PUBLIC OR PRIVATE AI: UN (NUOVO) DILEMMA PER LA TUTELA DEI DIRITTI FONDAMENTALI ALLA PROVA DELLA ATTUAZIONE DELLA STRATEGIA ITALIANA PER L'INTELLIGENZA ARTIFICIALE

Abstract

La distinzione tra *Public* e *Private AI* inerente all'approccio da adottare per i modelli fondazionali si sta diffondendo nelle scienze aziendalistiche, ma è inedita nel campo del diritto pubblico e costituzionale. Eppure, in riferimento al ruolo della PA c'è comunque un'organizzazione che fornisce servizi all'utenza, anche se entro un orizzonte di interesse pubblico e di rispetto dei diritti fondamentali. Il saggio, partendo dalla distinzione tra i due tipi di approccio, mostra i vantaggi e gli svantaggi di ognuno e come la scelta da parte della PA non sia affatto scontata, né debba essere necessariamente unitaria. Alla luce della dicotomia tra *Public* e *Private AI* vengono esaminate le scelte strategiche contenute nella pianificazione nazionale sull'intelligenza artificiale illustrando l'importanza di un approccio ibrido e diversificato per un'ottimale tutela dei diritti fondamentali, a partire dalle esigenze di tutelare privacy e sicurezza nella gestione del dato personale.

The distinction between Public and Private AI inherent to the approach to be adopted for foundational models is spreading in business sciences, but is unprecedented in the field of public and constitutional law. Yet, in reference to the role of the PA there is however an organization that provides services to users, even if within a horizon of public interest and respect for fundamental rights. The essay, starting from the distinction between the two types of approach, shows the advantages and disadvantages of each and shows how the choice by the public administration is not at all obvious, nor must it necessarily be unitary. In light of the dichotomy, the strategic choices contained in the national planning on Artificial Intelligence are examined, illustrating the importance of a hybrid and diversified approach for optimal protection of fundamental rights, starting from privacy and security in the management of personal data.

Keywords: Personal Data, Foundation Models, Public Administration, Fundamental Rights, National Strategy.

1. La digitalizzazione della PA e la nuova visione europea sui dati

Il nostro tempo si caratterizza per un'ininterrotta rivoluzione tecnologica, con il succedersi di tecnologie *disruptives*. Sono trascorsi solo pochi lustri da quando negli anni '90 arrivavano nelle case delle famiglie i primi veri Home Computer e qualche anno dopo i Browser e l'uso di largo consumo del World Wide Web. Da allora numerose tecnologie rivoluzionarie si sono succedute,

sovrapposte e hanno interagito a cambiare le nostre vite private, la nostra vita sociale e il nostro rapporto con il potere pubblico. Da ultimo l'intelligenza artificiale generativa, esplosa con ChatGPT della società OpenAI (lanciato nel novembre 2022), sta trasformando già il nostro modo di vivere e lavorare, nonostante alcune difficoltà di business, e sorprende per la capacità di penetrazione nella vita quotidiana e lavorativa. Una piccola società (originariamente) *no profit* è riuscita a generare notevole interesse e valore grazie ad un uso avanzato del *deep learning*, consistente nell'addestrare un modello su grandi quantità di dati per riconoscere *pattern* e fornire risposte utili, in modo che le informazioni apprese generino risposte plausibili e pertinenti per gli utenti. Ciò avviene, nello specifico, con un'architettura costituita da una rete neurale chiamata "transformer" (introdotta appena nel 2017 da Vaswani et al. nel famoso paper *Attention is All You Need*), con cui la rete è stata addestrata utilizzando una grande quantità di testo al fine di apprendere modelli di linguaggio e rispondere con elaborazione in linguaggio naturale a domande o partecipare a conversazioni in modo coerente. Una tale intelligenza artificiale, detta appunto generativa o, più in generale, fondazionale (v. subito *infra*), è il simbolo del nuovo modo di essere nell'intelligenza artificiale. L'intelligenza artificiale storica era basata su operatori di tipo logico-simbolico, sintattici; una dimensione tecnologica che continua ad esistere per aspetti importanti del funzionamento del digitale, ma che è stata abbandonata quale programma generale. La seconda intelligenza artificiale, a cui appartengono i LLM (Large Language Model), è una tecnologia che sfrutta un enorme potere computazionale e che sarebbe sbagliato ridurre integralmente alla metafora della forza bruta, se consideriamo, intelligenza o meno che sia, la finezza con cui è in grado di sostenere la conversazione, nonostante le ben note "allucinazioni". In particolare i modelli fondazionali sono quei sistemi di grandi dimensioni (*large*) in grado di svolgere un'ampia gamma di compiti specifici, come la generazione di video, testi, immagini, l'elaborazione o la generazione di codice informatico nonché la citata conversazione in linguaggio naturale, e la normativa europea (regolamento UE 2024/1689) definisce tali modelli «ad alto impatto» quando siano addestrati con una grande quantità di dati e con complessità, capacità e prestazioni elevate.

