. Gli obiettivi del progetto sono: • Sviluppo di un prototipo di IA generativa specifica per Istat caratterizzata da una gestione robusta dei dati residenti in *data center* sul territorio italiano, da una governance trasparente del *Large Language Model* e dalla conformità a tutte le regole italiane ed europee applicabili. • Formazione del proprio personale nel campo dell'IA mediante applicazione nei processi di sviluppo di modelli LLM.	• Accrescimento e miglioramento delle competenze interne in campo tecnologico, metodologico e legale. • Specializzazione del modello sulla banca dati Istat.
Soluzione per semplificare l'analisi di scontrini per l'Indagine sulle spese delle famiglie	Realizzazione di una soluzione per l'analisi dei dati inviati attraverso il caricamento di immagini dai rispondenti all'Indagine sulle spese delle famiglie.	Riduzione del *burden* per i rispondenti alle indagini, aumento del tasso di risposta, aumento della qualità dei dati rilevati, maggiore standardizzazione nella raccolta delle informazioni.
Progetto europeo "ESSNET One-stop-shop for Artificial Intelligence and Machine Learning for Official Statistics Project (AIML4OS)"	Sviluppo di metodi e strumenti innovativi più efficienti ed efficaci basati su IA/ML. Tra i casi di uso/WP previsti, Istat/DCME co-coordina con Statistics Austria il WP13: *Generation of synthetic data in official statistics*: *techniques and applications*.	Valutazione di soluzioni per la produzione di dati sintetici, finalizzata alla gestione di dati personali e sensibili con dati generati da algoritmi IA.
Utilizzo dell'IA per l'accesso a IstatData	Definizione di una soluzione che supporti gli utenti nella navigazione guidata dei dati IstatData, attraverso l'uso di IA generativa.	Valorizzazione del patrimonio informativo prodotto da Istat attraverso l'uso controllato di una soluzione basata su IA generativa.

Fonte: Elaborazione degli autori

Il Prospetto 4.3 sintetizza i vari casi di uso presentati e li classifica rispetto alle aree di azione della Strategia IA per le pubbliche amministrazioni 2024-2026.

Prospetto 4.3 - I casi di uso identificati per l'Istat e loro posizionamento rispetto alla "Strategia Italiana per l'Intelligenza artificiale 2024-2026"

CASI DI USO	AREA DELLA STRATEGIA IA PER LA PUBBLICA AMMINISTRAZIONE 2024-2026
Ricerca soluzioni IA per Istat	PA.1 | Linee guida per promuovere l'adozione dell'IA nella Pubblica amministrazione
Soluzione produzione comunicati stampa	PA.5 | Efficientamento della Pubblica amministrazione
Proposta per assistente virtuale	PA.5 | Efficientamento della Pubblica amministrazione
Punto unico di contatto	PA.4 | Semplificazione per cittadini e imprese
Sperimentazione di soluzioni di IA generativa per la formazione	PA.5 | Efficientamento della Pubblica amministrazione
Motore di ricerca semantico	PA.4 | Semplificazione per cittadini e imprese
Soluzione produzione multilingue di *report*	PA.5 | Efficientamento della Pubblica amministrazione
Sperimentazione introduzione di *chatbot* su sito Istat	PA.4 | Semplificazione per cittadini e imprese
Partecipazione allo sviluppo di un LLM italiano con *Unint* e *Fastweb*	PA.3 | Linee guida per la realizzazione di applicazioni di IA nella Pubblica amministrazione
Soluzione per semplificare l'analisi di scontrini per l'indagine su spese famiglia	PA.3 | Linee guida per la realizzazione di applicazioni di IA nella Pubblica amministrazione
Progetto europeo ""ESSNET One-stop-shop for Artificial Intelligence and Machine Learning for Official Statistics Project (AIML4OS)"	PA.5 | Efficientamento della Pubblica amministrazione
Sperimentazione per l'utilizzo dell'IA per l'accesso a IstatData	PA.5 | Efficientamento della Pubblica amministrazione
GenAI4IOP	PA.4 | Semplificazione per cittadini e imprese e PA.5 | Efficientamento della Pubblica amministrazione

Fonte: Elaborazione degli autori

4.3 Intelligenza artificiale per l'interoperabilità

Un caso di uso dell'IA di crescente interesse per l'Istat e per la Presidenza del Consiglio dei ministri è quello di supporto all'interoperabilità tra gli enti della Pubblica amministrazione, ove l'interoperabilità venga intesa nell'accezione molto ampia dell'*Interoperable Europe Act* (entrato in vigore il giorno 11 aprile 2024), cioè come la capacità delle organizzazioni di interagire verso obiettivi reciprocamente vantaggiosi. L'utilizzo delle tecnologie dell'IA può offrire al settore pubblico strumenti per superare, ad esempio, le situazioni in cui i sistemi siano incompatibili, i dati frammentati e gli standard non univoci.

In questo ambito, spicca nel contesto italiano il progetto "Catalogo nazionale della semantica dei dati" (esposto pubblicamente in https://schema.gov.it/), nato nell'ambito del PNRR (Piano nazionale di ripresa e resilienza), che vede il Dipartimento per la trasformazione digitale come titolare e l'Istat come responsabile della realizzazione del Catalogo.

Il Catalogo rientra nella quinta parte della "Strategia italiana per l'intelligenza artificiale 2024-2026" e ha un grande potenziale di applicazione dell'IA per l'interoperabilità: un nuovo campo di utilizzo dell'IA denominato AI4IOP dalla Commissione europea[6].

Una *survey* completa dei possibili utilizzi dell'IA per l'interoperabilità in Europa è offerta dall'ultimo *report* della Commissione europea sull'interoperabilità nel settore pubblico europeo[7]: il *report* fornisce un'analisi quantitativa di 189 casi di uso dell'IA nel settore pubblico, selezionati per la loro rilevanza, e uno studio qualitativo che approfondisce alcuni casi. In ben il 26 per cento dei casi l'IA è impiegata proprio per contribuire all'interoperabilità del settore pubblico.

