Sulle Linee Guida per l'adozione di IA nella pubblica amministrazione

Le risposte dell'Osservatorio della Fondazione Astrid sulle dinamiche dell'intelligenza artificiale alla consultazione promossa dall'Agenzia per l'Italia Digitale

In via preliminare, si ritiene opportuno segnalare che una consultazione pubblica non dovrebbe essere svolta su un testo già definito, ma dovrebbe avere lo scopo di raccogliere dai soggetti interessati indicazioni utili alla successiva stesura preliminare del documento. La modalità di consultazione, così come proposta, rischia di risultare poco effettiva e poco partecipata.

Parimenti, è da ritenere contrario alle buone pratiche internazionali l'utilizzo di un portale differente dalla piattaforma "ParteciPA", la quale dovrebbe fungere da punto unico di accesso alle consultazioni pubbliche italiane. Ancora più critica l'impossibilità di trovare un rinvio alla piattaforma "Forum Italia" partendo da "ParteciPA".

Sarebbe stato auspicabile il coinvolgimento nel gruppo di lavoro delle autorità indipendenti, con specifico riferimento all'Autorità garante della concorrenza e del mercato, al Garante per la protezione dei dati personali e all'Autorità per le garanzie nelle comunicazioni.

Da un punto di vista strutturale, sarebbe utile inserire, per ogni capitolo, una sintesi finale in forma tabellare delle raccomandazioni/attività pratiche richieste alle PA, in modo da rendere più semplice l'effettiva comprensione e attuazione.

Esaurite le considerazioni preliminari, i commenti sono riportati seguendo la suddivisione del documento poiché così è strutturata la consultazione sul portale "Forum Italia".

1. **Ambito di applicazione**

Nessuna osservazione.

2. **Riferimenti e sigle**

2.1 Le linee guida dovrebbero fornire alle pubbliche amministrazioni (PA) delle indicazioni chiare e fugare possibili dubbi interpretativi. A tal fine, non si ritiene opportuno l'uso di termini quali "dovrebbero" o "possono". Il documento, inoltre, dovrebbe guidare le PA in termini operativi. Così come presentato appare invece di difficile lettura, teorico e privo di esempi concreti potenzialmente utili ai destinatari del documento.

2.2 I riferimenti a disposizioni normative esterne al documento (es. art. 14-bis, comma 2, lettera a) CAD) dovrebbero essere spiegati, senza rimandi ad altri testi (elemento che complica ulteriormente la lettura e l'utilizzo delle linee guida in termini operativi). Inoltre, il paragrafo dedicato alla struttura dovrebbe indicare anche le modalità e le tempistiche di aggiornamento degli allegati menzionati.

3. L'intelligenza artificiale

3/3.1 Una spiegazione di cos'è un sistema di intelligenza artificiale e del ciclo di vita dello stesso appare superflua in un documento tecnico rivolto a delle autorità pubbliche. Si consiglia di mantenere unicamente il riferimento alle definizioni di fornitore e utilizzatore (ora 3.2). In ogni caso, nell'esercizio definitorio sarebbero da considerare anche le linee guida della Commissione sulla definizione di un sistema di intelligenza artificiale (C (2025) 924 final del 6.02.2025). Si ricorda, inoltre, che la versione italiana del Regolamento UE non traduce il termine *deployer* in utilizzatore. Se si vuole mantenere tale impostazione, la descrizione del ciclo di vita dovrebbe essere più precisa per risultare utile (es. con riferimenti ai responsabili, ai rischi e agli obblighi per ogni fase del ciclo). Il tema dei *bias*, qui introdotto, dovrebbe essere approfondito con strategie di mitigazione. Nell'allegato F (casi d'uso), si torna sul punto traslando l'onere sulle PA che dovrebbero invece essere guidate su questo aspetto, attraverso esempi pratici e possibili strategie di mitigazione.

3.2 Sarebbe utile, per una pubblica amministrazione, una spiegazione dettagliata dei casi in cui – come *deployer* – assume le responsabilità del fornitore, attraverso esempi specifici.