Si tratta di una significativa novità che si pone in scia con la creazione degli algoritmi di personalizzazione che hanno fatto la fortuna di Google. Fino ad oggi, nonostante l'indubbia utilità o comunque successo dei social network, la potenza computazionale è stata sfruttata fondamentalmente per trasformare aziende estremamente innovative in grandi agenzie pubblicitarie, come avvenuto con la stessa Google. Il caso Cambridge Analytica ha dimostrato quanto la ricerca

comportamentale fosse andata avanti, grazie una disponibilità di una sconfinata mole di dati, capaci di creare miliardi di parametri. Per questa ragione i conversatori basati su modelli di LLM, pur con i loro limiti, sembrano per molti aspetti una novità positiva rispetto alle inquinanti, e per qualche aspetto inquietanti, derive dei social network, anche se la tendenza a fare di questi modelli i nuovi oracoli può produrre cocenti delusioni, se non preoccupazioni, alla luce del modo in cui l'informazione è elaborata e si trasforma, più che problematicamente, in conoscenza. Uno dei punti deboli o almeno problematici di queste intelligenze generative è che sono fondate su dataset, raccolte di dati, "a strascico", la cui origine non sempre viene chiarita adeguatamente (sul punto torneremo; ufficialmente ChatGPT ad es. era costruita su dati pubblici, addestrata su 45 TB di dati dal set *Common Craw*, e poi successivamente addestrata su dati ulteriori). Tuttavia, se il discorso sull'apporto in termini di autorevolezza di queste innovazioni può restare aperto (anche a sviluppi degli algoritmi e dei dataset), un certo successo funzionale di questi modelli è innegabile.

La pubblica amministrazione ha avuto un'enorme difficoltà a tenere il passo con queste innovazioni. Quando i dipendenti pubblici hanno preso iniziative autonome si sono prodotti numerosi incidenti, non diversamente da quanto avvenuto nel settore privato (si pensi al clamoroso caso Samsung). La tecnologia è una risorsa estremamente preziosa per la pubblica amministrazione e per rinforzare l'effettività dei principi della Costituzione, ma richiede competenze e risorse.

La pubblica amministrazione è chiaramente sotto stress, e fatica a reggere il passo delle innovazioni e delle opportunità che queste innovazioni aprono. Negli anni scorsi le pubbliche amministrazioni, centrali, periferiche e territoriali, hanno dovuto adeguare le loro strutture al passaggio da cartaceo a digitale, poi si sono dovute adattare alla "piattaformizzazione" della rete, quindi alla nascita dei cloud giustificata (anche) dalla mole di dati circolanti, e in generale alla digitalizzazione della nostra esistenza e delle nostre attività. Come sappiamo, neanche un istituto relativamente semplice come l'autocertificazione, la quale confida sul previo possesso da parte delle pubbliche amministrazioni di dati relativi all'utente e il loro facile recupero e riutilizzo (ma si potrebbe addure ad esempio anche il processo telematico e la porta certificata), dimostra quanto sia difficile stare sulla cresta dell'innovazione.

In ogni caso, gradualmente il tema del dato personale è transitato dalla sola questione della sua garanzia a quello dell'opportunità in ragione del valore economico e della capacità in una società dell'informazione di accelerare i processi di innovazione. Di recente sono state adottate misure per favorire la circolazione degli stessi dati personali, e si è chiarito meglio il concetto di "dato pubblico"

quale dato posseduto e generato dalla pubblica amministrazione nell'esercizio della sua funzione istituzionale. Con provvedimenti appositi, l'Unione europea ha inteso favorire il riuso dei dati pubblici (in particolare con il regolamento UE 2022/868), quale utilizzo di dati già esistenti per scopi diversi da quelli per cui sono stati originariamente raccolti, alla luce del loro rilievo economico e preziose risorse per il mercato e per gli stessi operatori istituzionali. Nonostante le previsioni volte a conciliare tali prospettive di un mercato aperto dei dati con i diritti fondamentali dell'individuo, a partire dalla privacy, questo approccio è entrato in tensione con il tradizionale approccio garantistico italiano in tema di dato personale, essendo del resto coerente con lo sviluppo di una strategia istituzionale sulla conoscenza e sull'economia digitale intrapresa da tempo. L'Unione europea ha adottato anche una specifica strategia europea di dati (COM(2020) 66 final). Il nostro ordinamento si è dovuto anche aprire al concetto di "open data", uno dei cardini della dottrina dell'*open government*, secondo il quale l'apertura dei dati pubblici può contribuire all'aumento della trasparenza e della partecipazione dei cittadini all'esercizio del potere.

Il riutilizzo dei dati, compresi i dati personali, richiede però la presenza di ambienti di elaborazione sicuri e di tecniche di anonimizzazione, come la cosiddetta *differential privacy* e la creazione di dati sintetici, cioè generati artificialmente. Questi dati sono per diversi aspetti preferibili a quelli non sintetici, ma hanno d'altra parte numerose limitazioni.

