Digeat N.6 - 19 Giugno 2025

La gestione documentale a supporto dell’interoperabilità e open data, tra certezza del dato e sicurezza delle informazioni

Di Andrea Piccoli

Rubrica: Gestione documentale 2027

Abstract

Il sesto articolo di questa rubrica accoglie alcune riflessioni sulla gestione dei dati e documenti nello scenario attuale della interoperabilità e specializzazione dei verticali applicativi, ragionando prima sulla frammentazione e dispersione dell’archivio e quindi sulla certezza del dato e della sua disponibilità nel tempo. La riflessione, quindi, prosegue sul dato incerto, sul modello di gestione dei rischi derivanti dall’utilizzo di dati incerti o alterati.

Indice

- La gestione documentale e l’interoperabilità
- Il dato incerto
- Il rischio di alterazione
- Conclusioni

La gestione documentale e l’interoperabilità

Il ruolo della **gestione documentale informatizzata** ha avuto origine con la adozione di **soluzioni per la gestione del protocollo informatico** in conseguenza dell’applicazione del DPR 445 del 2000. Il ruolo principe del protocollo informatico è stato, e lo è tutt’ora, quello di **gestione delle comunicazioni in entrata e in uscita** (esteso poi anche per le comunicazioni interne) **dell’ente che hanno un effetto giuridico** ovvero che producono una attività amministrativa nell’ente.

Le soluzioni si sono, poi, estese alla **gestione dei flussi** (workflow) per la formazione dei documenti informatici e hanno spesso introdotto la **gestione di repertori e serie** per raccogliere la documentazione non soggetta a protocollazione, come ad esempio i repertori di atti o di pubblicazioni in trasparenza.

La richiesta normativa e di buona pratica della **gestione dell’archivio con la fascicolazione delle informazioni** secondo il piano di fascicolazione, ha portato alla introduzione di funzionalità nella gestione documentale dedicate **alla gestione dei fascicoli** e del loro ciclo di vita.

Nelle **pubbliche amministrazioni** tuttavia sono, in modo parallelo e spesso non coordinato, nati ed evoluti **sistemi verticali specializzati** al contesto amministrativo di riferimento, come ad esempio

quello della **gestione del personale, degli approvvigionamenti o della gestione dello sviluppo dell'edilizia sul territorio o di gestione della sanità pubblica**, secondo l'ordinamento e lo scopo giudico dell'ente. Tali verticali finiscono per gestire spesso un archivio dedicato di informazioni giuridicamente rilevanti.

Al contempo, l'integrazione con l'ecosistema digitale nazionale ha ulteriormente stimolato la raccolta di dati e documenti rilevanti nelle attività amministrative che spesso sono archiviati e gestiti nei diversi verticali specializzati. Ad esempio, sono poche le integrazioni con SEND o con PagoPA che archiviano tutte le informazioni rilevanti nell'archivio della gestione documentale e ancora meno nel protocollo.

La dispersione delle informazioni rilevanti in diversi archivi non sempre integrati pone evidenti problemi sia nella ricostruzione e raccolta di tutti gli atti e fatti giuridicamente rilevanti per l'attività o procedimento amministrativo sia, ancora di più, quando si intende utilizzare la metadatazione (prevista dalla normativa per i documenti informatici) per applicare le logiche di "record in context"[1] come nel caso della ricerca di tutti i documenti che riguardano un cittadino o un dipendente.

> Dal punto di vista degli obiettivi di archiviazione e conservazione nel tempo dei documenti e delle informazioni, la frammentazione delle informazioni, e la loro replicazione in diversi archivi sono due tematiche che devono essere indirizzate nella gestione documentale dell'archivio corrente.

Se non si indirizza correttamente la raccolta in un archivio digitale unico dell'ente, con l'archiviazione delle informazioni e dei dati rilevanti e dei loro collegamenti di contesto, non si può certo pensare di indirizzare il problema dal lato del servizio di conservazione (eArchiving), dato che spesso, in senso peggiorativo rispetto alla frammentazione, verticali diversi versano in conservazione utilizzando conservatori diversi.

> Questa dispersione dei documenti e delle informazioni rilevanti pone problematiche che vanno oltre l'ambito archivistico e di conservazione delle informazioni[2], come quella sulla certezza e autenticità del dato.

Se in un procedimento o in un'attività amministrativa uso un dato rilevante prelevato da un altro verticale (e quindi un altro archivio) posso essere certo che quest'ultimo sia integro, autentico e aggiornato?

Stessa riflessione si estende anche al tema dell'interoperabilità realizzata con la PDND **e con l'utilizzo degli open data**: i dati pubblicati possono dirsi integri, autentici e aggiornati? Sugli aspetti di integrità e autenticità la PDND offre evidenti garanzie e misure tecniche adeguate mentre sul versante Open Data, sebbene siano da richiamare le prescrizioni delle Linee Guida sugli Open Data di AgID[3], le garanzie di integrità e autenticità sono demandate alle singole realizzazioni e alle modalità di utilizzo.

Il dato incerto

Dalle riflessioni precedenti si evince la **necessità di gestire in modo corretto il rischio** di utilizzo nelle attività e procedimenti amministrativi di dati non certi, ed in modo analogo il rischio di diffusione di dati non aggiornati ovvero non certi.

> Riprendendo le varie metodologie di gestione dei rischi[4], si tratta di individuare un modello che indirizzi in modo opportuno i fattori di rischio, collegati alle vulnerabilità

> che li possono rendere manifesti, e per ciascuno di essi valutare la probabilità di accadimento e impatto e quindi il rischio effettivo. Valutando, poi, le misure tecniche e organizzative attuate per pianificare, dunque, le attività necessarie alla mitigazione o accettazione del rischio residuo. In queste valutazioni, le considerazioni sulla certezza della fonte e sulla sua periodicità di aggiornamento, unite alla sicurezza del canale di comunicazione ed integrazione, sono centrali.

