AI ACT E

PUBBLICA AMMINISTRAZIONE

Elementi essenziali e impatto

Reg. UE 2024/1689 · Province · Comuni · Regioni

Di cosa vorrei parlare con voi oggi

1	Cosa è l'Intelligenza Artificiale	10 min
2	I rischi dell'AI	10 min
3	La strategia dell'AI Act	20 min
4	La PA come Deployer	20 min
5	La PA come Producer	20 min
6	Appendice: AI Act e GDPR	10 min

PARTE 1

Cosa è l'Intelligenza Artificiale

1

Quali sono le logiche di funzionamento alla base dei sistemi che chiamiamo AI?

AI ≠ software tradizionale: la differenza che conta

SOFTWARE TRADIZIONALE

Il programmatore scrive le regole

SE reddito < 15.000 **E** ISEE valido
ALLORA: ammetti la domanda

La regola giuridica (nota, esplicita) è tradotta in un codice software che ne riproduce il senso e le conseguenze

INTELLIGENZA ARTIFICIALE

Il sistema trova le regole nei dati

analisi di 10.000 pratiche storiche in applicazione di una medesima procedura
→ il modello impara quale regola è stata applicata, senza conoscerla previamente (addestramento)
→ sull base di quanto appreso, il modello applica questa regola ai nuovi casi

Le regole NON sono state indicate alla macchina: sono indivuate come implicite, sulla base dall'analisi delle regolarità e della correlazioni presenti nei casi.
(Come vedremo, questo è un caso di deep learning c.d. supervisionato).

Come il modello impara: correlazione e addestramento

Il meccanismo fondamentale

Il modello riceve grandi quantita di dati, identifica correlazioni statistiche ricorrenti tra variabili di input e output, e aggiusta progressivamente i propri parametri interni (pesi) per ridurre l'errore di previsione. Non "capisce": ottimizza una funzione matematica finche le correlazioni nei dati di addestramento si riflettono nelle previsioni.

APPRENDIMENTO SUPERVISIONATO

Dati etichettati: ogni esempio porta la risposta attesa.

Il modello apprende la correlazione input -> label e la generalizza a nuovi casi non visti.

Esempi nella PA:

- 10.000 domande ISEE con esito storico (ammessa/ rigettata) -> previsione automatica su nuove domande
- Atti storici etichettati per tipo -> classificazione automatica degli atti in arrivo

APPRENDIMENTO NON SUPERVISIONATO

Dati senza etichette: il modello trova strutture e raggruppamenti autonomamente.

Non prevede un output predefinito: scopre pattern latenti che nessuno aveva cercato esplicitamente.

Esempi nella PA:

- Fatture di fornitori senza categorie -> il modello raggruppa anomalie (antifrode)
- Segnalazioni dei cittadini -> clustering automatico per tema senza categorie predefinite

Come funziona la GenAI: tre caratteristiche fondamentali

ChatGPT, Copilot, Gemini non cercano in un archivio: producono testo nuovo ogni volta, combinando pattern appresi durante l'addestramento.

1

Generatività

Il sistema non “recupera”: compone. Ogni risposta è sintetizzata al momento, combinando miliardi di pattern statistici appresi nei dati di addestramento.

2

Variabilità

Lo stesso prompt in momenti diversi produce output diversi. Non esiste una "risposta giusta" univoca: il modello campiona tra possibilità probabilisticamente equipollenti.

La variabilità dipende anche da chi pone la domanda, dal contest in cui è posta, dal momento in cui è posta, etc.

3

Opacità

Il sistema non sa spiegare perché ha prodotto quell'output. Non esiste un "ragionamento" interno tracciabile: il risultato emerge da pesi numerici, non da regole leggibili.

Queste caratteristiche non sono difetti: sono la natura del modello. Capirle è il primo passo per usarlo correttamente.

Le utilità concrete per gli enti locali

Automazione pratiche

Classificazione automatica, instradamento, data entry da documenti scansionati.

Supporto decisionale

Analisi di grandi volumi di dati, rilevazione anomalie, prioritizzazione controlli.

Chatbot cittadini

Risposta automatica a domande standard: orari, procedure, moduli.

Supporto redazionale

Sintesi di documenti, bozze di comunicazioni, assistenza nella ricerca normativa.