Negli anni scorsi sia i dati dei privati che quelli detenuti dalla pubblica amministrazione che ha fatto ricorso alle tecnologie delle *big tech* sono stati investiti da imbarazzanti vicende maturate nell'ambito dei tormentati sviluppi degli accordi tra Unione europea e Stati Uniti in merito al trattamento dei dati. In particolare, si ricordi nel luglio del 2022 l'intervento del Garante per la protezione dei dati personali con cui si è deciso che tutti i siti web pubblici e privati che utilizzavano il servizio Google Analytics o analoghi senza le garanzie previste dal regolamento UE 679/2016 violavano la suddetta normativa in quanto idonei a trasferire negli Stati Uniti i dati degli utenti. Questa posizione giungeva a seguito della sentenza C-311/18 del 16 luglio 2020, c.d. Schrems II, con la quale la Corte di giustizia europea invalidava il quadro normativo noto come Privacy Shield, e a pochi mesi di distanza dalla presa di posizione dell'omologa autorità di garanzia austriaca che affermava che il servizio Google Analytics fornito da Google L.L.C. VW non era conforme alle normative europee. In effetti in quel periodo è stato attestato che senza alcuna ombra di dubbio Google raccoglieva mediante cookies informazioni relative alle modalità di interazione degli utenti con siti, pagine e servizi proposti, che i dati raccolti erano di vario genere e che solo apparentemente erano

anonimi, perché in realtà le opzioni tecnologiche entro cui circolavano li rendevano pseudoanonimi. Pertanto, Google era potenzialmente in grado di creare un profilo completo dell'utente associando l'indirizzo IP ad altre informazioni aggiuntive già in suo possesso: un grande problema di violazione della normativa e in particolare una macroscopica violazione nel principio di accountability, alla luce della mancata adozione da parte del titolare del trattamento delle misure tecniche organizzative necessarie per assicurare il corretto trattamento dei dati personali.

Bisogna ricordare anche che AgID, nelle sue *Linee guida di design per i siti Internet e i servizi digitali della PA* adottate con determina n. 224/2022, aveva suggerito di sostituire Google Analytics con un analogo servizio italiano, Web Analytics Italia, una versione open source in grado di produrre in tempo reale statistiche e report dettagliati sull'utilizzo dei dati della pubblica amministrazione. Tuttavia migliaia di amministrazioni pubbliche per carenza di sensibilità, per incapacità o altro non l'avevano e tuttora non l'hanno fatto.

Non era la prima volta. Dopo l'adozione del regolamento generale del 2016 e la sua entrata in vigore, a causa della contemporanea Brexit si realizzò una dislocazione di dati degli utenti dell'Unione europea dall'Irlanda agli Stati Uniti e si accertò che i server di Google erano localizzati al di fuori dell'Unione europea. Apparve anche chiaro che Google non applicava la normativa europea e che per di più scansionava sistematicamente i documenti salvati in drive e allegati alle mail, archiviando i contenuti di testo. Tutto avveniva all'insaputa dei clienti. Questa accumulazione di informazioni, naturalmente, era funzionale al funzionamento degli algoritmi di personalizzazione. Un tema che negli anni successivi diventava ancora più delicato.

Il vero problema emerso al tempo, sottolineato da un "terroristico" messaggio di posta elettronica certificata inoltrata da un gruppo di attivisti (collegati, probabilmente, all'attivismo austriaco) è che la pubblica amministrazione italiana complessivamente intesa, e ancor più specificamente gli enti locali, avevano una assai scarsa consapevolezza e, prima ancora, formazione in relazione alle tematiche inerenti al trattamento dei dati e ai diritti degli interessati, a prescindere dall'operare all'interno di essi di un DPO e di un registro dei trattamenti.

L'adozione del principio "cloud first", secondo cui le pubbliche amministrazioni devono in via prioritaria adottare il paradigma cloud prima di qualunque altra opzione tecnologica per nuovi progetti e servizi, si traduceva in una storica débâcle della privacy dei cittadini italiani.

Non si tratta, in questa sede, di criminalizzare alcuna azienda, la quale fornisce servizi a suo modo preziosi alle amministrazioni pubbliche italiane. Si tratta invece di sottolineare un'arretratezza

culturale e di innovazione che ha avuto e che ha pesanti costi in termini di diritti fondamentali, a partire dalla trasparenza, di cui la privacy è solo uno dei possibili aspetti.

2. *Public or Private AI*: una distinzione non indifferente per la dimensione della tutela della persona

In questo quadro cade la nuovissima questione relativa all'alternativa tra IA pubblica e privata, che riguarda ogni organizzazione pubblica o privata ma che allo stato è discussa soprattutto in campo privato. Si tratta invece di una scelta che anche le nostre amministrazioni pubbliche debbono compiere.

Come abbiamo già ricordato, negli anni scorsi sono stati affidati ai servizi delle big tech (in particolare Microsoft e Google) sistemi di posta, didattica, cloud e altri servizi delle nostre pubbliche amministrazioni. Ciò ha consentito di utilizzare tecnologie all'avanguardia, ma al contempo ha determinato una certa perdita di controllo sui processi e, come abbiamo visto, sui dati.

La distinzione *Public/Private AI* sul piano giuridico va chiarita bene. Si tratta di un dilemma relativo al modo in cui infrastrutture, software e programmi di IA vengono sviluppati, distribuiti e gestiti. La distinzione non riguarda il soggetto che utilizzata il sistema di IA, ma il modo in cui questa è costruita e utilizzata, a prescindere dalla natura dell'organizzazione.

Per IA pubblica si fa riferimento ad algoritmi di intelligenza artificiale che sono disponibili pubblicamente, cioè a tutti. Pertanto l'addestramento dei dati si basa su un'ampia serie di dati che magari in una prima fase vengono tratti da archivi pubblici ma che successivamente sono spesso provenienti dagli utenti o clienti del servizio. È noto, infatti, che sovente i fornitori di intelligenza artificiale addestrano i loro modelli utilizzando i dati degli stessi utilizzatori al fine di migliorare i servizi. Ciò trasforma la natura del dato personale, che in qualche misura diventa pubblico perché entra a far parte di dataset con i quali il modello viene addestrato per nuovi *output*.

