Digeat N.6 - 19 Giugno 2025

Il paradosso dell’archiviazione digitale: conservare l’autentico o l’artificiale verosimile?

Di Giusy Distratis

Abstract

L’avvento dell’Intelligenza Artificiale generativa ha modificato i paradigmi tradizionali dell’archiviazione digitale, fondati sulla conservazione di documenti autentici. La capacità dell’AI di creare contenuti indistinguibili da quelli reali pone un interrogativo: dobbiamo limitarci a preservare l’autentico, rischiando di ignorare una parte crescente della realtà digitale, o dobbiamo estendere i nostri sistemi per includere anche l’”artificiale verosimile”?

Indice

La crisi d’identità dell’archivio digitale e il valore dell’autentico

Da sempre gli archivi sono stati ritenuti importanti per preservare la veridicità storica e l’identità culturale dei popoli: manoscritti, fotografie e opere d’arte hanno contribuito a mantenere viva la memoria e restano testimonianze dell’epoca che le ha prodotte, trasmettendo valori e conoscenze alle future generazioni. Tradizionalmente confinata all’ambito archivistico, **l’autenticità si riferiva principalmente alla verifica dell’origine e della genuinità dei documenti**. Tale concetto ha lentamente abbandonato la sua storica collocazione **per assumere una valenza onnicomprensiva fino a diventare un pilastro fondamentale della cybersecurity e della gestione dell’informazione digitale**.

> Non si tratta più solo di certificare l’origine di un documento, ma di garantire l’integrità dell’intero ecosistema informativo: dalla creazione all’archiviazione, passando per la trasmissione e l’utilizzo dei dati. L’archiviazione digitale persegue l’obiettivo di garantire la conservazione nel tempo di documenti e dati **autentici, integri, affidabili, leggibili e reperibili**.

L’attuale quadro normativo sulla conservazione delineato dal CAD e dalle Linee Guida AgID, insieme a standard internazionali quali la ISO 15489 sulla gestione dei documenti o il modello OAIS (Reference Model for an Open Archival Information System) per la conservazione di archivi digitali[1], è focalizzato sulla **garanzia nel tempo dell’autenticità e integrità del documento informatico**

“originario”, inteso come prodotto dell’azione umana o di processi automatizzati ma predeterminati.

> La complessità dell’attuale sistema rende sempre più difficile discernere il vero dal falso, l’originale dal contraffatto, il sicuro dal compromesso. L’esistenza di contenuti AI indistinguibili *ictu oculi* da quelli autentici, mina la presunzione di affidabilità su cui si regge l’attuale sistema. Come potrà un giudice, o qualsiasi soggetto terzo, fare affidamento su un documento digitale conservato se sussiste la concreta possibilità che esso sia un artefatto AI non dichiarato?

Le conseguenze processuali dell’utilizzo di documenti informatici, di cui si dubiti dell’origine o si contesti l’autenticità, possono essere significative, tra tutte un **potenziale aggravamento dell’onere probatorio** a carico della parte che li produce in giudizio. Tale parte potrebbe essere tenuta a provare, oltre all’integrità del documento, anche la sua derivazione da attività umana ovvero, in alternativa, a dichiararne la natura artificiale. A ciò potrebbe aggiungersi la crescente esigenza di ricorrere a consulenze tecniche specialistiche per l’individuazione di artefatti di intelligenza artificiale, con conseguente incremento dei costi e della complessità dei procedimenti. Il potenziamento delle capacità generative dell’AI potrebbe rendere necessaria una revisione della nozione stessa di “originale” digitale e delle modalità di formazione del documento informatico giuridicamente rilevante.

> Ci troviamo di fronte a un paradosso: mentre affiniamo tecniche per preservare l’autenticità del passato digitale, una nuova ondata di contenuti “artificialmente verosimili” popola il nostro presente e futuro digitale. Deepfake video e audio indistinguibili dagli originali, testi sintetici capaci di imitare stili di scrittura specifici, campagne di disinformazione automatizzate e personalizzate: l’IA offre strumenti potenti per alterare la realtà percepita, minando alla base la fiducia nelle fonti informative, nelle istituzioni e nelle interazioni interpersonali online.

Cybersecurity come presidio di autenticità: dalla conservazione al nuovo paradigma archivistico

Se l’IA rappresenta una delle maggiori minacce all’autenticità, la normativa europea e nazionale sulla cybersecurity ne costituisce uno dei principali baluardi. La conservazione a norma, secondo le regole tecniche AgID, garantisce valore legale e autenticità nel lungo periodo ai documenti informatici. Il **Cybersecurity Act (Reg. UE 2019/881)**, introducendo schemi di certificazione europei per prodotti, servizi e processi ICT, mira a creare un ambiente digitale più affidabile. Un prodotto certificato offre maggiori garanzie sulla sua sicurezza intrinseca, contribuendo indirettamente all’autenticità dei dati che processa.