Efficienza gestionale

Ottimizzazione raccolta rifiuti, manutenzione predittiva, pianificazione personale.

Analisi predittiva

Previsione di rischi, ottimizzazione allocazione risorse, gestione del traffico.

PARTE 2

I rischi dell'Intelligenza Artificiale

Questo modo di produrre ed organizzare la conoscenza non è privo di rischi, vediamo perché

2

Caso concreto: le allucinazioni e le loro conseguenze

Un fatto accaduto negli USA — e già riprodotto in contesti professionali italiani

📋 **CASO — Mata v. Avianca (2023)**

Un avvocato deposita un atto giudiziario citando sei sentenze a supporto della sua tesi. Le sentenze però non esistono nel mondo reale: sono state "inventate" da ChatGPT, con numero di causa, anno e massima plausibile.

Il giudice commina ha comminato una pesante sanzione allo studio legale. Il caso è stato pubblicato su tutte le testate giuridiche internazionali ed è diventato un punto di riferimento sul rischio GenAI in ambito professionale.

Il rischio per la PA è analogo:

Parere giuridico

Un funzionario usa un LLM per cercare normativa: il sistema cita una circolare ministeriale inesistente.

Determine e atti

Una determina cita precedenti giurisprudenziali inventati: l'atto è viziato.

Relazioni tecniche

Un tecnico usa GenAI per redigere una relazione: i dati numerici sono plausibili ma errati.

I sei fattori che generano discriminazione algoritmica

La discriminazione non ha un'unica sorgente. Questi fattori spesso si combinano.

1 Data Bias

Dati storici che rispecchiano discriminazioni passate

2 Algorithmic Bias

Scelte di design che privilegiano efficienza sull'equità

3 Labeling Bias

Pregiudizi degli annotatori umani nei dati di training

4 Variabili proxy

Il codice postale "deduce" l'etnia, il linguaggio la provenienza

5 Team omogenei

"Punti ciechi" di chi ha sviluppato il sistema

6 Black Box

Opacità che impedisce di vedere e correggere la discriminazione

Esempio: data bias nei controlli ISEE

📋 **SCENARIO**

*Un Comune usa un sistema di scoring per **prioritizzare i controlli ISEE**.*

Il sistema è addestrato sui dati storici degli accertamenti passati: ma in passato i controlli erano stati condotti più frequentemente su alcune zone della città e su alcune categorie di lavoratori autonomi, per ragioni storiche e sociali.

Il sistema "apprende" queste correlazioni e continua a segnalare le stesse categorie (zone della cittò, tipologia di contribuente) come più rischiose — creando un circolo vizioso di discriminazione statistica.

Cosa accade:

Il sistema discrimina

Non per regola scritta, ma perché i dati storici erano già distorti.

Il bias si amplifica

Più controlli → più segnalazioni → più "conferme" → il sistema si rafforza.

È invisibile

Senza audit specifici, nessuno se ne accorge. I numeri sembrano corretti.

Il problema delle variabili proxy

Anche escludendo esplicitamente i dati sensibili, l'AI può ricostruirli indirettamente.

Codice postale	→	Etnia, reddito, livello istruzione
Tipo di scuola	→	Condizione socioeconomica
Linguaggio usato nella richiesta	→	Origine geografica, livello istruzione
Orario di presentazione della domanda	→	Tipo di lavoro, disponibilità tecnologica

***Caso documentato**: il sistema COMPAS (USA) per predire la recidiva criminale non usava la razza come variabile esplicita — eppure penalizzava sistematicamente le persone afroamericane attraverso variabili proxy correlate.*

Il problema più acuto per la PA: la black box e la motivazione

COME RAGIONA L'AI

Correlazione

"Se A è presente,
c'è il 90% di probabilità
che B sia vero."

Non sa il perché.
Non ha causalità (solo correlazione).
Non può spiegare.

VS

DIRITTO AMMINISTRATIVO

Motivazione causale

"La domanda è rigettata
perché [norma X] prevede
che [condizione Y]..."

L. 241/1990: ogni provvedimento deve essere motivato in modo razionale e sindacabile.

Automation bias: il pericolo della fiducia cieca

I sistemi AI sono percepiti come più affidabili di quanto non siano. L'Art. 14 AI Act esiste proprio per questo.