6 Cfr. Davide 2024.
7 Cfr. Tangi *et al.* 2023.

Tra i vari strati di interoperabilità, rubricati secondo l'*European Interoperability Framework*, è lo strato semantico dell'interoperabilità a essere in gioco nella stragrande maggioranza (91 per cento).

L'IA è quindi utilizzata in combinazione con ontologie e tassonomie per creare un linguaggio comune e un significato condiviso dei dati, aiutando notevolmente l'interoperabilità semantica tra organizzazioni e sistemi differenti.

Questo è esattamente il campo di azione del Catalogo nazionale dati, l'ambito in cui si sta implementando in Italia il regolamento di interoperabilità e si sta trasformando il settore pubblico europeo in un ecosistema più connesso, più agile e più efficiente nel fornire i servizi.

Il Catalogo appartiene al 36 per cento dei casi di AI4IOP in cui si cerca di contribuire alla gestione dei processi interni del settore pubblico, fornendo vantaggi come l'identificazione delle incoerenze normative, la semplificazione dei processi amministrativi, la velocizzazione dei processi decisionali, la creazione di vocabolari e la facilitazione del collegamento dei dati in domini specifici.

Entrando nel merito più operativo, le classi di azioni che i sistemi basati su IA compiono sui dati per conseguire l'interoperabilità sono soprattutto: Rilevazione (42 per cento), Strutturazione (22 per cento) e Classificazione (16 per cento).

Queste azioni si avvantaggiano della capacità dell'IA di elaborare automaticamente quantità enormi di dati, e sono soprattutto volte a trasformare i dati in formati standard mentre li filtrano per migliorarne la qualità, e quindi facilitarne la condivisione tra sistemi differenti.

4.4 Impiego di sistemi di intelligenza artificiale nell'attività gestionale

Il ricorso dell'IA è incentivato anche nell'espletamento delle funzioni amministrative. Sul fronte dei contratti, il codice del 2023 disciplina l'uso di procedure automatizzate nel ciclo di vita dei contratti pubblici: l'articolo 30, comma 1 del d.lgs. 31 marzo 2023, n. 36, stabilisce infatti che, per migliorare l'efficienza, le stazioni appaltanti e gli enti concedenti provvedono, ove possibile, "ad automatizzare le proprie attività ricorrendo a soluzioni tecnologiche, ivi incluse l'intelligenza artificiale e le tecnologie di registri distribuiti".

Nello svolgimento della propria attività, l'Istituto, come ogni altra organizzazione, effettua dei trattamenti per finalità amministrative di supporto alla *mission* istituzionale e in tale contesto l'IA offre un enorme potenziale per migliorare l'efficienza, l'efficacia e l'equità nell'assunzione e nella gestione del personale.

Tuttavia, è importante utilizzarla in modo responsabile ed etico, affrontando le sfide e le preoccupazioni che ne derivano.

4.5 Modelli architetturali

Come riferimento comune per confrontare le possibili soluzioni architetturali si può ricorrere a uno schema di classificazione (Figura 4.2) basato su tre "dimensioni caratteristiche":

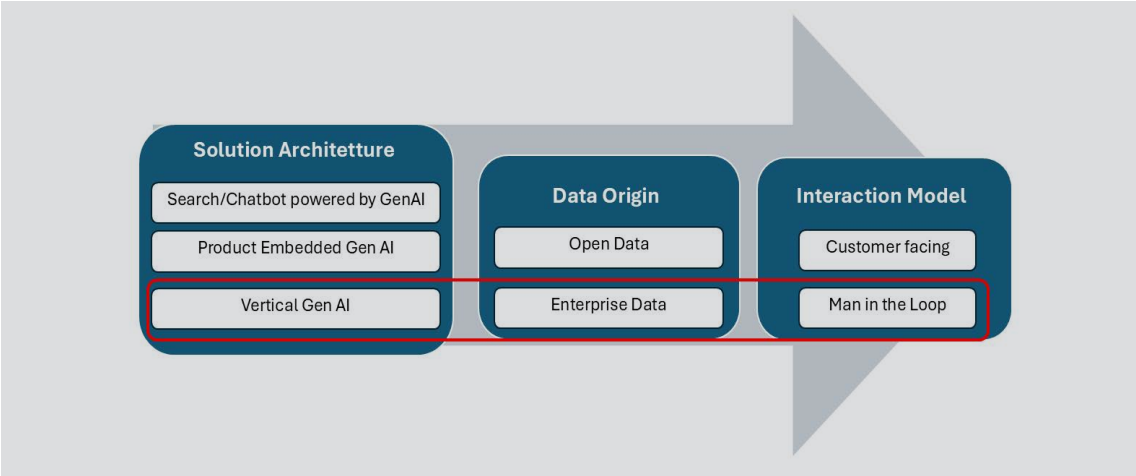
a. *Solution architecture*. È la dimensione che rappresenta la modalità con la quale i *Large Language Model* (LLM) sono integrati all'interno della soluzione informatica complessiva. In prima analisi si possono riconoscere le tre seguenti modalità, di complessità tecnologica crescente:

1. *Search/Chatbot powered by Gen AI*: architetture generiche basate su *chat engine* oppure motori di ricerca Gen AI nativi, prodotti da fornitori esterni, integrati con *plugin* forniti dai fornitori stessi per l'interfacciamento con il patrimonio informativo interno (come nel caso degli strumenti di *office automation*).
2. *Product embedded Gen AI*: qualsiasi soluzione di integrazione che integra tecnologie Gen AI *native* con l'aggiunta di *middleware* e/o connettori proprietari;
3. *Vertical Gen AI*: una soluzione *make*, con *tuning* sui dati privati dell'Istat e realizzata con minima dipendenza dal fornitore (*lock-in*), utilizzando alcune specifiche API (*Application Programming Interface*) esposte dai fornitori IA.

b. *Data origin*. Questa dimensione indica se i dati, testuali o multimediali, sono di ambito *enterprise*, cioè provenienti solo dal patrimonio informativo dell'Istat, oppure se la soluzione consente l'utilizzo di dati provenienti da fonti esterne.

c. *Interaction model*. Descrive l'interazione diretta (*customer facing*) o intermediata (*man in the loop*) degli utenti finali con le tecnologie generative.