Anche rispetto alla adozione di soluzioni di IA nella pubblica amministrazione e le relative Linee Guida di AgID[5] il rischio di utilizzo di informazioni incerte[6], nella fase di addestramento dei modelli e nella fase di loro utilizzo, è orientato richiamando la UNI ISO 42001:2023, che individua le metodologie per finalizzare l'analisi del rischio e di impatto sia rispetto ai diritti degli interessati[7] sia rispetto ai diritti fondamentali dei soggetti coinvolti e destinatari[8].

A parte gli aspetti tecnologici e realizzativi delle soluzioni, il fattore umano, rappresentato dal funzionario pubblico che utilizzando le soluzioni informatiche a disposizione opera nelle attività e procedimenti amministrativi, risulta essere una fonte di rischio da valutare attentamente. Al di là della indispensabile formazione continua va sicuramente affrontata e promossa un'impostazione fondata su una piena consapevolezza e una ben strutturata organizzazione delle attività.

Infine, quando si costruiscono delle integrazioni di interoperabilità basate su dati presenti su altre piattaforme e open data, **si deve riflettere sul tema della disponibilità del dato**, ovvero della resilienza del proprio servizio digitale rispetto alla mancata disponibilità delle altre sorgenti dei dati utilizzati. **Le attività della Agenzia per la Cibersicurezza Nazionale (ACN) in attuazione della direttiva NIS 2[9] indirizzano** le valutazioni di rischio e di impatto per la resilienza dei servizi digitali offerti dalla pubblica amministrazione **estendendo la valutazione alla intera filiera dei soggetti coinvolti compresi i fornitori di dati e loro piattaforme**.

Il rischio di alterazione

A completare questo articolo si vuole condividere la riflessione, legata alle considerazioni sulla certezza dei dati utilizzati nelle attività e procedimenti, relativa al rischio di utilizzo di dati alterati.

In un momento di sempre maggiore diffusione degli attacchi informatici la certezza dei dati può essere oggetto di manipolazioni malevole. In particolare, la sicurezza sia dei punti di comunicazione[10] che dei canali di comunicazione è fondamentale richiedendo non solo l'utilizzo di canali sicuri ma anche la verifica della reale identità del soggetto che pubblica il dato e del soggetto che lo richiede a dimostrazione della validità e integrità dei processi di identificazione messi in atto.

Sempre in questo contesto di rischio di alterazione si legano altre considerazioni sui dati errati creati da allucinazioni generate dai modelli delle soluzioni di intelligenza artificiale in uso e dall'utilizzo dei dati falsi o manipolati[11] nelle fasi di apprendimento e di utilizzo dei modelli stessi.

Conclusioni

I punti fondamentali da ricordare sono:

- **La gestione della frammentazione e dispersione dei dati e documenti deve essere indirizzata nella gestione documentale dell'archivio corrente**, raccogliendo **in modo interoperabile** le informazioni e i documenti certi con i loro metadati dai diversi verticali interni ed esterni che li formano e gestiscono.

- **L’utilizzo delle informazioni e documenti esterni**, sia attraverso le piattaforme nazionali sia, a maggior ragione, attraverso l’uso degli open data, **deve essere oggetto delle opportune valutazioni di rischio** sulla certezza, oltre che autenticità e integrità, e delle misure applicate e della sicurezza sia degli end-point che dei canali di trasmissione.
- Le valutazioni sulla resilienza dei servizi digitali offerti, che richiede disponibilità dei dati esterni rientrano nelle specifiche di attuazione del regolamento NIS 2.
- La certezza delle informazioni nell’ambito di utilizzo delle soluzioni basate su intelligenza e la certezza dei dati rappresentano il fulcro da analizzare sia nella fase di addestramento che di uso dei modelli.

NOTE

[1] Records in Contexts (RiC): a standard for archival description developed by the ICA Experts Group on Archival Description – ICA

[2] Come condiviso l’archivio conservato avrà una frammentazione delle informazioni non inferiore a quella dell’archivio corrente a causa della perdita di informazione nel processo di conservazione.

[3] Con Determinazione n. 183/2023 AgID ha adottato e pubblicato le “Linee Guida recanti regole tecniche per l’apertura dei dati e il riutilizzo dell’informazione del settore pubblico”, Open Data: AgID adotta le Linee Guida | Agenzia per l’Italia digitale.

[4] ISO/IEC 31000, La gestione del rischio: il percorso modulare – UNI – Ente Italiano di Normazione.

[5] Con Determinazione n. 17/2025 AgID ha avviato la consultazione pubblica delle “Linee guida per l’adozione dell’Intelligenza Artificiale nella Pubblica Amministrazione”, Intelligenza artificiale | Agid.

[6] Identificato come l’utilizzo di fonti di rischio correlate all’apprendimento automatico.

[7] DPIA, Data Protection Impact Analisys, in applicazione al GDPR.

[8] FRIA, in applicazione al IA Act.

[9] DECRETO LEGISLATIVO 4 settembre 2024, n. 138; NIS – Network Information Security – ACN.

[10] Intesi come gli “endpoint” dei servizi pubblicati e dei client di consultazione/integrazione.

[11] Ci si riferisce al tema delle “fake-news” o di attacchi informatici mirati ad iniettare dati manipolati nei modelli di IA in uso.