📋 **STUDIO SUI RADIOLOGI**

*Un esempio documentato: in uno studio su radiologi che affiancati da un sistema AI nella lettura di lastre e referti, **i medici commettevano più errori in presenza dell'AI che senza**. Quando il sistema sbagliava diagnosi, i radiologi tendevano a fidarsi dell'output e a non rilevare il problema — pur essendo professionisti perfettamente in grado di riconoscerlo autonomamente.*

Il punto non è che (solo) l'AI fosse (più o meno) inaffidabile: è che la sua presenza cambiava il comportamento umano. L'esperto smetteva di ragionare in modo autonomo e diventava, di fatto, un validatore passivo della macchina.

Nella PA significa:

Il funzionario deve evadere 100 pratiche al giorno. Il sistema AI classifica il 90% correttamente.
Tende inevitabilmente a fidarsi del sistema anche sul 10% sbagliato.

La supervisione umana "nominale" — un timbro sull'output AI —
non soddisfa l'AI Act.

Altri rischi: impiego e governance

Lock-in tecnologico

Dipendenza da un unico fornitore senza clausole di uscita. Nel tempo: potere contrattuale del fornitore, perdita di autonomia della PA.

Perdita di competenze

Se il Comune smette di fare graduatorie manualmente, tra dieci anni nessuno saprà più farlo. L'AI può "atrofizzare" le competenze interne.

Vuoto di accountability

"L'ha deciso l'algoritmo." "L'ha certificato il fornitore." Il cittadino non sa a chi ricorrere. Questo è esattamente ciò che l'AI Act vuole impedire.

Privacy e GenAI pubblica

Un funzionario che incolla dati personali del cittadino in ChatGPT (versione non configurata) viola il GDPR e mette a rischio il segreto d'ufficio.

PARTE 3

La strategia dell'AI Act

20 minuti · Il Regolamento UE 2024/1689 e cosa cambia per la PA

3

L'approccio: regolazione proporzionale al rischio

Reg. UE 2024/1689 · In vigore dall'agosto 2024 · Direttamente applicabile in Italia, senza recepimento

Un'analogia con il diritto amministrativo: modulare il controllo in relazione al livello del rischio

AVVISO Comunicazione	SCIA Dichiarazione	AUTORIZZAZIONE Piena istruttoria	DIVIETO Assoluto
Rischio MINIMO	Rischio LIMITATO	Rischio ALTO	Rischio INACCETTABILE

Più è alto il rischio → più stringenti sono gli obblighi

Rischio inaccettabile: i divieti già in vigore

In vigore dal 2 febbraio 2025. Nessuna eccezione per la PA. Chi viola, viola la legge oggi.

Social scoring

Sistemi che valutano la "meritevolezza sociale" dei cittadini per determinare accesso a servizi o opportunità. Il riferimento esplicito è al modello cinese di credito sociale.

Manipolazione subliminale

Sistemi che influenzano le scelte delle persone senza che ne siano consapevoli, sfruttando vulnerabilità psicologiche.

Sfruttamento vulnerabilità

AI che sfrutta età, disabilità, condizione economica o stato emotivo per influenzare decisioni delle persone.

Biometria in tempo reale

Identificazione biometrica real-time in spazi pubblici per fini di sicurezza, salvo tassative eccezioni per forze dell'ordine (terrorismo, persone scomparse).

Sistemi ad alto rischio: le applicazioni concrete per gli enti locali

Sistema AI in uso	Decisione supportata	Diritto a rischio	All. III
Scoring ISEE / antifrode	Prioritizzazione dei controlli fiscali sui dichiaranti	Parita di trattamento	Cat. 5
Graduatorie asili nido e mense scolastiche	Ammissione o esclusione dal servizio	Accesso a servizi essenziali	Cat. 5
Scoring per edilizia residenziale pubblica (ERP)	Posizione in graduatoria alloggi popolari	Diritto all'abitazione	Cat. 5
Valutazione domande di assistenza domiciliare	Accesso a cure e sostegno per anziani e disabili	Dignita, salute, autonomia	Cat. 5
Selezione e valutazione del personale PA	Assunzione, progressione, mobilita interna	Accesso al lavoro	Cat. 4
Videosorveglianza intelligente in spazi pubblici	Segnalazione comportamenti "anomali", identificazione persone	Liberta di circolazione, presunzione innocenza	Cat. 1/6
Chatbot per pre-screening domande di contributo	Filtro di ammissibilita prima della valutazione umana	Accesso a prestazioni sociali	Cat. 5

Ogni sistema che contribuisce — anche in modo consultivo — a una decisione con effetti giuridici su una persona specifica rientra nell'alto rischio.