La **Direttiva NIS 2 (UE 2022/2555)**, recepita in Italia con il recente **D. Lgs. 18 maggio 2024, n. 138**, ampliando il novero dei soggetti obbligati e innalzando i requisiti di sicurezza e notifica degli incidenti, impone misure tecniche e organizzative volte a garantire la **disponibilità, l’integrità e la riservatezza** dei sistemi e dei dati. Proteggere l’integrità dei sistemi informativi significa proteggere anche l’autenticità delle operazioni e delle informazioni che essi gestiscono.

L’**archiving digitale sostenibile** deve andare oltre la mera conservazione passiva. Deve integrare tecnologie e processi che garantiscano l’autenticità fin dalla creazione del dato o documento digitale (es. firme digitali qualificate, marche temporali, metadati standardizzati e sicuri). Tecnologie emergenti come la blockchain possono essere utilizzate per creare registri distribuiti e immutabili, volti a certificare l’esistenza e l’integrità di un asset digitale in un dato momento. L’emergere del quantum computing, ad esempio, potrebbe minacciare i sistemi di crittografia attualmente utilizzati per garantire l’autenticità, richiedendo lo sviluppo di **crittografia post-quantistica** o “quantum-resistant”. Inoltre,

l’espansione dell’Internet of Things (IoT) ha moltiplicato i potenziali punti di ingresso per attacchi che potrebbero compromettere l’autenticità dei dati, rendendo necessario un ripensamento delle strategie di sicurezza.

Verità Sintetica, Memoria Fragile: come Cambia l’Archivio?

L’AI generativa deve farci ripensare le fondamenta stesse dell’archiviazione digitale. La difficoltà di fondo risiede nella capacità di distinguere un contenuto creato *ex novo* da un essere umano da uno generato o alterato mediante sistemi di AI.

La diffusione deliberata di informazioni false o fuorvianti non è un fenomeno nuovo, pensiamo alla propaganda politica che affonda le sue radici in epoche remote, ben prima dell’Impero Romano, dove veniva impiegata per consolidare il potere dei sovrani, costruire il consenso attorno a decisioni importanti e mantenere una certa stabilità sociale.

L’autenticità non riguarda solo il contenuto dell’informazione, ma anche la sua provenienza (chi l’ha creata? è davvero chi dice di essere?) e la sua integrità (è stata manipolata durante la diffusione?). La lotta alle fake news richiede non solo strumenti di fact-checking, ma anche meccanismi per tracciare l’origine e garantire l’immutabilità delle informazioni legittime.

> Occorre chiedersi se i sistemi di conservazione debbano aprirsi anche all’archiviazione di contenuti generati da AI. Archiviare esempi significativi di contenuti AI-generated (es. deepfake, testi manipolati) potrebbe servire per sviluppare tecniche di rilevamento, per studiare l’evoluzione delle tecnologie generative e organizzare campagne di contrasto alla disinformazione. L’archivio potrebbe diventare una risorsa per comprendere e contrastare la manipolazione dell’informazione, oltre ad avere un **valore storico e sociologico,** in quanto prodotto del nostro tempo.

Inoltre, in determinati contesti (es. processi decisionali assistiti da AI, output di sistemi automatizzati), conservare l’output generato dall’AI potrebbe essere necessario per garantire la trasparenza e l’accountability degli algoritmi e dei loro utilizzatori, anche al fine di conservarne la tracciabilità e delineare eventuali profili di responsabilità. Manca, tuttavia, **una disciplina esplicita per la gestione e conservazione di contenuti la cui natura “artificiale” è nota *ab origine* e la cui conservazione è intenzionale** (ad esempio, per scopi di ricerca, analisi storica, documentazione di processi basati su AI, o anche per finalità forensi).

> Quali criteri adottare per selezionare i “falsi” che meritano di essere conservati a lungo termine? Si potrebbe rischiare di creare archivi di simulacri senza un chiaro valore documentale intrinseco, senza meccanismi robusti di distinzione, o di includere contenuti artificiali che potrebbero inquinare l’archivio con un sovraccarico dei sistemi di archiviazione.