Figura 4.2 - Modelli architetturali



Fonte: Elaborazione degli autori

Per gli scenari di utilizzo prima descritti nel contesto dell'Istituto, possiamo anticipare alcune considerazioni utili per la classificazione delle soluzioni da implementare.

Per il modello di interazione, nel settore pubblico in generale risulta preferibile il tipo *man in the loop*, poiché consente di gestire e controllare il rischio di allucinazione e tossicità dei *Large Language Model*, cioè la generazione di risposte e contenuti non conformi a quanto atteso o contrari alle *policy* dell'Istat.

Relativamente alla *solution architecture*, da una prima analisi degli scenari di utilizzo è emerso che questi ricadono principalmente in soluzioni di tipo *vertical Gen AI* e in particolare convergano in modo significativo su due *pattern* architetturali appartenenti alla categoria RAG (*Retrieval Aumented Generation*) di seguito elencati:

- *AI summarisation chain*: a fronte di un sistema documentale contenente i dati dell'Istituto (*enterprise*), la soluzione consente di realizzare una catena di azioni più o meno complessa di *search, reorganisation, formatting, prompting* in grado di generare un testo o schemi riassuntivi di uno o più concetti semantici da trasmettere all'utente finale dopo intermediazione umana.

- *AI data analytics*: a fronte di dati rilevati dall'esterno (ad esempio su base censimento), la soluzione consente di svolgere un *pre-staging* della produzione statistica, di prevedere risultati finali prima che la raccolta sia completa, di individuare anomalie nel processo di raccolta e produzione.

Relativamente alla *data origin*, l'Istat per sua natura e scelta limita gli scenari a casi di dati *enterprise*.

Moltissimi casi di uso sono stati ricondotti a questi *pattern* architetturali, nella cui prototipazione l'Istat sta attivamente investendo per perfezionarne l'applicazione in molteplici processi, tra cui quelli dell'ottimizzazione all'accesso delle principali sorgenti informative dell'Istat, preservando gli aspetti di privacy/GDPR e *compliance*.

Tra i molti possibili *pattern* architetturali generati dallo schema della Figura 4.2, resta prevalente, nel caso dell'Istituto, il *pattern vertical Gen AI - enterprise data - man in the loop*.

4.6 Considerazioni finali

In queste sezioni sono stati, dunque, illustrati alcuni scenari di adozione della tecnologia alla base dei *Large Language Model* e la strategia che si sta adottando in Istat per censire, raggruppare, filtrare e implementare i casi di uso di Gen AI che siano al contempo di impatto, in termini di ritorno dei benefici sui costi, e di utilità per una molteplicità di funzioni dell'Istat.

Sebbene gli scenari di interesse promettano ritorni importanti in termini di aumento della produttività dei lavoratori, si pongono al contempo sfide di tipo tecnologico, normativo e anche di rischio reputazionale, da affrontare in modo organico, al fine di governare e mettere a valore le opportunità offerte da questa tecnologia.

GLOSSARIO[1]

Alfabetizzazione in materia di intelligenza artificiale (IA)
Le competenze, le conoscenze e la comprensione che consentono ai fornitori, ai *deployer* e alle persone interessate, tenendo conto dei loro rispettivi diritti e obblighi nel contesto del Regolamento (UE) 2024/1689, di procedere a una diffusione informata dei sistemi di IA, nonché di acquisire consapevolezza in merito alle opportunità e ai rischi dell'IA e ai possibili danni che essa può causare.

Autorità di notifica
L'autorità nazionale responsabile dell'istituzione e dell'esecuzione delle procedure necessarie per la valutazione, la designazione e la notifica degli organismi di valutazione della conformità e per il loro monitoraggio.

Autorità di vigilanza del mercato
L'autorità nazionale che svolge le attività e adotta le misure a norma del Regolamento (UE) 2019/1020.

Autorità nazionale competente
Un'autorità di notifica o un'autorità di vigilanza del mercato; per quanto riguarda i sistemi di IA messi in servizio o utilizzati da istituzioni, organi e organismi dell'Unione europea, i riferimenti alle autorità nazionali competenti o alle autorità di vigilanza del mercato contenuti nel Regolamento (UE) 2024/1689 si intendono fatti al Garante europeo della protezione dei dati.

Capacità di impatto elevato
Capacità che corrispondono o superano le capacità registrate nei modelli di IA per finalità generali più avanzati.

Categorie particolari di dati personali
Le categorie di dati personali di cui all'articolo 9, par. 1, del Regolamento (UE) 2016/679, all'articolo 10 della Direttiva (UE) 2016/680 e all'articolo 10, par. 1, del Regolamento (UE) 2018/1725.

Componente di sicurezza
Un componente di un prodotto o di un sistema di IA che svolge una funzione di sicurezza per tale prodotto o sistema di IA o il cui guasto o malfunzionamento mette in pericolo la salute e la sicurezza di persone o beni.

1 Le definizioni sono quelle riportate nell'art. 3 del Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024 (*AI Act*) e nell'art. 3 del Regolamento (CE) n. 223/2009 del Parlamento europeo e del Consiglio dell'11 marzo 2009.

Consenso informato
L'espressione libera, specifica, inequivocabile e volontaria di un soggetto della propria disponibilità a partecipare a una determinata prova in condizioni reali, dopo essere stato informato di tutti gli aspetti della prova rilevanti per la sua decisione di partecipare.

Dati
Qualsiasi rappresentazione digitale o non digitale di atti, fatti o informazioni e qualsiasi raccolta di tali atti, fatti o informazioni sulle unità osservate.

Dati biometrici
I dati personali ottenuti da un trattamento tecnico specifico relativi alle caratteristiche fisiche, fisiologiche o comportamentali di una persona fisica, quali le immagini facciali o i dati dattiloscopici.