Applicazioni a basso rischio: cosa puo fare la PA da subito

Regola pratica: *se l'output AI non arriva mai direttamente al cittadino e non produce effetti giuridici su una persona specifica, il rischio è presumibilmente basso.*

1 Supporto redazionale

Sintesi di verbali, bozze di comunicati, traduzioni. Il funzionario rivede e approva: l’AI fa risparmiare tempo, ma non decide nulla.

2 Chatbot informativi

Risposte automatiche su orari, documenti necessari, scadenze. Il chatbot informa, non valuta situazioni individuali.

3 Smistamento posta

Classificazione e instradamento di email in arrivo all'ufficio competente. Funzione puramente organizzativa (se gestisse la PEC non sarebbe a basso rischio)

4 Ricerca documentale

Motore semantico sugli atti storici dell'ente. Nessun effetto verso terzi: trova l’informazione che può essere utile, poi il funzioanrio decide cosa farne.

5 Manutenzione predittiva

Analisi dei sensori su strade, impianti, edifici per segnalare quando programmare interventi. L'AI segnala, il tecnico decide.

6 Trascrizione automatica

Verbali di consigli comunali, commissioni, audizioni. Strumento di lavoro interno, senza effetti su posizioni individuali.

I due ruoli chiave: Provider e Deployer

PROVIDER

Chi sviluppa il sistema

Art. 3(3) AI Act

- → Aziende IT che costruiscono il software
- → La PA stessa, se sviluppa internamente
- → La PA, se commissiona su misura

Obblighi: progettazione, testing, certificazione, documentazione tecnica, registrazione nel DB europeo.

DEPLOYER

Chi usa il sistema

Art. 3(4) AI Act

- → La PA, in moltissimi casi
- → Acquisto software gestionale con AI
- → Uso di ChatGPT/Copilot in ufficio

Obblighi: FRIA, supervisione umana, monitoraggio, formazione del personale, informazione al cittadino.

Gli obblighi del provider: cosa deve garantire chi costruisce il sistema

*Art. 3(3) AI Act — provider e chi sviluppa, commissiona o modifica sostanzialmente **un sistema AI ad alto rischio**.*

Art. 9

Risk Management System

Processo documentato e continuativo per identificare, valutare e mitigare i rischi per tutta la vita del sistema.

Art. 10

Governance dei dati

Dati di addestramento pertinenti, rappresentativi e privi di bias. Documentazione di provenienza e trattamento.

Art. 11

Documentazione tecnica

Fascicolo completo: architettura, metriche di performance, limitazioni note, istruzioni per l'uso sicuro.

Art. 12

Logging automatico

Ogni operazione del sistema e registrata in modo inalterabile. Serve per audit e contestazioni.

Art. 13

Trasparenza verso deployer

Istruzioni d'uso chiare: per cosa il sistema e stato validato, per cosa non deve essere usato, soglie di affidabilità.

Art. 14

Human oversight by design

L'interfaccia deve rendere facile la supervisione umana, non scoraggiarla. Il deployer deve poter fermare il sistema.

Art. 43

Conformity assessment

Autovalutazione di conformità (nella maggioranza dei casi) e registrazione nel database europeo dei sistemi ad alto rischio.

Art. 72

Monitoraggio post-mercato

Performance monitorate nel tempo. Incidenti gravi segnalati all'autorita nazionale di vigilanza (AGID per la PA).

Gli obblighi del deployer: cosa deve fare la PA che usa il sistema

Art. 3(4) AI Act — deployer e chi usa un sistema ad alto rischio nel proprio contesto operativo. La PA e quasi sempre deployer.