3.3 La classificazione dei sistemi di IA sulla base del rischio non rispetta propriamente quanto previsto dal Regolamento UE. La divisione è: pratiche vietate, sistemi ad alto rischio, obblighi di trasparenza, modelli di IA per finalità generali (questi poi suddivisi in rischio sistemico e non). Inoltre, menzionare in questo paragrafo degli obblighi

generici per i sistemi ad alto rischio appare poco funzionale per il lettore e confusionario.

3.4 Per dei dipendenti di una pubblica amministrazione sarebbe certamente più utile leggere i principi elencati per l'adozione dell'IA rispetto alle disposizioni del Regolamento UE e con indicazioni pratiche sull'attuazione. Così come ora elencati, i principi appaiono poco utili ad una autorità pubblica che è già consapevole, in termini generici così come delineati, di dover essere conforme alla normativa, di dover rispettare i valori fondamentali dell'UE o la normativa in materia di protezione dei dati personali. Si rischia, inoltre, una duplicazione rispetto a quanto previsto dal Regolamento europeo in materia.

4. Modello di adozione dell'IA

Se si vuole mantenere la suddivisione tra condotte che le PA "dovrebbero" o "possono" porre in essere, sarebbe più utile una struttura in forma tabellare o un raggruppamento per categorie. La lettura è così poco chiara.

Le indicazioni fornite appaiono, inoltre, largamente generiche e rischiano di non contribuire a spiegare alle pubbliche amministrazioni come attuare quanto indicato. Sarebbe utile un paragrafo dedicato all'uso dell'*open source*, con vantaggi e rischi.

4.2 Interessante la raccomandazione di definire una strategia comune per l'IA (es. comuni o università). Sarebbe però opportuno fornire maggiori indicazioni che guidino le PA. Ad esempio, si potrebbe raccomandare alle PA di favorire l'uso di *dataset* pubblici, sia per la fase di addestramento che per la raccolta "dinamica" delle informazioni. L'IA dovrebbe essere in grado di operare con dati legittimamente formati e raccolti, provenienti principalmente da altri enti regolatori e piattaforme pubbliche (ad esempio, *database* pubblici esistenti, portali e risorse online che forniscono accesso alla normativa vigente, nonché articoli scientifici, sebbene questi ultimi con le necessarie precauzioni). La qualità dei dati deve essere intesa anche in termini di "contestualizzazione": i dati di addestramento devono essere raccolti tenendo conto dell'uso che il sistema ne farà e delle possibili interpretazioni che ne deriveranno nel contesto normativo di riferimento. Con riferimento alla raccomandazione di "valutare esperienze e sperimentazioni già effettuate da altre PA" occorrerebbe ricordare la possibilità di avvalersi del riuso previsto dal CAD (art. 69).

4.4 Come già menzionato per altre parti del testo, un generico richiamo per le PA a tenere in considerazione le norme tecniche appare superfluo e poco utile nell'ambito di uno strumento come le linee guida, che dovrebbe favorire in termini operativi l'attività di *compliance* da parte delle PA e la pianificazione strategica interna degli usi.

4.5.2 Anche con riferimento a tale paragrafo, elencare i requisiti per le PA rimandando ai capitoli del Regolamento appare poco utile, là dove le PA avrebbero più che altro bisogno di una guida alla *compliance*.

5. Conformità delle soluzioni di IA

In luogo a una mera indicazione per le PA di verificare se i sistemi di IA rientrano nell'ambito di applicazione del Regolamento, sarebbe stato più utile dare indicazioni in tal senso attraverso una *check list* dedicata o attraverso una valutazione svolta da un singolo ente, al fine di assicurare una interpretazione standardizzata a livello nazionale.

5.4 Così come redatte, le indicazioni sulla conservazione della documentazione rischiano di non fornire alle PA alcuna indicazione aggiuntiva, rinviando ciecamente alle disposizioni del CAD e alle linee guida AgID o facendo riferimento a un "periodo congruo" di conservazione. L'utilità di tale paragrafo, così come impostato, è dunque dubbia.