Dati di addestramento
I dati utilizzati per addestrare un sistema di IA adattandone i parametri che può apprendere.

Dati di convalida
I dati utilizzati per fornire una valutazione del sistema di IA addestrato e per metterne a punto, tra l'altro, i parametri che non può apprendere e il processo di apprendimento, al fine tra l'altro di evitare lo scarso (*underfitting*) o l'eccessivo (*overfitting*) adattamento ai dati di addestramento.

Dati di input
I dati forniti a un sistema di IA o direttamente acquisiti dallo stesso, in base ai quali il sistema produce un output.

Dati di prova
I dati utilizzati per fornire una valutazione indipendente del sistema di IA al fine di confermarne le prestazioni attese prima della sua immissione sul mercato o messa in servizio.

Dati non personali
Dati diversi dai dati personali di cui all'articolo 4, punto 1), del Regolamento (UE) 2016/679.

Dati operativi sensibili
Dati operativi relativi ad attività di prevenzione, accertamento, indagine o perseguimento di reati, la cui divulgazione potrebbe compromettere l'integrità dei procedimenti penali.

Dati personali
I dati personali quali definiti all'articolo 4, punto 1), del Regolamento (UE) 2016/679.

Deep fake
Un'immagine o un contenuto audio o video generato o manipolato dall'IA che assomiglia a persone, oggetti, luoghi, entità o eventi esistenti e che apparirebbe falsamente autentico o veritiero a una persona.

Deployer
Una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che utilizza un sistema di IA sotto la propria autorità, tranne nel caso in cui il sistema di IA sia utilizzato nel corso di un'attività personale non professionale.

Distributore
Una persona fisica o giuridica nella catena di approvvigionamento, diversa dal fornitore o dall'importatore, che mette a disposizione un sistema di IA sul mercato dell'Unione europea.

Finalità prevista
L'uso di un sistema di IA previsto dal fornitore, compresi il contesto e le condizioni di uso specifici, come dettagliati nelle informazioni comunicate dal fornitore nelle istruzioni per l'uso, nel materiale promozionale o di vendita e nelle dichiarazioni, nonché nella documentazione tecnica.

Fornitore
Una persona fisica o giuridica, un'autorità pubblica, un'agenzia o un altro organismo che sviluppa un sistema di IA o un modello di IA per finalità generali o che fa sviluppare un sistema di IA o un modello di IA per finalità generali e immette tale sistema o modello sul mercato o mette in servizio il sistema di IA con il proprio nome o marchio, a titolo oneroso o gratuito.

Fornitore a valle
Un fornitore di un sistema di IA, compreso un sistema di IA per finalità generali, che integra un modello di IA, indipendentemente dal fatto che il modello di IA sia fornito dallo stesso e integrato verticalmente o fornito da un'altra entità sulla base di relazioni contrattuali.

Identificazione biometrica
Il riconoscimento automatizzato delle caratteristiche umane fisiche, fisiologiche, comportamentali o psicologiche allo scopo di determinare l'identità di una persona fisica confrontando i suoi dati biometrici con quelli di individui memorizzati in una banca dati.

Immissione sul mercato
La prima messa a disposizione di un sistema di IA o di un modello di IA per finalità generali sul mercato dell'Unione europea.

Importatore
Una persona fisica o giuridica ubicata o stabilita nell'Unione europea che immette sul mercato un sistema di IA recante il nome o il marchio di una persona fisica o giuridica stabilita in un paese terzo.

Incidente grave
Un incidente o malfunzionamento di un sistema di IA che, direttamente o indirettamente, causa una delle conseguenze seguenti: a) il decesso di una persona o gravi danni alla salute di una persona; b) una perturbazione grave e irreversibile della gestione o del funzionamento delle infrastrutture critiche; c) la violazione degli obblighi a norma del diritto dell'Unione europea intesi a proteggere i diritti fondamentali; d) gravi danni alle cose o all'ambiente.

Infrastruttura critica
Infrastruttura critica quale definita all'articolo 2, punto 4), della Direttiva (UE) 2022/2557.

Infrazione diffusa
Qualsiasi azione od omissione contraria al diritto dell'Unione europea che tutela gli interessi delle persone:

a) che abbia arrecato o possa arrecare un danno agli interessi collettivi di persone che risiedono in almeno due Stati membri diversi dallo Stato membro in cui:
    i) ha avuto origine o si è verificata l'azione o l'omissione in questione;
    ii) è ubicato o stabilito il fornitore interessato o, se del caso, il suo rappresentante autorizzato; oppure
    iii) è stabilito il *deployer*, quando la violazione è commessa dal *deployer*;

b) che abbia arrecato, arrechi o possa arrecare un danno agli interessi collettivi di persone e che presenti caratteristiche comuni, compresa la stessa pratica illecita e lo stesso interesse leso e che si verifichi simultaneamente, commessa dal medesimo operatore, in almeno tre Stati membri.

Istruzioni per l'uso
Le informazioni comunicate dal fornitore per informare il *deployer* in particolare della finalità prevista e dell'uso corretto di un sistema di IA.

Marcatura CE
Una marcatura mediante la quale un fornitore indica che un sistema di IA è conforme ai requisiti stabiliti al capo III, sezione 2 del Regolamento (UE) 2024/1689, e in altre normative di armonizzazione dell'Unione europea applicabili e che ne prevedono l'apposizione.

Messa a disposizione sul mercato
La fornitura di un sistema di IA o di un modello di IA per finalità generali per la distribuzione o l'uso sul mercato dell'Unione europea nel corso di un'attività commerciale, a titolo oneroso o gratuito.

Messa in servizio
La fornitura di un sistema di IA direttamente al *deployer* per il primo uso o per uso proprio nell'Unione europea per la finalità prevista.

Metadati
Qualsiasi informazione che definisce e descrive dati e processi.