Per intelligenza artificiale privata ci si riferisce, per il nostro piano del discorso, a quelle soluzioni organizzative volte ad addestrare gli algoritmi attraverso dati specifici di un utente o di una organizzazione. Ciò vuol dire che il modello viene utilizzato soltanto dall'organizzazione che lo implementa. Spesso vi è comunque un fornitore della piattaforma, ma in questo caso il fornitore non utilizza i dati generati dal sistema operante presso l'organizzazione per addestrare i propri modelli. Se sul piano commerciale ciò produce il vantaggio di non contribuire a creare una intelligenza che si può ritorcere contro l'utilizzatore (come avviene, ad esempio, nella fruizione di servizi relativi

a tariffe aeree) o a beneficio dei concorrenti, per le pubbliche amministrazioni il discorso è diverso, anche in ragione degli specifici parametri giuridici posti in Costituzione.

Per restare ancora su un livello di definizione, va specificato che l'intelligenza artificiale privata può essere creata o internamente (disponendo del know how giusto) oppure coinvolgendo soggetti esterni, come un fornitore. In questo caso, ad esempio, un soggetto privato crea una piattaforma con modelli di machine learning che rendono possibili varie funzionalità e che utilizzano i dati privati del cliente, e dei suoi utenti, senza tuttavia mai utilizzare questi dati per addestrare l'algoritmo. Sembra che un modello di questo genere sia idoneo a garantire una forte riservatezza dei dati dei clienti senza necessità di supportare i pesanti costi di personale per creare dall'interno l'infrastruttura, che viene offerta alla organizzazione.

Appare evidente, pertanto, che software come ChatGPT sono esempi di (servizi di) intelligenza artificiale pubblica in quanto addestrati su dati disponibili al pubblico e reperiti (presumibilmente) sulla rete e su archivi (come biblioteche) disponibili al pubblico, e comunque addestrati con dati degli utenti che servono a migliorare ulteriormente il sistema a beneficio anche di altri utenti. Un sistema di questo tipo pone alcune questioni di pubblicizzazione di un dato personale, e non ha certo le garanzie interne proprie di un sistema con una cornice legale a tutela della persona. Per quanto la casa produttrice dichiari di non recuperare o memorizzare informazioni personali sugli utenti, restano dubbi in tale proposito, in quanto non v'è sufficiente chiarezza e trasparenza a riguardo. D'altra parte, un sistema siffatto è reso disponibile al pubblico attraverso diverse piattaforme e API (Interfacce di Programmazione delle Applicazioni), e ciò consente a sviluppatori, aziende e individui di integrare le sue capacità nelle loro applicazioni, servizi o progetti, come è tipico della IA pubblica. Pertanto una IA pubblica lavora su dati pubblici, è integrabile e adattabile da applicazioni da parte del pubblico, ed è accessibile a chiunque tramite i canali offerti dalla società che la gestisce e ne è proprietaria (nel caso, OpenAI). Se ChatGPT è destinata a ripercorrere le orme del successo di Google Suite, Microsoft Teams e di consimili sistemi, non si può escludere che si ripropongano *mutatis mutandis* le questioni di tutela dei diritti fondamentali dei cittadini di cui abbiamo detto sopra.

È affermazione corrente che l'intelligenza artificiale pubblica aumenta l'intelligenza collettiva e dunque potrebbe dirsi pubblica da un ulteriore punto di vista, cioè nell'accrescimento del valore di informazione e conoscenza che sono, appunto, anche beni pubblici. Per questa ragione, alcuni autori ritengono questi sistemi più adatti ad un utilizzo da parte delle istituzioni pubbliche, in

quanto in questi casi occorre favorire la partecipazione e l'arricchimento a beneficio della collettività. Tuttavia ci sembra che la considerazione sia eccessivamente semplificatrice, misurata alla luce della distinzione tra intelligenza artificiale pubblica e privata che abbiamo tratteggiato e che è corrente nell'ambito della sorgente questione. Anche una organizzazione pubblica, e non solo un'azienda, ha questioni di tutela, e uso, del dato personale del cittadino (o utente) nell'erogazione dei suoi servizi ed anzi mentre per il privato si pone la questione dell'efficacia orizzontale della tutela dei diritti fondamentali, in rapporto con il potere pubblico il privato è tutelato immediatamente alla luce di un quadro di diritti fondamentali posti in Costituzione.