Art. 26(1)	Rispettare le istruzioni d'uso del provider. Non usare il sistema per scopi non previsti.
Art. 26(2)	Garantire **supervisione umana effettiva**: assegnare il compito a persone competenti e formate.
Art. 26(5)	Monitorare le performance nel tempo. Se il sistema degrada o produce risultati anomali, sospenderlo.
Art. 26(6)	Conservare i log automatici per almeno 6 mesi. Documentare ogni decisione AI-assistita.
Art. 27	Condurre il FRIA (Fundamental Rights Impact Assessment) prima di attivare il sistema.
Art. 26(8)	Informare i cittadini che la decisione che li riguarda è assistita da un sistema AI.

Esempio pratico

Un Comune adotta un sistema di scoring per la graduatoria degli asili nido.

Prima di attivarlo

Conduce il FRIA: verifica che il sistema non svantaggi sistematicamente famiglie straniere o monoparentali.

Designa un responsabile

Un funzionario formato conosce i limiti del sistema e puo sovrascrivere l'esito per casi particolari.

Informa i cittadini

L'avviso nella domanda specifica che la graduatoria e calcolata con supporto AI e che e possibile chiedere revisione umana.

Monitora ogni anno

Verifica che la distribuzione degli esiti non si sia spostata in modo anomalo per alcune categorie.

Il FRIA: la valutazione obbligatoria prima di usare AI

Fundamental Rights Impact Assessment — Art. 27 AI Act — Obbligatorio per ogni PA che usa sistemi ad alto rischio

Non è un modulo da compilare: è una valutazione sostanziale che risponde a queste domande:

Quali diritti?

Quali diritti fondamentali può impattare questo sistema? Accesso a servizi, privacy, non discriminazione?

Chi è a rischio?

Quali categorie di persone sono più vulnerabili: anziani, stranieri, persone con disabilità, famiglie in difficoltà economica?

Danno potenziale?

Qual è il danno se il sistema sbaglia? Negazione di un sussidio, segnalazione errata alle forze dell'ordine?

Rimedi?

Esistono meccanismi adeguati di ricorso, contestazione e correzione per il cittadino interessato?

Fase 1: Prima dell'acquisto — le domande al fornitore

La fase più importante (e più trascurata).

- ? Il sistema ha il conformity assessment? Dove è la documentazione?
- ? È registrato nel database europeo dei sistemi ad alto rischio?
- ? Quali dati sono stati usati per l'addestramento? Come è stata gestita la governance dei dati?
- ? Quali sono i limiti noti del sistema e i casi in cui non deve essere usato?
- ? Il sistema prevede logging automatico di tutte le decisioni assistite?
- ? Il fornitore si impegna a notificare la PA in caso di incidenti gravi o aggiornamenti significativi?

Fasi 2 e 3: Implementazione e uso continuo

FASE 2 — Prima di “accendere” il sistema

- Condurre il FRIA
  (valutazione impatto sui diritti fondamentali)
- Designare il responsabile della supervisione umana
  (nome e cognome, non solo ruolo)
- Formare il personale: limiti del sistema, come riconoscere errori, come documentare l'override
- Configurare il logging inalterabile di tutte le decisioni
- Definire la procedura di informazione al cittadino

FASE 3 — Uso operativo continuo

- Monitoraggio mensile delle performance: il sistema funziona ancora come atteso?
- Verifica sistematica di eventuali differenze di esito per gruppi demografici diversi
- Audit periodici: campione di decisioni AI-assistite con verifica del sovrascrittura umana
- Incident reporting: documentare, notificare al fornitore, informare i cittadini interessati
- Record keeping: conservare tutta la documentazione per almeno 7 anni

La supervisione umana: reale vs. nominale

Art. 14 AI Act + obbligo costituzionale di responsabilità decisionale (Letztentscheidungsverantwortung)

❌ Supervisione NOMINALE

Il funzionario firma ogni output AI.
Non capisce il sistema.
Non sa riconoscerne gli errori.
Non soddisfa l'AI Act.

✓ Supervisione REALE

Il responsabile conosce i limiti del sistema.
Sa quando l'output non va accettato.
Documenta ogni override.
Può fermare il sistema.

Come calibrare l'intensità della supervisione

Reversibilità:	Decisioni difficili da correggere richiedono più controllo umano.
Impatto sui diritti:	Più la decisione incide su vita, libertà, salute → più supervisione.
Discrezionalità:	Dove la PA ha margini di apprezzamento, l'AI può supportare ma non sostituire.
Errori noti del sistema:	Se il sistema sbaglia su certi profili, quei casi vanno supervisionati sistematicamente.