6. Governance etica dell'IA

Positiva la raccomandazione per le PA di dotarsi di codici etici e di comportamento, anche in collaborazione con altre amministrazioni o tramite modelli redatti da enti terzi (es. associazioni di categoria).

Visto lo sforzo di proporre una lista di fonti o iniziative europee o internazionali, la quale sarà presto obsoleta, si consiglia di valutare la creazione di una pagina periodicamente aggiornata – pubblicata sul sito AgID – che raccolga le fonti utili allo sviluppo e al consolidamento dei codici etici.

7. Comunicazione

7.2 Sarebbe stata utile, soprattutto per questa sezione, un dialogo di AgID con il Garante per la protezione dei dati personali, al fine di validare il contenuto ed evitare il più possibile eventuali interventi del Garante difformi da quanto qui rappresentato.

8. Formazione e sviluppo delle competenze

Tale paragrafo manca di definire chiaramente il concetto di "AI literacy". Oltre a ricostruire le motivazioni per le quali occorre investire sulla formazione e lo sviluppo delle competenze e identificare alcune figure rilevanti, questa sezione potrebbe essere

arricchita con raccomandazioni più puntuali su come impostare la formazione negli uffici (es. piani di formazione, valutazione annuale sul livello di "AI literacy"; valutazione annuale sulle carenze degli uffici in termini di figure professionali specializzate e sviluppo di piani di intervento, ecc.).

Utile, in ogni caso, il richiamo ai programmi formativi già esistenti o in corso di sviluppo (es. piattaforma *Syllabus*).

9. Gestione e qualità dei dati

In tale sezione sarebbe stato utile raccomandare alle pubbliche amministrazioni l'addestramento dei sistemi di IA su dataset istituzionali (e.g. IstatData), al fine di garantire una maggiore qualità e certezza della fonte utilizzata e quindi del sistema che la alimenta.

10. Protezione dei dati personali

Tale sezione non apporta niente di nuovo per le pubbliche amministrazioni, in termini di indicazioni operative e strategie di mitigazione per assicurare la loro conformità alla normativa, ma riporta sinteticamente gli obblighi generali cui sono sottoposte le PA per la protezione dei dati personali. Inoltre, come già menzionato, appare preoccupante il mancato coinvolgimento del Garante per la protezione dei dati personali nella redazione di tale paragrafo. Forse proprio tale mancanza porta ad un mancato approfondimento del tema della DPIA che non può essere trattato solo in termini generici.

11. Sicurezza cibernetica

Sezione utile per i destinatari del documento in consultazione e ben strutturata.

A. Valutazione del livello di maturità nell'adozione di IA

Il contenuto è certamente interessante. Sarebbe però utile la predisposizione di una *check list* per permettere alle PA di autovalutare il proprio livello di maturità, con conseguente raccomandazione di pubblicazione sul proprio sito istituzionale.

B. Valutazione del rischio

Allegato utile. Sarebbe, tuttavia, utile l'aggiunta di esempi pratici per agevolare la comprensione (es. con riguardo ai criteri di classificazione e agli indicatori di prestazione – KPI applicati ad un caso concreto). La valutazione dovrebbe tenere anche conto dell'impatto sull'ambiente (solo parzialmente trattato) e dei costi da sostenere per l'adozione e lo sviluppo del sistema rispetto ai benefici.