Appare piuttosto evidente che un approccio *Public AI* contempli svantaggi connessi alla necessità di avere un quadro giuridico certo di tutela dei diritti fondamentali e maggiormente improntato al principio della trasparenza e dell'accountability, che costituiscono anch'esse importanti dotazioni pubbliche e un veicolo per accrescere la fiducia nel sistema, e tra cittadini e pubblica amministrazione. Da questo punto di vista sistemico, per quanto applicazioni come ChatGPT affermino esplicitamente di rispettare le normative sulla privacy in tutti i loro aspetti (nessuna raccolta attiva di dati personali, impossibilità di memorizzazione di informazioni personali identificabili, non conservazione dei dati a lungo termine, diritto all'oblio, trasparenza e accesso ai dati, minimizzazione, anonimizzazione e pseudonimizzazione, sicurezza e protezione dei dati), non sembrano offrire tutte le garanzie che ciò sia realmente vero. Ad esempio, se le interazioni sono generalmente gestite in tempo reale, senza conservazione di dati a lungo termine, tuttavia la società proprietaria (nel caso, OpenAI) può analizzare le conversazioni in forma anonima e aggregata per migliorare i modelli, sia pure – si asserisce – nel rispetto delle normative sulla privacy. Oppure la raccolta di dati personali e sensibili non v'è, salvo che non sia l'interlocutore stesso a fornirli durante le conversazioni, nel qual caso tali dati vengono utilizzati e viene dichiarato soltanto che le informazioni personali non vengono memorizzate dopo la chiusura della sessione. Ma queste sono, ancora, dichiarazioni ufficiali relative ad una società cresciuta troppo in fretta e, come attestano le cronache, che attraversa da tempo una crisi di management e una emorragia di dipendenti per ragioni ancora non del tutto chiare ma che hanno certamente a che fare con la trasformazione del suo statuto iniziale e la creazione di un modello di business che ruota completamente sui dati. Troppo spesso è stato appurato che le grandi società tecnologiche adottavano comportamenti molto lontani dalle dichiarazioni pubbliche e ufficiali.

Ad ogni modo, avviandoci ad una conclusione, nel modello *Public AI*, quanto all'accesso e alla distribuzione, lo sviluppo avviene per essere utilizzato da parte di un vasto pubblico, spesso su piattaforme cloud o server condivisi. Gli utenti inviano i loro dati ai server gestiti da aziende che offrono il servizio (come OpenAI, Google, Microsoft, etc.). In termini di condivisione, i dati possono essere elaborati su server remoti, il che implica che una certa quantità di informazioni transita su reti pubbliche o viene archiviata su server controllati da terze parti. Il vantaggio evidente è la flessibilità e la scalabilità, nel senso che questo tipo di AI è pensata per offrire accesso a strumenti di intelligenza artificiale potenti e su larga scala a chiunque ne abbia bisogno, con ampie risorse computazionali e infrastrutture cloud.

I modelli di *Private AI*, quanto all'accesso, sono invece progettati per essere utilizzati in contesti specifici e chiusi, come all'interno di reti aziendali private, su dispositivi locali o in ambienti con requisiti di privacy particolarmente elevati. In relazione alla elaborazione dei dati, essa avviene localmente o in ambienti controllati, riducendo i rischi legati alla trasmissione di informazioni sensibili su internet. Quanto ai profili relativi al controllo, le aziende o gli utenti hanno un maggiore controllo su come vengono gestiti i dati, minimizzando le possibilità che informazioni personali o sensibili possano essere esposte o sfruttate da terze parti. I dati non escono dalla infrastruttura aziendale. Sono modelli sviluppati e implementati proprio con l'obiettivo di garantire la privacy e la sicurezza dei dati degli utenti, e che elaborano dati sensibili direttamente sui dispositivi degli utenti o all'interno di reti chiuse, senza inviare informazioni personali a server remoti o cloud pubblici.

Dal punto di vista della sicurezza e del controllo dei dati, la pubblica amministrazione nazionale avrebbe diverse ragioni per orientarsi verso modelli di *Private AI*, dalle migliori garanzie in termini di protezione dei dati dei propri cittadini, ad una migliore tutela degli interessi nazionali. Le interazioni e l'elaborazione dei dati avverrebbero su un dispositivo locale oppure su un server dedicato e controllato da un'organizzazione specifica. Se, come ragionevole, questa (o queste) organizzazione/-i fosse(-ro) nazionale/-i le ragioni di sicurezza nazionale sarebbero meglio garantite rispetto ad un'elaborazione che avviene su server esterni normalmente dislocati all'estero, visto che la massima parte dei server è negli Stati Uniti. Sul piano giuridico, questo avverrebbe con un contratto di appalto per la fornitura di servizi digitali che dovrebbe verificare il possesso dei requisiti e delle capacità, inclusi gli standard di sicurezza. Uno Stato preserverebbe maggiormente la propria "sovranità digitale", mantenendo un maggiore controllo sui dati sensibili dei propri cittadini. Un Paese che sviluppa tecnologie *Private AI* può mantenere il controllo sui dati sensibili dei cittadini,

riducendo la dipendenza da attori esteri o grandi multinazionali. Si ridurrebbero inoltre i rischi di violazioni della privacy, garantendo che i dati rimangano confinati a livello locale o protetti secondo normative nazionali. Sarebbe anche favorita la competitività tecnologica, in quanto un paese che investe in *Private AI* può creare un ecosistema industriale innovativo, posizionandosi come leader in tecnologie avanzate e sicure, attirando investimenti internazionali. Ciò favorirebbe, in particolare, lo sviluppo di soluzioni di AI private da parte di aziende e startup nazionali, creando posti di lavoro e aumentando il valore delle competenze tecnologiche locali.