I diritti del cittadino nelle decisioni AI-assistite

L'AI Act non elimina il diritto amministrativo: si aggiunge ad esso. I diritti ordinari restano.

Diritto di sapere

Il cittadino ha diritto di essere informato che la decisione che lo riguarda è stata assistita da un sistema AI.

Diritto alla revisione umana

Ha diritto di chiedere che un essere umano riesamini la decisione. L'AI non può avere l'ultima parola.

Diritto al ricorso

Ricorso gerarchico, TAR, tutti i rimedi ordinari del diritto amministrativo restano pienamente applicabili.

Diritti GDPR (Art. 22)

Non essere soggetto a decisioni basate esclusivamente su trattamento automatizzato che producano effetti giuridici significativi.

PARTE 5

La PA come Producer

20 minuti · Quando la PA sviluppa, commissiona o adatta sistemi AI

5

Quando la PA diventa provider

Art. 3(3) AI Act: è provider chiunque sviluppi o faccia sviluppare un sistema AI — anche solo per uso interno.

1	**Sviluppo interno**	Il Comune sviluppa con il proprio IT un sistema AI (chatbot, sistema di classificazione pratiche). Anche una piccola applicazione conta.
2	**Commissioning su misura**	La PA commissiona a una software house un sistema costruito sui propri dati. La PA è co-provider e non può scaricare tutti gli obblighi sul fornitore.
3	**Fine-tuning**	La PA prende un modello open source (LLaMA, Mistral) e lo addestra sui propri documenti. Per l'AI Act è sviluppo di un nuovo sistema: la PA è provider.
4	**Modifica sostanziale**	La PA modifica in modo significativo un sistema già acquistato, cambiandone funzionalità o scopo. Acquisisce la posizione di provider per la versione modificata.

I tre scenari pratici per la PA (Weerts, 2025)

A seconda di come la PA accede all'AI, il suo ruolo giuridico cambia radicalmente.

I

Scenario I

Sviluppo in-house

La Regione crea internamente un sistema AI per valutare le pratiche di finanziamento.

PA = PROVIDER
Obblighi massimi

II

Scenario II

Commissioning

Il Comune commissiona un sistema di scoring per le graduatorie degli asili nido, costruito sui propri dati storici.

PA = CO-PROVIDER
Obblighi condivisi

III

Scenario III

Strumenti disponibili al pubblico

I funzionari usano ChatGPT (versione pubblica) per redigere comunicati e cercare normativa.

PA = DEPLOYER
⚠ Alto rischio privacy

Se la PA è provider: gli obblighi principali (sistemi alto rischio)

Art. 9 — **Risk Management System**

Processo continuo e documentato per tutta la vita del sistema.

Art. 10 — **Governance dati di training**

Dati pertinenti, rappresentativi, bias rilevati e corretti, documentati.

Art. 11 — **Documentazione tecnica**

Manuale completo: architettura, dati, metriche, limitazioni note.

Art. 12 — **Logging automatico**

Ogni decisione registrata in modo inalterabile. Conservazione adeguata.

Art. 13 — **Trasparenza verso deployer**

Istruzioni d'uso chiare: scopo, prestazioni attese, casi esclusi.

Art. 14 — **Human oversight by design**

L'interfaccia deve rendere la supervisione facile, non scoraggiarla.

Art. 43 — **Conformity assessment**

Autovalutazione (nella maggioranza dei casi) + registrazione nel DB UE.

Art. 72 — **Monitoraggio post-mercato**

Performance monitorate nel tempo. Incidenti gravi segnalati ad AGID.

Riassumendo

1

L'AI Act è già legge

I divieti sono in vigore da febbraio 2025. Non è una normativa futura.

2

Il contratto è il primo strumento

Le clausole AI Act vanno nel capitolato, prima della firma.

3

Supervisione umana reale

Un responsabile nominato non basta. Serve formazione e cultura organizzativa.

4

I diritti del cittadino si rafforzano (?)

Motivazione, contraddittorio, ricorso: restano. Si aggiunge il diritto alla revisione umana.

5

L'AI Act non risolve tutto

Black box, **motivazione** e **automation bias**: nodi aperti che il diritto amm. deve ancora risolvere.

