

La posta elettronica negli archivi di persona: conservazione e accesso*

ANNA ROVELLA**

ABSTRACT: The increasingly widespread use of email both as tool for personal communication and work, requires answers about the access to and the digital preservation of these kind of digital objects. This paper aims to analyze some critical aspects related to the long-term digital preservation of email archives in personal papers.

Keywords: Email, Digital preservation, Personal records, Lifecycle emails, Knowledge extraction.

1. Introduzione

Gli archivi di persona, la loro gestione e conservazione rappresentano un tema articolato e complesso che da diversi anni anima la discussione negli ambienti archivistici. La caratterizzazione che tali complessi documentari assumono nel corso delle loro formazione sedimentazione e trasmissione unitamente all'elevata possibilità di smembramento, di eccesso di selezione e di perdita di materiale incoraggiano la ricerca incessante di metodologie archivistiche adeguate¹. Il quadro di precarietà degli archivi di persona, si complica ancor più per gli archivi prodotti nel Ventunesimo secolo a causa delle criticità connesse

* Il presente lavoro è la rielaborazione di un intervento tenuto nel “Convegno Personal Digital Memories: i fondi di persona dall’analogico al digitale”, Ravenna 14 e 15 dicembre 2021.

** Dipartimento di Culture, Educazione e Società, Università della Calabria, Rende (CS), Italia. anna.rovella@unical.it.

¹ Antonio Romiti (2011, 16) afferma «nel periodo postunitario, si intensificarono le vicende disgregative, senza dubbio sempre più rapide, che passarono dalle episodiche a procedure di sistematicità» (Guarasci e Pasceri 2011). Secondo Romiti (2011, 14) lo smembramento della documentazione privata «ha rappresentato una preoccupazione da parte del legislatore italiano il quale, così come aveva già fatto nel testo unico dei Beni Culturali del 1999, si è preoccupato di proteggere il materiale anche nel Codice dei Beni culturali del 2004» (Guarasci e Pasceri 2011). Per una rassegna di archivi di persona e degli istituti di conservazione si consiglia la lettura di Barrera (2006, 617-57).

al crescente utilizzo di documenti nativi informatici. La posta elettronica, ad esempio, testimonia come la transizione dall'analogico al digitale possa realizzarsi rapidamente allorquando entrino in gioco tecnologie ubique e pervasive. Le lettere, i carteggi, forse più che altre forme di documenti consentono di rinvenire le relazioni che il soggetto produttore ha nel tempo intrattenuto con il mondo esterno. Esse costituiscono uno straordinario strumento per collocare il personaggio nel suo contesto, per comprenderne il rilievo politico, sociale, istituzionale e culturale, per delineare una sorta di via d'accesso privilegiata al pensiero e all'indole del produttore, utile a svelare lati della personalità e del carattere ma anche a fornire una chiave di lettura di aspetti organizzativi e/o contenutistici dell'archivio. Se la corrispondenza ha un forte impatto sulla memoria del soggetto produttore e presenta un notevole potenziale informativo giova ricordare che essa è anche la più fragile delle componenti di un archivio di persona. Basti pensare alle diverse implicazioni che spingono, i produttori o gli eredi (biologici, culturali ecc.) ad operare, ancor prima del versamento, forme di selezione, di controllo e di frammentazione della documentazione da sottoporre a conservazione (Navarrini 2005). A volte tali scelte sono dettate dalla volontà di tutela delle vicende personali del soggetto produttore, non di rado strettamente intrecciate anche con quelle di altri membri della famiglia o a soggetti terzi, oppure vi è l'esigenza di una trasmissione selettiva della memoria finalizzata, ad esempio, alla diffusione di una determinata immagine del produttore. In questi casi la scelta deliberata di non lasciare traccia di documenti connessi con quanto viene a vario titolo considerato molto riservato, comporta un depauperamento dell'archivio e della trasmissione di conoscenza. Si tratta di interventi che possono giungere alla distruzione parziale o totale del complesso documentario a causa di una scarsa sensibilità verso la conservazione del patrimonio o di una mancata consapevolezza della fragilità dei documenti, come avviene, ad esempio, nel caso dei messaggi di posta elettronica che in assenza di idonei processi preservativi possono andare inesorabilmente perduti.

Su un piano molto generale si può affermare che la trasmissione del corpus delle email di un archivio di persona è una procedura complessa e problematica per la presenza di molteplici fattori che la letteratura ha, in parte, evidenziato a più riprese. Così malgrado la conservazione della posta elettronica rappresenti un aspetto molto dibattuto si può affermare che soluzioni definitive e sistematiche sono ancora certamente lontane da realizzarsi.

L'obiettivo del presente contributo è quello di apporre alcune sintetiche note a margine di un tema vasto e certamente complesso con la finalità di focalizzare elementi significativi per il processo di conservazione.