Sarebbe importante, in tale contesto, la definizione di una *check list* con matrice di rischio ai fini della valutazione. Le *check list* sono, in concreto, un foglio di calcolo, diviso anch'esso tra criteri statici e dinamici. Lo strumento è utile a valutare la conformità afferente a situazioni specifiche, sino ad arrivare al grado di rischio complessivo e a stabilire e commisurare l'azione corrispondente alla situazione di eventuale non conformità. In particolare, si dovrebbe impostare un sistema di categorizzazione che includa principalmente due fasi. La prima, utile a definire una categorizzazione preliminare sulla base di criteri statici (afferenti alle caratteristiche del sistema, ai rischi noti, al settore specifico in cui opera la PA interessata allo sviluppo o all'adozione del sistema, all'esposizione, alla pericolosità e alla capacità di mitigazione del rischio identificato). La seconda, utile a definire una categorizzazione definitiva (salvo poi da sottoporre ad aggiornamento periodico), sulla base di criteri dinamici e tecnici, tra cui un'analisi graduale degli impatti del sistema nei primi tempi di utilizzo. Si ricorda, sul punto, che il Regolamento UE in materia, al considerando 81, indica di tenere conto delle specificità del settore come pure delle competenze e dell'organizzazione dell'autorità pubblica interessata.

Si tratta infatti di valutazioni che presuppongono una forte componente scientifica, ma anche evidenze empiriche connesse all'esperienza delle amministrazioni nel settore di interesse. È noto infatti che il concetto di rischio sia contraddistinto da una particolare dinamicità, la quale permette allo stesso tempo di assumere un significato differente in base al periodo storico, all'avanzamento scientifico e alla, maggiore o minore, specializzazione del personale cui è affidata l'analisi. Non esiste infatti una semplice concezione di regime basato sul rischio, ove questo incontra differenze in base al settore e alla formazione. Si tratta, in ogni caso, di prospettive che si sovrappongono e che richiedono un bilanciamento per una valutazione il più possibile fondata per la conseguente categorizzazione dei sistemi di IA. Tale discorso si inserisce nel doveroso presupposto dell'impossibilità di raggiungere un rischio zero, laddove lo scopo è dunque di mitigare i rischi più importanti, gestire o talvolta accettare quelli di minore rilevanza. Si tratta di un'attività che richiede, inoltre, un costante monitoraggio e una

rivalutazione dei criteri considerati per la valutazione e per il bilanciamento degli interessi, privati e pubblici.

C. Valutazione d'impatto

Tutti gli allegati, ma questa sezione nello specifico, dovrebbero essere più chiari a livello terminologico. Se qui si vuole far riferimento al tema centrale della valutazione di impatto sui diritti fondamentali (FRIA), questa dovrebbe essere approfondita con maggiore dettaglio, presentando un quadro per la conduzione di una FRIA. Si dovrebbe rappresentare un approccio che integri strumenti qualitativi e quantitativi per la valutazione dei rischi. In particolare, le linee guida dovrebbero fornire un questionario per raccogliere informazioni contestuali e operative sul sistema di IA e, inoltre, una matrice per mappare sistematicamente i potenziali impatti su specifici diritti e attribuire un punteggio quantitativo agli impatti. Sul punto, si dovrebbe prendere ispirazione dal lavoro di Andrea Cosentini et al, "Assessing the impact of artificial intelligence systems on fundamental rights", pubblicato su "MediaLaws". Si dovrebbe, da ultimo, considerare di ricordare quanto previsto dall'art. 27 del Regolamento europeo in materia, con riguardo alla possibilità di utilizzare la FRIA come integrazione della DPIA.

D. Modello di codice etico

Nessun commento.

E. Norme tecniche in ambito IA

Riferimenti utili e utile anche l'indicazione di una pagina web aggiornata sullo stato di avanzamento delle attività di normazione tecnica.

F. Casi d'uso

Utile il modello, sarebbe più funzionale con un esempio di applicazione ad un caso concreto.

G. Funzionalità dell'IA

Tale sezione potrebbe essere inclusa nelle precedenti per non appesantire ulteriormente il documento e soprattutto la valutazione delle pubbliche amministrazioni.

H. Procedure di governance

Tale sezione potrebbe essere inclusa nelle precedenti per non appesantire ulteriormente il documento e soprattutto la valutazione delle pubbliche amministrazioni.

I. Indicatori di prestazione

Sarebbe utile spostare questa parte all'inizio degli allegati o applicarla ad ognuna delle sezioni specifiche.

Marzo 2025