Domande?

Per approfondire:

Reg. UE 2024/1689 (AI Act)

Koivisto, Koulu, Larsson (2024) — Maastricht Journal

Weerts (2025) — Cambridge Forum on AI: Law and Governance

Krönke (2025) in: Buying AI — Edward Elgar

APPENDICE

AI e tutela dei dati personali

GDPR e AI Act: due regolamenti, due logiche, due assi paralleli

APPENDICE — PARTE 1

La tensione latente tra AI e GDPR

1

Due logiche che confliggono by design

Il deep learning ha fame di dati. Il GDPR impone minimizzazione.

LOGICA DELL'AI

Piu dati = modello migliore

- I dati personali sono i più abbondanti e i più "segnalatori".
- Più sono rappresentativi della realtà, più il modello generalizza.
- La qualità del modello e proporzionale alla quantità e varietà dei dati.

LOGICA DEL GDPR

Meno dati = meno rischio

- Minimizzazione: trattare solo i dati strettamente necessari (Art. 5(1)(c)).
- Limitazione della finalità: i dati raccolti per uno scopo non possono essere riusati liberamente per scopi ulteriori non compatibili (Art. 5(1)(b)).
- Proporzionalità: il trattamento deve essere adeguato, pertinente e non eccedente rispetto alla finalità.

Il GDPR nasce nel 2016 e viene applicato dal 2018, quando i grandi modelli linguistici non esistevano ancora.
Non è stato progettato per questo problema — e si vede.

APPENDICE — PARTE 2

Le fasi del ciclo di vita AI e la tutela dei dati

Il rischio cambia radicalmente a seconda di dove si trova il dato

2

Fase 1: addestramento del modello

Questa fase avviene prima che la PA usi il sistema. Ma i rischi che produce ricadono sulla PA comunque.

1 Base giuridica per il training

I testi del web contengono dati personali. Il fatto che siano pubblici non significa che possano essere usati per addestrare modelli AI senza base giuridica. Il "legittimo interesse" invocato dai provider e contestato da più autorita europee, incluso il Garante italiano.

2 Dati sensibili impliciti

Nei testi compaiono condizioni di salute, opinioni politiche, orientamento sessuale, etnia -- anche senza che siano dichiarati. Il modello li "assorbe" e puo riprodurli o inferirli. Il loro trattamento richiede base giuridica rafforzata (Art. 9 GDPR).

3 Memorizzazione e extraction attacks

I modelli non dimenticano. Con tecniche apposite è possibile estrarre dal modello frammenti di testo esatti appresi durante il training, inclusi dati personali reali (nomi, indirizzi, numeri di telefono).

4 Il diritto all'oblio (Art. 17 GDPR)

Se i dati di una persona sono stati usati per addestrare il modello, come si cancellano? I dati sono "disciolti" nei pesi della rete: non esiste oggi un meccanismo standard per rimuovere l'influenza di un singolo dato senza riaddestrare da capo.

Fase 2: utilizzo di un modello LLM

Ogni volta che un funzionario invia un prompt, si apre una serie di questioni GDPR.

1. **Il funzionario inserisce dati nel prompt**
2. **Il dato viaggia sui server del provider**
3. **Il modello elabora e restituisce output**
4. **L'output contiene dati o inferenze**

Chi riceve i dati del prompt?

Il fornitore del modello. Se non c'e DPA, non è qualificabile come "responsabile del trattamento": è un terzo che riceve dati senza vincoli. Varie violazioni del GDPR.

I dati vengono usati per training?

Alcuni provider, nella versione pubblica, usano le conversazioni per migliorare il modello. I dati dei fascicoli dei cittadini potrebbero entrare nel corpus di training futuro.

Dove sono i server?

I principali LLM sono gestiti da aziende USA. Un eventuale trasferimento extra-UE richiede SCC o adeguatezza (Art. 44-49 GDPR). Spesso non verificata.

Cosa produce l'output?

Il modello può riprodurre dati reali appresi nel training, o generare dati falsi ma plausibili su persone reali (allucinazione su persone reali = violazione + diffamazione).