Modello di IA per finalità generali
Un modello di IA, anche laddove tale modello di IA sia addestrato con grandi quantità di dati utilizzando l'autosupervisione su larga scala, che sia caratterizzato da una generalità significativa e sia in grado di svolgere con competenza un'ampia gamma di compiti distinti, indipendentemente dalle modalità con cui il modello è immesso sul mercato, e che può essere integrato in una varietà di sistemi o applicazioni a valle, a eccezione dei modelli di IA utilizzati per attività di ricerca, sviluppo o prototipazione prima di essere immessi sul mercato.

Modifica sostanziale
Una modifica di un sistema di IA a seguito della sua immissione sul mercato o messa in servizio che non è prevista o programmata nella valutazione iniziale della conformità effettuata dal fornitore e che ha l'effetto di incidere sulla conformità del sistema di IA ai requisiti di cui al capo III, sezione 2 del Regolamento (UE) 2024/1689, o comporta una modifica della finalità prevista per la quale il sistema di IA è stato valutato.

Norma armonizzata
La norma armonizzata di cui all'articolo 2, punto 1), lettera c), del Regolamento (UE) n. 1025/2012.

Operatore
Un fornitore, un fabbricante del prodotto, un *deployer,* un rappresentante autorizzato, un importatore o un distributore.

Operazione in virgola mobile
Qualsiasi operazione o assegnazione matematica che comporta numeri in virgola mobile, un sottoinsieme dei numeri reali generalmente rappresentati sui computer mediante un numero intero con precisione fissa avente come fattore di scala un esponente intero di una base fissa.

Organismo di valutazione della conformità
Un organismo che svolge per conto di terzi attività di valutazione della conformità, incluse prove, certificazioni e ispezioni.

Organismo notificato
Un organismo di valutazione della conformità notificato in conformità del Regolamento (UE) 2024/1689 e di altre pertinenti normative di armonizzazione dell'Unione europea.

Piano dello spazio di sperimentazione
Un documento concordato tra il fornitore partecipante e l'autorità competente in cui sono descritti gli obiettivi, le condizioni, il calendario, la metodologia e i requisiti relativamente alle attività svolte all'interno dello spazio di sperimentazione.

Piano di prova in condizioni reali
Un documento che descrive gli obiettivi, la metodologia, l'ambito geografico, della popolazione e temporale, il monitoraggio, l'organizzazione e lo svolgimento della prova in condizioni reali.

Prestazioni di un sistema di IA
La capacità di un sistema di IA di conseguire la finalità prevista.

Profilazione
La profilazione quale definita all'articolo 4, punto 4), del Regolamento (UE) 2016/679.

Prova in condizioni reali
La prova temporanea di un sistema di IA per la sua finalità prevista in condizioni reali al di fuori di un laboratorio o di un ambiente altrimenti simulato al fine di raccogliere dati affidabili e solidi e di valutare e verificare la conformità del sistema di IA ai requisiti del Regolamento (UE) 2024/1689 e che non è considerata immissione sul mercato o messa in servizio del sistema di IA ai sensi di questo Regolamento, purché siano soddisfatte tutte le condizioni di cui all'articolo 57 o 60.

Rappresentante autorizzato
Una persona fisica o giuridica ubicata o stabilita nell'Unione europea che ha ricevuto e accettato un mandato scritto da un fornitore di un sistema di IA o di un modello di IA per finalità generali al fine, rispettivamente, di adempiere ed eseguire per suo conto gli obblighi e le procedure stabiliti dal Regolamento (UE) 2024/1689.

Richiamo di un sistema di IA
Qualsiasi misura volta a ottenere la restituzione al fornitore, la messa fuori servizio o la disabilitazione dell'uso di un sistema di IA messo a disposizione dei *deployer*.

Rischio
La combinazione della probabilità del verificarsi di un danno e la gravità del danno stesso.

Rischio sistemico
Un rischio specifico per le capacità di impatto elevato dei modelli di IA per finalità generali, avente un impatto significativo sul mercato dell'Unione europea 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.

Ritiro di un sistema di IA
Qualsiasi misura volta a impedire che un sistema di IA nella catena di approvvigionamento sia messo a disposizione sul mercato.

Set di dati di convalida
Un *set* di dati distinto o costituito da una partizione fissa o variabile del *set* di dati di addestramento.

Sistema di categorizzazione biometrica
Un sistema di IA che utilizza i dati biometrici di persone fisiche al fine di assegnarle a categorie specifiche, a meno che non sia accessorio a un altro servizio commerciale e strettamente necessario per ragioni tecniche oggettive.

Sistema di IA
Un sistema automatizzato progettato per funzionare con livelli di autonomia variabili e che può presentare adattabilità dopo la diffusione e che, per obiettivi espliciti o impliciti, deduce dall'input che riceve come generare output quali previsioni, contenuti, raccomandazioni o decisioni che possono influenzare ambienti fisici o virtuali.

Sistema di IA per finalità generali
Un sistema di IA basato su un modello di IA per finalità generali e che ha la capacità di perseguire varie finalità, sia per uso diretto sia per integrazione in altri sistemi di IA.

Sistema di identificazione biometrica remota
Un sistema di IA finalizzato all'identificazione di persone fisiche, senza il loro coinvolgimento attivo, tipicamente a distanza mediante il confronto dei dati biometrici di una persona con i dati biometrici contenuti in una banca dati di riferimento.

Sistema di identificazione biometrica remota a posteriori
Un sistema di identificazione biometrica remota diverso da un sistema di identificazione biometrica remota «in tempo reale».

Sistema di identificazione biometrica remota in tempo reale
Un sistema di identificazione biometrica remota in cui il rilevamento dei dati biometrici, il confronto e l'identificazione avvengono senza ritardi significativi, il quale comprende non solo le identificazioni istantanee, ma anche quelle che avvengono con brevi ritardi limitati al fine di evitare l'elusione.

Sistema di monitoraggio successivo all'immissione sul mercato
Tutte le attività svolte dai fornitori di sistemi di IA al fine di raccogliere e analizzare l'esperienza maturata tramite l'uso dei sistemi di IA che immettono sul mercato o che mettono in servizio, al fine di individuare eventuali necessità di immediate azioni correttive o preventive.