2. Considerazioni sulla natura delle email

A poco meno di un decennio dall'avvio dei primi processi organici di conservazione dei documenti informatici è possibile constatare che pratiche consapevoli e coerenti di creazione e di gestione dei documenti operate durante la prima parte del loro ciclo di vita condizionano positivamente la qualità del materiale conservato e la possibilità di utilizzarlo nel tempo, salvaguardando quelle garanzie di completezza e di evidenza probatoria che normalmente si richiedono alle fonti archivistiche. Se le politiche, le regole e i sistemi di gestione documentale rappresentano un assunto fondamentale per gli archivi correnti delle pubbliche amministrazioni e delle imprese – pur nella consapevolezza delle numerose difficoltà di un'applicazione compiuta e capillare in tal senso – nel caso della documentazione prodotta da persone fisiche le possibilità di intervento *ex ante* si riducono a sporadici esempi non esaustivamente rappresentativi di un universo molto composito. Una siffatta situazione di partenza pone un primo importante pregiudizio in tema di conservazione dei messaggi di posta elettronica negli archivi di persona. In molti casi si arriva addirittura a non aver percezione alcuna dell'esistenza delle email e della necessità di tutela di tale tipo di documentazione. Non capita poi tanto raramente che gli archivi di persona contemporanei versati a istituzioni culturali di vario genere, non presentino corpus di email al loro interno².

Stessa sorte tocca alla messaggistica istantanea e, non di rado, anche alle fotografie e ai video presenti nei dispositivi mobile (Allegrezza 2022). Altre volte il deposito di tali documenti è reso impossibile dalla perdita accidentale dei dati che il produttore o gli eredi hanno subito per cause legate alla tecnologia, alla sicurezza, o più semplicemente ad un fattore culturale. Così una parte importante dei documenti di personalità che hanno avuto un ruolo nella società del Ventunesimo secolo viene perduta definitivamente senza che produttori, donatori o archivisti ne accusino neppure il danno.

Nel nostro Paese la letteratura archivistica in materia di conservazione di email nei fondi di persona è ancora molto esigua a testimonianza della scarsa percezione generale del problema (Allegrezza e Gorgolini 2016; Carassi 2018; Weston et al. 2019; AIB 2019). Per meglio comprendere le ragioni della poca attenzione archivistica verso le email può essere utile, in primis, formulare alcune considerazioni sulla natura e sul ciclo di vita di questo particolare oggetto digitale.

In inglese come in italiano l'uso corrente metonimico del termine email può indicare il vettore o il messaggio e talvolta anche l'account. La polisemia di significato viene evidenziata da Richard Pearce-Moses, in *A Glossary of Archival & Records Terminology*, che definisce le email (electronic mail) come «1. An asynchronous message, especially one following the RFC 2822 or MIME

² Stessa sorte tocca alla messaggistica istantanea e, non di rado, anche alle fotografie e ai video presenti nei dispositivi mobile (Allegrezza 2022).

standards, including a header, body, and - optionally - attachments, sent via a computer network held in online accounts to be read or downloaded by the recipients. 2. A system for transmitting email» (Pearce-Moses 2005). Anche per Tullio De Mauro la e-mail è «posta elettronica o messaggio inviato tramite posta elettronica» (De Mauro 2016). Solo apparentemente, sotto il profilo archivistico, il messaggio rappresenta l'oggetto da gestire, conservare e rendere fruibile nel tempo. Le implicazioni probatorie³ e informative dell'account, della relazione tra account e messaggio, dell'eventuale documento allegato, delle informazioni di processo di trasmissione, ecc. costituiscono elementi fondamentali la cui indubbia ricaduta sull'affidabilità del messaggio obbliga anche alla scelta di specifiche strategie conservative. La posta elettronica al pari di altri oggetti digitali complessi (ad esempio i database o le piattaforme di gestione documentale (Landino e Marzotti 2018)) rimanda a sfide di digital preservation metodologicamente e tecnologicamente più articolate rispetto alla semplice conservazione di file PDF e metadati descrittivi⁴. In tali sistemi è necessario conservare gli oggetti, le relazioni, le informazioni relative ai flussi e ai processi⁵. La completezza informativa che il digitale pone come condizione necessaria al recupero e all'uso successivo dei documenti richiede un approccio di conservazione di archivi di posta elettronica, di account, piuttosto che di soli messaggi. Ecco allora una prima importante criticità ma anche una sfida verso una nuova pratica documentale che obbliga archivisti e conservatori di

³ Sul valore probatorio delle email l'art. 2712 Codice Civile recita «Le riproduzioni fotografiche, informatiche o cinematografiche, le registrazioni fonografiche e, in genere, ogni altra rappresentazione meccanica di fatti e di cose formano piena prova dei fatti e delle cose rappresentate, se colui contro il quale sono prodotte non ne disconosce la conformità ai fatti o alle cose medesime». A precisazione degli elementi stabiliti nel citato articolo del Codice Civile la I sez. civile della Corte di Cassazione con l'ordinanza n. 19155 del 17 luglio 2019 afferma: «il messaggio di posta elettronica (cd. e-mail) costituisce un documento elettronico che contiene la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti che, seppure privo di firma, rientra tra le riproduzioni informatiche e le rappresentazioni meccaniche di cui all'art. 2712 c.c. e, pertanto, forma piena prova dei fatti e delle cose rappresentate se colui contro il quale viene prodotto non ne disconosca la conformità ai fatti o alle cose medesime. ... 1) l'e-mail rappresenta senza dubbio un documento informatico, nell'accezione fornita dall'articolo, 1, comma 1, lettera b), del D.P.R. 445/2000, a mente del quale per documento informatico si intende la rappresentazione informatica di atti, fatti o dati giuridicamente rilevanti; 2) in particolare, l'e-mail costituisce documento informatico sottoscritto con firma elettronica "leggera", "in quanto il mittente, per poter creare ed inviare detta mail, deve eseguire un'operazione di validazione inserendo il proprio username e la propria password"; 3) tale processo di validazione è da considerare equivalente alla firma elettronica leggera, così come definita in precedenza».

⁴ «Therefore, any programmatic attempt to preserve email must begin not only with an understanding of the specific technologies used in an email ecosystem, but a detailed knowledge of how server administrators and end users have configured software and hardware» (Prom 2011, 10).

⁵ Sulla natura delle email e la complessità ad esse associata (Pontevolpe e Salza 2009).

gitali italiani ad una rilettura del modello OAIS (ISO 2012), dei pacchetti informativi (di versamento e di archiviazione e di distribuzione), dello standard UNI SInCRO (UNI 2020), degli schemi nazionali di metadati e soprattutto delle strategie di interoperabilità degli oggetti della conservazione⁶.

Un altro elemento di difficoltà nella conservazione di archivi di posta elettronica è rappresentato dai diversi account che un soggetto produttore può possedere ed utilizzare. Si va dalla posta ordinaria a quella certificata⁷, quest'ultima largamente utilizzata, ad esempio, dai professionisti. L'urgenza della conservazione della PEC è certamente più sentita visto il carattere di recapito certificato e di opponibilità a terzi che la contraddistingue. Non sempre gli eredi che versano i documenti sono al corrente dei diversi account o hanno possibilità di accedervi o decidono di versare tutti gli account di posta del produttore ad un unico soggetto conservatore. Se lo scenario di frammentazione degli account di posta elettronica di un soggetto produttore costituisce, ai fini della conservazione, un potenziale ulteriore rischio di perdita di documenti e di difficoltà di visione di insieme è anche vero che gli archivi di email di personalità stanno assumendo progressivamente una loro dimensione convergente con un ruolo informativo anche all'interno di raccolte di altri oggetti culturali. Gli archivi, i documenti non sono avulsi dal tempo in cui si producono, dal loro contesto e inevitabilmente gli archivi contemporanei vivono una forte sollecitazione verso l'interconnessione, anelano ad esprimersi in una rete della conoscenza che per esplicitarsi deve necessariamente trovare il suo fondamento nella rappresentazione coerente degli oggetti e nell'interoperabilità di metadati e informazioni. In tale ottica il digitale manifesta le sue potenzialità, la sua capacità di restituire una nuova ricchezza conoscitiva multidimensionale, che i sistemi interculturali (musei, archivi e biblioteche) provano a teorizzare con risultati applicativi ancora non del tutto incoraggianti (Valacchi 2015, 54-7, 237-41; Carassi 2018, 11). Fin quando la conservazione digitale resterà, in Ita-

⁶ Ad oggi la conservazione di interi archivi di email pone criticità metodologiche e tecnologiche. Le sfide aperte sono principalmente riconducibili alla formazione di pacchetti di versamento esaustivi anche relativamente a metadati di processo, o all'utilizzo di tecnologie di emulazione o di migrazione sicure e performanti in termini di quantità e qualità dei dati. Per una panoramica dei diversi progetti, tool e scenari di workflow dalla donazione alla conservazione si consiglia la lettura di CLIR (2018).

⁷ In Italia i dati AgID (Agenzia per l'Italia Digitale) della Posta elettronica certificata nell'anno 2021 indicano circa 14 milioni di caselle PEC attivate e 2 miliardi e mezzo di messaggi scambiati. <https://www.agid.gov.it/it/piattaforme/posta-elettronica-certificata/statistiche-utilizzo-pec> (ultima consultazione: 26/05/2022). Ovviamente i dati AgID sulle PEC sono in parte rilevante riferibili alla pubblica amministrazione ma, certamente, vi è un numero non trascurabile di professionisti che fa largo uso di questo strumento per lo svolgimento delle attività lavorative. Al trend di crescita della produzione di documenti digitali si contrappone un dato preoccupante di documenti conservati digitalmente per i quali si registra una battuta di arresto nel 2021 con numeri stazionari rispetto ai due anni precedenti (Dati AssoConservatori 2019–2021, Stati generali della Conservazione, Università della Calabria).

lia, confinata ad una prospettiva a medio termine, senza la modellazione di un quadro evoluto, complessivo e condiviso di *long term digital preservation*, ogni tentativo di sistematizzazione rischia di rimanere confinato a definiti domini e specifiche casistiche.

Quanto evidenziato mette in luce anche l'urgenza di una rilettura, secondo una logica più inclusiva, della conservazione digitale che è chiamata a superare il mero concetto di adempimento riferito quasi esclusivamente a tipologie documentali ben determinate.

3. Considerazioni sul ciclo di vita delle email in un archivio di persona

La posta elettronica è un servizio a diffusione molto ampia a causa del continuo utilizzo nella sfera privata e nell'ambito lavorativo. L'essenzialità di tale servizio unitamente alle caratteristiche di ubiquità e pervasività della tecnologia di base rimandano ad una produzione estremamente copiosa di messaggi⁸. I numeri incrementali che caratterizzano il fenomeno delle email⁹ sollecitano un approccio sistematico e consapevole nella scelta di strategie di sicurezza, di tecnologia, di selezione, di recupero, di accesso, di rappresentazione delle informazioni, di mantenimento di evidenze documentali e giuridiche. Ciò anche in ragione del ciclo di vita delle email le cui fasi e relative attività sono condizionate anche dalla quantità e dalla struttura dei messaggi presenti in ogni account. Definire e tracciare il ciclo di vita delle email, in presenza di così tante variabili, non è un'operazione semplice soprattutto negli archivi di persona, nei quali il processo di conservazione subentra in una fase in cui i messaggi sono stati formati, gestiti e consolidati con modalità che non neces-

⁸ Malgrado il dilagare dei social e dei sistemi di messaggistica l'uso della posta elettronica si conferma quale importante sistema di comunicazione con numeri importanti: 306 miliardi di email vengono scambiate ogni giorno con un totale annuo di oltre 101.000 miliardi di messaggi di posta elettronica. Un settore in crescita che vede stime per il 2025 con più di 376 miliardi di email al giorno e che coinvolge oltre 4 miliardi di persone in tutto il mondo utilizzano email, di cui la metà circa tramite smartphone. <https://www.key4biz.it/50-anni-fa-la-prima-email-oggi-se-ne-inviano-piu-di-300-miliardi-al-giorno-mercato-in-mano-alle-big-tech/352180/> (ultima consultazione: 26/05/2022).

⁹ «Recent trends show that email usage is growing and becoming more deeply embedded in our work and personal lives, blurring the line between business and personal communication[...] Taken as a whole, these trends present an archival paradox. Email and other electronic messages are both ubiquitous and ephemeral, documenting people's professional and personal lives in a chaotic stream of messages and relationships, calling for a new documentary practice[...] But in comparison, email preservation has taken a backseat, with few grant-funded projects and relatively little institutional support[...] Each of these works place attention on the entire range of cultural, legal, ethical, professional and technical considerations that must be addressed if an organization wishes to identify email of permanent value, preserve it in an authentic form, and render it for future use» (Prom 2011, 4).

sariamente permettono una successiva ricostruzione fedele della formazione e sedimentazione. Tuttavia, la definizione, sia pur generica, delle fasi del ciclo di vita delle email aiuta ad individuare attività, politiche, responsabilità e processi che si rivelano fondamentali per la conservazione. Provando a schematizzare il ciclo di vita delle email si possono individuare le seguenti fasi: formazione e uso, trattamento preconservativo, conservazione, recupero e accesso (Fig. 1).



Figura 1: Fasi del ciclo di vita delle email.

Formazione e uso

Negli archivi di persona è difficile e talvolta impossibile esercitare un controllo della posta elettronica fin dal momento della formazione¹⁰. Tuttavia, come accade per tutti i documenti informatici, anche nel caso delle email, la fase di formazione rappresenta un processo delicato e fondamentale nel quale si determinano le condizioni essenziali e le informazioni necessarie per le successive attività di gestione e di conservazione. La scelta delle tecnologie, le azioni in chiave di sicurezza, la gestione dell'account anche mediante applicativi dedicati installati in locale, l'organizzazione dei messaggi e la loro tutela predeterminano, quindi, il corso del ciclo di vita delle email inclusa la loro conservazione e le garanzie di recupero e validità nel tempo. Gli archivisti nord americani e inglesi, consapevoli di tale difficoltà, hanno da tempo proposto una forma di salvaguardia più ampia della posta elettronica negli archivi di persona, che rimanda anche all'instaurazione di un rapporto tra archivistica e soggetto produttore da stabilire fin dalle fasi di sedimentazione e selezione del materiale¹¹. La possibilità di raccogliere direttamente dal soggetto produttore informazioni utili alla scelta e alla conservazione dei messaggi di posta è certamente un'opportunità rilevante che consente di sviluppare un progetto condi-

¹⁰ Il controllo della formazione degli account, dei messaggi email e in generale dell'uso della posta elettronica nel caso di pubbliche amministrazioni e o di grandi aziende può essere risolto mediante la definizione di policy preventive e regolamenti specifici che, ad esempio, possono ricondurre la gestione della posta all'interno della gestione documentale con la possibilità di tracciare, di recuperare e di sottoporre a conservazione almeno le comunicazioni connesse a processi lavorativi.

¹¹ Tale pratica va affermandosi anche in Italia. "Ecco allora la prima riflessione, ovvia ma non scontata: quando è possibile, è meglio collaborare con gli esperti della persona di cui si riordina l'archivio fin dall'inizio dell'intervento" (Tealdi 2018, 7-9).

viso di conservazione¹². Tuttavia, sebbene non sia infrequente che i produttori decidano di menzionare nel proprio testamento l'archivio è assai più raro che essi si rendano disponibili, in vita, ad una collaborazione con gli archivisti. Probabilmente i tempi non sono maturi per una sensibilità ed una cultura di tal genere. L'approccio culturale ha un riflesso anche sulle procedure di selezione dei messaggi da sottoporre a conservazione. Un passaggio che si rende necessario in ragione del proliferare delle comunicazioni pubblicitarie, delle copie generate dall'inoltro di messaggi da un account di posta ad un altro di un medesimo produttore o dalla ricezione di materiale di nessuna utilità ai fini della conservazione. Tali messaggi dovrebbero essere periodicamente eliminati manualmente o attraverso il settaggio di regole di selezione automatica o funzioni analoghe normalmente presenti negli applicativi di posta elettronica. Tali tecniche possono supportare anche una gestione più sicura della posta filtrando, ai fini dell'eliminazione, messaggi di spam, phishing, malware ecc. eventualmente non intercettati dal provider.

Il soggetto produttore dovrebbe, dunque, operare uno sfoltoimento periodico della posta al fine di eliminare gli elementi superflui o addirittura pericolosi. Un'attività di selezione, più analitica e metodica dovrebbe essere poi praticata almeno alla fine di ogni anno, ad esempio, prima della rituale ma necessaria archiviazione dei messaggi di posta prodotti e ricevuti. Il processo dell'archiviazione della posta è certamente una forma di tutela dei messaggi e diventa ancora più efficace se, nel corso dell'anno i messaggi vengono via via classificati in apposite cartelle secondo un particolare ordine ad esempio per mittente, o argomento. Tale classificazione ulteriore, in genere, facilita il successivo recupero della posta ricercata. In sintesi una buona classificazione della posta elettronica potrebbe prevedere un'ordine cronologico annuale e un ordine per materia (mittente, argomento, ecc.). Un modello che richiama metodologie consolidate di classificazione della corrispondenza analogica in cui unità archivistiche cronologicamente ordinate presentano sottodivisioni funzionali organizzate in sottofascicoli per mittente. In ambiente digitale, anche la sola archiviazione per argomenti o corrispondenti potrebbe risultare efficace evitando eventualmente la frammentazione degli affari o degli argomenti su base annuale. La preventiva classificazione e archiviazione della posta elettronica ad opera del soggetto produttore produce un vantaggio nella successiva fase di conservazione e valorizzazione della stessa soprattutto, quando il soggetto produttore, ha anche la lungimiranza di predisporre strumenti atti a documentare la metodologia di classificazione definita ed applicata e a raccordare la posta elettronica con le altre componenti dell'archivio digitale e/o analogico¹³.

¹² <http://www.ariadne.ac.uk/issue/55/williams-et-al/> (ultima consultazione: 26/05/2022).

¹³ Antonio Romiti commentando il saggio di Stefano Moscadelli "L'Archivio di Fabrizio De Andrè: osservazioni a volo d'uccello", afferma «In sostanza si nota come non sia stato possibile reperire tra le carte un vincolo naturale, evidenziando così una situazione che mostrava

Anche nel caso della tutela dei dati personali un approccio *risk based* mediante classificazione, metadattazione, consapevole valutazione e selezione del materiale da sottoporre a conservazione evita situazioni di *data breach*.

Trattamento preconservativo

La natura di documento informatico dei messaggi di posta elettronica richiama l'urgenza della conservazione al fine di evitare situazioni di perdita, di mancato recupero o di impossibilità di lettura di tali oggetti. Da un punto di vista teorico il processo di conservazione digitale della posta elettronica deve, quindi, essere reso quanto più possibile vicino al momento della formazione dei messaggi, allegati inclusi. Negli archivi di persona tali procedure incontrano ostacoli di varia natura (culturale, tecnologica, economica ecc.) e il rinvio della conservazione diventa una prassi normale se non addirittura scontata. Conseguentemente la fase di trattamento preconservativo, avviene spesso a distanza di molti anni dalla produzione dei messaggi generando scenari molto diversificati. In questa fase del ciclo di vita delle email è necessario verificare l'organicità e la congruenza del materiale da conservare, così come è fondamentale definire metodologie e implementare strumenti che possano supportare una conservazione coerente e consistente. Come già evidenziato le situazioni che possono presentarsi all'archivista a seguito di un processo di donazione o di versamento di un archivio di posta elettronica possono essere molto eterogenee e, in un elevato numero di casi, gli attori coinvolti sono diversi dal soggetto produttore (eredi, archivista ecc.). L'archivista è chiamato, così, a conferire sistematizzazione al materiale da conservare, in situazioni in cui la mancanza di evidenze documentali o informative sufficienti difficilmente permette, malgrado l'impegno meticoloso e responsabile, di restituire archivi di posta completi, documentati e affidabili. La letteratura scientifica di riferimento ritorna più volte su questo aspetto sottolineando la solitudine dell'archivista di fronte a scelte strategiche o di responsabilità quali ad esempio la selezione del materiale, la tutela dei dati personali e delle informazioni riservate, l'autenticità di quanto conservato, in un contesto interdisciplinare in cui mancano modelli stabili di riferimento. Si può affermare che per le criticità citate la fase di trattamento preconservativo è strettamente connessa con processi di qualità in cui la valutazione del rischio, la definizione delle responsabilità e la gestione dei dati assumono particolare rilievo. L'esito positivo di tale fase presuppone un lavoro di squadra in cui sono chiamate in causa diverse

il ricorso ad antecedenti operazioni mirate a compiere consistenti operazioni di scarto. Questi comportamenti sono frequenti nella gestione delle carte degli archivi personali e spesso sono effettuati da persone che non hanno la minima idea del significato dell'operazione che stanno conducendo» (Guarasci, e Pasceri 2011).

professionalità e competenze (documentali, giuridiche, informatiche, storiche, di dominio ecc.).

La complessa e intricata vicenda della posta elettronica di Hilary Clinton durante il suo mandato di Segretario di Stato degli Stati Uniti d'America rappresenta un esempio delle problematiche legate alla gestione e conservazione della posta elettronica anche in ragione del complicato intreccio tra le vicende private e professionali di un soggetto produttore (Salahub and Oldford 2018; Salahub and Oldford 2017; Vasundharaand and Mueller 2016).

La presenza di messaggi privati negli account istituzionali o aziendali è, ad esempio, un tema che ritorna frequentemente negli archivi di posta elettronica. La commistione di documentazione pubblico/privata è abbastanza usuale negli archivi di persona specie se il soggetto produttore ha ricoperto incarichi pubblici e istituzionali. Non di rado la documentazione pubblica confluisce nel corpus dell'archivio privato o viceversa i documenti privati presenti negli account istituzionali vengono eliminati senza che il soggetto produttore o gli eredi ne siano informati (Rovella et al. 2020). Tali fenomeni hanno assunto una dimensione ancor più significativa nel periodo della pandemia, durante il quale a seguito del lockdown e della diffusione dello smartworking i dispositivi personali hanno ospitato sempre più frequentemente anche documentazione istituzionale o aziendale.

La costruzione di un buon team con esperienze e competenze diverse produce ricadute positive in termini di organicità, affidabilità e futuro recupero di quanto potrà essere conservato anche a vantaggio della salvaguardia della documentazione e del rispetto della tutela dei dati riservati (personali, sensibili ecc.) presenti nella posta elettronica. Significativi risultati, in tale contesto, arrivano, ad esempio, dall'applicazione di tecniche di Natural Languages Processing (NLP)¹⁴ utilizzate per la modellazione degli oggetti, all'interno delle collezioni, con la finalità di identificare documenti che richiedono restrizioni circa i dati personali o sensibili (Hutchinson 2018, 2251–55).

Dal punto di vista della cybersecurity le email sono particolarmente vulnerabili e anche questo è un aspetto rilevante ai fini della loro conservazione. Una risposta interessante arriva dall'applicazione di tecniche di classificazione automatica delle email (Rovella et al. 2020). I provider di posta elettronica, utilizzano largamente tali strumenti per identificare, ad esempio, i messaggi di spam, phishing ecc. La classificazione automatica della posta può essere utilizzata per dare ordine ai messaggi fin dalla loro formazione permettendo una sedimentazione organizzata per tipologie di messaggi o per contenuto con un impatto positivo sulla consultazione dei contenuti dei messaggi. L'ingente e costante investimento nella costruzione di strumenti per la sicurezza delle mail ha permesso di sviluppare tool e algoritmi sempre più sofisticati in grado

¹⁴ «Natural Language Processing (NLP) is the application of computational methods to model and extract information from human language» (Kamath, Liu and Whitaker 2019).

di lavorare sul piano funzionale e su quello semantico. Tali sistemi partendo da un corpus di email possono associare automaticamente i messaggi a classi pre-determinate o generate dinamicamente. Probabilmente su tale tema potrebbe essere aperta una riflessione con il coinvolgimento degli attori istituzionali non solo italiani ma anche dell'Unione Europea. È evidente che Google, Microsoft e Apple, proprietari dei primi 10 client di posta e, di fatto, capaci di controllare il 90% del mercato delle email, non avvertano nessuna pressione della domanda di strumenti a supporto della gestione dei messaggi di posta elettronica. Una concertazione europea, sui ruoli, sulle regole e sulle metodologie di tutela delle mail nel dialogo tra istituzioni e grandi player potrebbe rappresentare un tema di interesse anche in una logica di cittadinanza digitale.

Gli elementi finora considerati sono rilevanti anche ai fini della creazione del pacchetto di versamento (SIP) (ISO 2012) da inviare al sistema di conservazione (AIP). Tutte le scelte operate in merito alle strategie di conservazione che vanno dagli accordi di donazione, alla definizione di schemi di metadati e dei formati tecnologici fino alla redazione del contratto di conservazione, avranno un ruolo molto importante nella composizione del pacchetto che dovrà presentare caratteristiche di coerenza e di interoperabilità, proprietà fondamentali anche per la formazione del successivo e conseguente Pacchetto di Archiviazione. Durante la fase di preconservazione sarà allora necessario predisporre un accurato piano di conservazione, un piano di gestione dei dati corredato da adeguati schemi di metadati, utili a descrivere e documentare le email anche in relazione ai processi di trasmissione, provenienza e agli elementi di autenticità (Dinnen and Kartalić 2020, 413–24). Il modello organizzativo/gestionale così elaborato costituirà una solida base per l'accordo di conservazione.

Conservazione, recupero e accesso ai messaggi

Da quanto finora affermato è facilmente desumibile che anche la fase di conservazione riserva una serie di criticità che richiamano una programmazione attenta di tutti i processi. Per tale ragione è fondamentale che nel sistema di conservazione risultino chiare le responsabilità dei diversi attori che contribuiscono alla creazione dei pacchetti informativi e alla gestione degli oggetti digitali. La formazione dell'AIP richiede al conservatore soluzioni che avranno una ricaduta decisiva sulla possibilità di estrazione dei messaggi e sull'autoconsistenza informativa espressa tramite i metadati.

I possibili scenari di conservazione delle email possono contemplare situazioni diversificate in cui il SIP può essere composto da singoli messaggi completi di metadati o dall'intero account di posta. Il focus si sposta allora sull'AIP che, a prescindere dallo scenario che ha guidato la formazione del SIP, dovrebbe ospitare le email sia in formato originale che in un formato normalizzato (generalmente eXtensible Markup Language – XML). I metadati, i log, l'oggetto

del messaggio, il mittente, gli allegati, le informazioni relative alla stabilità e l'integrità dovrebbero essere salvate in un database XML, o in una struttura XML flessibile da archiviare (come file) all'interno del pacchetto. La struttura dell'indice del pacchetto di archiviazione indicata nello standard UNI SInCRO (2020) permetterebbe di ospitare tutti gli oggetti, i metadati e le informazioni da conservare. I processi di *digital preservation*, gli agenti, gli eventi e le eventuali informazioni aggiuntive potranno essere documentate da metadati specifici in accordo con il modello PREMIS completando il set di metadati anche negli elementi che documentano la provenienza e l'autenticità degli oggetti digitali. Il risultato sarà un nuovo AIP, in cui sono presenti email e metadati ampliati. I pacchetti di archiviazione così creati, con la caratteristica di autoconsistenza potrebbero essere collocati in un repository certificato del tipo Trustworthy.

L'attenzione verso la completezza e l'autoconsistenza dell'AIP facilita la formazione di pacchetti di distribuzione coerenti favorendo l'accesso sicuro ai contenuti conservati nel sistema. Le architetture per l'interoperabilità e la cooperazione applicativa¹⁵ tra i sistemi di conservazione previste nel modello OAIS hanno effetto anche sull'accesso ai documenti conservati. Nel caso degli archivi a lunga conservazione sarebbe auspicabile, come propone Stefano Pigliapoco, «realizzare architetture miste, ad esempio, combinando la *Federated archives* con la *Shared resources archives* per conseguire il duplice obiettivo di ridurre i costi dell'impianto tecnologico e attivare forme avanzate di accesso ai contenuti digitali» (Pigliapoco 2016). In quest'ultimo aspetto assumono rilevanza i diversi sistemi di classificazione e estrazione automatica dei contenuti quali strumenti qualificanti nel processo di accesso alle email. Un discorso complesso e articolato che può chiamare in gioco tecnologie di Machine Learning e Deep Learning, in grado di fornire risultati su un piano generale di organizzazione ed estrazione di contenuti ma anche ad un livello molto dettagliato di analisi dei concetti, di estrazione di dati e metadati anche nel caso di testi che con grafici, tabelle (Rovella et al. 2022), immagini e audio. All'interno del messaggio è possibile trovare diversi elementi, talvolta anche sotto forma di link verso altre risorse, in tal caso molta attenzione deve essere prestata al formato digitale e alla sua capacità di autocontenimento. Negli ultimi anni è stata considerata la possibilità di utilizzare il formato FITS per la conservazione delle email e una sperimentazione condotta in tal senso potrebbe verificare l'eventuale l'efficacia di tale scelta (UNI 2022).

¹⁵ Le tre diverse architetture definite nello standard OAIS sono: *Independent archives*; *Cooperating archives*; *Federated archives*; *Shared resources archives* (ISO 2012).

4. Alcune considerazioni conclusive

Le poche note richiamate in questo lavoro rimandano ad una letteratura certamente ampia ma non definitiva sul tema della conservazione delle email negli archivi di persona. Malgrado i numerosi progetti internazionali¹⁶ al momento è comunque difficile tracciare un quadro di insieme (Standford Libraries. n.d; Schneider et al. 2017; Schneider et al. 2019). Manca uno sforzo ulteriore di sistematizzazione verso un approccio inclusivo delle diverse sfaccettature che connotano un tema complesso ed esteso che, a diversi livelli, evidenzia la necessità di un raccordo tra i molteplici attori del sistema in una logica di cooperazione e di accountability necessaria per la buona riuscita del processo di conservazione. Se tali considerazioni sono valide in un contesto internazionale, in Italia, vista anche la scollatura che si è andata generando tra la conservazione digitale e la conservazione digitale a lungo termine, si coglie la traccia di un allontanamento progressivo da soluzioni integrate e definitive.

Riferimenti bibliografici

- Allegrezza, Stefano, e Luca Gorgolini, a cura di, 2016. *Gli archivi di persona nell'era digitale. Il caso dell'archivio di Massimo Vannucci*. Bologna: il Mulino.
- Allegrezza, Stefano. 2022. *Recent developments on email preservation: towards the ultimate solution?* IRCDD. http://ircdl2022.dei.unipd.it/downloads/papers/IRCDDL_2022_paper_24.pdf.
- Associazione Italiana Biblioteche (AIB). 2019. *Linee guida sul trattamento dei fondi personali*. Milano. <https://www.aib.it/struttura/commissioni-e-gruppi/gbaut/strumenti-di-lavoro/linee-guida-sul-trattamento-dei-fondi-personali/>.
- Barrera, Giulia. 2006. *Archivi di Persona*. Vol. 2 di *Storia d'Italia nel secolo ventesimo. Strumenti e fonti*, a cura di Claudio Pavone. Roma: Ministero per i beni e le attività culturali, Dipartimento per i beni archivistici e librari, Direzione generale per gli archivi.
- Carassi, Marco, a cura di. 2018. "Sugli Archivi Di Persona: Esperienze a Confronto". *Il mondo degli archivi*. I Quaderni 5. http://www.ilmondodegliarchivi.org/images/Quaderni/MdA_Quaderni_n5.pdf.
- Council on Library and Information Resources (CLIR). 2018. *The future of email Archives. A Report from the Task Force on Technical Approaches for Email Archives*. Washington DC.

¹⁶ Per una panoramica sui principali progetti internazionali si consiglia la lettura di CLIR (2018).

- De Mauro, Tullio. 2016. *Nuovo vocabolario di base della lingua italiana*.
- Dinnen, Jesse David, and Maja Kartalić. 2020. "E-Mail as Legacy: Managing and Preserving E-Mail as a Collection." *Libraries and the Academy* 20 (3): 413-24. Doi:10.1353/pla.2020.0022.
- Guarasci, Roberto, e Erika Pasceri, a cura di. 2011. *Archivi privati: Studi in onore di Giorgetta Bonfiglio-Dosio*. Roma: Consiglio Nazionale delle Ricerche, SEGID.
- Kamath, Uday, John Liu, and James Whitaker. 2019. *Deep Learning for NLP and Speech Recognition*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-14596-5>.
- Hutchinson Tim. 2018. "Protecting privacy in the archives: Preliminary explorations of topic modeling for born-digital collections." In 2017 IEEE International Conference on Big Data (Big Data) 2251-55. <https://doi.org/10.1109/BigData.2017.8258177>.
- International Standard Organization (ISO) 14721:2012. *Space data and information transfer systems – Open archival information system (OAIS) – Reference model*.
- Landino, Costantino, e Pasqualina A. Marzotti. 2018. *Memorie dinamiche. La conservazione dei data base e il web archiving*. Roma: Edizioni ANAI.
- Navarrini, Roberto. 2005. *Gli Archivi Privati*, Lucca: Civita Editoriale.
- Pearce-Moses, Richard. 2005. *A Glossary of Archival & Records Terminology*, Chicago: SAA.
- Pigliapoco, Stefano. 2016. "Il modulo ACCESS del modello OAIS. Architetture e funzionalità avanzate." *AIDAinformazioni*, 34 (1-2): 49-59.
- Pontevolpe, Gianfranco, and Silvio Salza. 2009. "Keeping and Preserving Email report." In *Progetto InterPARES*. http://www.interpares.org/ip3/ip3_products.cfm?cat=10.
- Prom, Christopher J. 2011. *Preserving Email*. Digital Preservation Coalition. <http://dx.doi.org/10.7207/twr11-01>.
- Romiti, Antonio. 2011. "Gli archivi privati visti da più prospettive". In *Archivi privati: Studi in onore di Giorgetta Bonfiglio-Dosio*, a cura di Roberto Guarasci e Erika Pasceri, 7-28. Roma: Consiglio Nazionale delle Ricerche, SEGID.
- Rovella, Anna, Roberto Guarasci, e Francesca Parisi. 2020. "Gestione e conservazione dei messaggi di posta elettronica." *AIDAinformazioni*, (1-2):149-75.

- Rovella, Anna, Alexander Murzaku, Eugenio Cesario, Martin Critelli, Armando Bartucci, and Francesca M.C. Messiniti. 2022. "Analysis, evaluation and comparison of knowledge extraction tools in the Environmental and Health domain. A holistic approach." In *Knowledge Organization and Management in the domain of Environment and Earth Observation (KOMEEEO)*, *Proceedings of the International KOMEEEO Conference 2021*, edited by Antonietta Folino, and Roberto Guarasci. ERGON VERLAG.
- Salahub, Christopher, and Wayne Oldford. 2017. *Interactive Filter and Display of Hillary Clinton's Emails: A Cautionary Tale of Metadata*. <https://www.researchgate.net/publication/315876309>.
- Salahub, Christopher, and Wayne Oldford. 2018. "About 'Her Emails.'" *Significance* 15, no. 3: 34–37. <https://doi.org/10.1111/j.1740-9713.2018.01148.x>.
- Schneider, Josh, Peter Chan, Glynn Edwards e Sudheendra Hangal. 2017. "ePADD: Computational Analysis Software Facilitating Screening, Browsing, and Access for Historically and Culturally Valuable Email Collections". *D-Lib Magazine* 23, n. 5/6 (maggio). <https://doi.org/10.1045/may2017-schneider>.
- Schneider, John, Chance Adams, Sally De Bauche, Reid Echols, Callum Mckean, Jessica Moran, and Dorothy Waugh. 2019. "Appraising, Processing, and Providing Access to Email in Contemporary Literary Archives". *Archives and Manuscripts* 47(3):305-326 DOI:10.1080/01576895.2019.1622138.
- Tealdi, Elena. 2018. "Piero Martinetti, filosofo. (Pont Canavese 21 agosto 1872 – Cuorgnè, Torino 22 marzo 1943)". In *Sugli Archivi Di Persona: Esperienze a Confronto*, a cura di Marco Carassi, 7-9. *Il Mondo degli archivi*. I Quaderni 5. http://www.ilmondodegliarchivi.org/images/Quaderni/MdA_Quaderni_n5.pdf.
- UNI 11386:2020. *Supporto all'Interoperabilità nella Conservazione e nel Recupero degli Oggetti digitali (SInCRO)*.
- UNI 11845:2022. *Processi di gestione della conservazione a lungo termine di immagini digitali con l'uso del formato FITS*.
- Valacchi, Federico. 2015. *Diventare archivisti. Competenze tecniche di un mestiere sul confine*. Milano: Editrice bibliografica.
- Vasundharaan, Dehiya, and Mueller Klaus. 2016. "Analyzing Hillary Clinton's Emails." In *Poster Abstracts of IEEE VIS*, October 2016. <https://www3.cs.stonybrook.edu/~mueller/papers/Analyzing%20Hillary%20Clinton%20Emails%20VIS%202016.pdf>.
- Weston, Paul Gabriele, Primo Baldini, Emmanuela Carbé, e Laura Pusterla. 2019. "PAD - Pavia Archivi Digitali e gli archivi degli scrittori." In *Digitalia*, XIV(1).