Sebbene il watermarking digitale e le firme crittografiche per output AI siano tecnicamente validi, permangono vulnerabilità quali la potenziale rimozione e, soprattutto, la mancanza di standard universali e vincolanti. Una soluzione potrebbe essere l’introduzione di un “marchio d’origine” digitale obbligatorio che certifichi inequivocabilmente la natura artificiale dei contenuti AI. Parallelamente, i metadati potrebbero rivestire un ruolo cruciale nell’assicurare l’autenticità dei contenuti AI archiviati. Oltre ai set di metadati descrittivi, gestionali e strutturali esistenti, **sarebbe opportuno definire e standardizzare set specifici**, includendo un indicatore esplicito ***“Generato/Manipolato da AI”,*** l’identificativo univoco del modello AI, il prompt di generazione, la marcatura temporale certa, la finalità di generazione e conservazione.

L’autenticità come impegno continuo per un’archiviazione ibrida ma consapevole

> Il dominio dell’autenticità digitale non è uno stato da raggiungere una volta per tutte, ma un **processo continuo di gestione del rischio e di adattamento**. La sfida dell’autenticità nell’era digitale non è solo tecnica ma anche etica e sociale, richiedendo una riflessione continua sul significato stesso di “realtà” e sulle responsabilità di chi gestisce e diffonde informazioni in un ecosistema sempre più vulnerabile alla manipolazione.

È necessario un approccio olistico e multidisciplinare, nel quale aziende e professionisti (legali, informatici, archivisti, responsabili della comunicazione) devono collaborare per adottare una **visione strategica** in cui l’autenticità sia considerata un asset aziendale critico, integrato nelle strategie di risk management, compliance e comunicazione. Occorre un quadro strategico per la tutela dell’integrità e della sicurezza delle informazioni digitali. È necessario:

- **implementare misure tecniche robuste** quali l’utilizzo di sistemi di identità digitale sicuri e autenticazione a più fattori, adottare firme elettroniche e marche temporali qualificate per i documenti rilevanti, implementare sistemi di logging sicuri e immutabili per tracciare accessi e modifiche, investire in strumenti di cybersecurity avanzati per rilevare compromissioni che potrebbero alterare dati o processi, valutare tecnologie di watermarking digitale o blockchain per tracciare la provenienza di contenuti sensibili, monitorare e utilizzare, ove possibile, strumenti per il rilevamento di contenuti generati da IA (deepfake detection);
- **definire policy e procedure chiare,** stabilendo linee guida interne sulla creazione, gestione, condivisione e conservazione sicura delle informazioni, sviluppare piani di **incident response** che includano scenari di compromissione dell’autenticità (es. diffusione di fake news sull’azienda, alterazione di database critici), integrare i requisiti della NIS 2 e del Cybersecurity Act nei processi aziendali;
- **promuovere una cultura della verifica,** attraverso la formazione del personale sul riconoscimento delle minacce (phishing, social engineering, fake news), incoraggiare il pensiero critico e la verifica delle fonti prima di condividere informazioni;
- **curare l’archiviazione digitale** con l’implementazione di sistemi di gestione documentale e conservazione digitale a norma, definire metadati accurati e completi per garantire contesto e reperibilità, pianificare la conservazione a lungo termine, considerando l’obsolescenza tecnologica.

Il paradosso tra conservare l’autentico e gestire l’artificiale verosimile esclude soluzioni binarie.

> Escludere i contenuti AI precluderebbe utili strumenti analitici, mentre la loro indiscriminata inclusione comprometterebbe la fiducia nell’archivio come garante della verità fattuale. Superare tale dicotomia impone un impegno sinergico tra legislatori, enti di standardizzazione e professionisti dell’informazione. Archiviare sia dati autentici che sintetici esige consapevolezza e rigore, trasformando gli archivi digitali da sterili depositi o musei di finzioni in ambienti dinamici dove ogni dato è univocamente identificato come “reale” o “artificiale”, con motivazione esplicita. Affrontare questa sfida richiede un cambio di paradigma verso un approccio multi-livello.

Occorre **potenziare la conservazione dell’autentico** attraverso il perfezionamento di tecniche e normative atte a garantirne autenticità, integrità e provenienza certa, incrementando la resilienza alla falsificazione. Parallelamente, **è necessaria una conservazione selettiva e qualificata dell’artificiale verosimile**, definendo protocolli chiari per l’archiviazione intenzionale di contenuti

generati da AI, scelti per specifiche finalità di ricerca, analisi forense e documentazione storica.

NOTE

[1] L’OAIS – approvato come standard ISO 14721:2003 – è un archivio inteso come struttura organizzata di persone e sistemi, la cui responsabilità è conservare nel lungo periodo, inteso come tempo abbastanza ampio da essere interessato da cambiamenti tecnologici, i documenti e i relativi metadati al fine di renderli disponibili ad una comunità di riferimento.