Sistema di riconoscimento delle emozioni
Un sistema di IA finalizzato all'identificazione o all'inferenza di emozioni o intenzioni di persone fisiche sulla base dei loro dati biometrici.

Soggetto
Ai fini della prova in condizioni reali, una persona fisica che partecipa a prove in condizioni reali.

Spazio accessibile al pubblico
Qualsiasi luogo fisico di proprietà pubblica o privata accessibile a un numero indeterminato di persone fisiche, indipendentemente dal fatto che possano applicarsi determinate condizioni di accesso e indipendentemente dalle potenziali restrizioni di capacità.

Spazio di sperimentazione normativa per l'IA
Un quadro controllato istituito da un'autorità competente che offre ai fornitori o potenziali fornitori di sistemi di IA la possibilità di sviluppare, addestrare, convalidare e provare, se del caso in condizioni reali, un sistema di IA innovativo, conformemente a un piano dello spazio di sperimentazione per un periodo di tempo limitato sotto supervisione regolamentare.

Specifiche comuni
Un insieme di specifiche tecniche quali definite all'articolo 2, punto 4), del Regolamento (UE) n. 1025/2012, che forniscono i mezzi per soddisfare determinati requisiti stabiliti a norma del Regolamento (UE) 2024/1689.

Titolare dei dati
Una persona fisica o giuridica o qualsiasi altra entità che ha il diritto, conformemente al diritto dell'Unione europea o nazionale applicabile, e la capacità di gestire e mettere a disposizione i dati ottenuti nell'ambito della propria attività.

Ufficio per l'IA
La funzione della Commissione europea volta a contribuire all'attuazione, al monitoraggio e alla supervisione dei sistemi di IA e dei modelli di IA per finalità generali, e della governance dell'IA prevista dalla decisione della Commissione del 24 gennaio 2024. I riferimenti all'ufficio per l'IA contenuti nel Regolamento (UE) 2024/1689 si intendono fatti alla Commissione.

Uso improprio ragionevolmente prevedibile
L'uso di un sistema di IA in un modo non conforme alla sua finalità prevista, ma che può derivare da un comportamento umano o da un'interazione con altri sistemi, ivi compresi altri sistemi di IA, ragionevolmente prevedibile.

Valutazione della conformità
La procedura atta a dimostrare se i requisiti di cui al capo III, sezione 2 del Regolamento (UE) 2024/1689, relativi a un sistema di IA ad alto rischio, sono stati soddisfatti.

Verifica biometrica
La verifica automatizzata e uno a uno, inclusa l'autenticazione, dell'identità di persone fisiche mediante il confronto dei loro dati biometrici con i dati biometrici forniti in precedenza.

RIFERIMENTI BIBLIOGRAFICI

Agenzia per la cybersicurezza nazionale - ACN. 2024*a*. *Linee Guida per il rafforzamento della protezione delle banche dati rispetto al rischio di utilizzo improprio*. Roma, Italia: ACN. https://www.acn.gov.it/portale/w/online-le-linee-guida-per-il-rafforzamento-della-protezione-delle-banche-dati-rispetto-al-rischio-di-utilizzo-improprio-.

Agenzia per la cybersicurezza nazionale - ACN. 2024*b*. *La tassonomia cyber dell'ACN. Definizione della tassonomia cyber dell'Agenzia per la cybersicurezza nazionale*. Roma, Italia: ACN. https://www.csirt.gov.it/contenuti/la-tassonomia-cyber-dellacn.

Agenzia per la cybersicurezza nazionale - ACN. 2024*c*. *Linee guida per il rafforzamento della resilienza e referente per la cybersicurezza*. Roma, Italia: ACN. https://www.acn.gov.it/portale/linee-guida-rafforzamento-resilienza.

Agenzia per l'Italia digitale - Agid. *Linee guida*. Pagina web sul sito Agid. Roma, Italia: Agid. https://www.agid.gov.it/it/linee-guida.

Bellitti, G., e M. Fedeli (*a cura di*). 2023. *Trasformazione digitale della Pubblica amministrazione. Metodi per l'interoperabilità per lo sviluppo di e-service*. Letture Statistiche - Metodi. Roma, Italia: Istat. https://www.istat.it/produzione-editoriale/trasformazione-digitale-della-pubblica-amministrazione-metodi-per-linteroperabilita-per-lo-sviluppo-di-e-service/.

Bellitti, G., e M. Fedeli (*a cura di*). 2022. *Regole e strategie nel trattamento digitale e nella produzione dei dati quantitativi e qualitativi*. Letture Statistiche - Metodi. Roma, Italia: Istat. https://www.istat.it/produzione-editoriale/regole-e-strategie-nel-trattamento-digitale-e-nella-produzione-dei-dati-quantitativi-e-qualitativi/.

Bellitti, G., e C. Colasanti (*a cura di*). 2021. *Manuale sui principali adempimenti in materia di trattamento di dati personali: il caso dell'Istat*. Letture Statistiche - Metodi. Roma, Italia: Istat. https://www.istat.it/produzione-editoriale/manuale-sui-principali-adempimenti-in-materia-di-trattamento-di-dati-personali-il-caso-dellistat/.

Benanti, P., e S. Maffettone. 2024. *Noi e la macchina. Un'etica per l'era digitale*. Roma, Italia: Luiss University press.

Commissione europea (Gruppo indipendente di esperti ad alto livello sull'intelligenza artificiale). 2019. *Orientamenti etici per un'IA affidabile*. Lussemburgo: Ufficio delle pubblicazioni dell'Unione europea. https://op.europa.eu/it/publication-detail/-/publication/d3988569-0434-11ea-8c1f-01aa75ed71a1.

Davide, F. 2024. "Semic 2024: verso un'Europa digitale e interoperabile". *Agenda Digitale*. https://www.agendadigitale.eu/industry-4-0/semic-2024-verso-uneuropa-digitale-e-interoperabile/.