Le vicende che hanno riguardato il trasferimento dei dati dall'Unione europea all'estero (quasi sempre Stati Uniti), di cui abbiamo detto, fornisce un'ulteriore ragione a favore di uno sviluppo di una *Private AI*. Alle preoccupazioni circa la protezione dei dati e all'utilità della promozione di un ecosistema nazionale innovativo si aggiunge la possibile alea derivante da possibili decisioni giudiziarie future che dovessero concludere per un contrasto con il diritto europeo anche da parte del terzo accordo di trasferimento dati con gli Stati Uniti.

Vi è una ragione ulteriore che sembra giocare maggiorente a favore di una *Private AI*, consistente nella maggiore flessibilità del sistema, che consente di lavorare con dataset costruiti allo scopo, quindi più puliti e trasparenti. Si sa che una delle maggiori criticità dell'intelligenza artificiale generativa è la pulizia del dato, che laddove sia carente rischia di compromettere il risultato. È noto che la *Public AI* si è sviluppata in modo disordinato, con un saccheggio generalizzato di dati dalla rete e da database (senza che si sappia neanche bene quali) e che pertanto vi sia anche un problema di aggiornamento (che OpenAI adduce, per qualche aspetto, quale elemento di forza o almeno di garanzia). Una *Private AI* potrebbe offrire prestazioni maggiormente calibrate derivanti da dataset costruiti su obiettivi specifici, e non derivanti da un sistema costruito come a fruizione pubblica e senza particolari obiettivi (se non, genericamente, conversazionali).

Una *Private AI* si sviluppa infatti proprio su dataset costruiti in ragione dello scopo e contigui (culturalmente, geograficamente, etc.), per così dire, all'utenza finale.

3. L'approccio della strategia italiana per l'intelligenza artificiale ai sistemi di IA "fondazionali"

La pubblica amministrazione, una espressione che racchiude servizi e missioni assai varie, potrebbe anche adottare un sistema ibrido, che a seconda delle necessità faccia ricorso ad un sistema di intelligenza artificiale di natura pubblica (tipicamente un sistema di risposta al cittadino di tipo

chatbot) o privata (ricorrendo per quest'ultimo, verosimilmente, alla forma dell'appalto). Non si tratta di una scelta necessariamente escludente né tantomeno senza ritorno, anche se lo Stato deve effettuare delle valutazioni in ordine alla razionalità e all'economicità dei sistemi prescelti. Proprio la preminente tutela dei diritti fondamentali dei cittadini, dalla privacy all'accountabillity dei processi pubblici, può condurre ad una scelta più articolata in ragione della specificità dei singoli settori (es. difesa, sanità, fisco, etc.) e, perché no, anche in relazione a specificazioni ulteriori. Più in generale, infatti, non c'è un modello di IA giusta o sbagliata rispetto ad una fornitura di sevizi. Certamente i servizi maggiormente basati su "categorie particolari di dati" (già "dati sensibili") vedono la IA privata maggiormente capace di assicurare servizi personalizzati e mantenere un alto livello di protezione e sicurezza. D'altra parte, laddove si tratti di un servizio destinato ad una larga fruizione e dove si ravvisino particolari ragioni di tutela, la IA pubblica, più facilmente accessibile tramite API o servizi basati su cloud e capace di includere librerie di apprendimento automatico, modelli pre-addestrati o servizi basati su cloud forniti da una *big tech*, sembra maggiormente adatta ove si tratti di offrire dei punti di accesso che consentano ai fruitori di utilizzare le capacità di una IA in maniera più accessibile e dove serva una fruizione più decentralizzata. Per esempio, si può dire che una IA pubblica si presta tendenzialmente meglio ad un uso da mobile, cioè dove la IA funziona direttamente sui devices senza inviare dati a server. Va però detto che la natura privata o pubblica di sistemi di intelligenza artificiale può dipendere anche dalle modalità prescelte: ad esempio assistenti vocali privati (es. Siri e Alexa) che notoriamente pongono problemi di riservatezza in relazione alla utilizzazione dei dati degli utenti per il miglioramento del sistema a beneficio di tutti possono essere ben impostati con modalità più aderenti al modello privato, in modo che le richieste vengano elaborate localmente senza essere inviati ai server. Lo stesso può dirsi dei servizi di messaggistica, dove impostazioni (ma più spesso il modo stesso in cui sono strutturati) possono utilizzare o meno tecniche di crittografia avanzata per garantire la sicurezza delle comunicazioni. Per venire ad un settore tra quelli di maggiore impatto potenziale per la pubblica amministrazione, le stesse diagnosi mediche, che ovviamente molto beneficiano da modelli di *Public AI*, possono essere prodotte da dispositivi medici che analizzano dati sanitari sensibili direttamente sull'apparecchio senza invio di dati a server esterni (c.d. diagnosi mediche locali). Pertanto dal momento che la pubblica amministrazione gestisce grandi quantità di dati personali sensibili su campi come sanità, previdenza e servizi fiscali (dove già esistono, come diremo presto servizi di IA) l'adozione di un'ottica *Private AI* consente di migliorare la sicurezza perché i dati possono essere elaborati localmente sui dispositivi

degli utenti senza invio a server centrali, peraltro con forte limitazione del rischio di violazione di dati o attacchi informatici e con minimizzazione nella quantità di informazioni trasferite o conservate in cloud o server esterni (e, per quel poco che conta per il piano della nostra riflessione, con una minore latenza, cioè una capacità di risposta molto più rapida).