APPENDICE — PARTE 3

Il ruolo della PA nei due sistemi di regolazione

AI Act e GDPR: due assi diversi, combinazioni non banali

3

Due sistemi di ruoli che non coincidono

AI ACT — Reg. UE 2024/1689

Asse funzionale al Sistema

PROVIDER	Chi sviluppa, commissiona o modifica il sistema AI	VS	**DEPLOYER**	Chi usa il sistema nel proprio contesto

GDPR — Reg. UE 2016/679

Asse funzionale al Dato

TITOLARE	Chi determina finalita e mezzi del trattamento	VS	**RESPONSABILE**	Chi tratta su istruzione del titolare

I due assi si incrociano: un soggetto che è "provider" nell'AI Act puo essere "titolare", "responsabile" o "contitolare" nel GDPR, a seconda di come è strutturato il rapporto con i dati.

Due esempi: cosa può succedere quando manca l'allineamento

SCENARIO A — PA usa LLM pubblico senza DPA

Esempio concreto

Un funzionario dell'ufficio servizi sociali deve redigere una relazione su un nucleo familiare in carico. Per velocizzare, incolla nel prompt di ChatGPT (versione gratuita) nome, codice fiscale, composizione del nucleo, reddito ISEE e note sui minori presenti. Ottiene una bozza di relazione in pochi secondi.

Cosa è successo

Nessun DPA	OpenAI non e "responsabile del trattamento": e un terzo autonomo.
Trasferimento extra-UE	Dati su server USA senza Standard Contractual Clauses.
Dati sensibili non protetti	Situazione familiare e reddito: categorie ad alto rischio.
Possibile uso per training	ChatGPT gratuito puo usare le conversazioni per migliorare il modello.

SCENARIO B — Il fornitore usa i prompt della PA per migliorare il proprio modello

Esempio concreto

Un Comune stipula un contratto enterprise con un fornitore di LLM per assistere i funzionari nella redazione di atti. Il DPA è firmato. Ma nelle condizioni tecniche -- non nel DPA -- è previsto che i prompt vengano usati "in forma aggregata e anonimizzata" per migliorare il modello. Il Comune non se ne accorge al momento della firma.

Cosa e successo

Viola Art. 28(3)(a):	Il responsabile non può trattare i dati per finalita proprie.
Il fornitore diventa titolare autonomo:	Per quella finalità, sfugge al controllo della PA.
Il DPA non copre:	La clausola tecnica non è nel DPA: è nelle condizioni d'uso.

Come si incrociano i ruoli: la matrice pratica per la PA

La PA è spesso Deployer (AI Act). Il suo ruolo ai fini del GDPR dipende da cosa fa il fornitore.

Scenario	PA: ruolo AI Act	PA: ruolo GDPR	Fornitore: ruolo GDPR
PA usa sistema con **DPA**; fornitore non usa i dati per scopi propri	**Deployer**	**Titolare**	**Responsabile del trattamento**
PA usa LLM pubblico senza DPA	**Deployer**	**Titolare (che ha perso il controllo)**	**Titolare autonomo — PA viola il GDPR**
Fornitore usa i prompt per migliorare il proprio modello	**Deployer**	**Titolare (inadempiente)**	**Titolare autonomo per quella finalità**
PA commissiona sistema sui propri dati storici	**Co-provider**	**Titolare per i dati di training**	**Responsabile nel training; poi titolare se riusa il modello**
PA sviluppa il sistema internamente	**Provider**	**Titolare**	**Nessun fornitore: tutto in capo alla PA**

In ogni contratto con fornitori AI *— verificare se esiste DPA, vietare esplicitamente l'uso dei dati per training proprio, verificare la localizzazione dei server.*

3 regole operative

1

Prima di adottare qualsiasi sistema AI

Verificare se esiste DPA con il fornitore e se i server sono in UE o ci sono SCC. Senza queste garanzie, non inserire dati personali dei cittadini.

2

Il DPA da solo non basta

Deve contenere il divieto esplicito di uso dei dati per training, analytics propri o riuso del modello per altri clienti. Leggere le clausole, non fermarsi alla firma.

3

DPIA e FRIA vanno coordinati

La DPIA (GDPR, Art. 35) valuta il rischio per i dati personali. Il FRIA (AI Act, Art. 27) valuta il rischio per i diritti fondamentali. Sono due valutazioni diverse, da condurre insieme prima di attivare il sistema.