Decreto legislativo 4 settembre 2024, N. 138. "Recepimento della Direttiva (UE) 2022/2555, relativa a misure per un livello comune elevato di cibersicurezza nell'Unione, recante modifica del regolamento (UE) n. 910/2014 e della Direttiva (UE) 2018/1972 e che abroga la Direttiva (UE) 2016/1148". *Gazzetta ufficiale Serie generale N. 230 del 1° ottobre 2024*. https://www.gazzettaufficiale.it/eli/id/2024/10/01/24G00155/SG.

Decreto legislativo 30 giugno 2023, N. 196. "Codice in materia di protezione dei dati personali". *Gazzetta ufficiale Serie generale N. 174 del 29 luglio 2003, Supplemento ordinario N. 123*. https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2003-07-29&atto.codiceRedazionale=003G0218.

Decreto legislativo 31 marzo 2023, N. 36. "Codice dei contratti pubblici in attuazione dell'articolo 1 della legge 21 giugno 2022, n. 78, recante delega al Governo in materia di contratti pubblici". *Gazzetta ufficiale Serie generale N. 77 del 31 marzo 2023, Supplemento ordinario N. 12*. https://www.gazzettaufficiale.it/dettaglio/codici/contrattiPubblici.

Decreto legislativo 7 marzo 2005, N. 82. "Codice dell'amministrazione digitale". *Gazzetta ufficiale Serie generale N. 112 del 16 maggio 2005, Supplemento ordinario N. 93*. https://www.gazzettaufficiale.it/atto/serie_generale/caricaDettaglioAtto/originario?atto.dataPubblicazioneGazzetta=2005-05-16&atto.codiceRedazionale=005G0104.

Dipartimento per la trasformazione digitale, e Agenzia per l'Italia digitale - Agid. 2024. *Strategia Italiana per l'Intelligenza artificiale 2024-2026*. Roma, Italia: Dipartimento per la trasformazione digitale; Agid. https://innovazione.gov.it/notizie/articoli/strategia-italiana-per-l-intelligenza-artificiale-2024-2026/.

Direttiva (UE) 2022/2557 del Parlamento europeo e del Consiglio del 14 dicembre 2022 del Parlamento europeo e del Consiglio relativa alla resilienza dei soggetti critici e che abroga la Direttiva 2008/114/CE del Consiglio (Testo rilevante ai fini del SEE). *Gazzetta ufficiale dell'Unione europea, L 333, 27 dicembre 2022*: 164-198. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32022L2557.

Direttiva (UE) 2022/2555 del Parlamento europeo e del Consiglio del 14 dicembre 2022 relativa a misure per un livello comune elevato di cybersicurezza nell'Unione, recante modifica del regolamento (UE) n. 910/2014 e della Direttiva (UE) 2018/1972 e che abroga la Direttiva (UE) 2016/1148 (Direttiva NIS 2). *Gazzetta ufficiale dell'Unione europea, L 333, 27 dicembre 2022: 80-152*. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32022L2555.

Direttiva (UE) 2016/1148 del Parlamento europeo e del Consiglio, del 6 luglio 2016, recante misure per un livello comune elevato di sicurezza delle reti e dei sistemi informativi nell'Unione. *Gazzetta ufficiale dell'Unione europea, L 194, 19 luglio 2016*: 1-30. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32016L1148.

Direttiva 2006/54/CE del Parlamento europeo e del Consiglio del 5 luglio 2006, riguardante l'attuazione del principio delle pari opportunità e della parità di trattamento fra uomini e donne in materia di occupazione e impiego (rifusione). *Gazzetta ufficiale dell'Unione europea*, *L 204*, *26 luglio 2006*: 23-36. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX%3A32006L0054.

European Data Protection Board - EDPB. 2024. *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models*. Bruxelles, Belgium: EDPB. https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_it.

European Data Protection Supervisor - EDPS. 2024. *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*. Bruxelles, Belgium: EDPS. https://www.edps.europa.eu/data-protection/our-work/publications/guidelines/2024-06-03-first-edps-orientations-euis-using-generative-ai_en.

European Union Agency for Cybersecurity - Enisa. 2022. *European Cybersecurity Skills Framework (ECSF). User Manual*. Athens, Greece: Enisa. https://www.enisa.europa.eu/publications/european-cybersecurity-skills-framework-ecsf.

Eurostat. 2017. *Codice delle statistiche europee. Per le autorità statistiche nazionali ed Eurostat (autorità statistica dell'UE)*. Lussemburgo: Ufficio delle pubblicazioni dell'Unione europea. https://www.sistan.it/fileadmin/Repository/Home/EUROPA/CoP_IT.pdf.

Gabaglio, A. 1880. *Storia e teoria generale della statistica*. 1880. Milano, Italia: Hoepli.

Garante per la protezione dei dati personali - GPDP. 2021. "Nuove regole deontologiche per trattamenti a fini statistici o di ricerca scientifica effettuati nell'ambito del Sistema Statistico nazionale". Registro dei provvedimenti n. 133 del 15 aprile 2021. *Gazzetta ufficiale Serie generale N. 104 del 3 maggio 2021*: 6-9. https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9582086.

Geretto, P. 2008. *Mezzi automatici di elaborazione dati e informatica 1926-2005*. Roma, Italia: Istat. https://www.istat.it/it/files/2010/09/informatica.pdf.

Giannini, M.S. 1979. *Rapporto sui principali problemi della amministrazione dello Stato*. Roma, Italia: Senato della Repubblica italiana. https://www.tecnichenormative.it/RapportoGiannini.pdf.

Istituto Nazionale di Statistica - Istat. *Linee guida per la qualità*. Pagina web sul sito dell'Istat. Roma, Italia: Istat. https://www.istat.it/classificazioni-e-strumenti/strumenti-per-la-qualita/linee-guida/.

Istituto Nazionale di Statistica - Istat. 2024. *Istat, Fastweb e UNINT di Roma: accordo di collaborazione sull'Intelligenza Artificiale*. Notizia, 12 luglio 2024. https://www.istat.it/notizia/istat-fastweb-universita-studi-internazionali-di-roma-accordo-collaborazione-intelligenza-artificiale/.