D'altra parte, se è vero che una IA privata per sua natura ha standard elevati di sicurezza, occorre sempre che siano conformi alla normativa italiana e a quella dell'Unione europea e allo stesso modo una IA pubblica può integrare misure di sicurezza variabili a seconda degli standard normativi che adotta. Mentre una IA pubblica è una forma di scambio tra dati e servizi e i dati vengono utilizzati per rendere i servizi, una IA privata addestra propri dati (di cui resta proprietaria) e controlla meglio gli output. Quando si parla di pubblica amministrazione, tuttavia, il problema è gestire i dati dei cittadini-utenti. Bisogna approfondire, nella IA privata, in che termini l'organizzazione (che può essere la stessa PA che costruisce all'interno l'infrastruttura, o una organizzazione privata in appalto) utilizzi propri dati per erogare il servizio e in che termini occorre che l'amministrazione li fornisca. È chiaro che se c'è immedesimazione, con l'ipotesi della onerosa costruzione di un modello interno, non c'è differenza; diversamente, fermo restando che i dati restano all'interno di questo rapporto organizzativo (cosa che definisce la IA privata: che limita le *query* e le richieste all'interno dell'azienda che costruisce l'architettura o comunque a fonti private) bisogna comprendere in che termini una strategia di IA privata oggetto di un appalto di servizi ad un'impresa privata esterna alla PA porti alla costruzione di un modello basato su dati. Se in una IA pubblica le informazioni vengono non solo archiviate su database di terze parti ma anche rese disponibili a fornitori di terze parti, in quella privata viene meno questo secondo aspetto (costruendo una sorta di intranet aziendale) ma resta la questione del primo. D'altra parte, questa questione esiste sempre quando la PA non fa tutto in house, ma come sappiamo finora sono stati utilizzati servizi privati per la PA che avevano le caratteristiche di una IA pubblica, nel senso che i dati prodotti dagli utenti addestravano (o addestrano) ulteriormente, pur con tutte le (presunte) cautele in termini di privacy, i modelli a beneficio di chiunque (si pensi alle suite in uso in molte organizzazioni educative e d'istruzione).

Il Piano triennale per l'Informatica nella pubblica amministrazione dell'Agid (2024-2026) prevede a tale proposito che le PA si assicurano che i *foundation models* adottino adeguate misure di trasparenza che chiariscono l'attribuzione delle responsabilità e dei ruoli e che le PA che acquistano servizi di intelligenza artificiale tramite API, valutano con attenzione le modalità e le

condizioni con le quali il fornitore del servizio gestisce i dati. A tal fine è previsto un investimento in formazione e sviluppo delle competenze necessarie per gestire e applicare l'intelligenza artificiale in modo efficace nell'ambito dei servizi pubblici. In ogni caso le pubbliche amministrazioni forniscono informazioni adeguate agli utenti al fine di consentire loro di prendere decisioni informate riguardo all'utilizzo dei servizi che sfruttano l'intelligenza artificiale e adottano elevati standard di sicurezza e protezione della privacy per garantire che i dati dei cittadini siano gestiti in modo sicuro e responsabile, assicurando in ogni caso che le tecnologie utilizzate rispettino i principi di equità, trasparenza e non discriminazione. A tale proposito sono stati avviati vari progetti pilota in settori come la salute, l'istruzione e la gestione dei dati, per testare l'impatto dell'intelligenza artificiale e si prevede che la pubblica amministrazione collabori con università e aziende tecnologiche per sviluppare soluzioni innovative.

L'uso della intelligenza artificiale, ed in particola dei modelli fondazionali, nella PA potrebbe avere un notevole impatto soprattutto nella automatizzazione delle attività di ricerca e analisi delle informazioni semplici e ripetitive (in tal modo destinando il tempo di lavoro risparmiato ad attività a maggior valore), ad aumentare le capacità predittive, in tal modo migliorando il processo decisionale pubblico basato sui dati, e a supportare la personalizzazione dei servizi incentrata sull'utente, aumentando l'efficacia dell'erogazione dei servizi pubblici anche attraverso meccanismi di proattività. Nel suo *Piano Coordinato sull'Intelligenza Artificiale* (COM (2021) 205) del 21 aprile 2021, la Commissione europea si propone l'ambizioso obiettivo di «rendere il settore pubblico un pioniere nell'uso dell'IA». Naturalmente è un obiettivo estremamente ambizioso, in particolare per un ordinamento come il nostro. Anche in Italia sono state realizzare alcune precoci esperienze di utilizzo di modelli fondazionali (Agenzia delle Entrate, Inps, Istat) ma si tratta di alcune tra le più grandi organizzazioni operanti nel pubblico. Le pubbliche amministrazioni, prima di adottare *foundation models* "ad alto impatto" (di cui abbiamo detto *supra* al par. n. 1), devono assicurarsi che essi adottino adeguate misure di trasparenza che chiariscano l'attribuzione delle responsabilità e dei ruoli, non solo degli utenti ma anche dei fornitori di sistema di IA.

A sviluppo del piano triennale di cui si è detto in questi mesi è in atto un percorso normativo e organizzativo volto a dotare la PA di sistemi di IA, anche di tipo fondazionale, che passa preliminarmente attraverso l'adozione di linee guida su più piani (per lo più entro il 2024) e che prevede per i prossimi anni l'attivazione di un certo, elevato, numero di progetti di innovazione mediante IA, di avvio di iniziative di acquisizione di servizi di IA (c.d. *procurement*, cioè approvvigionamento

mediante appalti di forniture), di progetti di sviluppo di soluzioni IA, di realizzazione di applicazioni di IA a valenza nazionale e di sviluppo e implementazione di soluzioni basate su IA finalizzate al miglioramento della qualità dei servizi pubblici, con l'obiettivo di garantire uniformi livelli di servizio su tutto il territorio nazionale, anche attraverso la identificazione delle soluzioni nazionali fondate sull'IA (target 2024), lo sviluppo delle soluzioni nazionali (target 2025) e, infine, il dispiegamento nei territori delle soluzioni realizzate (target 2026). Pertanto dall'anno venturo partiranno progetti di innovazione, iniziative di sviluppo interno e *procurement* di soluzioni di IA nella piena conformità con la normativa italiana ed europea.

In particolare, per quanto riguarda i dati, l'AgID deve completare entro il 2024 una ricognizione delle basi di dati strategiche per identificare il *corpus* di dati necessario per addestrare i Large Language Models (c.d. IA generative) e ottenere soluzioni adeguate alla specificità dei vari domini e settori interni della pubblica amministrazione, mentre negli anni successivi la medesima Agenzia per l'Italia Digitale dovrà digitalizzare le basi di dati strategiche e, a partire dal 2026, promuoverle per l'addestramento dei sistemi di IA. Ciò avverrà in parallelo alla realizzazione di applicazioni di IA "a valenza nazionale", attraverso un processo che muove dalla identificazione delle soluzioni già esistenti a livello nazionale basate sull'IA, procedendo allo sviluppo di queste soluzioni, per trasformarle dal 2026 in prototipi per servizi estendibili su scala nazionale e per dispiegare successivamente sui territori le soluzioni realizzate. In tale ottica andrà affrontata anche la questione della localizzazione e del dimensionamento dei servizi mediante l'uso del cloud.

Pertanto, la scelta attuale maturato nell'ambito della strategia nazionale è che la pubblica amministrazione dovrà non solo acquisire competenze nel *procurement* di soluzioni già disponibili sul mercato, ma anche saper maturare competenze attivandosi concretamente per sviluppare proprie soluzioni. Viene pertanto perseguita un'attiva collaborazione tra il settore pubblico e privato, tra il mondo della ricerca e le realtà produttive, e si esplicita l'intenzione di favorire l'implementazione di soluzioni interoperabili in modo da ridurre il carico di dati e mitigare la congestione, investendo nella standardizzazione di soluzioni di rete per tecnologie di IA.

Sul piano dell'impatto sui diritti dei cittadini e degli utenti, la Strategia italiana per la IA detta indicazioni specifiche in tema di dataset e modelli, da costruirsi «secondo principi di trasparenza e *fairness*, che siano eticamente affidabili *by design* e che siano riusabili per accelerare le soluzioni delle aziende italiane. Sul piano metodologico si definirà un protocollo nazionale per garantire che i dataset siano *trustworthy-by-design* e *trusthworthy-by-default* sia sul piano legale che tecnico,

definendo approcci per la mitigazione di rischi (in termini etici e di cyber sicurezza) e i progetti che riceveranno finanziamenti pubblici, nell'ambito della strategia e al di fuori, saranno tenuti a riportare i dataset utilizzati e i modelli prodotti nel registro, in accordo a linee guida che definiranno i livelli di accesso e le modalità di riuso»[1].

La creazione di modelli fondazionali «safe» dovranno rispondere appieno ai valori e alle regolamentazioni europee in termini di trasparenza sui dati di training, garantendo il rispetto delle leggi sulla non discriminazione, privacy, tutela dei diritti umani, e fornendo informazioni affidabili sulle fonti in base alle quali vengono generati i contenuti; protezione dai contenuti generati falsi (c.d. allucinazioni; nonché proteggere i diritti degli autori e dei creatori le cui opere sono utilizzate nei dati di addestramento); e contenere meccanismi di tracciamento dei contenuti generati dall'IA (es. *watermarking*), nonché farsi carico della sostenibilità ambientale, puntando a tecniche innovative di riduzione delle dimensioni dei modelli (anche con approcci di *incremental* e *federated learning*) e degli impatti socio-economici a medio e lungo termine. A tal fine, ogni iniziativa dovrà essere supervisionata da un Comitato etico che ne approvi le linee generali, le metodologie realizzative e evidenzierà i rischi connaturati ad ogni iniziativa, orientando le scelte progettuali verso approcci che garantiscano la sicurezza e l'affidabilità delle soluzioni[2].

Le premesse e anche i primi sviluppi lasciano ben sperare in riferimento alla creazione, adattamento o implementazione di sistemi di intelligenza artificiale, anche fondazionali (sia linguistici che multimediali), che assicurino non solo la privacy, ma anche la tracciabilità delle fonti, la credibilità, accuratezza e pertinenza dei dataset al fine di creare un rapporto di fiducia tra utenti e sistemi, e di conseguenza, tra il cittadino e il potere pubblico, guidato dall'operare dei principi costituzionali.

MARCO PLUTINO
Università degli Studi di Cassino
e del Lazio Meridionale

[1] Cfr. *Strategia italiana per l'IA 2024-2026*, p. 15.
[2] *Ibid.*, p. 19.