Legge 28 giugno 2024, N. 90. "Disposizioni in materia di rafforzamento della cybersicurezza nazionale e di reati informatici". *Gazzetta ufficiale Serie generale N. 153 del 2 luglio 2024*. https://www.gazzettaufficiale.it/eli/id/2024/07/02/24G00108/SG.

Organisation for Economic Co-operation and Development - OECD. 2024. *Recommendation of the Council on OECD Legal Instruments Artificial Intelligence*. Paris, France: OECD. https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449.

Regolamento (UE) 2024/3018 del Parlamento europeo e del Consiglio del 27 novembre 2024, che modifica il regolamento (CE) n. 223/2009 relativo alle statistiche europee. *Gazzetta ufficiale dell'Unione europea*, *L*, *16 dicembre 2024*. https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX:32024R3018.

Regolamento (UE) 2024/1689 del Parlamento europeo e del Consiglio del 13 giugno 2024, che stabilisce regole armonizzate sull'intelligenza artificiale e modifica i regolamenti (CE) n, 300/2008, (UE) n, 167/2013, (UE) n, 168/2013, (UE) 2018/858, (UE) 2018/1139 e (UE) 2019/2144 e le direttive 2014/90/UE, (UE) 2016/797 e (UE) 2020/1828 (regolamento sull'intelligenza artificiale). *Gazzetta ufficiale dell'Unione europea, L, 12 luglio 2024*. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=OJ%3AL_202401689.

Regolamento (UE) 2024/903 del Parlamento europeo e del Consiglio del 13 marzo 2024, che stabilisce misure per un livello elevato di interoperabilità del settore pubblico nell'Unione (regolamento su un'Europa interoperabile). *Gazzetta ufficiale dell'Unione europea, L, 22 marzo 2024*. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:32024R0903.

Regolamento (UE) 2022/2065 del Parlamento europeo e del Consiglio del 19 ottobre 2022 relativo a un mercato unico dei servizi digitali e che modifica la Direttiva 2000/31/CE (regolamento sui servizi digitali). *Gazzetta ufficiale dell'Unione europea, L 277, 27 ottobre 2022*: 1-102. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:32022R2065.

Regolamento (UE) 2022/1925 del Parlamento europeo e del Consiglio del 14 settembre 2022 relativo a mercati equi e contendibili nel settore digitale e che modifica le direttive (UE) 2019/1937 e (UE) 2020/1828 (regolamento sui mercati digitali). *Gazzetta ufficiale dell'Unione europea, L 265, 12 ottobre 2022*: 1-66. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex:32022R1925.

Regolamento (UE) 2022/868 del Parlamento europeo e del Consiglio del 30 maggio 2022 relativo alla governance europea dei dati e che modifica il regolamento (UE) 2018/1724 (Regolamento sulla governance dei dati). *Gazzetta dell'Unione europea, L 152, 3 giugno 2022*: 1-44. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=CELEX:32022R0868.

Regolamento (UE) 2018/1725 del Parlamento europeo e del Consiglio del 23 ottobre 2018, sulla tutela delle persone fisiche in relazione al trattamento dei dati personali da parte delle istituzioni, degli organi e degli organismi dell'Unione e sulla libera circolazione di tali dati, e che abroga il regolamento (CE) n. 45/2001 e la decisione n. 1247/2002/CE. *Gazzetta ufficiale dell'Unione europea, L 295, 21 novembre 2018*: 39-98. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex%3A32018R1725.

Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la Direttiva 95/46/CE (regolamento generale sulla protezione dei dati). *Gazzetta ufficiale dell'Unione europea, L 119, 4 maggio 2016*: 1-88. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex%3A32016R0679.

Regolamento (CE) n. 223/2009 del Parlamento europeo e del Consiglio dell' 11 marzo 2009, relativo alle statistiche europee e che abroga il regolamento (CE, Euratom) n. 1101/2008 del Parlamento europeo e del Consiglio, relativo alla trasmissione all'Istituto statistico delle Comunità europee di dati statistici protetti dal segreto, il regolamento (CE) n. 322/97 del Consiglio, relativo alle statistiche comunitarie, e la decisione 89/382/CEE, Euratom del Consiglio, che istituisce un comitato del programma statistico delle Comunità europee. *Gazzetta ufficiale dell'Unione europea*, *L 87, 31 marzo 2009*: 164-173. https://eur-lex.europa.eu/legal-content/IT/TXT/?uri=celex%3A32009R0223.

Risoluzione del Parlamento europeo del 21 gennaio 2021 recante raccomandazioni alla Commissione sul diritto alla disconnessione. *Gazzetta ufficiale dell'Unione europea, C 456, 10 novembre 2021*: 161-176. https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=CELEX%3A52021IP0021.

Tangi, L., M. Combetto, J. Martin Bosch, and A.P. Rodriguez Müller. 2023. *Artificial Intelligence for Interoperability in the European Public Sector. An exploratory study*. JRC (Joint Research Centre) technical report. Luxembourg: Publications Office of the European Union. https://publications.jrc.ec.europa.eu/repository/handle/JRC134713.

United Nations Economic Commission for Europe - UNECE, High Level Group for the Modernisation of Official Statistics (HLG-MOS). 2023. *Large Language Models for Official Statistics. HLG-MOS White Paper*. Geneva, Switzerland: UNECE. https://unece.org/sites/default/files/2023-12/HLGMOS%20LLM%20Paper_Preprint_1.pdf.

United Nations Economic Commission for Europe - UNECE. 2020. *Implementation of the new role of national statistical offices at the time of expanded possibilities*. Geneva, Switzerland: UNECE. https://documents.un.org/doc/undoc/gen/g20/088/28/pdf/g2008828.pdf.

US Office of Management and Budget - OMB. 2024. *Memorandum for the heads of executive departments and agencies. Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence*. Washington, DC, U.S.: OMB. https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf.