



Università
Ca' Foscari
Venezia

Corso di Laurea Magistrale

in Storia e gestione del patrimonio archivistico e
bibliografico (Interateneo)

Tesi di Laurea

**Sistemi di archiviazione digitale nativo:
ricognizione di carattere logico, tecnologico e
procedurale**

Relatore

Ch. Prof. Paul Gabriele Weston

Correlatore

Ch. Prof. Riccardo Ridi

Laureanda

Yeshihareg La Ragione

Matricola 858976

Anno Accademico

2019/ 2020

“La mancanza di sensibilità e di cura filologica verso il digitale, la fragilità dei supporti e l’obsolescenza sia dell’hardware sia del software per la lettura dei contenuti aumentano il rischio di vuoti di memoria¹” (ALDiNa, 2020)

¹ <http://www.aiucd.it/aldina-archivi-letterari-digitali-nativi/>

Ringraziamenti

Prima di proseguire con la trattazione, vorrei riservare qualche riga a chi mi ha sostenuta nel mio percorso di ricerca e mi ha supportata nella concretizzazione di questo elaborato.

Desidero innanzitutto ringraziare il mio relatore prof. Weston per avermi seguita durante la stesura della tesi, per la sua immensa pazienza e per gli indispensabili consigli, correzioni e indicazioni fondamentali per la realizzazione di questo lavoro.

Un doveroso ringraziamento va poi alla dott.ssa Mancinelli e prof.ssa Tomasi per i loro suggerimenti, consigli sul tema della tesi.

Inoltre, ringrazio la mia famiglia che mi ha sempre sostenuta, supportata e ha reso possibile la realizzazione di questo importante risultato.

Infine, ringrazio tutti gli amici che mi hanno fatto sentire la loro vicinanza.

INDICE

ABSTRACT	1
----------	---

INTRODUZIONE	2
--------------	---

CAPITOLO I: Generalità

1.1 Conservazione digitale	4
----------------------------	---

1.2 I materiali “Born Digital”	10
--------------------------------	----

1.3 I Metadati	14
----------------	----

1.4 Il modello OAIS	19
---------------------	----

CAPITOLO II: Sistemi di archiviazione digitale

2.1 Sistemi di conservazione	24
------------------------------	----

2.2 Casi di studio italiani	28
-----------------------------	----

2.3 Regole, Procedure, Tecnologie	31
-----------------------------------	----

CAPITOLO III: Esperienze di conservazione del digitale nativo

3.1 Casi di studio	44
--------------------	----

3.2 Progetti	46
--------------	----

3.3 Archivi orali e audiovisivi	65
---------------------------------	----

IV CONCLUSIONI

BIBLIOGRAFIA	73
--------------	----

SITOGRAFIA	76
------------	----

ABSTRACT

Questo elaborato tratta in linea generale la conservazione digitale, partendo dalle sue diverse definizioni e giungendo alla definizione condivisa della conservazione digitale che è: “l’insieme dei principi, delle politiche, delle disposizioni e delle strategie finalizzate a prolungare l’esistenza di un documento o di una risorsa digitale grazie alla sua tenuta in condizioni adatte all’uso nella forma originale e/o in un formato persistente che garantisca l’integrità della configurazione logica e del contenuto²”.

In questo lavoro viene fatta una ricognizione di carattere logico-tecnologico e procedurale dei diversi sistemi di archiviazione digitale fino ad analizzare specifici sistemi di conservazione del digitale nativo. Saranno trattati il modello di riferimento (OAIS), gli standard, le Linee Guida, le buone pratiche e i manuali di conservazione per l’archiviazione del digitale nativo.

I progetti coinvolti in questa ricognizione sono PAD, Archivio Vi.Vo, AISO, Archivio Massimo Vannucci, ALDiNa, facendo un confronto con alcuni archivi personali internazionali.

Gli obiettivi della tesi sono: quello di individuare standard, Linee Guida, buone pratiche e Manuali per la conservazione del nativo digitale, partendo dallo standard di riferimento e quello di studiare i sistemi di archiviazione digitali internazionali e nazionali.

² https://e-l.unifi.it/pluginfile.php/227895/mod_resource/content/1/ConsevezioneDigitale_Guerrini_Giglio.pdf;
Standard ISO 14721:2012 – Open Archive Information System;
http://media.regesta.com/dm_0/ANAI/anaiCMS/ANAI/000/0188/ANAI.000.0188.0017.pdf.

INTRODUZIONE

Il sistema di archiviazione del digitale nativo elabora i materiali da preservare in pacchetti informativi che si suddividono in pacchetti di: versamento, archiviazione e distribuzione.

Lo scopo del sistema è tenere inalterato il valore giuridico, l'autenticità e permettere l'accesso al materiale digitale nel tempo. Inoltre, assicura integrità, affidabilità, leggibilità, reperibilità della risorsa, secondo le modalità indicate nelle linee guida, negli standard internazionali e nelle buone pratiche. Gli elementi di conservazione qui in oggetto sono documenti nati digitali, dematerializzati e digitalizzati. Gli oggetti nativi digitali, oltre ad essere di facile diffusione grazie ad Internet, presentano molte sfide nel piano di archiviazione.

Preservare le memorie collettive e personali è difficoltoso, poiché il sistema di conservazione digitale è soggetto a danneggiamento ed evoluzione, quindi bisogna mettere in campo diverse competenze e prendere in considerazione l'obsolescenza della tecnologia.

In questo elaborato si fa una ricognizione logica, tecnologica e procedurale dei sistemi di archiviazione del digitale nativo per tracciare un quadro sui sistemi operativi, applicativi e gli strumenti esistenti in Italia.

Obiettivo primario di questo elaborato è stata la ricerca di sistemi di archiviazione del nativo digitale nazionale e internazionale. Ed attraverso la disponibilità dei sistemi di conservazione si vuole tentare di dimostrare che l'archiviazione letteraria del nativo digitale italiana è molto indietro rispetto alle altre nazioni dove si possono trovare con maggior facilità linee guide, buone pratiche, consigli e raccomandazioni ben sviluppate e documentate.

La tesi parte da una considerazione di fondo sul concetto di conservazione, archiviazione, preservazione digitale delle informazioni nel lungo periodo e prosegue con una spiegazione di cosa sia il materiale digitalizzato e il materiale nativo digitale.

Inoltre, a partire dall'analisi del modello di riferimento OAIS, si fornisce una definizione e descrizione dei metadati.

Nella seconda parte dell'elaborato vengono introdotte le regole, le norme, le procedure, i manuali, le tecnologie e le strategie dei sistemi operativi, divengono quindi tema centrale

dell'elaborato i sistemi di conservazione digitali con esempi sia nazionali che internazionali, quindi viene fatto un focus tra le normative italiane e gli standard internazionali.

Inoltre, viene fornito uno schema in cui sono elencati i formati di riferimento importanti per l'archiviazione delle risorse nativo digitali.

Infine, nel terzo capitolo, vengono trattati i casi di studio riguardanti le esperienze di conservazione del digitale nativo letterario. Vengono descritti i progetti come PARADIGM, PAD, ALDiNa, Phaidra, web DOMAIN.UK, che possono considerarsi progetti pilota per la preservazione. Di questi vengono anche affrontati le linee guida, i consigli, le raccomandazioni, i formati, gli standard etc.

CAPITOLO I: Generalità

1.1 Conservazione digitale

“Digital preservation refers to the management of digital information over time. Unlike the preservation of paper or microfilm, the preservation of digital information demands ongoing attention. This constant input of effort, time, and money to handle rapid technological and organisational advance is considered the main stumbling block for preserving digital information beyond a couple of years³. Indeed, while we are still able to read our written heritage from several thousand years ago, the digital information created merely a decade ago is in serious danger of being lost⁴.

Digital preservation can therefore be seen as the set of processes and activities that ensure the continued access to information and all kinds of records, scientific and cultural heritage existing in digital formats⁵.

Digital preservation is defined as: long-term, error-free storage of digital information, with means for retrieval and interpretation, for all the time span that the information is required for. “Retrieval” means obtaining needed digital files from the long-term, error-free digital storage, without possibility of corrupting the continued error-free storage of the digital files⁶”.

“Basic digital preservation is the maintenance of digital material over the long-term with a view to ensuring continued accessibility. Digital material refers to any material processed by a computer and includes both that which is “digitised” (reformatted to digital) as well as those resources that are “born digital”. Long-term in this context should be taken to mean long enough to be concerned with the impact of changing technologies- and should include timescales of decades and even centuries⁷.

Digital preservation is the series of actions and interventions required to ensure continued and reliable access to authentic digital objects for as long as they are deemed to be of value. This encompasses not just technical activities, but also all of the strategic

³HALBERT (2010), p. 8;
ROBERT W. WOODRUFF LIBRARY (2007), pp.1-35.
Per specifiche indicazioni sulla conservazione digitale vedi il link:
http://en.wikipedia.org/wiki/Digital_preservation.

⁴ Ibid.

⁵ROBERT W. WOODRUFF LIBRARY (2007), pp.1-35.

⁶HALBERT (2010), p. 8;

⁷<http://www.ukoln.ac.uk/nof/support/manual/digital-preservation/>

and organisational considerations that relate to the survival and management of digital material⁸”.

“Preservation is therefore a more immediate issue for digital than for traditional resources. Digital resources will not survive or remain accessible by accident: pro-active preservation is needed. Digital Preservation requires not only the maintenance and disaster recovery procedures needed for securing the media and its contents (often referred to as 'archiving') but also strategies and procedures to maintain its accessibility and authenticity over time. This is challenging as often it can require collaboration between different stakeholders, good licensing practice, the association of appropriate metadata and creation of sufficient documentation before technical issues can be addressed⁹”.

“The phrase digital preservation refers to the preservation of digital materials and to the preservation of paper based materials and other artefacts through their digitization¹⁰”.

La conservazione digitale riguarda la gestione e preservazione delle informazioni digitali nel lungo periodo. Il più grande ostacolo di questa conservazione è l'obsolescenza tecnologica dell'hardware e del software.

La definizione base della conservazione digitale è il mantenimento, l'archiviazione di lungo termine e priva di errori nelle informazioni digitali. Inoltre, ha i mezzi per recuperare, ottenere e interpretare le informazioni digitali. Quindi la conservazione, insieme di processi e attività (tecniche, considerazioni strategiche e organizzative) garantisce l'accesso continuo ai materiali digitali fino a quando sono ritenuti di valore. Il materiale digitale, di cui si parla nella pagina precedente, è qualsiasi materiale elaborato da un computer e include anche quello “digitalizzato”.

Le risorse digitali per essere accessibili necessitano di una conservazione proattiva, cioè richiede procedure di manutenzione, ripristino per il supporto e il contenuto, ma anche sistemi e norme per mantenerli nel tempo¹¹.

⁸ <https://www.dcc.ac.uk/sites/default/files/documents/publications/digitalpreservationbp.pdf>

⁹ Joint Information Systems Committee (JISC); http://www.jisc.ac.uk/publications/publications/pub_digipreservationbp.aspx

¹⁰ <http://www.bibliotecheoggi.it/2000/20000806401.pdf>

¹¹ HALBERT (2010), p. 8;
ROBERT W. WOODRUFF LIBRARY (2007), pp. 1-35.

La definizione condivisa della conservazione digitale è “l’insieme dei principi, delle politiche, delle disposizioni e delle strategie finalizzate a prolungare l’esistenza di un documento o di una risorsa digitale grazie alla sua tenuta in condizioni adatte all’uso nella forma originale e/o in un formato persistente che garantisca l’integrità della configurazione logica e del contenuto¹² (Standard ISO 14721 – Open Archive Information System). La conservazione digitale, nel glossario allegato alle Linee guida Agid 2020, è definita come: “l’insieme delle attività finalizzate a definire ed attuare le politiche complessive del sistema di conservazione e a governarne la gestione in relazione al modello organizzativo adottato¹³, garantendo nel tempo le caratteristiche di autenticità, integrità, leggibilità, reperibilità dei documenti.

Questa definizione della preservazione digitale è tale solo per la conservazione a lungo termine di informazioni digitali ed è accompagnata dall’inserimento di metadati di conservazione.

La conservazione è l’attività che protegge e custodisce nel tempo gli archivi di documenti e dati informatici, infatti, il termine conservazione digitale si riferisce sia alla conservazione di documenti digitali creati da un originale cartaceo che alla conservazione di documenti nativi digitali.

Il termine ampio materiali digitali comprende sia i “surrogati digitali” sia i materiali “nati digitali”. I primi sono materiali digitalizzati, acquisiti con l’uso di apparecchiature come scanner, banchi ottici o fotocamere per una copia digitale di un “documento” analogico. Questa attività di digitalizzazione è, in generale, il progetto della biblioteca digitale e si usa per l’acquisizione di file di tipo grafico, sonoro e testuale¹⁴.

Secondo Stefano Pigliapoco, la “digitalizzazione” è: “l’uso intensivo ed esteso delle informazioni in forma digitale. Esso è un processo che ha avuto inizio con la nascita degli

¹² GUERRINI- GIGLIO (2006), p.16. Accessibile all’URL:

https://e-unifi.it/pluginfile.php/227895/mod_resource/content/1/ConsevazioneDigitale_Guerrini_Giglio.pdf;
http://media.regesta.com/dm_0/ANAI/anaiCMS/ANAI/000/0188/ANAI.000.0188.0017.pdf.

¹³ GUERRINI-GIGLIO (2006), p.16;

https://www.agid.gov.it/sites/default/files/repository_files/allegato_1_glossario_dei_termini_e_degli_acronimi.pdf;

https://www.agid.gov.it/sites/default/files/repository_files/manuale_conservazione_microdata_1.6_1.pdf;

https://www.agid.gov.it/sites/default/files/repository_files/documentazione/manuale_di_conservazione_microdata_0.pdf;

https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione_v6.4.pdf.

¹⁴ RIDI (2007), pp.85-89.

Si veda *Conservazione del digitale: definizioni, strategie, approcci* al link:

<https://www.iccu.sbn.it/export/sites/iccu/documenti/Santoro.pdf>

elaboratori elettronici, con la diffusione del personal computer ed infine con internet ha avuto una notevole accelerazione. La digitalizzazione ha interessato tutti i settori della società e consiste nell'impiego di tecnologie informatiche¹⁵.

Affianco alla digitalizzazione troviamo un'altra tecnica che è la “digitazione” usata per l'acquisizione di una copia digitale da un documento cartaceo, che prevede la riscrittura per mezzo della tastiera del computer del volume intero, creando una copia dell'originale in formato digitale. Questa tecnica, però ha costi elevati e tempi lunghi per la riproduzione del documento, inoltre, si perdono gli elementi di impaginazione, i caratteri usati, eventuali immagini e c'è il rischio di errori di digitazione da parte degli operatori¹⁶ se fatto manualmente.

Le edizioni digitali come Liber Liber¹⁷ oppure il progetto Gutenberg¹⁸ si fanno attraverso la digitazione che coinvolgono due tecniche: la digitalizzazione (cioè si fa la fotografia digitale del testo pagina per pagina) e poi si passa al OCR (sistema di riconoscimento dei caratteri) in quanto materiale a stampa.

Fare un'edizione digitale significa partire da un testo che già esiste, e fare la digitalizzazione e poi grazie ad OCR, si ottiene una ricostruzione di un testo che diventa a quel punto trattabile, manipolabile perché non c'è più la fissità della fotografia, ma si parte comunque da qualcosa che non è totalmente fatto a mano.

Quindi si fa la scansione, dopo di che si sottopone ad OCR ovvero al programma, software di riconoscimento dei caratteri. Qualora questo lavoro venga fatto con i progetti Liber Liber, progetto Gutenberg, Biblioteca Italiana o con macchine, programmi, software professionali e sofisticati si ottiene un'immagine di alta qualità. Dopo di che anche il software che fa l'OCR è molto valido, poiché fa riferimento a dei vocabolari, nel senso che alla macchina viene dato una serie di informazioni, come il lessico del testo che si vuole digitalizzare, poiché la macchina possa verificare i caratteri di difficile comprensione. Poi l'operatore fa la comparazione tra il testo originale e la trascrizione del software OCR, ed infine, si fa anche la marcatura con dei marcatori.

¹⁵ PIGLIAPOCO (2010), a cura di., pp. 15-17.

¹⁶ RIDI (2007), pp.85-89.

¹⁷ Liber Liber: è un'Associazione Culturale che coordina il Progetto Manuzio (biblioteca digitale), creata con i volontari che hanno trascritto elettronicamente i testi.

¹⁸ Progetto Gutenberg: archivio di testi, scaricabili gratuitamente. I volumi sono in lingua inglese, poi aggiunte anche edizioni in lingue straniere. Anche di questo progetto si occupano i volontari con l'aiuto economico di alcune accademie.

Come afferma Vrana: “Digitization is the process of taking traditional library materials, typically in the form of books and papers, and converting them to electronic form where they can be stored and manipulated by a computer¹⁹.”

Uno dei vantaggi dell’uso della digitalizzazione è l’ampio accesso alle risorse, la tutela e la preservazione degli originali. Infatti, l’oggetto digitale, tramite la rete è accessibile a un vasto numero di utenti che lo visualizzano in qualsiasi momento.

Poi questa permette la preservazione nel tempo del materiale originale, riducendo l’utilizzo del documento cartaceo²⁰.

Per i progetti di digitalizzazione sono previste delle fasi da seguire per la buona riuscita ed efficacia, quali ad esempio:

- la pianificazione del progetto;
- la digitalizzazione del materiale;
- la creazione dei metadati;
- il caricamento del materiale su una piattaforma fruibile agli utenti;
- la conservazione del materiale digitalizzato²¹.

Mentre i materiali nati digitali, secondo il ricercatore di OCLC Ricky Erway, sono “elementi creati e gestiti in forma digitale”²² e l’accesso continuo a questi elementi digitali è assicurato dall’insieme delle attività della Digital Preservation (o Conservazione del digitale). Per accesso si intende la fruibilità continua e continuativa di una risorsa digitale, conservando tutte le qualità di autenticità, accuratezza e funzionalità ritenute essenziali per gli scopi per cui il materiale digitale è stato creato e / o acquisito.

La conservazione può essere:

1. conservazione di lungo termine ovvero l’accesso continuo ai materiali digitali, o almeno alle informazioni in essi contenute, a tempo indeterminato;
2. conservazione di medio termine cioè l’accesso continuo ai materiali digitali oltre i cambiamenti tecnologici per un periodo di tempo definito ma non indefinitamente.

¹⁹ VRANA (2010), p.326.

²⁰ ZHANG-GOURLEY (2008), pp.3-5.

²¹ ZHANG-GOURLEY (2008), p.7.

²² ERWAY (2010),
accessibile all’URL:

<https://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf>

3. conservazione di breve termine cioè l'accesso ai materiali digitali per un periodo di tempo definito mentre è previsto l'uso, ma che non si estende oltre il prevedibile futuro e / o fino a quando non diventa inaccessibile a causa dei cambiamenti nella tecnologia.

I materiali digitali sono minacciati dall'obsolescenza tecnologica di hardware e software, dall'invecchiamento dei supporti (deterioramento fisico) e dall'obsolescenza dei formati (Hedstrom e Montgomery 1998). Infatti, l'accesso alle risorse digitali dipende dall'hardware e dal software che diventano obsoleti a causa della velocità dei cambiamenti tecnologici. Inoltre, l'arretratezza della tecnologia è considerata la più grande minaccia tecnica per l'accesso continuo al materiale digitale. Il secondo problema della preservazione riguarda la fragilità e l'instabilità dei supporti di archiviazione a causa di condizioni di conservazione e gestione non adeguati. L'altra causa è l'antichità dei formati, cioè l'evoluzione nel tempo dei formati e dei programmi di elaborazione dei testi. Per far fronte a queste obsolescenze vengono proposte strategie, tecniche di preservazione come: refreshing, migrazione ed emulazione.

Il refreshing è la tecnica che si occupa di “mantenere i bit in buona salute”, consiste nel copiare le informazioni su nuovi supporti in tempi brevi.

La migrazione si occupa del trasferimento sistematico dei documenti su nuovi hardware e software prima che diventino obsoleti, e grazie a questa tecnica si conserva il contenuto intellettuale degli oggetti digitali e si mantiene la capacità di recuperarli, mostrarli, usarli a causa dei mutamenti della tecnologia.

L'emulazione è la tecnica che ha come obiettivo di mantenere insieme contenuto e contesto; “È un approccio che permette l'emulazione di sistemi obsoleti su sistemi futuri e non ancora noti, in modo che un software originale di un documento digitale possa essere usato in futuro malgrado sia ormai obsoleto”²³.

Maria Guercio propone tre strategie, tipologie di intervento per garantire la conservazione di documenti digitali:

- *“un approccio passivo basato sull'archiviazione dei dati garantita da una sequenza ininterrotta di operazioni di backup;*

²³ ROTHENBERG (2000), pp.1-85.

- *il ricorso al cosiddetto “museo delle tecnologie”, che consente di mantenere in vita l’intero ambiente originario di sistema. Ad esempio, scegliendo questo metodo, dovendo “aprire” una versione di Microsoft Word non utilizzabile da un certo calcolatore perché obsoleta, bisognerà procurarsi un hardware e un sistema operativo adatti a supportare gli strumenti software in grado di lavorare con l’applicazione sopra indicata;*
- *una modalità attiva che prevede emulazioni e migrazioni, ovvero trasformazioni anche precoci dei dati medesimi e/o dei processi di gestione²⁴”.*

1.2 I materiali “Born Digital”

I Born digital sono materiali digitali che non sono destinati ad avere un equivalente analogico, né come fonte di origine né come risultato della conversione in forma analogica. Infatti, Rick Erway nel suo saggio, “Defining Born Digital”, fornisce una definizione di nativo digitale e dà un elenco di tipologie di materiale “nativamente digitale”²⁵.

Secondo Erway, le risorse digitali native sono “elementi creati e gestiti in forma digitale”.

I tipi di materiali considerati digitali da Erway sono elencati nel riquadro sottostante²⁶.

Materiali digitali	Note
Fotografie digitali	le fotografie digitali sono una delle forme in più rapida crescita di contenuto digitale nato. La loro custodia dovrebbe essere garantita dal loro essere nei formati correnti e tradizionali e che siano copiati su media contemporanei e durevoli. L’integrità delle foto può essere minacciata dalla compressione.
Documenti digitali	quasi tutti i documenti sono attualmente creati in formato digitale. Per i documenti mantenuti in formato digitale, dovrebbero essere utilizzati formati standard come il Portable Document Format (PDF) per mantenere la formattazione, separando i documenti dal software che li ha

²⁴ GUERCIO (2013), pp. 1-232.

²⁵ ERWAY, 2020: <https://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf>

²⁶ <https://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf>
<https://hangingtogether.org/?p=823>

	<p>creati. I numerosi sforzi per catturare e preservare la produzione intellettuale di un'università in un archivio istituzionale stanno sviluppando competenze in questo settore.</p>
<p>Contenuti web raccolti</p>	<p>Internet Archive acquisisce istantanee del Web, ma le istituzioni potrebbero assumersi la responsabilità di eseguire un'archiviazione più mirata in modo più completo. Una biblioteca nazionale può archiviare i siti Web della sua nazione. Un'università può archiviare il proprio dominio. Gli archivi potrebbero essere raccolti da siti Web relativi a un particolare argomento o evento. Gli strumenti open-source sviluppati da Internet Archive possono essere utilizzati per eseguire la ricerca per indicizzazione e fornire l'accesso al contenuto. I dati possono essere conservati nel formato di file WARC (WebARChive) standard ISO. L'approccio alla raccolta Web è abbastanza stabile.</p>
<p>Manoscritti digitali</p>	<p>i “documenti” personali possono arrivare come manoscritti digitali nati, che possono accumularsi mentre gli archivisti pianificano cosa farne. La maggior parte delle raccolte garantisce un trattamento minimo incentrato esclusivamente sulla conservazione del contenuto.</p>
<p>Record elettronici</p>	<p>questa categoria include documenti governativi e archivi aziendali, istituzionali e organizzativi. Questo tipo di raccolta potrebbe consistere principalmente in documenti in formati di elaborazione testi o includere una serie di e-mail, database, fogli di calcolo, presentazioni e altri tipi di file, alcuni dei quali possono essere letti solo utilizzando software proprietario. Nella maggior parte dei casi è meglio ottenere il contenuto da formati proprietari. Gli archivisti dovrebbero essere coinvolti nella definizione delle politiche per le loro istituzioni e non solo nella pulizia. Fortunatamente, esistono strumenti open source per</p>

	l'acquisizione e la gestione di documenti elettronici e la formazione è disponibile presso la Society of American Archivists e tramite i programmi di formazione regionali finanziati dalla National Historical Publications and Records Commission.
Set di dati statici	i set di dati vengono creati nel corso della ricerca e possono essere la base per la ricerca futura, ma spesso vengono creati senza tener conto della conservazione o dell'accesso futuro. Alcuni set di dati richiedono software e documentazione speciali per renderli utilizzabili e il sistema potrebbe dover essere conservato o emulato. Il contesto, inclusa la natura del campione, l'approccio alla raccolta dei dati e il software utilizzato, dovrebbero essere mantenuti.
Dati dinamici	questa tipologia include set di dati aggiunti nel tempo, basati sul tempo o che includono sequenziamento genetico o progettazione assistita da computer (CAD). Può includere dati privi di significato fino a quando non vengono applicati e può esserci un numero infinito di azioni e risultati. In molti casi il software e l'hardware devono essere mantenuti o emulati. I dati dinamici possono includere anche ambienti sociali. La Library of Congress ha ripreso l'archivio di Twitter.
Arte digitale	l'arte digitale può essere semplice come la fotografia digitale o può essere molto più complessa in quanto potrebbe essere un mezzo misto, dinamico o potrebbe richiedere la ricreazione di un'intera installazione per renderla in modo efficace.
Pubblicazioni sui media digitali	questi sono materiali che vengono regolarmente pubblicati in formato digitale. Pubblicazioni commerciali come CD musicali, film su DVD e videogiochi si trovano su supporti abbastanza stabili e quando tali supporti vengono sostituiti,

	il contenuto viene spesso rilasciato nuovamente in nuovi formati. Le biblioteche tendono a tenere il passo con i formati desiderati dai loro utenti. C'è poca preoccupazione immediata qui e la licenza e il copyright rendono difficile per le biblioteche o gli archivi agire. Ma, come con i primi film, a un certo punto il contenuto perderà il suo valore commerciale e, a meno che qualcuno non si assuma la responsabilità della custodia, andrà perso per obsolescenza o decadimento.
--	--

I materiali visivi e audiovisivi (come fotografie, audio e video) sono i formati digitali nativi più frequentemente raccolti, seguiti da documenti istituzionali, manoscritti e altri archivi. Vengono raccolti formati come la posta elettronica, tesi, materiali cartografici digitali, storie orali, onorificenze universitarie, prodotti accademici e blog.

I materiali "nati digitali" utilizzano la tecnologia per fornire un livello di praticità e funzionalità che non è possibile nell'ambiente analogico. Ad esempio, continuano a proliferare database dinamici che vengono costantemente aggiornati, per produrre mappature su larga scala o pubblicazioni su richiesta, che utilizzano la tecnologia in modo molto efficace per l'accesso corrente ma pongono notevoli sfide in termini di capacità di mantenerne l'accesso nel tempo e anche la capacità di confrontare i dati in diversi momenti nel tempo.

Entrambi i surrogati digitali degli originali analogici e le risorse "nate digitali" porranno alla fine sfide simili in termini di garanzia della loro sopravvivenza, sebbene queste ultime possano essere considerate le più vulnerabili in quanto non esiste un originale analogico in caso di smarrimento. In generale, più complessi sono i materiali, più difficile sarà garantire che rimangano accessibili e mantengano la stessa funzionalità nel tempo²⁷.

La capacità di preservarne l'accesso anche in futuro dei materiali digitali dipende dal coinvolgimento di un'ampia gamma di parti interessate. I principali tra questi sono i creatori di contenuti digitali, il cui coinvolgimento nella conservazione potrebbe

²⁷ Preservation Management of Digital Materials: The Handbook;
<https://www.dpconline.org/handbook/introduction>

comportare, ad esempio, la considerazione di standard in termini di formato, media e garantire la disponibilità di una documentazione sufficiente per consentirne la gestione da parte di altri.

Un'altra parte interessata sono le istituzioni, le quali fungono da depositi a lungo termine per i materiali digitali e devono stabilire un dialogo continuo con i creatori ma anche un approccio proattivo alle potenziali adesioni future.

Le grandi istituzioni coinvolte nella creazione di materiali digitali possono essere più sensatamente quelle che li conservano nel tempo, garantendo così il massimo ritorno sull'investimento iniziale di creazione.

I modelli cooperativi per la conservazione a lungo termine potrebbero includere un certo numero di organizzazioni, alcune delle quali potrebbero avere esperienza nell'assicurare la conservazione di materiali cartacei e cercare logicamente di estendere questo mandato alle loro controparti digitali, mentre altre potrebbero specializzarsi in particolari argomenti aree e / o particolari tipologie di materiali digitali.

In tutto il mondo, alcune istituzioni pubbliche come archivi, biblioteche e musei hanno assunto un ruolo nella conservazione dei materiali digitali, diventando leadership nell'affrontare le implicazioni pratiche della preservazione.

1.3 I Metadati

Nel sistema di conservazione digitale per organizzare e recuperare i contenuti sono utilizzati i metadati²⁸, che sono informazioni che descrivono, spiegano, localizzano una risorsa. Inoltre, riassumono il contenuto dei dati, la struttura e la provenienza. Essi sono spesso chiamati dati sui dati²⁹ o informazioni sulle informazioni poiché aggiungono importanza e scopo per «l'identificazione di dati simili in diverse raccolte di dati». Consentono, anche, agli archivi di conservazione di tracciare la catena di custodia, confermare la fissità e garantire l'integrità del documento digitale.

²⁸ RIDI, 2010, pp.1-206.

[http://lettere2.unive.it/ridi/metadati.htm#9.0;](http://lettere2.unive.it/ridi/metadati.htm#9.0)

[http://eprints.rclis.org/19696/1/informatica_biblioteche.pdf.](http://eprints.rclis.org/19696/1/informatica_biblioteche.pdf)

²⁹ http://www.di.uniba.it/~gentile/DispensePPCD1112/PPCD2011-2012_3.pdf

I metadati sono considerati fondamentali per la descrizione degli oggetti digitali, poiché hanno un ruolo importante nei progetti di gestione degli archivi di materiale digitale e soprattutto nella digitalizzazione.

Sono numerose le definizioni di metadati che sono state formulate da varie organizzazioni nel corso degli anni, che riflettono la funzione per la quale vengono esaminati.

Ad esempio nel Manuale d'uso per lo standard MAG (Metadati amministrativi-gestionali), che è stato promosso ed elaborato dall'Istituto Centrale del Catalogo Unico per la realizzazione della Biblioteca Digitale Italiana, i metadati vengono definiti come “una risorsa digitale è inseparabilmente composta dal contenuto informativo (una sequenza di bit) e da una serie di informazioni (metadati) tali da rendere quella sequenza di bit significativa, individuabile, e accessibile per la fruizione, l'archiviazione, la conservazione, la disseminazione e le altre operazioni gestionali”³⁰. Nel mondo americano, l'ufficio, il soggetto e l'organizzazione che si occupa della definizione degli standard fornisce quest'altra descrizione, dove vengono messe in evidenza le funzioni dei metadati, le quali sono informazioni strutturate che descrivono, spiegano l'utilizzo o la gestione di una risorsa di informazioni: “*Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information*”³¹.

Riccardo Ridi, nella sua pubblicazione “Il mondo dei documenti”, fornisce una differente definizione, egli infatti afferma che: “*i metadati sono rappresentazioni semplificate e standardizzate delle entità che si vogliono organizzare, più semplici da ordinare e ricercare rispetto agli oggetti che rappresentano*”³². E' un'affermazione molto sintetica rispetto alle precedenti, ma in essa sono riassunte molto bene tutti i concetti riguardanti i metadati.

Un' istituzione, un ente collettivo, la National Science Foundation, suggerisce questa diversa descrizione, secondo la quale i metadati riassumono il contenuto dei dati, la struttura, le interrelazioni e la provenienza. Inoltre, aggiunge rilevanza e scopo ai dati e

³⁰ <https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html>

³¹ National Information Standards Organization, Understanding Metadata, an introduction to metadata that includes an overview of leading metadata contenders and examples of practical applications, NISO Press, 2004, p. 20.

³² RIDI (2010), p.15.

consente l'identificazione di dati simili in diverse raccolte di dati³³. In fine, Edward, W. Corrado, Heather, L. M. Sandy, nel *“Digital Preservation for Libraries, Archive, & Museums”*, danno quest'altra definizione, secondo la quale *“i metadati sono una delle tecnologie fondamentali che i curatori digitali utilizzano per organizzare e recuperare i contenuti nel sistema di conservazione digitale. Come i sistemi di conservazione digitale, i metadati sono un argomento complesso e le decisioni sui metadati possono influenzare drasticamente l'accesso a lungo termine. I metadati, tuttavia non devono essere visti come una minaccia poiché le decisioni sbagliate sui metadati sono reversibili se è stata conservata una buona documentazione e i metadati sono stati applicati in modo coerente. [...] I metadati consentono agli archivi di conservazione di tracciare la catena di custodia, confermare la fissità e garantire l'integrità del documento digitale*³⁴.

I metadati come si è potuto evincere dalle precedenti definizioni costituiscono il linguaggio della biblioteca digitale e possono essere distinti/classificati in diverse tipologie funzionali³⁵:

- **metadati descrittivi**: permettono di rappresentare la struttura delle risorse contenute nella biblioteca digitale e descrivono una risorsa per scopi di presentazione (*discovery*), identificazione, recupero (*retrieval*) delle informazioni degli oggetti digitali con caratteristiche come i titoli, gli autori, e i soggetti. *“Sono costituiti da descrizioni normalizzate dei documenti fonte (o dei documenti digitali nativi), risiedono generalmente nelle basi dati dei sistemi di Information Retrieval all'esterno degli archivi degli oggetti digitali, e sono collegati a questi ultimi tramite appositi link [...]”*³⁶. Un metadato descrittivo è il Dublin Core³⁷ che ha lo scopo di fornire elementi per descrivere il contenuto dell'oggetto digitale in maniera standardizzata.

³³ National Science Foundation, Cyberinfrastructure Vision for the 21st Century Discovery, Cyberinfrastructure Council (CIC), 2007, P. 22.

³⁴ CORRADO-HEATHER (2017), pp. 177-178.

³⁵ Per approfondimenti vedere CRUPI (2015), pp. 398-399.

³⁶ <https://www.iccu.sbn.it/export/sites/iccu/documenti/MetaAGVZintroduzione.pdf>;
<http://dspace.unive.it/bitstream/handle/10579/14214/855497-1224942.pdf?sequence=2>;

³⁷ Sviluppato dal DCMI (Dublin Core Metadata Initiative), il set di marcatori Dublin Core è il più diffuso standard di metadati a livello internazionale. Il Dublin Core deve il suo successo da un lato alla sua semplicità e dall'altro alla sua estrema flessibilità. Il progetto si è sviluppato in ambito OCLC (On line Computer Library Center), la grande rete di servizi americana per le biblioteche. Nel marzo 1995 si è tenuta una conferenza nella città americana di Dublin (Ohio), durante la quale i partecipanti (bibliotecari, archivisti, editori, ricercatori e sviluppatori di software) hanno riconosciuto all'unisono la necessità di creare un insieme di strumenti condivisi per l'accesso alle risorse digitali.

- **metadati strutturali:** servono “*per descrivere la struttura interna dei documenti (es. introduzione, capitoli, indice di un libro) e gestire le relazioni fra le varie parti componenti degli oggetti digitali [...]. Sono separati [dai metadati descrittivi] solo per le specifiche proprietà dell’ambiente digitale, mentre nel mondo tangibile la struttura o è parte integrante della descrizione (ad esempio per gli archivi o per gli spogli delle riviste), o non è affatto (ad esempio nell’articolazione in capitoli o pagine di un oggetto bibliografico, che non è oggetto di catalogazione)*”³⁸;
- **metadati tecnici:** forniscono informazioni sulla tipologia di software utilizzato per creare il documento, descrivono la composizione dei file e sono utili per accedere all’elemento elettronico descritto. Essi identificano il formato e la dimensione del file, l’hardware e il software necessari per raggiungere i contenuti;
- **metadati amministrativi-gestionali:** sono utilizzati “*per le svariate operazioni di gestione degli oggetti digitali all’interno dell’archivio*”³⁹. Inoltre, “*sono utilizzati principalmente in ambito amministrativo per compiti quali acquisizione, proprietà intellettuale, diritti e restrizioni all’accesso delle informazioni, caratteristiche tecniche relative alla storia del trattamento, della provenienza e della conservazione dei contenuti digitali. Esempi di questa tipologia di metadati sono i riferimenti topografici, i numeri di catalogo, i riferimenti alla responsabilità della descrizione, il copyright, e molti altri*”⁴⁰. MAG è composto da sezioni di metadati descrittivi e strutturali fondamentali per la gestione del materiale digitale, poi esso è uno standard aperto ed è utilizzabile per molti tipi di oggetti digitali;
- **metadati d’utilizzo:** sono metadati collegati all’uso delle collezioni, alle risorse informative ed al loro livello;
- **metadati di conservazione:** sono metadati inerenti la gestione di collezioni e altre risorse legate alla conservazione⁴¹.

Lo schema METS (Metadata Encoding and Transmission Standard)⁴² è lo standard che codifica i metadati descrittivi, amministrativi e strutturali degli elementi digitali. Inoltre,

³⁸ <https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#manuale>.

³⁹ FELICIATI (2010), pp. 203-204.

⁴⁰ http://dspace.unive.it/bitstream/handle/10579/14214/855497-1224942.pdf?sequence=2;https://bibliotecadigitale.cab.unipd.it/chi_siamo/i-progetti/Appunti.pdf.

⁴¹ http://eprints.rclis.org/19696/1/informatica_biblioteche.pdf.

⁴² <http://www.loc.gov/standards/mets/>.

fornisce regole per codificare in XML i metadati e informare sull'organizzazione delle risorse digitali.

I metadati in ambito di conservazione possono essere classificati in:

- Metadati identificativi, per il soggetto produttore e i formati elettronici abilitati;
- Metadati sul contesto storico-istituzionale, solo per il soggetto produttore;
- Metadati sul contesto organizzativo, solo per il soggetto produttore;
- Metadati sul contesto tecnologico, solo per il soggetto produttore;
- Metadati sul contesto archivistico, solo per il soggetto produttore;
- Metadati sugli oggetti informativi digitali⁴³.

Nel modello OAIS troviamo definite alcune qualità dei metadati fondamentali per la conservazione e *“...questi requisiti hanno avuto una profonda influenza sul successivo sviluppo di un certo numero di schemi di metadati per la conservazione. I metadati per la conservazione sono metadati che supportano il processo di conservazione digitale a lungo termine, e comprendono informazioni sulla provenienza, sui diritti di proprietà intellettuale e sugli ambienti tecnici e interpretativi di un oggetto digitale archiviato. In questo senso, la finalità dei metadati per la conservazione si sovrappone notevolmente con i tipi di metadati definiti nel modello informativo OAIS e quindi OAIS è un punto di partenza naturale e un riferimento obbligato nello sforzo per sviluppare schemi di metadati per la conservazione. Sebbene le iniziative riguardanti i metadati per la conservazione siano state numerose, una – PREMIS⁴⁴ – è emersa come lo standard de facto. [...] Sebbene l'ambito del Data Dictionary di PREMIS non si sovrapponga esattamente con i requisiti per i metadati descritti nel modello informativo OAIS (per esempio, il concetto OAIS di Informazioni descrittive non è contenuto in PREMIS), i concetti chiave di OAIS di Informazione sulla rappresentazione e di Informazioni sulla conservazione sono fondamentali per PREMIS, come già lo erano stati per la struttura di metadati per la conservazione su cui PREMIS si è basato. PREMIS può davvero essere visto come una implementazione di questi concetti, anche se non necessariamente copre*

⁴³ PIGLIAPOCO (2007), pp. 100-103.

⁴⁴ <http://www.documento-elettronico.it/glossario/112-premis-preservation-metaddata-implementation-strategies>

l'intera gamma di informazioni che si prevede che questi debbano comprendere. Il punto chiave è che il modello informativo OAIS funge da fondamento per ciò che attualmente è lo standard prevalente per i metadati per la conservazione. Nello stesso modo in cui il modello funzionale OAIS ha esercitato una profonda influenza sullo sviluppo dei sistemi e dei servizi di deposito di conservazione digitale, il modello informativo OAIS ha influenzato la progettazione degli oggetti informativi che i depositi devono gestire"⁴⁵.

1.4 Il modello OAIS

Lo standard OAIS (Open Archive Information System) è il modello di riferimento per la conservazione digitale ed è stato realizzato dal Consultative Committee for Space Data Systems (CCSDS) con l'obiettivo di sviluppare specifiche tecniche per la condivisione di dati. OAIS è approvato e pubblicato come standard ISO 14721:2003 nel 2003 e poi è stato aggiornato dal CCSDS nel 2012 come ISO 1421:2012. Lo scopo del modello è la conservazione di lungo periodo di informazioni digitali di qualsiasi oggetto digitale. Inoltre, OAIS è un sistema informativo aperto per l'archiviazione, dove per "aperto" si intende lo sviluppo della progettazione dello standard fatto in collaborazione con tutti i soggetti interessati⁴⁶.

Difatti, OAIS (ISO 1421:2012) è:

*"un archivio, costituito da un'organizzazione, che può essere parte di una più ampia organizzazione, di persone e sistemi, che ha accettato la responsabilità di conservare informazioni e renderle disponibili per una comunità di riferimento"*⁴⁷.

In questa definizione è fatto risaltare il compito principale che prendono la conservazione e l'accesso nella struttura ideata dal modello. La descrizione del sistema può essere usata per rappresentare qualsiasi istituto di archiviazione (tipo archivio, biblioteca, museo...),

⁴⁵ LAVOIE (2016), pp. 1-68.

Per approfondimenti su PREMIS vedere l'URL: <http://www.loc.gov/standards/premis/>.

⁴⁶PIGLIAPOCO (2010), pp. 123-137; MICETTI (2008), pp. 32-49;

Conservazione digitale - Centro di eccellenza italiano sulla conservazione digitale, *OAIS Reference Model*, 2014, oppure <<https://www.iso.org/standard/57284.html>> in data 28 dicembre 2020;

GUERCIO (2013), pp. 59-83.

CORRADO-HEATHER (2017), pp. 54-63.

CECONE (2019), pp. 9-30.

WESTON (2013), a cura di., pp. 23-24.

⁴⁷PIGLIAPOCO (2014), a cura di., p.301.

anche se vincolato a precise responsabilità ed a una comunità di riferimento. “Un archivio, inteso come struttura organizzata di persone e sistemi, la cui responsabilità è conservare a lungo termine [...] i documenti e i relativi metadati di uno o più soggetti produttori, per renderli disponibili per una comunità di riferimento⁴⁸”.

OAIS indica come devono essere conservate le risorse digitali per una comunità di utenti dalla fase di acquisizione (*ingest*) del materiale digitale fino alla creazione del pacchetto di informazioni utili per usufruire del materiale digitale preservato (*dissemination package*) per l'utente finale. Inoltre, è un modello flessibile che può essere usato in vari contesti e questo può avere dei risvolti negativi per quanto riguarda la gestione degli aspetti tecnici che sono lasciate alle istituzioni. Esso fornisce un vocabolario concettuale e un quadro metodologico (modello informativo e funzionale) oltre ad essere uno strumento per ideare un progetto di *repository* digitale e valutare gli archivi digitali già creati. Una parte integrante del modello è l'*information package* (IP) che è formato dal materiale digitale da archiviare, dai metadati e dai *packaging information* che raccolgono in un unico pacchetto. Tra i pacchetti informativi, OAIS ne distingue tre che sono: il pacchetto di versamento SIP (*Submission Information Package*), il pacchetto di archiviazione AIP (*Archival Information Package*) e il pacchetto di distribuzione DIP (*Dissemination Information Package*).

Il fulcro dello standard OAIS sono due modelli:

- il modello funzionale, che definisce compiti, funzioni e flussi di informazioni di archiviazione;
- il modello informativo, che definisce le diverse componenti degli oggetti che un archivio riceve, cura e diffonde.

Il modello informativo esamina più da vicino i tipi di informazioni richiesti per la conservazione e la diffusione a lungo termine, inclusi i metadati necessari per accedere e comprendere un oggetto archiviato. Esso specifica anche come i diversi tipi di informazioni si relazionano tra loro e come sono strutturati. Nel modello funzionale sono incluse le informazioni sulle modalità di lavoro trattato da ogni funzione. OAIS permette di illustrare i dati delle funzioni: acquisizione,

⁴⁸PIGLIAPOCO (2014), pp.287-310.

archiviazione, gestione dei dati, pianificazione della conservazione, accesso, amministrazione e servizi generali⁴⁹.

Il fine del modello OAIS è quello della conservazione a lungo termine e permettere al consumatore l'accesso all'informazione. Oltre a questi obiettivi sono elencate le responsabilità che un archivio OAIS deve osservare, che sono:

- *“negoziare e accettare dai produttori le informazioni da conservare;*
- *ottenere il controllo dell'informazione archiviata a un livello tale da assicurarne la conservazione a lungo termine;*
- *definire la “comunità di riferimento” dell'archivio;*
- *assicurare che l'informazione conservata sia comprensibile autonomamente dalla comunità degli utenti, e cioè che l'informazione possa essere usata senza ricorrere al produttore dell'informazione stessa;*
- *seguire linee di comportamento e procedure documentate tali da assicurare che le informazioni siano conservate in sicurezza rispetto a qualunque rischio prevedibile;*
- *rendere l'informazione conservata disponibile per la comunità degli utenti e consentire la distribuzione di copie autenticate dell'informazione conservata nella sua forma originale o in una forma riconducibile all'originale⁵⁰”.*

Lo scopo dello standard è preservare “la stringa di bit”⁵¹ e gli altri dati essenziali per definire quel materiale digitale. Inoltre, il modello prevede alcune entità di riferimento che interagiscono con l'archivio che sono il produttore di oggetti digitali (*producer*), il destinatario/l'utente (*consumer*) e il responsabile della gestione del deposito (*management*). Il produttore è l'entità che deposita il materiale all'interno dell'archivio per la conservazione e qui si stabiliscono i tipi di informazioni, i formati, i metadati e la tipologia di trasferimento della risorsa. L'utente è colui che interagisce con i contenuti dell'archivio per trovare e ottenere le informazioni. Infine, il responsabile si dedica alle politiche globali dell'archivio identificando i sistemi di pianificazione, svolgendo revisioni all'archivio e ricercando finanziamenti per la conservazione.

⁴⁹ PIGLIAPOCO, op. cit. (2010);

<http://ww12.conservazionedigitale.org>

⁵⁰ STEFANI (2017), p. 31.

⁵¹ <http://digitalia.sbn.it/article/view/441/281>

OAIS è preso come modello di riferimento da molti progetti di ricerca nel campo della preservazione digitale.

Tra i progetti che hanno preso come modello di riferimento OAIS, ricordiamo: NEDLIB⁵² (*Networked European Deposit Library*) prima esperienza di collaborazione tra diverse biblioteche europee per la costruzione di un deposito librario condiviso per la conservazione di pubblicazioni in formato digitale. Nel 2000 viene avviato il progetto DSPACE⁵³ dal *Massachusetts Institute of Technology* per un archivio digitale basato sul modello OAIS. Affine il progetto DRS (*Digital Repository Service*) creato nell'Università di Harvard⁵⁴ e negli stessi anni CEDARS⁵⁵ (*CURL Exemplars in Digital Archives*) progetto di ricerca che si basa sui problemi della conservazione digitale. Similare a quest'ultimo è il progetto CAMiLEON (*Creative Archiving at Michigan & Leeds: Emulating the Old on the New*)⁵⁶, poi PANDORA (*Preserving and Accessing Networked Documentary Resources of Australia*)⁵⁷, ERA (*Electronic Records Archive*, progetto americano)⁵⁸, ARSBNi (Arricchimento dei servizi della bibliografia nazionale italiana)⁵⁹, BIBIT (Biblioteca italiana)⁶⁰, il progetto CASPAR (*Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval, 2006-2009*)⁶¹, InterPARES 1 (1999-2001), InterPARES 2 (2002-2006)⁶², InterPARES 3 (2007-2012), InterPARES Trust (2012-2019), ERPANET- DPE (2001-2007, 2006-2009), DELOS (2004-2007), PLANET (2006-2009), JISC, CAPITAL PROGRAMME (2006- 2009)⁶³, DPC (*Digital Preservation Coalition* creata nel 2001 UK), DCC (*The Digital Curation Centre*, 2003, UK), PARADIGM (UK, DAL 2006), ReDigIt (Italia), NESTOR (Germania), APARSEN (una rete europea per i repository della ricerca scientifica, 2011-2014).

⁵² <https://cordis.europa.eu/project/id/LB5648/it>

⁵³ [https://duraspace.org/dspace/;](https://duraspace.org/dspace/)

<http://www.bibliotecheoggi.it/2005/20050502701.pdf>.

⁵⁴ <https://library.harvard.edu>

⁵⁵ <https://www.webarchive.org.uk>

⁵⁶ <https://www.dpconline.org/docs/miscellaneous/events/270-future-r-d-wheatley/file>

⁵⁷ <http://pandora.nla.gov.au/pandas.html>

⁵⁸ <https://www.archives.gov/era>

⁵⁹ <https://www.minervaeurope.org/structure/nrg/npp/italy/ongoingproject.htm#ars;>

<https://www.iccu.sbn.it/it/pagina/Lista-dei-progetti-che-utilizzano-i-metadati-in-Italia/>

⁶⁰ <http://www.bibliotecaitaliana.it>

⁶¹ <https://www.casparpreserves.eu>

⁶² https://www.academia.edu/29966336/I_progetti_di_conservazione_digitale_una_proposta_per_lItalia

⁶³ https://www.academia.edu/29966336/I_progetti_di_conservazione_digitale_una_proposta_per_lItalia

La conservazione mantiene la memoria del passato e rende gli oggetti accessibili nel presente e soprattutto nel futuro. Infatti, il materiale digitale necessita di essere preservato, protetto perché non diventi inaccessibile o si danneggi. La conservazione oltre che avere il fine di conservare la memoria del patrimonio culturale, può essere anche una preservazione personale legata all'affetto. Questo modello (OAIS) standard delinea una metodologia basandosi su un sistema aperto di conservazione digitale, "aperto" ovvero l'elaborazione di uno standard tramite i forum aperti. Il modello si sofferma sulla distinzione degli oggetti tra quelli acquistati, quelli archiviati e quelli che vengono mostrati all'utente. E per questa suddivisione lo standard OAIS si basa su tre concetti che sono:

- *“l’archiviazione di un contenuto informativo presuppone la formazione di un pacchetto informativo che, oltre al contenuto medesimo, contiene i metadati che lo identificano, lo qualificano sotto il profilo dell’integrità e lo collocano nel contesto di provenienza;*
- *la struttura dei pacchetti informativi deve essere preventivamente concordata tra l’archivio OAIS e il soggetto produttore il quale dovrà garantire, nelle fasi di produzione e gestione dei documenti, la valorizzazione di tutti i metadati previsti;*
- *i contenuti informativi digitali sono trattati come sequenze di bit alle quali sono associate le relative informazioni di rappresentazione, e questo è possibile solo se si adottano formati elettronici aperti e ben documentati⁶⁴”.*

⁶⁴ PIGLIAPOCO, op.cit. (2010) p. 133;
PIGLIAPOCO, op.cit. (2014), p 304.

CAPITOLO II: Sistemi di archiviazione digitale

2.1 Sistemi di conservazione

Il sistema di conservazione è il “luogo” in cui archiviare i documenti informatici per conservarne nel tempo l'autenticità, l'integrità e l'accesso.

Il sistema è l'insieme di regole, procedure e tecnologie che assicurano la conservazione dei documenti informatici previsto dall'art.44, comma 1, del CAD (Codice dell'Amministrazione Digitale)⁶⁵. Il sistema garantisce le caratteristiche di autenticità (si riferisce al fatto che un documento garantisce di essere ciò che dichiara di essere, senza aver subito alterazioni o modifiche ed è l'insieme di identità e integrità, in cui identità significa identificazione e provenienza), integrità (quando il documento è intatto, completo, inalterato, cioè il documento non ha subito modifiche), affidabilità (capacità di un documento di rappresentare i fatti di cui tratta, livello di fiducia), leggibilità (fruibilità delle informazioni durante l'intero ciclo di gestione documentale) e reperibilità (capacità di reperire ed esibire il documento)⁶⁶. Inoltre, assicura la conservazione degli oggetti digitali come:

- i documenti informatici e i documenti amministrativi informatici con i rispettivi metadati;
- le aggregazioni documentali informatiche e i propri metadati;
- gli archivi informatici e i metadati⁶⁷.

Il sistema di archiviazione assicura l'accessibilità all'oggetto preservato per la durata programmata dal piano di conservazione del proprietario dell'oggetto digitale conservato e dalla normativa vigente, a prescindere dall'evoluzione della tecnologia⁶⁸. Le funzioni primarie del sistema di archiviazione sono: quello di conservare l'informazione,

⁶⁵ Il Codice dell'Amministrazione Digitale (CAD) è un testo unico che riunisce e organizza le norme riguardanti l'informatizzazione della Pubblica Amministrazione nei rapporti con i cittadini e le imprese. Istituito con il decreto legislativo 7 marzo 2005, n. 82, è stato successivamente modificato e integrato prima con il decreto legislativo 22 agosto 2016 n. 179 e poi con il decreto legislativo 13 dicembre 2017 n. 217 per promuovere e rendere effettivi i diritti di cittadinanza digitale.

<https://www.agid.gov.it/it/agenzia/strategia-quadro-normativo/codice-amministrazione-digitale>

⁶⁶ <https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/conservazione.html#sistema-di-conservazione>

⁶⁷ https://www.agid.gov.it/sites/default/files/repository_files/linee_guida_sul_documento_informatico_.pdf

⁶⁸ https://www.agid.gov.it/sites/default/files/repository_files/linee_guida_sul_documento_informatico_.pdf

assicurando la sua continuità nel lungo periodo e quello di dare accesso all'oggetto digitale (informazione) conservato agli utenti di un sistema OAIS o "comunità di riferimento"⁶⁹. Oltre a questo, il sistema restituisce intatto sia il valore giuridico che l'autenticità dell'informazione nel lungo periodo. I sistemi di conservazione devono garantire un elevato livello di interoperabilità con i sistemi di gestione documentale e tra di loro.

Le funzioni principali di un sistema (basato sul modello OAIS) includono:

- Acquisizione, ricezione delle informazioni e degli oggetti digitali dal produttore al conservatore per l'archiviazione;
- Archiviazione (memorizzazione) delle informazioni;
- Gestione dei dati, controllo dei metadati descrittivi e dei dati amministrativi delle risorse digitali conservate;
- Accesso, aiuta gli utenti ad identificare gli oggetti conservati e le informazioni utili per il loro utilizzo;
- Amministrazione dell'archivio, gestione delle attività quotidiane;
- Pianificazione dell'archiviazione, mette in luce il lavoro di conservazione e le loro evoluzioni in unione con il progresso della tecnologia.

Un sistema di conservazione digitale completo è Archivematica, che si basa su standard internazionali e su regole archivistiche indicate dal progetto InterPARES. È sviluppato secondo il modello OAIS, quindi è *open source*, inoltre, gli standard adoperati sono Premis (per i metadati di conservazione), Dublin Core (per l'identificazione e descrizione degli oggetti digitali) e METS (uno standard per la struttura dell'oggetto digitale a cui si agganciano gli altri set di metadati come: DC o altri metadati per la descrizione; MIX per i metadati tecnici; ed altri per le informazioni sui diritti, sul comportamento e sull'amministrazione).

Il sistema Archivematica usa BagIt⁷⁰ (è un insieme di convenzioni di layout di file gerarchico progettato per supportare l'archiviazione e il trasferimento di contenuti digitali

⁶⁹ LAVOIE (2016).

⁷⁰ KUNZE, et al. (2019), p.2.

«BagIt is a set of hierarchical file layout conventions designed to support storage and transfer of arbitrary digital content. A "bag" consists of a directory containing the payload files and other accompanying

arbitrari) per creare AIP⁷¹ (*archival information package*) per l'archiviazione. Il "sistema supporta la conservazione di tipologie di oggetti rendendo disponibili e analizzando le "proprietà significative" dei formati dei file, inoltre, si basa sulla creazione e integrazione per la funzione conservativa di *micro-servizi* indipendenti, ma interoperabili, facili da sviluppare, migliorare o sostituire"⁷². L'applicativo detiene i formati originali di tutti gli oggetti digitali ottenuti per strategie di conservazione diverse come migrazione e/o emulazione. Il sistema è flessibile e aperto ai cambiamenti tecnologici poiché concepito come insieme di tecnologia, risorse umane, procedure e strumenti software⁷³. Altri sistemi più usati a livello internazionale, oltre a quest'ultimo sono: Fedora Commons, Dspace, Phaidra, Islandora, Xena, ArchivesSpace, Rosetta ExLibris, Greenstone, LOCKSS.

In questo elaborato oltre ad Archivematica affrontiamo alcuni sistemi come Fedora Commons, LOCKSS e Dspace come esemplificativi.

- **LOCKSS** (*Lots of Copies Keep Stuff Safe*, o per "molte copie rendono il materiale sicuro"): il sistema sviluppato presso la biblioteca dell'Università di Stanford dal 1999, è un *software open source* che raccoglie, conserva e diffonde ai suoi utenti una copia dei materiali (periodici) archiviati che sono ad accesso aperto⁷⁴. Il sistema collabora con altri sistemi in una rete *peer-to-peer*, inoltre, LOCKSS opera su un numero insieme di piattaforme *hardware* e *software*. Il sistema è indipendente dai formati e contiene un piccolo *server web* tramite il quale il contenuto è reso accessibile⁷⁵. LOCKSS può archiviare materiali non accessibili ad Internet Archive⁷⁶.

metadata files known as "tag" files. The "tags" are metadata files intended to facilitate and document the storage and transfer of the bag. Processing a bag does not require any understanding of the payload file contents, and the payload files can be accessed without processing the BagIt metadata.

The name, BagIt, is inspired by the "enclose and deposit" method [ENCDEP], sometimes referred to as "bag it and tag it"».

<https://tools.ietf.org/html/rfc8493>

⁷¹«Il pacchetto di archiviazione (AIP) può consistere di una singola unità (archival information unit) o essere in realtà un'aggregazione di più pacchetti archivistici (archival information collection): ciò rende possibile la gestione uniforme di risorse accomunate da caratteristiche logiche o fisiche simili». Descrizione tratta da "Rivista del digitale nei beni culturali", III, 1, 2008, scritta da Giovanni Michetti (il modello OAIS).

⁷² GUERCIO (2013), p. 95.

⁷³ GUERCIO (2013), p.97;

http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/VC_Van_Garderen_et_al_26_Workshop1.pdf pp.5-6;

<https://wiki.archivematica.org/index.php?title=Overview>

⁷⁴ https://www.digitalpreservationeurope.info/publications/briefs/it_LOCKSS.pdf

⁷⁵ Ibid.

⁷⁶ Internet Archive: è una biblioteca digitale non profit che consente l'accesso universale alla conoscenza. <https://archive.org/> https://it.wikipedia.org/wiki/Internet_Archive

- **Fedora Commons** (*Flexible Extensible Digital Object Repository Architecture*): è un sistema di *repository* flessibile, modulare e *open source* per la gestione e diffusione dei materiali digitali⁷⁷. All'inizio Fedora è stato creato da Sandy Payette e Carl Lagoze nel 1997 alla Cornell University, ma nel 2009 il sistema è diventato un progetto DuraSpace quando le organizzazioni Fedora Commons e DSpace si sono fuse per formare DuraSpace.

“L'architettura modulare di Fedora si basa sul principio che l'interoperabilità e l'estensibilità si ottengono meglio fornendo un insieme limitato di servizi di repository stabili e basati su standard e modelli comuni per l'integrazione con altri sistemi e applicazioni best practice. Questi servizi vengono forniti tramite API RESTful in conformità con i moderni standard web. Fedora fornisce una base su cui costruire molti tipi di framework di repository, inclusi i popolari progetti Hydra e Islandora. Il sistema operativo è scritto in Java, testato su Linux, Windows e Mac OSX⁷⁸”.

Il sistema Fedora è usato da organizzazioni accademiche e patrimoni culturali, biblioteche, archivi, università per un accesso circoscritto ad oggetti digitali di tipo storico, culturale e scientifico.

“The Fedora repository system is an open source, digital object repository system using public APIs exposed as web services. The Fedora repository system consists of three layers: the Web Services Exposure Layer, the Core Subsystem Layer, and the Storage Layer. The Web Services Exposure Layer is comprised of three related web services described using Web Services Definition Language (WSDL).

- *“Management Service (API-M) - Defines an open interface for administering the repository. It includes operations necessary for clients to create and maintain digital objects and their components. API-M is implemented as a SOAP-enabled web service⁷⁹”.*
- *“Access Service (API-A) - Defines an open interface for accessing digital objects stored in the repository. It includes operations necessary for clients to perform disseminations on objects in the repository (i.e., to access an object's*

⁷⁷ <https://duraspace.org/fedora/about/>.

⁷⁸ <https://duraspace.org/fedora/resources/technical-specifications/>.

⁷⁹ <http://mirror.dlib.org/dlib/april03/staples/04staples.html>

<http://www.ukoln.ac.uk/metadata/resources/digital-repositories/>.

<http://www.ukoln.ac.uk/metadata/resources/digital-repositories/index.html>.

content) and to discover information about an object using object reflection. API-A is implemented as a SOAP-enabled web service⁸⁰”.

- *“Access-Lite Service (API-A-Lite) - Defines a streamlined version of the Fedora Access Service that is implemented as an HTTP-enabled web service⁸¹”.*

- **DSpace**: è un sistema di conservazione digitale che preserva le produzioni intellettuali di un'università di ricerca ed è un sistema di deposito istituzionale che *“trova la sua origine in eventi importanti per l'approccio open access, come il convegno “Universal pre-print server” (1999), organizzato da diverse organizzazioni bibliotecarie e dalla Research Library del Los Alamos National Laboratory, la Open Archives Initiative, la Budapest Open Access Initiative e la Convenzione di Santa FE. Nel 2000 il Massachusetts Institute of Technology (MIT), gli Hewlett Packard Labs e un comitato di consulenza composto da ricercatori delle facoltà interessate hanno iniziato a lavorare sul software DSpace grazie a un finanziamento di 1,8 milioni di dollari. Nel novembre del 2002 il sistema è stato reso disponibile open source alle istituzioni di ricerca di tutto il mondo, e tra la primavera e l'estate 2002 si è svolta la fase di prova, durante la quale DSpace è stato adottato da quattro istituzioni di ricerca: la Sloan School of Management, il Department of Ocean Engineering, il Centre for Technology, Policy and Industrial Development e il Laboratory for Information and Decision Systems.*

Gli scopi di questo periodo sperimentale erano:

- 1) ottenere consapevolezza di come i collaboratori del MIT avrebbero usato il sistema;*
- 2) informare riguardo le caratteristiche che gli utenti avrebbero trovato utili o avrebbero preferito fossero meglio sviluppate;*
- 3) prevedere i costi al fine di tracciare un piano finanziario a lungo termine⁸²”.*

2.2 Casi di studio italiani

Due iniziative di deposito digitale nazionale sono ParER e Dax, rispettivamente, il primo è della Regione Emilia-Romagna, mentre il secondo è della Regione Toscana. Sono due

⁸⁰ <https://arxiv.org/pdf/1312.1385.pdf>; <http://mirror.dlib.org/dlib/april03/staples/04staples.html>

⁸¹ <http://mirror.dlib.org/dlib/april03/staples/04staples.html>;
<http://www.dlib.org/dlib/april03/staples/04staples.html>

⁸² <http://www.bibliotecheoggi.it/2005/20050502701.pdf>

progetti con la finalità di dar vita ad infrastrutture tecnologiche e organizzative complesse in grado di conservare diversi tipi di materiali⁸³.

ParER utilizza nel processo di conservazione il software *open source* SacER, il sistema di conservazione di ParER, per generare e gestire i pacchetti di archiviazione (AIP)⁸⁴. ParER ha una struttura tecnologica specifica e adatta all'elaborazione di grandi volumi di dati per la sicurezza di integrità, disponibilità e riservatezza. Inoltre, è provvisto di un sito di *Disaster Recovery* per eventi imprevedibili⁸⁵.

“I servizi del sistema di ParER sono testimoniati dalle certificazioni ISO 9001, 27001, 27017 e 27018 rilasciate e rinnovate ogni anno da certificatori indipendenti accreditati presso Accredia⁸⁶. Inoltre, i servizi erogati da ParER nella forma di Software as a Service sono qualificati presso AgID nel cloud PA⁸⁷”.

ParER ha come modello lo standard ISO 14721: 2012 (OAIS) e il progetto internazionale sulla conservazione InterPARES per la logica organizzativa e i rapporti con i produttori. Nel Manuale di Conservazione del sistema sono descritti i processi di preservazione e le responsabilità degli attori del processo⁸⁸.

Il Digital Archive eXtended (Dax) è il sistema di conservazione a lungo termine del materiale digitale. Dax è progettato logicamente unico, ma è diviso in due sottosistemi indipendenti con ruoli complementari, ma collegati tra loro⁸⁹. Il sistema assicura l'accesso, la reperibilità, la leggibilità, l'intelligibilità, la riproducibilità e la trasferibilità dei materiali digitali nel medio e lungo periodo⁹⁰.

Un'altra iniziativa italiana di sistema di conservazione è Pavia Archivi Digitali (PAD) nato nel 2009⁹¹, *“grazie a un'idea di Beppe Severgnini (allora presidente dell'Associazione Alunni dell'Università di Pavia) che propose l'istituzione di un*

⁸³ GUERCIO (2013), p.145.

⁸⁴ <https://poloarchivistico.regione.emilia-romagna.it/documentazione>

⁸⁵ <https://poloarchivistico.regione.emilia-romagna.it/conservazione/conservazione-parer>

⁸⁶ <https://poloarchivistico.regione.emilia-romagna.it/conservazione/conservazione-parer>

⁸⁷ <https://poloarchivistico.regione.emilia-romagna.it/conservazione/conservazione-parer>

<https://www.agid.gov.it/it/infrastrutture/cloud-pa/qualificazione-saas>

⁸⁸ <https://poloarchivistico.regione.emilia-romagna.it/conservazione/funzione-di-conservazione>

⁸⁹ <https://www.regione.toscana.it/documents/10180/13766808/Slides%20DAX%20-%20Sistema%20di%20conservazione%20a%20lungo%20termine%20e%20gestione%20degli%20archivi%20digitali/60b452a9-e5b7-4e5f-9d0b-0e4d8fe9f2f3>

⁹⁰ <https://www.regione.toscana.it/-/dax-infrastruttura-per-la-conservazione-a-norma-dei-documenti-informatici>

⁹¹ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf

archivio⁹² di memorie digitali di scrittori, giornalisti e intellettuali del nostro tempo. PAD è nato dunque per raccogliere, conservare a lungo termine e mettere a disposizione degli studiosi materiali digitali di diverso tipo: stesure di romanzi, appunti, bozze, inediti, prove di copertina, fotografie, stesure di articoli etc., nella consapevolezza tecnica e culturale che nelle elaborazioni delle fase scritte del testo e di conseguenza nella relativa archiviazione dei documenti a uso personale e professionale, l'imponente diffusione degli strumenti di videoscrittura ha portato via via, dalla metà degli anni Ottanta in avanti, e con rapidità sempre crescente, al graduale se non in alcuni casi definitivo abbandono del tradizionale supporto cartaceo. Un primo nucleo di lavoro dell'Università di Pavia nel 2010 ha iniziato a elaborare un progetto sostenibile che potesse durare nel tempo. Beppe Severgnini ha subito conferito oltre sedicimila file e sponsorizzato il progetto presso la sua casa editrice, che ha aderito all'iniziativa: da qui il passo alla donazione di altri autori Rizzoli è stato molto breve;

PAD ha infatti ottenuto l'adesione di Silvia Avallone e Gianrico Carofiglio. PAD ha potuto contare su importanti finanziamenti della Fondazione Alma Mater, sulla sponsorizzazione della banca UBI (che ha messo a disposizione un caveau per il deposito di DVD con una copia originale dei conferimenti) e, grazie a Beppe Severgnini, di Nokia e Vodafone.

Nel 2015 e 2016 altri autori hanno affidato il loro archivio digitale all'ateneo pavese: Francesco Pecoraro, Paolo Di Paolo e Franco Buffoni⁹³”.

PAD è un sistema di preservazione che nasce dopo lo studio e una attenta analisi di tutti gli altri sistemi sopracitati. È un sistema semplice da usare, infatti, richiede solo il codice di conferimento nella fase di acquisizione (*ingest*)⁹⁴. Lo svantaggio di PAD è l'esser stato creato per uno specifico progetto, infatti, è un sistema di software proprietario con una struttura semplice, interoperabile e implementabile con altri sistemi. Il sistema è stato strutturalmente ampliato e costruito in Java, mentre le applicazioni sono trattate da FileMaker. Inoltre, esso è modificabile per una futura migrazione dei dati⁹⁵.

⁹² <http://pad.unipv.it>

⁹³Tratto dalla tesi di laurea di Michele Stefani: “Gli archivi degli scrittori nativi digitali: verso la conservazione e oltre” (Stefani 2017), disponibile sul sito (link controllato il 12 gennaio 2021) <http://dspace.unive.it/bitstream/handle/10579/10134/855567-1204608.pdf?sequence=2> pp.103-105; <http://pad.unipv.it/storia/>
<http://www-5.unipv.it/users/pad/>
<http://digitalia.sbn.it/article/view/2274/1545>

⁹⁴ DIGITALIA (2019), p. 38.

⁹⁵ Ibid. pp.109-110.

Per la realizzazione del sistema PAD sono state prese in considerazione diverse tecnologie per la procedura dell'identificazione e l'analisi dei grandi dati, infatti, si è rivolti all'informatica forense. Per l'organizzazione, la conservazione, e la descrizione sono stati presi come esempio i repository istituzionali, i Big Data e i Digital Asset Management⁹⁶.

2.3 Regole, Procedure, Tecnologie

Le principali normative italiane fondamentali per il sistema di conservazione sono:

il DPCM⁹⁷ del 22 febbraio 2013 (Regole tecniche in materia di generazione, apposizione e verifica delle firme elettroniche avanzate, qualificate e digitali ai sensi degli articoli 20, comma 3, 24, comma 4, 28, comma 3, 32, comma3, lettera b), 35, comma 2, 36, comma 2, e 71)⁹⁸;

il DPCM 3 dicembre 2013 (Regole tecniche in materia di sistema di conservazione ai sensi degli articoli 20, commi 3 e 5-bis, 23-ter, comma 4, 43, commi 1 e 3, 44, 44-bis e 71, comma 1, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005)⁹⁹;

il DPCM 13 novembre 2014 (Regole tecniche in materia di formazione, trasmissione, copia, duplicazione, riproduzione e validazione temporale dei documenti informatici nonché di formazione e conservazione dei documenti informatici delle pubbliche amministrazioni ai sensi degli articoli 20, 22, 23-bis, 23-ter, 40, comma 1, 41, e 71, comma 1, del Codice dell'amministrazione digitale di cui al decreto legislativo n. 82 del 2005)¹⁰⁰;

infine, le Linee Guida sulla formazione, gestione e conservazione dei documenti informatici adottate da AgID con Determinazione del Direttore Generale n. 407/2020 del 9 settembre 2020¹⁰¹.

⁹⁶ Ibid.

⁹⁷ DPCM è Decreto del Presidente del Consiglio dei Ministri.

Queste normative sono tratte dal sito del Polo Archivistico della Regione Emilia-Romagna.
https://poloarchivistico.regione.emilia-romagna.it/documentazione/manuale_di_conservazione

⁹⁸ ibid.

https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione digitale cmt_v16.pdf

⁹⁹ https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione digitale cmt_v16.pdf; https://poloarchivistico.regione.emilia-romagna.it/documentazione/manuale_di_conservazione.

¹⁰⁰ https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione digitale cmt_v16.pdf; https://poloarchivistico.regione.emilia-romagna.it/documentazione/manuale_di_conservazione.

¹⁰¹ In questo sito si trovano tutte le Linee Guida: <https://www.agid.gov.it/it/linee-guida>

Secondo, le Linee guida AgID il responsabile della conservazione indica e attua le politiche del sistema di archiviazione. Infatti, il responsabile:

- a) *“definisce le politiche di conservazione e i requisiti funzionali del sistema di conservazione, in conformità alla normativa vigente e tenuto conto degli standard internazionali, in ragione delle specificità degli oggetti digitali da conservare (documenti informatici, aggregazioni informatiche, archivio informatico), della natura delle attività che il Titolare dell’oggetto di conservazione svolge e delle caratteristiche del sistema di gestione documentale adottato;*
- b) *gestisce il processo di conservazione e ne garantisce nel tempo la conformità alla normativa vigente;*
- c) *genera e sottoscrive il rapporto di versamento, secondo le modalità previste dal manuale di conservazione;*
- d) *genera e sottoscrive il pacchetto di distribuzione con firma digitale o firma elettronica qualificata, nei casi previsti dal manuale di conservazione;*
- e) *effettua il monitoraggio della corretta funzionalità del sistema di conservazione;*
- f) *effettua la verifica periodica, con cadenza non superiore ai cinque anni, dell’integrità e della leggibilità dei documenti informatici e delle aggregazioni documentarie degli archivi;*
- g) *al fine di garantire la conservazione e l’accesso ai documenti informatici, adotta misure per rilevare tempestivamente l’eventuale degrado dei sistemi di memorizzazione e delle registrazioni e, ove necessario, per ripristinare la corretta funzionalità; adotta analoghe misure con riguardo all’obsolescenza dei formati;*
- h) *provvede alla duplicazione o copia dei documenti informatici in relazione all’evolversi del contesto tecnologico, secondo quanto previsto dal manuale di conservazione;*
- i) *adotta le misure necessarie per la sicurezza fisica e logica del sistema di conservazione come previsto dal par. 4.11;*
- j) *assicura la presenza di un pubblico ufficiale, nei casi in cui sia richiesto il suo intervento, garantendo allo stesso l’assistenza e le risorse necessarie per l’espletamento delle attività al medesimo attribuite;*

- k) *assicura agli organismi competenti previsti dalle norme vigenti l'assistenza e le risorse necessarie per l'espletamento delle attività di verifica e di vigilanza;*
- l) *provvede per le amministrazioni statali centrali e periferiche a versare i documenti informatici, le aggregazioni informatiche e gli archivi informatici, nonché gli strumenti che ne garantiscono la consultazione, rispettivamente all'Archivio centrale dello Stato e agli archivi di Stato territorialmente competenti, secondo le tempistiche fissate dall'art. 41, comma 1, del Codice dei beni culturali;*
- m) *predispone il manuale di conservazione di cui al par. 4.7 e ne cura l'aggiornamento periodico in presenza di cambiamenti normativi, organizzativi, procedurali o tecnologici rilevanti¹⁰²”*

Il Sistema utilizza componenti tecnologiche *hardware* e *software* (infrastruttura tecnologica principale, la quale si suddivide in componenti logiche tecnologiche e fisiche/operative) per la preservazione¹⁰³, e gli altri strumenti di riferimento usati sono gli standard, le Linee guida, le buone pratiche, il manuale di conservazione e i formati.

Gli standard di riferimento per la conservazione¹⁰⁴ sono indicati nell'Allegato 4 al documento “*Linee guida sulla formazione, gestione e conservazione dei documenti informatici*”, alla voce “Standard e Specifiche tecniche”¹⁰⁵.

Nella tabella sottostante si fornisce una lista sommaria delle norme prescritte¹⁰⁶:

NORME	DENOMINAZIONE	DESCRIZIONE - NOTE
UNI 11386	Standard SInCRO	Interoperabilità del sistema di conservazione
ISO 14721	OAIS	Sistema informativo aperto per l'archiviazione

¹⁰²<https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/conservazione.html#sistema-di-conservazione>.

¹⁰³ <https://cloud.italia.it/marketplace/>
<https://cloud.italia.it/marketplace/supplier/market/index.html>.

¹⁰⁴ https://www.agid.gov.it/sites/default/files/repository_files/linee_guida/la_conservazione_dei_documenti_informatici.pdf;
https://www.postel.it/Documents/la_conservazione_dei_documenti_informatici_rev_def.pdf.

¹⁰⁵ <https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/allegati.html>;
https://www.gazzettaufficiale.it/do/atto/serie_generale/caricaPdf?cdimg=14A0209800300010110001&dg_u=2014-03-12&art.dataPubblicazioneGazzetta=2014-03-12&art.codiceRedazionale=14A02098&art.num=1&art.tiposerie=SG.

¹⁰⁶ <https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/allegati.html>; <file:///Users/yeshi/Desktop/Gazzetta%20Ufficiale.webarchive>.

ISO 15836	Dublin Core (DC)	Sistema di metadata Dublin Core
ISO 20652	Space data and information transfer systems – producer-archive interface ¹⁰⁷	identifica, definisce e fornisce una struttura alle relazioni e alle interazioni tra un produttore di informazioni e un archivio. Definisce la metodologia per la struttura delle azioni che sono richieste dal momento iniziale del contatto tra il produttore e l'archivio fino a quando gli oggetti di informazione vengono ricevuti e convalidati dall'archivio.
ISO 16363	Space data and information transfer systems - Audit and certification of trustworthy digital repositories ¹⁰⁸	definisce una pratica raccomandata per valutare l'affidabilità degli archivi digitali. È applicabile all'intera gamma di archivi digitali. La ISO 16363: 2012 può essere utilizzata come base per la certificazione.
ISO 16919	Space data and information transfer systems - Requirements for bodies providing audit and certification of candidate trustworthy digital repositories ¹⁰⁹	è destinato principalmente a coloro che istituiscono e gestiscono l'organizzazione che esegue l'audit e la certificazione dei repository digitali. Dovrebbe anche essere utile a coloro che lavorano o sono responsabili di archivi digitali che cercano una misurazione oggettiva dell'affidabilità del loro archivio e desiderano comprendere i processi coinvolti.
ISO 17068	Information and documentation - Trusted third party repository for digital records ¹¹⁰	La ISO 17068: 2017 specifica i requisiti per un archivio di terze parti attendibili (TTPR) a supporto del servizio di custodia autorizzato al fine di salvaguardare l'integrità e l'autenticità comprovabili dei record

¹⁰⁷ <https://www.iso.org/obp/ui/#iso:std:iso:20652:ed-1:v1:en>

¹⁰⁸ <https://www.iso.org/standard/56510.html>

¹⁰⁹ <https://www.iso.org/standard/57950.html>

¹¹⁰ <https://www.iso.org/standard/66760.html>

		<p>digitali dei clienti e fungere da fonte di prove affidabili. È applicabile ai servizi di conservazione o deposito di record digitali come fonte di prova durante i periodi di conservazione degli obblighi legali sia nel settore privato che in quello pubblico. Inoltre, ha la limitazione che la custodia autorizzata dei record archiviati è solo tra il TTPR e il cliente.</p>
ISO/IEC 27001	<p>Information technology - Security techniques - Information security management systems – Requirements, Requisiti di un' ISMS (Information Security Management System);</p>	<p><i>“definisce i requisiti per pianificare, attuare, operare, monitorare, riesaminare, mantenere e migliorare il sistema di gestione per la sicurezza delle informazioni delle aziende. La certificazione secondo la norma ISO/IEC 27001 è applicabile in particolare alle organizzazioni operanti nella gran parte dei settori commerciali e industriali, nonché alle pubbliche amministrazioni e dimostra che il sistema di gestione della sicurezza delle informazioni è conforme allo standard internazionale di riferimento¹¹¹”.</i></p>
ETSI TS 101 533-1 V1.2.1	<p>Technical Specification, Electronic Signatures and Infrastructures (ESI); Information Preservation Systems Security; Part 1: Requirements for Implementation and Management</p>	<p>Requisiti per realizzare e gestire sistemi sicuri e affidabili per la conservazione elettronica delle informazioni</p>

¹¹¹ <https://www.imq.it/it/certificazione-di-sistemi-di-gestione-aziendali/iso-iec-27001-sistemi-di-gestione-della-sicurezza-delle-informazioni>

ISO/CD TR 26102	Requirements for long-term preservation of electronic records.	Requisiti per la conservazione a lungo termine dei record elettronici.
SIARD	Software Independent Archiving of Relational Databases 2.0;	Un formato aperto sviluppato dall'Archivio federale svizzero, progettato per archiviare database relazionali in una forma indipendente dal fornitore. Un archivio SIARD è un pacchetto di file basato su ZIP basato su XML e SQL: 1999. Un file SIARD incorpora non solo il contenuto del database, ma anche metadati strutturali elaborabili dalla macchina che registrano la struttura delle tabelle del database e le loro relazioni.
METS	Metadata Encoding and Transmission Standard;	Lo schema METS è uno standard per la codifica di metadati descrittivi, amministrativi e strutturali relativi agli oggetti all'interno di una libreria digitale, espresso utilizzando il linguaggio dello schema XML del World Wide Web Consortium. Lo standard è mantenuto dal Consiglio METS in collaborazione con il Network Development e MARC Standards Office della Library of Congress, e avviato come iniziativa della Digital Library Federation.
PREMIS	PREservation Metadata Implementation Strategies;	<i>“PREMIS è uno standard di metadati per la conservazione degli oggetti digitali. È stato sviluppato, tra il 2003 e il 2005, da un gruppo di lavoro internazionale promosso da OCLC e RLG (Online Computer Library Center/Research Libraries Group) con l'obiettivo di definire un insieme di elementi</i>

		<i>basilari e facilmente implementabili per la conservazione degli oggetti digitali¹¹²”.</i>
EAD (3)/ISAD (G)	Encoded Archival Description/ General International Standard Archival Description	EAD si è proposto come uno strumento per la conversione e la pubblicazione in formato elettronico di strumenti di ricerca archivistici prodotti originariamente su supporto cartaceo e la loro pubblicazione in formato elettronico nonché per l'elaborazione e lo scambio di descrizioni archivistiche in formato nativamente digitale. Lo standard internazionale di descrizione archivistica, regole di descrizione ¹¹³ .
EAC-CPF/ISAAR (CPF)/NIERA;	Encoded Archival Context-Corporate Bodies, Persons and Families/ International Standard Archival Authority Records for Corporate Bodies, Persons and Families – ISAAR (CPF)	intende costituire uno standard per la codifica in XML e l'interscambio di record di autorità basati sull'International Standard Archival Authority Records for Corporate Bodies, Persons and Families
SCONS2/EAG/ISDIAH	International Standard for Describing Institutions with Archival Holdings	standard internazionale per la descrizione degli istituti conservatori di archivi ¹¹⁴
MAG	Metadati descrittivi, amministrativi e gestionali	<i>“è un “Application profile” perfettamente conforme agli standard internazionali e permette l'uso di metadati mantenuti e definiti da altri schemi (Dublin Core e NISO) in combinazione con metadati</i>

¹¹² <http://www.documento-elettronico.it/glossario/112-premis-preservation-metadata-implementation-strategies>

¹¹³ <https://www.icar.beniculturali.it/index.php?id=58>

¹¹⁴ Ibid.

		<i>specifici definiti per una particolare applicazione¹¹⁵</i> ”.
--	--	---

Per la conservazione a lungo termine sono consigliati linguaggi e formati standard. Un formato testo è formato da caratteri composti da combinazioni di otto 0 e 1(byte). “Un formato di codificazione è un elemento in un insieme di strumenti che lo rende leggibile¹¹⁶”. Questi formati di codifica possono essere aperti, chiusi, proprietari e standard (aperti e non proprietari). Per esempio, i formati dei testi si fondano sui caratteri alfanumerici codificati in ASCII e UNICODE, riconosciute come norme ISO che agevolano l’archiviazione dei dati digitali.

In questo riquadro sono riportati i formati di riferimento:

Tipi di formati¹¹⁷	Accettati
Documenti paginate	PDF/A, RTF, DOCX, ODT, IDML,
Ipertesti	XML, XSD, XSL, HTML, XSLT, CSS, MD,
Basi di dati	SQL, CSV, ACCDB, MDB, ODB, JSON-LD, JWT
Posta elettronica	MBOX, EML
Fogli di calcolo	XLSX, XLS, ODS
Presentazioni multimediali	PPTX, PPT, ODP
Immagini raster	JPEG, TIFF, PNG, GIF, EXR, JP2K, DICOM, PSD, DPX, ARI
Immagini vettoriali e modellazione digitale	SVG, AI, EPS,
Modelli digitali	STL, DWGTM, DXFTM, DWFTM, FBXTM
Fonti tipografiche	OTF, TTF, WOFF

¹¹⁵ https://www.iccu.sbn.it/export/sites/iccu/documenti/Reference.html#sez0_2

¹¹⁶ FRANZESE (2002), pp. 1-79;

http://media.regesta.com/dm_0/ANAI/anaiCMS/ANAI/000/0113/ANAI.000.0113.0012.pdf

¹¹⁷ “Formati di file e riversamento”, Allegato 4 del documento “Linee Guida”:

<https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/allegati.html>

Suono	WAV, BWF, RAW, PCM, SND, AIFF, AIFC, AIF, MUSIC.XML, MID
Video	MPEG2, MPEG4
Sottotitoli	TTML, DFXP, XML, EBU-TT
Contenitori multimediali	MXF, VOB, TS, PS, AVI, MP4, RIFF, MKV, MOV, QT, WEBM
Pacchetti multimediali	IMF, IMP, DCP, DCDM, EXR, DPX, WAV, AMF, XDCAM
Contenitori di archivi	TAR, ZIP, GZIP, 7Z, TGZ, T7Z, RAR, TAR, ISO, VDMK, DMG
Documenti amministrativi	SAML, SPID
Applicazioni e codice sorgente	EXE, COM, JAR, MSI, ANDROID, APK, PKG, IPA, A, LIB, SO, DLL, DYLIB, SH, BAT, CMD, PY, PERL, JS, GO, R, C, CPP, H, JAVA, ASM
Contenitori crittografici	CER, CRT, PEM, PKIX, TSR, TSD, TST, XML, P7M, P7S, PDF, ZIP, KDM.XML

Le componenti tecnologiche *hardware* e *software* appropriati garantiscono l'accesso ai materiali digitali. Ma, l'accesso alle risorse nel tempo è reso difficile da *software* e *hardware* non compatibili con le versioni precedenti o con le versioni future. Per garantire l'accessibilità ai materiali conservati il *software* deve mantenere i materiali così come si presentano originariamente. Inoltre, deve saper leggere i vecchi materiali nel formato in cui erano stati conservati e raffigurarli nella stessa forma originaria.

Il nuovo applicativo deve essere compatibile con il software precedente. Inoltre, dovrebbe essere capace di accettare e produrre file in diversi formati. È facile accedere e spostare i materiali su altri sistemi se il software è interoperabile, cioè, ha la capacità di interagire con altre tecnologie.

Le specifiche del software sono fondamentali e quindi bisogna mantenerle, infatti, questa documentazione essendo essenziale la si trova nel Manuale di conservazione e/o manuale del proprietario. In esso sono documentate la costruzione, la struttura e le funzioni del

sistema di conservazione. Ed è proprio qui che sono identificati le componenti essenziali come hardware, software e le periferiche.

Sono sempre registrate le informazioni necessarie per l'aggiornamento del sistema e sono descritti i formati da adottare, che devono essere indipendenti dalla piattaforma, non compressi e gratuiti, cioè "formati aperti"¹¹⁸.

Il Manuale¹¹⁹ di Conservazione in conformità con le Linee Guida spiega l'architettura, l'infrastruttura utilizzata, l'organizzazione, la descrizione del processo, le misure di sicurezza, i soggetti coinvolti e i loro ruoli, le modalità e i formati usati dal sistema di conservazione. Inoltre, il manuale (documento informatico) riferisce¹²⁰:

- a. *"i dati dei soggetti che nel tempo hanno assunto la responsabilità del sistema di conservazione, descrivendo in modo puntuale, in caso di delega, i soggetti, le funzioni e gli ambiti oggetto della delega stessa;*
- b. *la struttura organizzativa comprensiva delle funzioni, delle responsabilità e degli obblighi dei diversi soggetti che intervengono nel processo di conservazione;*
- c. *la descrizione delle tipologie degli oggetti digitali sottoposti a conservazione, comprensiva dell'indicazione dei formati gestiti, dei metadati da associare alle diverse tipologie di oggetti e delle eventuali eccezioni;*
- d. *la descrizione delle modalità di presa in carico di uno o più pacchetti di versamento, comprensiva della predisposizione del rapporto di versamento;*
- e. *la descrizione del processo di conservazione e del trattamento dei pacchetti di archiviazione;*
- f. *la modalità di svolgimento del processo di esibizione e di esportazione dal sistema di conservazione con la produzione del pacchetto di distribuzione;*
- g. *la descrizione del sistema di conservazione, comprensivo di tutte le componenti tecnologiche, fisiche e logiche, opportunamente documentate e delle procedure di gestione e di evoluzione delle medesime;*

¹¹⁸ [http://www.interpares.org/public_documents/ip2\(pub\)creator_guidelines_booklet.pdf](http://www.interpares.org/public_documents/ip2(pub)creator_guidelines_booklet.pdf)

¹¹⁹ Per approfondimenti sul Manuale di conservazione vedere:
https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione_v6.4.pdf

¹²⁰ <https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/conservazione.html#id10>

- h. *la descrizione delle procedure di monitoraggio della funzionalità del sistema di conservazione e delle verifiche sull'integrità degli archivi con l'evidenza delle soluzioni adottate in caso di anomalie;*
- i. *la descrizione delle procedure per la produzione di duplicati o copie;*
- j. *i tempi entro i quali le diverse tipologie di oggetti digitali devono essere trasferite in conservazione ed eventualmente scartate, qualora, nel caso delle Pubbliche Amministrazioni, non siano già indicati nel piano di conservazione allegato al manuale di gestione documentale;*
- k. *le modalità con cui viene richiesta la presenza di un pubblico ufficiale, indicando anche quali sono i casi per i quali è previsto il suo intervento;*
- l. *le normative in vigore nei luoghi dove sono conservati gli oggetti digitali¹²¹”.*

La conformità agli standard garantisce l'autenticità dei contenuti e le attività di migrazione; infatti, una *policy* per la conservazione deve individuare gli standard, le raccomandazioni e le regole di cui si richiede l'applicazione per ragioni di continuità operativa e amministrativa in relazione al dominio considerato¹²².

L'autenticità dei materiali digitali conservati è supportata da prove sostenute utilizzando tecnologie e procedure amministrative che assicurano l'identità e l'integrità continua e/o riducono i rischi di modifica delle risorse. Un meccanismo per affermare l'autenticità del materiale digitale è la crittografia, cioè la firma digitale che può essere utilizzata durante la trasmissione di documenti tra persone, sistemi o applicazioni per dichiarare la loro originalità in un determinato momento.

Come *l'hardware* e il *software*, anche le firme digitali, sono soggette a obsolescenza e a causa della loro breve durata non possono essere trasferiti insieme ai documenti a cui sono allegate su applicazioni software nuove o aggiornate.

Una delle raccomandazioni, nonché regola del sistema è adottare una strategia per affrontare i problemi di obsolescenza dei macchinari e dei sistemi operativi. Questa regola propone di trasferire le funzionalità dell'hardware al software, cioè usare un'applicazione

¹²¹<https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/conservazione.html#sistema-di-conservazione>

¹²² http://www.interpares.org/display_file.cfm?doc=ip2_dissemination_ws_duranti_venice_2006b.pdf

software per simulare le azioni di un *hardware*. Quest'azione serve per poter mantenere le funzionalità *dell'hardware* quando quest'ultimo diventa obsoleto¹²³.

Quando alcune componenti dell'ambiente tecnologico diventano vecchi, vengono aggiornate e i materiali digitali del sistema vengono spostati nella nuova tecnologia. Affinché questo avvenga nella maniera corretta le capacità del nuovo strumento devono essere uguali a quelli *dell'hardware* sostituito.

Le applicazioni *software* usano i metadati¹²⁴ per i dati sull'identità dei materiali digitali come le informazioni sull'identificazione e il recupero. Inoltre, i metadati sono fondamentali per valutare e mantenere l'autenticità nei *record*. I metadati dell'identità trasmettono l'identità degli oggetti elettronici e permettono la distinzione tra materiali digitali.

Altri set di metadati, detti di integrità consentono l'identificazione dei materiali originali da quelli modificati, corrotti. Il contenuto e i dati contenuti devono essere integri e ciò avviene se il valore e la presentazione del contenuto e dei dati sono inalterati.

L'integrità del materiale digitale dipende dalla manutenzione, supervisione e dall'obsolescenza tecnologica o dagli spostamenti da un sistema all'altro dei materiali¹²⁵. Questi metadati delle applicazioni *software* certificano l'integrità delle informazioni, assicurano la loro conservazione nel tempo e conservano l'accessibilità inalterata per una comunità di utenti.

Tra i metadati ricordiamo quelli amministrativi e gestionali che hanno il compito di rendere l'informazione digitale sicura e affidabile, registrando come viene generato, immesso, archiviato l'oggetto digitale nel sistema di conservazione.

I MAG (metadati amministrativi e gestionali) sono usati nei "modelli logico-funzionali dell'archivio degli oggetti digitali"¹²⁶ per una conservazione permanente. Inoltre, MAG è uno standard aperto, gratis, ma soprattutto è "indipendente da specifiche piattaforme *hardware* e *software*"¹²⁷ perciò può essere usato per molti fini.

¹²³ InterPARES, raccomandazione n.9;

[http://www.interpares.org/public_documents/ip2\(pub\)creator_guidelines_booklet.pdf](http://www.interpares.org/public_documents/ip2(pub)creator_guidelines_booklet.pdf)

Per approfondimenti vedere *InterPARES*: "Project Preservation Resources".

¹²⁴ I metadati sono informazioni che descrivono, spiegano, localizzano una risorsa, inoltre, riassumono il contenuto dei dati, la struttura e la provenienza. <https://www.dpconline.org/docs/digital-preservation-handbook/299-digital-preservation-handbook/file>

¹²⁵ InterPARES, http://www.interpares.org/book/interpares_book_f_part3.pdf

¹²⁶ Standard MAG, <https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#standard>.

¹²⁷ Ibid.

Una caratteristica fondamentale di MAG è che “interagisce ed interopera con diversi standard di codifica di metadati, infatti MAG può assumere altri schemi di codifica come Dublin Core, NISO e trasformarsi in diversi formati come METS e MPEG-21¹²⁸”.

Standard, *Case Studies*, Linee Guida internazionali e buone pratiche sono le fondamenta delle analisi, delle descrizioni delle strategie, delle politiche e delle tecnologie dell’archiviazione digitale. Sono strumenti fondamentali per la gestione di tutte le fasi del sistema di conservazione degli oggetti digitali.

Tra questi ricordiamo le *Guidelines* internazionali come *Preservation Management of Digital Materials (The Handbook della Digital Preservation Coalition)*¹²⁹, il Manuale di buone pratiche per la digitalizzazione del patrimonio culturale del progetto europeo Minerva¹³⁰, la *NINCH Guide to good practice in the Digital Representation and Management of Cultural Heritage Materials dell’Humanities Advanced Technology and Information Institute*¹³¹ e della *National Initiative for a Networked Cultural Heritage*¹³², la *Technical Guidelines for Digitizing Archival Materials for Electronic Access* degli statunitensi NARA (*National Archives and Record Administration*)¹³³, le *Guidelines for Digitization Projects* dell’IFLA (*International Federation of Library Associations*)¹³⁴ in collaborazione con l’ICA (*International Council on Archives*) e tutta la documentazione prodotta dall’iniziativa americana FADGI (*Federal Agencies Digitization Guidelines Initiative*)¹³⁵.

¹²⁸ Ibid.

¹²⁹ <https://www.dpconline.org>

¹³⁰ <https://www.minervaeurope.org/bestpractices/listgoodpract.htm>

¹³¹ <http://ninch.org>

¹³² <https://www.cni.org/wp-content/uploads/2013/06/ARL.pdf>

¹³³ <https://www.archives.gov>

¹³⁴ <https://www.ifla.org/publications/guidelines-for-digitization-projects-for-collections-and-holdings-in-the-public-domain>

¹³⁵ <http://www.digitizationguidelines.gov/term.php?term=object>

CAPITOLO III: Esperienze di conservazione del digitale nativo

3.1 Casi di studio

In questo capitolo ci concentreremo su alcuni progetti che possono considerarsi dei validi esempi della creazione di archivi per la conservazione di materiale letterario e di documenti letterari nativo digitali.

A livello globale il tema degli archivi letterari, formati da documenti nativi digitali, non è affrontato allo stesso modo e non ricopre la medesima importanza.

È nell'ambiente angloamericano che il tema è stato affrontato per la prima volta e solo successivamente vi è stata l'interesse anche di altri paesi come Inghilterra e Italia che come già detto precedentemente saranno il focus di questo elaborato. Da una attenta analisi si può notare che la struttura degli archivi inglesi e italiani mostrano delle differenze. Per esempio, prendendo in considerazione il progetto inglese PARADIGM¹³⁶, che si occupa del materiale digitale di due partiti politici (i conservatori e i laburisti), e si rivolge alla produzione di carattere testuale e non letterario, cioè la documentazione è di carattere politico e amministrativo richiede un linguaggio specifico. Poiché si tratta di progetti di legge, atti pubblici etc., la documentazione pubblica richiede una conservazione poiché oltre ad avere valore legale è sicuramente utile anche per lo studio della storia del paese.

PARADIGM stabilisce delle buone pratiche già dal momento in cui il documento viene confezionato, infatti esso fornisce una procedura costituita da regole “ferree” da seguire in base alla tipologia di documento che si vuole creare. Ad esempio, nella creazione di un file si devono seguire queste procedure:

- organizzare e denominare il file;
- organizzare la gestione della posta elettronica;
- selezionare formati e software adatti;
- eseguire il backup del file.

¹³⁶ <https://wayback.archive-it.org/org-467/20161101115538/http://www.paradigm.ac.uk/workbook/index.html>

L'archivio letterario è molto diverso, innanzitutto è una documentazione privata, può ad esempio trattarsi di una bozza, una fotografia, una intervista o un articolo. Quindi a differenza di PARADIGM non è possibile dettare delle condizioni agli autori, su come devono svolgere la loro produzione letteraria, poiché potrebbe ostacolare la libera creazione degli artisti.

Nel caso di PARADIGM si cerca di anticipare i problemi legati alla conservazione dando delle indicazioni su come creare i documenti e quali accorgimenti tecnici seguire. Inoltre, suggerisce che questi documenti debbano seguire un iter procedurale dal momento della loro produzione fino al passaggio nell'archivio dello Stato¹³⁷. Mentre nel caso degli archivi letterari non è previsto un iter e nemmeno un controllo poiché lo scrittore dispone di una libertà decisionale. Infatti, non è detto che l'autore decida di rendere pubblico il suo elaborato presente sul computer, essi potrebbe decidere di eliminare quello che ha scritto perché non più convinto della sua prima stesura.

Non essendo un documento di fonte pubblica, ma scrittura individuale e privata, l'autore non è vincolato a nessun obbligo nel mantenimento dell'opera o nella conservazione del materiale redatto.

Ogni scrittore ha un proprio stile di scrittura e solo alla fine, dopo che ha messo la sua produzione nel *cloud* e decide di conferire il proprio archivio digitale (ovvero il *cloud* proprio) ad un archivio letterario digitale nativo come, ad esempio PAD (Pavia archivi digitali) verranno messe in atto determinate politiche e accorgimenti atti alla conservazione del materiale. Il conferimento del *cloud* può essere eseguito dall'autore stesso, qualora sia ancora in vita oppure da parenti qualora dovesse essere deceduto.

Quando il conferimento è eseguito dall'autore stesso, l'archivio di riferimento di norma cerca di sensibilizzarlo ricordandogli di fare delle copie, salvare il documento in una cartella e di conservare/scrivere in un luogo sicuro le *password* del *cloud* (se viene utilizzata la tecnica della crittografia).

Per la raccolta delle informazioni preliminari per quanto riguarda gli archivi digitali nativi letterari, ci si può servire del questionario informativo, che è stato redatto da degli studenti dell'Università di Pavia, su idea di Simone Albonico, componente del Comitato di coordinamento di PAD¹³⁸.

¹³⁷<https://wayback.archive-it.org/org-467/20161101115943/http://www.paradigm.ac.uk/workbook/preservation-strategies/digital-archives.html>

¹³⁸BALDINI-CARBÉ-WESTON (2017), pp.154-172.
bibliothecae.unibo.it/.../7027/6761.

Il questionario, che è in forma digitale, viene proposto all'autore e viene messo con tutta la documentazione nell'area dedicata alle info. Al suo interno deve contenere i nomi degli operatori coinvolti e i dati cronologici relativi all'invio e alla ricezione. Il questionario costituito principalmente da domande multiple, presenta quindici aree tematiche diverse e una sezione dedicata ai commenti, suggerimenti o precisazioni dell'autore.

3.2 Progetti

PARADIGM

I materiali visivi e audiovisivi sono i formati digitali nativi più frequentemente raccolti, seguiti da documenti istituzionali, manoscritti e altri archivi. I materiali nativi digitali più diffusi sono negli archivi istituzionali¹³⁹.

Il progetto *Personal Archives Accessible in Digital Media* (PARADIGM), è stato finanziato tra il 2005 e il 2007, nell'ambito del programma JISC "*Supporting Institutional Digital Preservation and Asset Management*". Le principali biblioteche di ricerca dell'università di Oxford e Manchester si riunirono per esplorare le questioni coinvolte nella conservazione di documenti privati digitali.

PARADIGM includeva un modello per garantire l'accesso a lungo termine per le disponibilità istituzionali di documenti personali digitali, linee guida di buone pratiche (sotto forma di una cartella di lavoro sulle questioni relative all'archiviazione dei documenti personali in formato digitale), capacità istituzionale locale rafforzata per la conservazione digitale, modelli sviluppati per le politiche istituzionali per lo sviluppo, la conservazione e lo smaltimento delle raccolte, *test* pratico dei *software di repository* digitali DSpace e Fedora e strumenti correlati, risorse di ricerca per la storia politica del XX secolo sotto forma di nuove collezioni d'archivio¹⁴⁰.

PARADIGM ha proposto una serie di possibili approcci allo sviluppo di raccolte di archivi nativi digitali (o ibridi) personali e questi metodi di base sono:

- le adesioni istantanee regolari;

¹³⁹ DOOLEY-LUCE (2010), p.51.

<https://wayback.archive-it.org/org-467/20161101115817/http://www.paradigm.ac.uk/workbook/appendices/guidelines.html>

¹⁴⁰ PARADIGM: Digital repositories, Digital preservation strategies, capitolo 8; <http://paradigm.sers.ox.ac.uk/workbook/index.html>

- l'approccio post-custodia;
- una combinazione dell'approccio istantaneo e post-custodia;
- L'approccio tradizionale;
- il trasferimento periodico di record tramite *hardware* e supporti ritirati;
- l'auto-archiviazione.

Per questo progetto sono stati utilizzati gli archivi personali dei politici inglesi contemporanei dei conservatori e dei laburisti che comprendono documenti personali, occupazionali, professionali e di interesse speciale che sono noti anche come documenti privati e manoscritti. I politici che partecipano al progetto PARADIGM erano riluttanti a collocare materiale personale corrente, o documenti riservati relativi a terzi, in una biblioteca, soprattutto ai fini di un progetto di banco di prova. Il materiale esemplare a cui si è fatto accesso nel corso del progetto rifletteva i ruoli professionali, piuttosto che personali, dei politici partecipanti e in quanto tale, includeva *briefing* politici, bozze di discorsi e altri documenti interni generati dagli uffici privati dei politici.

Il materiale acquisito dei politici attivi comprendeva 20 diversi formati di file tra cui: e-mail, documenti elaborati tramite testi, fogli di calcolo, immagini digitali, presentazioni, pagine *web* personali e blog. PARADIGM ha anche lavorato con la componente digitale dell'archivio di Barbara Castle, che rappresentava tre generazioni di tecnologia informatica: *hardware*, *software* e formati.

PARADIGM crea linee guida per i creatori di archivi personali fornendo consigli pratici, definendoli come “fonte di consulenza”.

Questi suggerimenti, consigli vanno dal prendere decisioni consapevoli sulla denominazione e sul formato quando si crea un file, all'eliminazione di materiale di scarso valore per liberare spazio di archiviazione e rendere più facile trovare ciò che si sta cercando. Sono presenti consigli sul *backup*, sull'amministrazione e sulla cura dei computer; inoltre, ci sono suggerimenti per salvaguardare la propria *privacy* che quella degli altri.

La prima indicazione riguarda come organizzare e denominare i file e questi suggerimenti sono:

- Essere concisi;
- Selezionare nomi significativi;

- Sviluppare convenzioni di denominazioni standard;
- Evitare maiuscole e spazi;
- Utilizzare il formato “yyyymmdd” (giorno, mese, anno) per le date di registrazione;
- Adottare un sistema di controllo della versione per le bozze;
- Archiviare le informazioni relative insieme in cartelle ben denominate;
- Aggiungere informazioni al corpo dei documenti digitali. Informazioni semplici, come un registro degli autori, una cronologia del documento e una nota sullo scopo di un documento. Oppure aggiungere una semplice tabella schematica al documento;
- Registra metadati extra oltre a quelli automatici;
- Per le immagini: se la fotocamera o il *software* correlato consente di aggiungere metadati alle immagini (ovvero informazioni su quando e dove è stata scattata una fotografia e cosa o chi rappresenta) al momento dell'acquisizione o dell'esportazione, utilizzare questa funzione. Salvare le foto in cartelle appositamente create piuttosto che in modo casuale nella cartella predefinita "Immagini" e valutare la possibilità di salvare un semplice file di testo che descriva il contesto di sfondo delle immagini (data, luogo, persone, ecc.);
- Quando si salvano i documenti dal *web* a scopo di riferimento, bisogna prendere nota della loro fonte poiché gli indirizzi *web* possono cambiare.
- Eliminare ciò che non è importante¹⁴¹.

La seconda indicazione riguarda la gestione della posta elettronica e le e-mail, PARADIGM consiglia la scelta di un servizio che fornisca POP3 / IMAP gratuito in modo da poter leggere la posta elettronica offline ed estrarla dal servizio quando necessario. I suggerimenti sono:

- Identificare la fonte degli allegati di posta elettronica nel nome del file;
- Eliminare le e-mail che non hanno valore a lungo termine il prima possibile;
- Non ignorare la cartella della posta inviata: elimina la posta non pertinente e sposta regolarmente i messaggi inviati nelle cartelle pertinenti;
- Conservare nella posta in arrivo solo i messaggi a cui non si è data una risposta o che riguardano affari in corso. Archiviare o eliminare il resto il prima possibile;
- Organizzare la posta elettronica in cartelle di oggetti con titoli concisi e pertinenti.

¹⁴¹ Appendice 10, in PARADIGM Workbook on Digital Private Papers;
<https://ora.ox.ac.uk/objects/uuid:116a4658-deff-4b06-81c5-c9c2071bc6d0>.

- Fornire ai messaggi di posta elettronica un oggetto significativo e conciso che sia rilevante per il contenuto del messaggio. Se l'oggetto di uno scambio di e-mail cambia, modificare la riga dell'oggetto;
- Rispondere alle e-mail in un blocco di testo solido piuttosto che inserire testo in vari punti del messaggio originale;
- Separare le e-mail personale da quella professionale, se possibile. Se si utilizza un account e-mail basato sul lavoro per la corrispondenza e-mail personale, conservare la posta non correlata al lavoro in una cartella "Personale". Proteggi la privacy chiarendo cosa è privato;
- Se si riceve un'e-mail crittografata, bisogna decrittirla e poi salvarla; questo faciliterà l'accesso ad esso in futuro quando la *directory* di posta elettronica verrà archiviata. Le e-mail crittografate potrebbero diventare inaccessibili nel tempo poiché il metodo di crittografia diventa obsoleto;
- Includere informazioni sufficienti nella *directory* della posta elettronica per identificare le persone rappresentate. La rubrica o l'elenco dei contatti è uno strumento utile per registrare le informazioni sui corrispondenti frequenti;
- Utilizzare un filtro antispam e non aprire e-mail o allegati se si sospetta che si tratti di spam; gli allegati in particolare sono spesso fonte di virus informatici;
- Salvare le e-mail di copertura relative agli allegati che si desidera conservare; ciò fornirà informazioni contestuali sulla relazione tra messaggio e allegato¹⁴².

La terza indicazione riguarda la selezione di formati e *software* adatti. Infatti, un fattore chiave per aumentare la longevità digitale è la selezione del formato più semplice e supportato da più applicazioni. Si suggerisce quindi di evitare formati proprietari riservati, ma usare formati standard aperti.

L'ubiquità di alcuni pacchetti *software* (come *suite Office 2003* di *Microsoft*) li rende fin troppo facili da usare, ma le loro specifiche sono opache, rendendo impossibile una comprensione approfondita di come funzionano. Mentre alcuni sviluppatori di formati proprietari pubblicano le loro specifiche (come Adobe PDF), i formati proprietari riservati che possono essere letti solo insieme a *software* specifici sono meno affidabili dei formati disponibili pubblicamente che possono essere letti da più applicazioni. Quindi si consiglia

¹⁴² PARADIGM Workbook on Digital Private Papers;
<https://ora.ox.ac.uk/objects/uuid:116a4658-defb-4b06-81c5-c9c2071bc6d0>

di utilizzare *software open source* (OSS) i cui costi di acquisizione sono minori, il costo *dell'hardware* è ridotto (i sistemi operativi OS possono essere eseguiti su macchine con specifiche inferiori) così come l'obsolescenza *dell'hardware* (la trasparenza dei sistemi OS come GNU/ Linux possono essere utilizzati in ambienti *hardware* alternativi).

PARADIGM suggerisce alcuni formati che facilitano la conservazione e l'accesso a lungo termine. Per i documenti testuali di ufficio propone l'OASIS *Open Document Format* (ODF) che è uno standard ISO pubblicato (ISO / IEC26300) e quindi più "preservabile" una volta archiviato. È possibile utilizzare molte applicazioni, tra cui *OpenOffice.org*, per creare e salvare documenti in ODF ed è disponibile un *plug-in* per *Microsoft Word* che consente al programma di aprire e salvare documenti in formato ODF.

Per le Banche dati esempi di database *open source* come MySQL, PostgreSQL e Firebird funzionano su una vasta gamma di piattaforme, tra cui GNU / Linux, Mac OSX e Windows, e si confrontano favorevolmente con i loro equivalenti proprietari, in particolare in termini di velocità e stabilità.

Molte applicazioni desktop, incluso *OpenOffice.org*, possono essere utilizzate per accedere a questi motori di database.

Ed infine, PDF/A è un formato di file ampiamente utilizzato per le copie di presentazione dei documenti di *Office* che non possono essere modificati da coloro che visualizzano il file.

PDF è l'acronimo di *Portable Document Format* (PDF), che rappresenta i documenti elettronici come sono creati indipendentemente dagli strumenti e dai sistemi originariamente utilizzati per la creazione, l'archiviazione e il *rendering* dei file.

La "A" in PDF/A sta per "Archivio" e significa che il formato è stato limitato alle funzioni PDF di base per semplificarne la conservazione a lungo termine. L'adozione di questo formato favorisce la conservazione dei file prevenendo la crittografia, i meccanismi dei diritti digitali e altre caratteristiche che ne impediscono la conservazione.

PDF/A è stato approvato come standard ISO nel 2005 (come ISO 9005-).

Per le immagini raccomanda l'utilizzo del formato TIFF che è uno standard aperto e ben supportato mentre e per i siti *web* e *weblog* consiglia di rispettare le direttive del W3C¹⁴³.

¹⁴³ W3C (World Wide Web Consortium).

La quarta indicazione riguarda l'eseguire il *backup* dei file, in quanto quest'ultimo è un'attività essenziale se fatto regolarmente. PARADIGM suggerisce di fare copie su supporti portatili (utilizzo di CD-R o DVD-R), di archiviare una copia fuori sede (o archiviazione *off-site*) e utilizzare servizi o *software* di sincronizzazione dei dati.

La quinta raccomandazione riguarda la cura *dell'hardware* e dei supporti, infatti, PARADIGM propone di sostituire i PC, i server e i supporti di memorizzazione su base ciclica, prima che si rompano per evitare perdite di dati. Inoltre, indica di ridurre al minimo i guasti dei componenti del computer e dei supporti mantenendoli puliti. *L'hardware* e i supporti devono essere conservati in condizioni ambientali stabili per evitare il danneggiamento dei dati, le chiavi USB e altri dispositivi di archiviazione di massa come lettori MP3 e fotocamere digitali devono essere correttamente rimossi *dall'hardware* in base alle esigenze del sistema operativo. CD e DVD devono essere trattati con cura per evitare danni.

Il sesto consiglio riguarda l'amministrazione del sistema, qui si suggerisce di pensare e pianificare le azioni prima di eseguire importanti aggiornamenti di *hardware* e *software*, poiché vecchi file vanno persi con gli aggiornamenti. Quindi, PARADIGM propone di eseguire il *backup* dei file altrove prima di apportare modifiche.

Il settimo suggerimento interessa l'uso di *password* e dispositivi di crittografia per mantenere al sicuro i dati personali. PARADIGM consiglia di selezionare un *software* di crittografia *open source* per garantire una disponibilità continua, inoltre, indica un esempio di software di crittografia del disco *open source* (gratuito) per Windows XP / 2000/2003 e GNU / Linux è TrueCrypt.

L'ottavo riguarda la consapevolezza dei diritti di proprietà intellettuale e della *privacy*, si parla di *copyright* e licenze poiché ci sono materiali come articoli, immagini, musica... conservati che sono creati da altri autori. Il *copyright* si applica alle opere scritte inedite per 70 anni dopo la morte, altrimenti bisogna prendere in considerazione licenze come quelle sviluppate da Creative Commons.

Il nono raccomanda di tenersi aggiornati sui cambiamenti tecnologici che sono rapidi e che bisogna evolvere costantemente l'ambiente digitale per l'interoperabilità e la minaccia dell'obsolescenza di sistemi applicativi.

Il decimo consiglio riguarda la gestione dei file digitali *legacy* che possono essere di un membro della famiglia, di un predecessore del lavoro ma che sono stati abbandonati da tempo dopo un aggiornamento del *software*. Questi possono essere in formati meno recenti, non familiari o potrebbero essere semplicemente difficili da valutarne il contenuto.

La gestione di *hardware* e supporti meno recenti, essendo fragili, sono difficili, così come lo è l'amministrazione dei formati sconosciuti. PARADIGM suggerisce anche una valutazione dei file *legacy* per un'impressione generale sul materiale contenuto nella cartella e identificare i materiali di scarso valore che possono essere eliminati definitivamente e quelli che possono essere utilizzati a breve, medio o lungo termine.

L'undicesimo suggerisce di chiedere aiuto agli archivisti digitali e non fare affidamento solo sulle linee guida generiche.

La tabella seguente riporta schematicamente gli 11 suggerimenti del progetto PARADIGM¹⁴⁴: *Eleven top tips for preserving your personal data*¹⁴⁵

¹⁴⁴ Appendice 10, in PARADIGM Workbook on Digital Private Papers; <https://wayback.archive-it.org/org-467/20161101115817/http://www.paradigm.ac.uk/workbook/appendices/guidelines.html>

¹⁴⁵ Ibid. p.288.

	ACTION	KEY INFORMATION	FREQUENCY
1	Name and file appropriately	<ul style="list-style-type: none"> ✓ Develop meaningful and concise filenaming conventions. ✓ Establish a version control system. ✓ Make your files self-documenting. 	Ongoing
2	Manage your emails	<ul style="list-style-type: none"> ✓ Select your email client and webmail service carefully. ✓ Follow guidelines on managing your emails effectively. 	Ongoing
3	Select suitable formats and software	<ul style="list-style-type: none"> ✓ Use formats that are open, published, standards supported by multiple applications. ✓ Use applications that enable you to use these open formats. 	Ongoing
4	Back up files	<ul style="list-style-type: none"> ✓ Decide what should be backed up. ✓ Design a simple backup routine. ✓ Make copies on removable media. ✓ Store a copy offsite. ✓ If online backup/synchronisation services appeal, decide which to use carefully. 	Weekly
5	Look after your hardware and media	<ul style="list-style-type: none"> ✓ Replace your hardware before it fails. ✓ Keep hardware/media clean and cool. ✓ Treat removable media with care. ✓ Invest in an Uninterruptible Power Supply. 	Ongoing 5 year hardware replacement cycle
6	Administer your system	<ul style="list-style-type: none"> ✓ Ensure files are backed up before updating hardware and software. ✓ Ensure anti-virus software is installed and updated regularly. ✓ Beware of badware. 	Ongoing
7	Consider using passwords and encryption devices	<ul style="list-style-type: none"> ✓ Use passwords or encryption devices to protect data on laptops and removable media. ✓ Consider using a password safe to collate data about all your passwords and encryption keys. 	At point of record creation – review regularly
8	Be aware of intellectual property and privacy	<ul style="list-style-type: none"> ✓ Respect the intellectual property of others. ✓ Look after your own and others' privacy. 	Ongoing

Tabella 1 Appendice 10, in PARADIGM Workbook on Digital Private Papers¹⁴⁶.

9	Keep up to date	<ul style="list-style-type: none"> ✓ Keep in contact with digital archivists who can update you on developments in digital preservation. ✓ Stay aware of technical trends; evolve your digital environment but be aware of potential effects on your archive. 	Ongoing
10	Handling legacy digital files	<ul style="list-style-type: none"> ✓ Ensure records are readable and data is accessible. ✓ Undertake background research and conduct a record survey. ✓ Consult a digital archivist for advice on difficult material 	Case by case as needed
11	Ask digital archivists for advice	<ul style="list-style-type: none"> ✓ Don't just rely on generic guidelines; seek advice from digital archivists. ✓ Think about where you might deposit your archive in the future. 	As required

Tabella 2, seconda parte (Appendice 10, in PARADIGM Workbook on Digital Private Papers)¹⁴⁷.

Lo scenario e il modello organizzativo elaborati da PARADIGM per gli archivi nativi digitali sono molto più elaborati. Infatti, per sviluppare le buone pratiche e le linee guida

¹⁴⁶ Ibid.

¹⁴⁷ Ibid.

sono stati presi in considerazione il modello OAIS e il modello archivistico tradizionale. Lo scopo finale era quello di poter sovrapporre le due procedure al fine di determinare una serie di linee guida che avessero dei criteri tra loro armonizzate. Dal confronto è nato il *Workbook for Digital Private Papers*, uno strumento valido sia per le carte private che per i documenti digitali creati con standard, strumenti e tecnologie sconosciute al controllo dei conservatori¹⁴⁸.

PAD – Pavia Archivi Digitali

PAD è il primo archivio, in Italia, nato per conservare i documenti letterari nativi digitali degli scrittori, intellettuali e giornalisti contemporanei. Conserva appunti, bozze, romanzi inediti, fotografie, articoli, interviste radiofoniche, videoregistrazioni di presentazioni e conferenze, scansioni di copertine, di recensioni ma anche copie di siti raccolte nella memoria dei computer degli studiosi. PAD come detto nel capitolo precedente è una struttura autonoma, che si occupa anche dei *social network* che sono diventati luogo di scrittura, confronto, ispirazione degli scrittori.

Gli archivi letterari, formati da documenti nativi digitali, rientrano in una zona che non può essere riconosciuta distintamente, quindi possiamo dire che “ricade in una zona grigia”¹⁴⁹, a causa di alcune specificità degli archivi: sia da un punto di vista archivistico (poiché il materiale di proprietà dell’autore è sottoposto a continue revisioni e modifiche sia della struttura che del contenuto e della forma, nonché da un utilizzo prolungato nel tempo) sia da un punto di vista di interesse storico (infatti questa tipologia di archivio non fa parte della tutela dei beni culturali).

Al momento attuale le linee guida per il trattamento della documentazione nativa digitale sono scarse come anche le buone pratiche a cui fare riferimento per la conservazione. In ottica di trovare una metodologia comune nel 2012 “OCLC ha intrapreso un progetto di ricerca denominato *Demystifying Born Digital*¹⁵⁰”, in cui sono riusciti ad identificare come possibile iter: l’individuazione di strategie, identificazione di obiettivi, identificazione di risorse *hardware* e *software* e competenze tecniche adatte alla conservazione del digitale nativo.

¹⁴⁸ BALDINI-CARBÉ-WESTON (2017), pp.160-172.

¹⁴⁹ Ibid.

¹⁵⁰ Ibid.

Come suggerito nell'articolo "Se i bit non bastano: pratiche di conservazione del contesto di origine per gli archivi letterari nativi digitali" (2017) di Weston et al., uno scrittore dovrebbe sempre servirsi di una piattaforma in *cloud computing*, utilizzando delle funzioni automatiche di *versioning*. Altrimenti dovrebbe essere responsabilità del curatore digitale affiancare l'autore al fine di evitare eventuali problematiche future (suggerendogli i formati e i supporti giusti, come selezionare, scartare e mettere la corretta denominazione dei file nonché il mantenimento dello strumento usato per la scrittura). PAD utilizza lo strumento "questionario" per collezionare dati preliminari sul materiale che deve essere conservato. Il questionario informativo si compone di 15 aree tematiche così suddivise:

1. *"Rapporto tra archivio personale e archivio conferito: com'è organizzato l'archivio che ha intenzione di conferire e quali le differenze, se sono rilevanti, con quello originale?"*
2. *Struttura e contenuti dell'archivio: saranno apportate modifiche all'archivio prima del conferimento (ad es. nella struttura delle directory, nei nomi delle cartelle, nei contenuti o nei nomi dei file)?"*
3. *Consultazione: parte dei contenuti non saranno consultabili per un certo numero di anni? L'archivio sarà consultabile anche da remoto o vi si potrà accedere solo dalla sede di PAD?"*
4. *Privacy: l'archivio contiene dati sensibili, materiali generati da terzi o che potrebbero violare la privacy di altre persone?"*
5. *Ulteriori conferimenti: ha intenzione di incrementare l'archivio con successivi conferimenti? Acconsente a che PAD lo contatti per chiedere aggiornamenti?"*
6. *Altri conferimenti: ha intenzione di conferire materiale speciale (e-mail, pagine web, social network, newsletter, etc.)?"*
7. *Computer principale: che computer utilizza per lavorare? Quando è stato acquisito? È lo stesso da cui provengono i file che saranno conferiti?"*
8. *Macchine precedenti: con che frequenza vengono cambiati gli strumenti informatici, e in particolare il computer principale? Che computer venivano utilizzati in precedenza? Le macchine sostituite sono ancora nella sua disponibilità? I file che verranno conferiti sono stati generati anche da computer precedenti?"*

9. *Sistema operativo: qual è il sistema operativo installato sul computer principale? Utilizza o ha utilizzato anche altri sistemi operativi?*
10. *Altri computer e dispositivi: utilizza altri computer oltre a quello principale? Utilizza altri dispositivi per l'attività di scrittura e lavoro (ad es. tablet, smartphone)? Se li utilizza, provvede poi a trasferire sul computer principale i testi elaborati altrove? Pensa che alcuni suoi testi siano ora soltanto su dispositivi diversi dal computer principali? I file che ha intenzione di conferire sono stati generati anche da altri dispositivi (ad es. app di tablet e smartphone, macchine fotografiche)?*
11. *Organizzazione dei dati: conosce, all'incirca, la disposizione all'interno del computer dei documenti? Quando salva un file sceglie una collocazione particolare?*
12. *Software di scrittura e revisione: quali programmi di scrittura utilizza abitualmente? È una scelta pensata/consigliata/che dipende dal tipo di computer? Tende a salvare a più riprese i suoi file di lavoro, a cambiare il nome dei file per ottenere versioni diverse e progressive? Elimina i file di sue scritture senza tenerne copia? Quando lavora sui testi con i redattori delle case editrici, utilizza strumenti particolari (ad es. commenti su bozze in pdf, documenti in cui vengano evidenziate parti di testo, discussioni via e-mail, etc.)? Nell'archivio sono presenti file frutto di un lavoro condiviso con i redattori?*
13. *Processi di scrittura: parte del lavoro è cartaceo? Che rapporto c'è tra il lavoro condotto avvalendosi di un supporto cartaceo e quello svolto in ambiente informatico?*
14. *Scrittura e web: qual è il rapporto con le nuove tecnologie e il web? Legge e-book, si documenta via internet mentre scrive? Ritiene che i nuovi strumenti a disposizione stiano influenzando il suo lavoro? Ha mai tenuto un sito internet o un blog? Se sì, è ancora attivo? Ha mai scritto su siti/blog collettivi?*
15. *Social network: ritiene che i social network siano oggi rilevanti per promuovere al pubblico l'attività di un autore? È iscritto a social network? Se li utilizza abitualmente, pensa che questa attività influisca sui processi di scrittura? Ritiene che i materiali elaborati per i social network possano entrare nell'archivio PAD?¹⁵¹”.*

¹⁵¹ BALDINI-CARBÉ-WESTON (2017), pp.167-169.

Uno dei primi esempi di archivio ibrido (carta e bit) è l'archivio di Massimo Vannucci, contenente materiale manoscritto e dattiloscritto che rappresenta il primo tentativo di inventariazione e riordinamento di un archivio personale d'autore. È probabile che altri centri come archivi, biblioteche, università ed accademie stiano preservando oggetti letterari nativi digitali insieme alla documentazione cartacea¹⁵².

È il primo archivio privato composto da documenti digitali e non ad essere ritenuto di interesse storico dalla Sovrintendenza archivistica delle Marche. Un altro esempio di questo tipo molto noto è il fondo Franco Buffoni (poeta anglista che diede il suo materiale al Centro Manoscritti dell'Università di Pavia)¹⁵³.

PAD nel 2019 per sperimentare lo strumento costruito per conservare a lungo termine le risorse *web* chiese a Franco Buffoni di usare il suo sito *web* personale¹⁵⁴ poiché ritenuto adatto, in quanto pieno di contenuti con diversi formati (txt, pdf, mp3...).

Lo strumento realizzato per l'archiviazione delle risorse digitali è il sistema PAD *Web Archiving*¹⁵⁵. I siti *web* vengono aggiornati e modificati continuamente quindi bisogna concordare con l'autore per iniziare a conservare il materiale digitale periodicamente, in modo da archiviare le differenti versioni del sito.

“Attraverso un software per il web scraping, il sito dell'autore è stato riprodotto in locale, in modo da garantirne il browsing offline. Così l'utente futuro potrà navigare liberamente nella copia dell'intero sito¹⁵⁶”. “Per progettare questa implementazione, si è dovuto tenere conto della struttura anche molto complessa che i siti possono talvolta presentare, comprendente riferimenti numerosi ad altre pagine, interne o esterne nel web¹⁵⁷”. “Per questo di ogni pagina che compone il sito web, PAD memorizza, oltre alla pagina stessa, i link anche alle pagine esterne, con un'immagine della pagina a cui il link conduce, nonché i documenti allegati. In questo modo si può tenere meglio traccia dei path che il creatore del sito ha voluto valorizzare¹⁵⁸”. “Se, ad esempio, un link a una pagina esterna non fosse più funzionante o se la pagina non risultasse più

¹⁵² ALDiNA (2020);

<http://www.aiucd.it/aldina-archivi-letterari-digitali-nativi/>

¹⁵³ Franco Buffoni è uno degli scrittori che ha conferito il proprio archivio digitale a PAD. In PAD fu riversato 1065 elementi di cui documenti testuali, video, immagini, link, audio.

¹⁵⁴ www.francobuffoni.it.

¹⁵⁵ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf.

¹⁵⁶ <http://pad.unipv.it/il-progetto/web-archiving/>; https://convegni.unicatt.it/aiucd-Weston_et_al.pdf.

¹⁵⁷ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf.

¹⁵⁸ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf.

esistente, una parte di ciò che l'autore intendeva comunicare, una componente probabilmente significativa del suo pensiero, andrebbe perduta¹⁵⁹”.

Il formato di conservazione a lungo termine usato è il Web ARChive (WARC), che divenne standard nel 2009 e conosciuto come ISO 28500:2017, mentre si usa il *software* Heritrix creato da Internet Archive per il Web Archiving di PAD. Inoltre, viene usato anche il *software* HTTrack (*Web crawler open source*) per la formazione delle pagine. Questo permette di salvare il sito *web* in un deposito locale, procurandosi l'HTML, i file, le immagini da internet al PC, mantenendo intatta la struttura del sito e i vari collegamenti tramite link.

Le principali procedure preliminari adottate sono quelle di creare diverse copie dell'archivio e collocarle su più dispositivi (su quello di PAD, su quello dell'università di Pavia e in quello dell'università di Cremona); fare una copia su un supporto di memorizzazione *hardware* esterno; passare l'archivio nell'area di *working* ed estrarre i metadati ed infine descrivere i file.

PAD è un progetto costruito per autori e istituti culturali selezionati, infatti, esso lavora a stretto contatto con gli scrittori che riversano il loro materiale nel sistema di conservazione. PAD conserva pagine personali come Facebook, Twitter e YouTube, oltre che le e-mail e secondo Weston et al., (2017) sappiamo che gli archivi personali d'autore sono presi in considerazione per una loro ulteriore conservazione dal progetto Magazzini Digitali nazionali¹⁶⁰. Inoltre, alcuni centri italiani iniziano ad archiviare oggetti nativi digitali d'autori contemporanei come il Centro Interdipartimentale di Ricerca (CIRTT) dell'università di Cassino, il Fondo Autografi Scrittori Sardi (FASS) dell'università di Sassari, oltre che Pavia Archivi Digitali (PAD). Mentre, Emmanuela Carbé, nel suo lavoro del 2018: “Digitale d'autore: un archivio possibile?”, afferma che la preservazione del digitale deve interessare anche gli archivi letterari nativi digitali e non solo quelli digitalizzati che hanno già delle buone pratiche e linee guida specifiche. Con questa

¹⁵⁹ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf;
https://aiucd2020.unicatt.it/aiucd-Weston_et_al.pdf;
<http://pad.unipv.it/il-progetto/web-archiving/>

¹⁶⁰ https://convegni.unicatt.it/aiucd-Weston_et_al.pdf

Il progetto fu avviato nel 2006 dalla Fondazione Rinascimento Digitale, dalla Biblioteca nazionale centrale di Firenze e dalla Biblioteca nazionale centrale di Roma.

affermazione si apprende che è all'inizio la designazione di modelli, standard, linee guida e buone pratiche per la conservazione di questi archivi¹⁶¹.

ALDiNa - Archivi Letterari Digitali Nativi

Il progetto ALDiNa nasce grazie a Emmanuela Carbé (assegnista, Università di Siena), Tiziana Mancinelli (assegnista, Università di Venezia) e Federico Boschetti (ricercatore, CNR-ILC) con la collaborazione di AIUCD e CLARIN-IT¹⁶².

L'obiettivo del progetto di ricerca è fare una mappatura dei materiali letterari nativi digitali conservati negli archivi, nelle biblioteche, nelle fondazioni, nelle università e nei centri di ricerca poiché in Italia non ci sono ancora gruppi di lavoro per l'individuazione, la definizione e la pubblicazione di buone pratiche e linee guida sulla documentazione archiviata. ALDiNa vuole raccogliere informazioni sui dati, formati, dispositivi *hardware* e *software* conservati e verificare il loro funzionamento per costruire ed elaborare un documento digitale di riferimento per altri archivi letterari digitali nativi.

Questo progetto sarebbe uno strumento fondamentale, nonché utile per ricercatori, archivisti e studiosi, in quanto mappa e cataloga tutti i materiali letterari nativi digitali contenuti negli archivi, ma soprattutto, elaborerebbe metodologie e buone pratiche per archivisti letterari digitali d'autore.

ALDiNa usa la modalità del questionario per avere le informazioni che necessità dalle istituzioni e le domande sono così strutturate:

- *“conserva oggetti digitali di ambito culturale/letterario;*
- *conserva hardware di ambito culturale/letterario (es. floppy disk, hard disk esterni, computer;*
Se sì 1 e/o 2:
- *ci sono dettagli (es. tipologia di computer, quantità di file, se l'hardware è pienamente o almeno in parte funzionante, se gli oggetti digitali presentano problemi, etc.);*

¹⁶¹ CARBÉ (2018), pp.209-216.

¹⁶² <http://www.aiucd.it/aldina-archivi-letterari-digitali-nativi/>

- *sono attivate strategie di conservazione a medio/lungo termine (es. refreshing, migrazione, etc.), ed eventualmente quali;*
- *il materiale è catalogato e, in caso affermativo, tramite quali strumenti;*
- *ci sono politiche di accesso al materiale;*
- *ci è consentito di darne notizia nel contesto della mappatura;*
- *ci è consentito un contatto in fasi successive del progetto¹⁶³”*

A proposito di buone pratiche e linee guida per il materiale digitalizzato, facciamo qui un accenno ad un sistema di conservazione molto usato da alcune università venete: Phaidra (*Permanent Hosting, Archiving and Indexing of Digital Resources and Assets*). Phaidra è usato a Padova e Venezia nel sistema bibliotecario dell'università, dove troviamo una guida che illustra la procedura per la creazione, conservazione e gestione di un oggetto digitale. Il sistema usa un *software open source* Fedora (sistema di preservazione affrontato nel capitolo precedente) detenuto dall'università di Vienna. Questo è possibile perché dal 2010 l'Università di Padova lavora insieme con quella di Vienna e dal 2014 anche l'Università Iuav e Cà Foscari usano Phaidra. Con questo sistema è possibile preservare oggetti come documenti PDF, audio, immagini, video, libri, collezioni e materiale *web*. Nella guida all'archiviazione, Phaidra, consiglia l'uso di formati aperti, formati non proprietari che si possono leggere da altri sistemi. Il materiale conservato è identificato da metadati che sono conosciuti come UWmetadata (Universität Wien Metadata) dell'Università di Vienna e i metadati Dublin Core per l'interoperabilità con altri archivi come Internet Archive ed Europeana¹⁶⁴.

Phaidra raccomanda i criteri per la scelta dei formati dei file per l'archiviazione a lungo termine che sono:

- apertura (formato aperto che può essere usato da chiunque);
- portabilità (uso del formato su molti piattaforme, hardware e software);
- qualità e funzionalità (formato che assicura la qualità dei materiali);
- supporto allo sviluppo (manutenzione e sviluppo dei prodotti informatici);

¹⁶³ ALDiNa (2020), prima bozza, conferitami dal professore Paul Gabriele Weston.

<http://www.aiucd.it/aldina-archivi-letterari-digitali-nativi/>

¹⁶⁴ <https://phaidra.cab.unipd.it/static/guida-all-archiviazione.pdf>

- diffusione (livello di utilizzo negli archivi del formato adottato);
- trasparenza (codifica del contenuto digitale con standard di codifica);
- auto-documentazione (metadati descrittivi, tecnici e amministrativi rendono il documento sostenibile, più gestibile e trasferibile in altri sistemi)¹⁶⁵.

Per la conservazione del testo è consigliato l'uso del formato PDF/A; per l'immagine JPEG, TIFF e SVG; per il video AVI, MXF e DPX; invece, per l'audio WAVE e FLAC. Inoltre, Phaidra, mette a disposizione dell'utente un documento dove sono descritte tutte le procedure per la digitalizzazione che guardano agli standard e buone pratiche internazionali, ma soprattutto nazionali¹⁶⁶.

L'archiviazione digitale personale riguarda sia i materiali digitalizzati che quelli *born-digital*. Una parte importante di questa conservazione consiste nel determinare, individualmente o con l'aiuto di un'istituzione, quali materiali digitali salvare in modo permanente.

Tre casi di studio forniscono un'esplorazione più dettagliata delle singole pratiche di archiviazione digitale personale e considerano il possibile ruolo che le istituzioni del patrimonio culturale potrebbero svolgere nel preservare i file digitali pubblicati sui social media come Facebook e YouTube.

Ad esempio, l'archiviazione *web* è stata guidata da archivi e biblioteche in tutto il mondo, infatti, le biblioteche in Australia e Nuova Zelanda hanno iniziato a sperimentare con l'archiviazione *web* alla fine degli anni '90 e da allora hanno sviluppato programmi robusti.

Nel 1996, Internet Archive, con sede negli Stati Uniti, ha sviluppato *Wayback Machine*, un servizio che fornisce alle persone l'accesso alle versioni archiviate dei siti *web*, e nel 2005 ha lanciato Archive-It, un servizio di raccolta del *web* per biblioteche e archivi.

Il primo coinvolgimento della *British Library* nella conservazione digitale e nella cura dei documenti digitali personali è iniziato nel 2001 con l'esperimento di archiviazione *web* DOMAIN.UK che ha portato allo sviluppo di un programma per acquisire selettivamente i siti *web* del Regno Unito (Tuck, 2004). Con la formazione del UK *Web*

¹⁶⁵ <https://phaidra.cab.unipd.it/static/formati-file-compresso.pdf>

¹⁶⁶ <https://www.ifla.org/publications/guidelines-for-digitization-projects-for-collections-and-holdings-in-the-public-domain>

Archiving Consortium nel 2003 e il successivo sviluppo del UK Web Archive, il Regno Unito è emerso come leader nell'archiviazione *web* (UK *Web Archiving Consortium*).

Alcune organizzazioni e progetti, nei primi anni del ventunesimo secolo, hanno iniziato a prendere forma e sostenere l'importanza di preservare i file digitali personali.

Alla fine del 2001, fu istituita la *Digital Preservation Coalition* (DPC) con l'obiettivo di "sviluppare un'agenda per la conservazione digitale del Regno Unito in un contesto internazionale" (Beagrie, 2002). Il DPC rimane una comunità e una risorsa vitale per le persone impegnate nell'archiviazione digitale personale e nella conservazione digitale più in generale.

Dal 2005 al 2007, la *Bodleian Library* di Oxford e la *John Rylands Library* di Manchester hanno intrapreso il progetto PARADIGM che ha esplorato i problemi legati alla conservazione dei documenti digitali personali e ha prodotto una cartella di lavoro che descrive in dettaglio le linee guida per le migliori pratiche (PARADIGM, 2005)¹⁶⁷.

Negli Stati Uniti, la *Library of Congress* ha offerto per la prima volta consigli sull'archiviazione digitale personale sul proprio sito *web* nel 2007; da allora, ha continuato a fornire preziose indicazioni su *The Signal*, un *blog* dedicato alla conservazione digitale¹⁶⁸.

Sempre nel 2007, la *British Library*, in collaborazione con l'*University College London* e l'Università di Bristol, ha lanciato il progetto *Digital Lives* finanziato dall'*Arts and Humanities Research Council* (AHRC) (2007-2009). Questa importante iniziativa ha esplorato le questioni legali, etiche, tecnologiche, comportamentali e di accesso relative al modo in cui le persone creano, acquisiscono e condividono gli equivalenti digitali dei documenti personali (John, Rowlands, Williams, Dean, 2010). Il progetto finanziato da NEH nel 2008 e il rapporto *Approaches to Managing and Collecting Born-Digital Literary Resources for Scholarly Use* (Kirschenbaum et al., 2009) hanno esplorato le sfide affrontate da archivisti, amministratori e / o studiosi presso tre istituzioni statunitensi le cui collezioni includono personale archivi digitali: la *Manuscript and Rare Book Library* (MARBL) presso la *Emory University*, Atlanta, Georgia, l'*Harry Ransom Center* presso l'Università del Texas ad Austin e la *Deena Larsen Collection* presso il *Maryland Institute for Technology in the Humanities* presso l'Università del Maryland *College Park*¹⁶⁹.

¹⁶⁷ <http://blogs.loc.gov/digitalpreservation/> <http://www.digitalpreservation.gov/>

¹⁶⁸ Ibid.

¹⁶⁹ Ibid.

Gli archivi personali contengono spesso un assortimento di articoli cartacei, nativi digitali e riformattati digitalmente, che esistono tutti in relazione l'uno con l'altro. La relazione, o il contesto, di questi file l'uno con l'altro e con l'archivio nel suo insieme, può spesso essere importante quanto il loro contenuto.

Alcune persone creano e salvano deliberatamente i propri file digitali; all'inizio degli anni 2000, Gordon Bell, un ex ricercatore presso *Microsoft Research*, ha intrapreso il progetto *MyLifeBits* per registrare ogni momento e interazione della sua vita e salvare questo record digitale completo per i posteri (O'Neill, 2013; Microsoft Research, senza data).

Altri evitano la tecnologia digitale, preferendo invece corrispondere per lettera piuttosto che e-mail, utilizzare un telefono fisso e non partecipare ai social media. La stragrande maggioranza di noi rientra tra questi due estremi, utilizzando tecnologie sia digitali che non digitali per condurre il nostro lavoro quotidiano e la vita personale.

Molti utenti hanno scoperto che creare file digitali è la parte più facile, mentre l'archiviazione e la gestione di questi file richiede più tempo e impegno.

Alcune persone organizzano attivamente ed eseguono il *backup* dei propri file, mentre altri dimenticano o si perdonano d'animo di fronte a ciò che sembra un compito insormontabile¹⁷⁰.

Nella sua indagine del 2008 sulle strategie di archiviazione digitale personale, l'ex ricercatrice di Microsoft Cathy Marshall ha descritto l'approccio casuale che la maggior parte delle persone adotta per gestire i propri artefatti personali come "negligenza benigna", una strategia a mani libere che può avere risultati disastrosi. Spesso le persone pensano di dover fare di più per gestire i propri file digitali, ma faticano a trovare il tempo. Il campo della gestione delle informazioni personali (PIM) si è sviluppato per fornire alle persone risorse e strategie pratiche per l'organizzazione e la gestione dei file personali.

Poiché gli oggetti privati esistono sempre più solo in formato digitale, le esigenze di coloro che creano e salvano file digitali hanno incoraggiato una nuova comprensione dell'aspetto di lettere, foto e altri ricordi nell'era digitale e allo stesso tempo hanno sollevato importanti domande su come gestire oggetti digitali personali da utilizzare oggi e come conservarli per le generazioni future¹⁷¹.

¹⁷⁰ Ibid.

¹⁷¹ Ibid.

Nell'ultimo decennio, il numero di servizi progettati per aiutare le persone a salvare, aggregare, condividere e archiviare file digitali è esploso. Solo alcuni dei molti servizi popolari disponibili negli ultimi anni includono Picasa, Flickr e Google Docs per la gestione e la condivisione di fotografie e altri file; YouTube per la condivisione di video; Dropbox, Box e Google Drive per condividere e archiviare file di ogni tipo; e servizi di social media come Facebook, Instagram, MySpace e Twitter che forniscono agli utenti un modo per creare, condividere e annotare diversi tipi di contenuti.

Tutti questi servizi hanno una cosa in comune: il *cloud storage*. Quando una persona carica i propri file su Google, YouTube o qualsiasi altro servizio online, tali file personali vengono archiviati su computer remoti di proprietà o affittati dalla società che offre il servizio. È importante tenere presente il fatto che, sebbene l'archiviazione "nel *cloud*" sia invisibile all'utente medio, questi server remoti sono entità fisiche che si trovano nello spazio e nel tempo e sono soggetti alla legge.

Oltre alla gestione dei file e ai servizi di condivisione, alcune aziende come Amazon, Google, Apple (*iCloud*), Microsoft (*OneDrive*) e *Dropbox* hanno iniziato a includere l'archiviazione nel *cloud* come parte del pacchetto che offrono. Sempre più spesso le questioni relative alla *privacy* e alla sicurezza sono state in prima linea nelle conversazioni sui file digitali personali, in particolare per quanto riguarda i servizi di *cloud storage*¹⁷².

Il lavoro che le istituzioni del patrimonio culturale hanno intrapreso negli ultimi tre decenni per preservare gli archivi digitali personali è stato recentemente integrato da una crescente consapevolezza tra le persone dell'importanza dei propri file digitali. Molti dei servizi e degli strumenti attualmente disponibili possono aiutare le persone a essere proattive nella gestione delle proprie "cose" digitali, ma queste risorse sollevano anche domande sempre più urgenti sulla *privacy* e sulla sicurezza. Gli archivi digitali personali sono fondamentali da preservare¹⁷³.

¹⁷² Ibid.

¹⁷³ REDWINE (2015), pp. 1-33.

<https://www.dpconline.org/docs/technology-watch-reports/1460-twr15-01/file>

3.3 Archivi orali e audiovisivi

Il Caso di Studio Archivio Vi.Vo

Il progetto Archivio Vi.Vo, è il modello per la conservazione dei documenti audio ed audiovisivi. Gli attori principali coinvolti sono l'Università di Siena, CLARIN-IT¹⁷⁴ e la soprintendenza archivistica e bibliografica di Toscana insieme all'Unione dei Comuni del Casentino i quali collaborano al mantenimento di questo progetto. Il progetto nasce come modello per l'archiviazione sonora e viene utilizzato per la preservazione dell'archivio sonoro di Caterina Bueno¹⁷⁵. L'archivio dell'artista ed etnomusicologa contiene materiali di ricerca, articoli, disegni, note, lettere, fotografie è il caso "prototipico" del progetto.

"Archivio Vi.Vo. mira a costituire uno studio pilota nell'ambito di CLARIN-IT per sperimentare metodi e offrire servizi alle discipline interessate alle fonti orali. All'interno di Archivio Vi.Vo. il repository ILC4CLARIN sarà migliorato attraverso un approccio sperimentale alla conservazione, alla gestione e all'accesso a dati e metadati audio e audio-video. L'attività sperimentale mirerà ad adottare il modello e i servizi di

¹⁷⁴ CLARIN sta per "Common Language Resources and Technology Infrastructure". È un'infrastruttura di ricerca che è stata avviata dalla visione che tutte le risorse e gli strumenti linguistici digitali provenienti da tutta Europa e oltre sono accessibili attraverso un ambiente online per il supporto dei ricercatori nel campo delle scienze umane e sociali. Nel 2012 CLARIN è stata fondata e ha assunto la missione di creare e mantenere un'infrastruttura per supportare la condivisione, l'uso e la sostenibilità dei dati linguistici e degli strumenti per la ricerca nelle scienze umane e sociali. Attualmente CLARIN fornisce un accesso facile e sostenibile ai dati linguistici digitali (in forma scritta, parlata o multimodale) per gli studiosi di scienze sociali e umanistiche, e oltre. CLARIN offre anche strumenti avanzati per scoprire, esplorare, sfruttare, annotare, analizzare o combinare tali set di dati, ovunque si trovino. Ciò è reso possibile da una federazione di centri in rete: archivi di dati linguistici, centri di servizi e centri di conoscenza, con accesso Single Sign-On per tutti i membri della comunità accademica in tutti i paesi partecipanti. Strumenti e dati provenienti da diversi centri sono interoperabili, in modo che le raccolte di dati possano essere combinate e strumenti da diverse fonti possano essere concatenati per eseguire operazioni complesse a supporto dei ricercatori nel loro lavoro. L'infrastruttura CLARIN è pienamente operativa in molti paesi e un gran numero di centri partecipanti offre servizi di accesso a dati, strumenti e competenze. Allo stesso tempo, CLARIN continua a essere costruito in alcuni paesi che hanno aderito più di recente e i set di dati e i servizi di CLARIN vengono costantemente aggiornati e migliorati. In Italia l'infrastruttura di ricerca è CLARIN-ERIC (CLARIN-IT); <https://www.clarin.eu/content/clarin-nutshell-0>; <https://www.clarin.eu/content/about-eric>.

¹⁷⁵ Caterina Bueno (1943-2007), artista ed etnomusicologa toscana che produsse documentazione sonora registrando contadini, uomini, donne, artigiani toscani che cantavano canzoni popolari (ninne nanne, canzoni narrative, stornelli...); <http://www.clarin-it.it/it/content/archivio-vivo>; <https://www.clarin-it.it/it/content/archivio-vivo>

elaborazione e archiviazione ad alte prestazioni della nuova infrastruttura di rete GARR, costruita secondo il paradigma CLOUD¹⁷⁶”.

Lo scopo di Archivio Vi.Vo. è quello di:

- realizzare un impianto (infrastruttura) per la preservazione del materiale sonoro e audiovisivo per la regione Toscana;
- legare il lavoro a un piano internazionale (materiale accessibile all'interno di CLARIN-IT);
- istituire un modello di descrizione audio, audiovisivo, generare dati e metadati FAIR (*findable, accessible, interoperable, reusable*)¹⁷⁷.

AISO (Associazione italiana di storia orale) si è formata nel 2006 a Roma per preparare una struttura volta a conservare, cogliere le fonti orali degli enti, degli istituti, delle associazioni e dei singoli.

“Nel 2014 l’AISO ha formato un gruppo di lavoro incaricato di stendere delle linee guida che potessero aiutare i ricercatori ad affrontare i nodi deontologici e giuridici legati al trattamento delle fonti orali. Il documento finale delle “Buone pratiche di storia orale”¹⁷⁸ è stato presentato pubblicamente in occasione del convegno “Buone pratiche di storia orale. Questioni etiche, deontologiche, giuridiche”¹⁷⁹ organizzato da AISO e Fondazione Museo storico del Trentino, con la collaborazione dell’Università di Trento e dell’Università Ca’ Foscari di Venezia, tenutosi nel novembre 2015 a Trento. E’, in corso, l’adeguamento delle Buone pratiche di storia orale al Regolamento generale per la protezione dei dati personali (GDPR)¹⁸⁰”.

I progetti che vogliono conservare, raccogliere e usare le fonti orali devono fare attenzione alle questioni giuridiche, etiche e deontologiche. La storia orale è l’incisione video o audio di un’intervista che deve essere resa accessibile agli studiosi. L’intervista

¹⁷⁶ <http://www.clarin-it.it/it/content/archivio-vivo>

¹⁷⁷ <http://www.soprintendenzaarchivisticatoscana.beniculturali.it/index.php?id=304>

¹⁷⁸ www.aisoitalia.org/.../lassociazione/presentazione;

[https://www.aisoitalia.org/lassociazione/.](https://www.aisoitalia.org/lassociazione/)

Buone pratiche di storia orale è un documento, strumento di informazione e sensibilizzazione, non impone standard, ma suggerisce buone pratiche.

¹⁷⁹ Buone pratiche di storia orale. Questioni etiche, deontologiche, giuridiche

¹⁸⁰ <https://www.aisoitalia.org/lassociazione/presentazione/>

viene depositata nell'archivio insieme a una scheda di corredo dove sono descritti le informazioni dell'intervistato, la data, il luogo ed altre caratteristiche fondamentali se il soggetto non è anonimo.

IV CONCLUSIONI

La definizione di regole, norme prodotte da archivi regionali di Firenze, Bologna, Ancona e Trento per l'archiviazione delle memorie digitali, sono messe a disposizione dei professionisti, delle istituzioni e degli utenti per una cooperazione, ma soprattutto per una maggiore consapevolezza sulla conservazione digitale. Le istituzioni, come l'Università di Roma (La Sapienza), i depositi di Firenze (DAX) e Bologna (ParER), gli archivi nazionali e le associazioni (Fondazione Rinascimento digitale), ma anche gli archivi dei comuni e gli enti più piccoli supportano una rete di pratiche, prassi, linee guida, standard. Inoltre, grazie al quadro giuridico nazionale, sono forniti modelli (standard) di preservazione internazionali, sostenibili e flessibili per un sistema di conservazione e gestione dei record digitali¹⁸¹.

La conservazione del digitale nativo letterario, quindi gli archivi di persona, di autori contemporanei, è un campo non del tutto esplorato che presenta delle criticità concettuali e pratiche. Ciò è dovuto anche dal fatto che i sistemi di gestione e conservazione informatica dei documenti o record, in Italia, ha avuto una maggiore difficoltà ad affermarsi in questo campo di indagine rispetto ad altre nazioni, infatti, sono a rischio le memorie individuali che non sono tutelate come gli archivi pubblici.

Come rammenta Mariella Guercio: *“il tema della sopravvivenza delle memorie personali digitali è rimasto a lungo assente nella riflessione della comunità archivistica anche a livello internazionale. Eppure, si tratta di documentazione che presenta molte più difficoltà e problemi di quella prodotta da qualunque organizzazione sia pubblica che privata*¹⁸²”.

¹⁸¹ <https://fondazionefeltrinelli.it/archivi-documenti-infrastrutture-digitali/>

¹⁸² GUERCIO (2016), pp. 133-134.

Infatti, per gli archivi letterari digitali nativi, ci sono pochi casi di studio e questo non facilita la produzione, individuazione di buone pratiche e linee guida adeguate alla conservazione a lungo termine dei materiali d'autore.

L'archiviazione di queste risorse, che hanno caratteristiche particolari, richiede soluzioni specifiche.

Un progetto di ricerca (*Demystifying Born Digital*) del 2012 intrapreso da OCLC ha visto la partecipazione di svariati professionisti nei differenti campi del digitale nativo per individuare obiettivi, strategie, risorse *hardware* e *software* e tecniche. La cooperazione ha dato una serie di rapporti necessari per la preservazione digitale¹⁸³.

Ricordiamo alcuni progetti pilota importanti come quello dell'Harry Ransom Center¹⁸⁴ e dell'Emory University's Manuscript Archives and Rare Book Library¹⁸⁵, quello delle biblioteche dell'Università di Oxford e Manchester ed infine citiamo anche i due casi italiani: l'archivio ibrido di Massimo Vannucci e Pavia Archivi Digitali (PAD)¹⁸⁶.

Uno strumento di lavoro molto significativo per la conservazione del digitale nativo è il “*Workbook for Digital Private Papers*”¹⁸⁷, raggiungibile *online*, dove si possono trovare buone pratiche, linee guida e standard per l'archiviazione delle risorse digitali, ma sono presenti anche altri strumenti e pubblicazioni fondamentali per la preservazione. Questa guida è il frutto del progetto PARADIGM.

Un altro progetto internazionale molto importante è AIMS, nato per fornire raccomandazioni, buone pratiche, strumenti per la gestione e conservazione del materiale nativo digitale. Il gruppo di lavoro AIMS sviluppa un modello basato su studi di casi reali di collezioni per ogni istituzione, adottando un approccio pratico (indagini, discussioni, disciplina) e fornendo una varietà di casi di studio con diversi formati di record e problemi *legacy*. Inoltre, il gruppo ha definito un contesto archivistico per identificare questioni etiche e altre sfide, dimostrando la necessità di soluzioni praticabili e sostenibili. Nell'

¹⁸³ La ricerca denominata “Demystifying Born Digital” dell'OCLC, è il progetto che aiuta il personale delle biblioteche e degli archivi ad acquisire la fiducia necessaria per intraprendere i primi passi per lanciare un programma di gestione digitale nato che può essere ampliato nel tempo.

<https://www.oclc.org/research/areas/research-collections/borndigital.html>

¹⁸⁴ Per approfondimenti vedere: <https://www.hrc.utexas.edu/>

¹⁸⁵ Per approfondimenti vedere: <https://rose.library.emory.edu/>

¹⁸⁶ Per curiosità vedere: <https://u-pad.unimc.it/retrieve/handle/11393/246604/55042/Westin%20et%20al%202018.pdf>

¹⁸⁷ <http://paradigm.sers.ox.ac.uk/workbook/index.html>

“AIMS *White Paper*” si apprende che non esiste un’unica soluzione per le raccolte del nativo digitale¹⁸⁸.

Un elemento fondamentale che risalta in tutti questi casi di studio è la necessità di collaborazione tra varie discipline/ figure professionali come informatici, archivisti, bibliotecari, politici, editori e produttori di *software* poiché questa è una realtà difficile¹⁸⁹.

Un altro concetto molto importante è la consapevolezza della pluralità e della complessità delle questioni legate alla conservazione dei sistemi informatici digitali.

La gestione delle collezioni dell’archivio nativo digitale non promette nulla di sicuro, poiché ci sono sempre sorprese a livello tecnico di *hardware* e *software* e per evitare ciò si consigliano sondaggi, interviste e altri tipi di valutazione approfonditi e chiari prima di procedere con l’acquisizione. Le politiche, le pratiche di acquisizione documentate e ben formate possono diminuire l’ambiguità sui dettagli del trasferimento di materiali digitali nativi, come tempistiche, standard di imballaggio, spedizione, frequenza degli accumuli e la presenza di "dati non autorizzati" trasferiti involontariamente o non completamente affrontati nel contratto di acquisizione. Inoltre, per migliorare la qualità dei trasferimenti e delle adesioni digitali, i “donatori” e gli archivisti dovrebbero lavorare per migliorare la comprensione e la documentazione delle abitudini informatiche dei benefattori¹⁹⁰.

Secondo Gabriela Redwine in “*Personal Digital Archiving*” (2015), l’imprevisto continuerà a sfidare e sorprendere i conservatori che gestiscono materiali nativi digitali, nonostante gli sforzi notevoli per creare politiche chiare e attuabili. Inoltre, afferma che le tattiche come l’apertura delle linee di comunicazione tra donatori, personale del *repository* e l’istituzione di pratiche di archiviazione trasparenti ed efficienti promettono di ridurre al minimo le spiacevoli sorprese e migliorare la qualità delle acquisizioni dei materiali nativi digitali¹⁹¹. I programmi di sensibilizzazione prima della conservazione, sono necessari per una archiviazione migliore, quindi una conversazione tra autore e conservatore¹⁹², dove si suggerisce l’adozione di una strategia che prevede la scelta del formato più adatto in fase di creazione della risorsa digitale e che proceda mantenendo

¹⁸⁸ <https://dcs.library.virginia.edu/aims/white-paper/>

¹⁸⁹ <http://www.bibliotecheoggi.it/2000/20000806401.pdf>

¹⁹⁰ REDWINE et.al. (2013), pp.1-28.

¹⁹¹ Ibid.

¹⁹² MCGARRIGLE (2015), pp. 164- 181.

una cura ed una considerazione continua sui materiali che si intende conservare, depositare e archiviare¹⁹³.

Secondo Tommaso Giordano¹⁹⁴ “*gli studi e le esperienze finora maturate nel campo della conservazione a lungo termine degli archivi personali nativi digitali mettono in evidenza l’assoluta necessità di un approccio ex ante; cioè, intervenire quando l’autore è ancora in vita (pre-custodial intervention), quando è ancora disponibile la tecnologia che ha generato i file da conservare. In altri termini, è necessario che l’archivio destinato a essere ceduto venga predisposto per l’archiviazione a lungo termine*”. Inoltre, propone delle esperienze come modelli a cui fare riferimento e cita la *Library of Congress*, in particolare il programma di aiuto per la preservazione degli archivi personali digitali del 2000 nell’ambito NDIIPP¹⁹⁵, dove sono presenti consigli per l’archiviazione di molti tipi di materiali digitali come fotografie, video, mail e propone pubblicazioni oltre a guide sulla conservazione di archivi personali.

Questo per rendere le persone attive nella conservazione dei propri materiali, quindi “archivisti di sé stessi”.

L’altro programma già citato nelle pagine precedenti e che qui disquisiamo è *l’Emory University Library* “*che entrò in possesso dell’archivio letterario personale dello scrittore Salman Rushdie. L’archivio conteneva sia documenti cartacei, ma soprattutto documenti nativi digitali, come foto, blog, mail, testi, video, corrispondenze e diari*¹⁹⁶”. “*La strategia di conservazione adottata si basa sulla tecnica dell’emulazione, che consiste nello sviluppo su macchine nuove di programmi che simulano le funzionalità delle macchine obsolete che hanno originato i file. Questa strategia è considerata la più affidabile nel caso degli archivi letterari, perché è in grado di salvaguardare l’integrità e mantenere la funzionalità dei file originali*¹⁹⁷”. Quindi la collaborazione tra donatore, archivisti, informatici e giuristi è importante per la riuscita del progetto.

In un articolo di Paul Gabriele Weston et al., (2017) in cui si parla di PAD, si può notare che la fase preliminare dell’acquisizione, cioè la parte precedente al conferimento

¹⁹³ ALLEGREZZA (2009), pp. 1-10.

¹⁹⁴ Tommaso Giordano è direttore della biblioteca dell’Istituto universitario europeo, Fiesole.

¹⁹⁵ NDIIPP è National Digital Information Infrastructure and Preservation Program.

https://en.wikipedia.org/wiki/National_Digital_Information_Infrastructure_and_Preservation_Program

¹⁹⁶ GIORDANO (2018), pp. 3-8.

<http://dspace.unive.it/bitstream/handle/10579/10134/855567-1204608.pdf?sequence=2>

¹⁹⁷ GIORDANO (2018), p. 6.

<http://www.bibliotecheoggi.it/rivista/article/view/797/833>

dell'archivio digitale personale, dove l'autore e il conservatore entrano in comunicazione è fondamentale, anzi cruciale. Questo è necessario per risolvere i problemi legati alla tecnologia, ma anche per ricostruire il quadro dell'autore¹⁹⁸.

Gli standard e le buone pratiche attuali non sono sufficienti, infatti, si stanno ancora facendo ricerche per soluzioni sostenibili ed affidabili tramite collaborazioni di molti centri di ricerca.

Inoltre, tutto questo è ancora più difficile, poiché le *policies* e le strategie di preservazione si diversificano per le diverse tipologie di materiali nativi digitali e le buone pratiche continuano ad evolversi nel tempo.

In questo ambito è fondamentale porre attenzione alla tecnologia, ai dati, all'integrità e identità dei documenti digitali ed è fondamentale osservare le regole, le raccomandazioni dei vari casi di studio presenti e fare riferimento al progetto InterPares che propone dei consigli, identificandole in otto regole.

Per l'archiviazione dell'audio digitale sono state approfondite le buone pratiche e le linee guida principali e tra questi ricordiamo “Linee guida sulla produzione e conservazione di oggetti audio digitali IASA-TC 04”, che elabora standard internazionali, “*Capturing Analog Sound for Digital Preservation: Report of a Roundtable Discussion of Best Practices For Transferring Analog Discs and Tapes*” per la digitalizzazione da analogico a digitale e per i formati adatti alla conservazione dell'audio digitale, inoltre, qui le procedure per la conversione del materiale sono molto definite.

Vi si possono trovare presenti anche suggerimenti e istruzioni su come creare nuove risorse digitali audio resa possibile, grazie alla *Collaborative Digitazion Program Digital Audio Working Group*. Infatti, viene pubblicato il “*Sound Directions: Best Practices for Audio Preservation*” dove sono descritte le procedure di lavoro e le migliori pratiche per la preservazione dell'audio. Quest'ultimo è stato preso come punto di riferimento per altri progetti dall' *International Association of Saund and Audiovisual Archives (IASA)*, dalla *Library of Congress*, dalla *Digital Library Federation*, e dalla *European Broadcasting Union*.

Ed infine ricordiamo anche l'*Audio Engineering Society* che diffonde linee guida e standard per la creazione di audio e i rispettivi metadati¹⁹⁹.

¹⁹⁸ BALDINI-CARBÉ-WESTON (2017), pp. 160-172.

¹⁹⁹ https://it.qaz.wiki/wiki/Digital_preservation

Per l'archiviazione dell'immagine in movimento (video, film) o audiovisive non c'è uno standard di formato per la preservazione di video o record digitali "unico per tutti". Sono noti alcuni casi di studio su linee guida, pratiche, metodi per la digitalizzazione di immagini ed audiovisivi ma si tratta di un'iniziativa delle agenzie federali (FADGI).

Per l'archiviazione del proprio archivio personale la *Library Congress* mette a disposizione una lista per la preservazione dove sono elencati programmi, linee guida per la conservazione di mail e social media. Tra i programmi è presente HTTrack (strumento *software* per scaricare un sito da internet in un *directory* locale) e Muse (*Memories Using E-mail, programma per far rivivere i ricordi*)²⁰⁰.

La conservazione del nativo digitale dovrebbe poggiare sulle pratiche sociali, cioè su prassi che iniziano con l'abitudine del singolo utente che può intraprendere azioni attive per documentare e proteggere il proprio materiale, contenuto.

In un saggio²⁰¹ di Rachel Donadio intitolato "*Literary Letters Lost and Found in Cyberspace*", si ha un esempio di scomparsa del patrimonio letterario, cioè la produzione digitale di una corrispondenza tra autori, editori, agenti ed editori, nel cyberspazio perché non conservata in tempo²⁰². Quindi bisognerebbe incoraggiare le persone ad assumersi la responsabilità di conservare gli oggetti digitali creati, che creeranno ed incontreranno.

²⁰⁰ Ibid.

²⁰¹ Donadio, *Literary Letters Lost and Found in Cyberspace*, New York Times Book Review, 2005.

²⁰² Kirschenbaum, *Mechanisms: New Media and the Forensic Imagination*, The Mit Press, Londra, 2008.

BIBLIOGRAFIA

ALLEGREZZA, Stefano [2009] La conservazione digitale: problematiche e strategie, in “Informatica di base. Conoscere e comprendere le risorse digitali nella società dell’informazione”, Macerata, SIMPLE, 2009.

BALDINI, Primo, CARBÉ, Emmanuela, WESTON, P. Gabriele [2017] Se i bit non bastano: pratiche di conservazione del contesto di origine per gli archivi letterari nativi digitali, in “Bibliothecae.it”, VI (2017), 1.

BALDINI, Primo, CARBÉ, Emmanuela, WESTON, P. Gabriele [2017] If bits are not enough: preservation practices of the original context for born digital literary archives, in Bibliothecae.it, VI (2017), 1.

BODLEIAN LIBRARIES, JOHN R. UNIVERSITY LIBRARY [2019] PARADIGM Workbook on Digital Private Papers, Oxford, University of Oxford, 2019.

CARBÉ, Emmanuela [2018] Digitale d’autore: un archivio possibile?, Bari, Archivi editoriali, 2018.

CECONE, Martina [2019] Conservazione del Materiale digitale: Il caso dell’Archivio della memoria delle donne, Venezia, Università Cà Foscari Venezia, 2019.

CORRADO, W. Edward, HEATHER, L. M. Sandy [2017] Digital Preservation for Libraries, Archive, & Museums, 2. Ed., Rowman & Littlefield, 2017.

CRUPI, Giancarlo [2015] Biblioteca digitale, in “Biblioteche e biblioteconomia”, a cura di Giovanni Solimine, Paul Gabriele Weston, Beni culturali 43, Carocci, Roma, 2015.

DIGITALIA [2019] Rivista del digitale nei beni culturali, XIV, I, 2019.

DOOLEY, Jackie, LUCE, Katherine [2010] Taking Our Pulse: The OCLC Research Survey of Special Collections and Archives, Dublin, OH: OCLC Research, 2010

ERWAY, Ricky, [2010] Defining Born Digital, Dublin, OCLC Research, 2010.

FELICIATI, Pierluigi [2010] Gestione e conservazione di dati e metadati per archivi: quali standard?, in Pigliapoco S. “Conservazione digitale”, EUM, 2010.

FRANZESE, Paolo [2002] (traduzione di) Archiviazione e conservazione delle risorse digitali, 2002.

GIORDANO, Tommaso (2018) Quando le carte diventano bit, in “Biblioteche oggi”, vol 36, 2018.

GUERCIO, Maria [2013], *Conservare il digitale: principi, metodi e procedure per la conservazione a lungo termine di documenti digitali*, Roma-Bari, Laterza, 2013.

GUERCIO, Maria [2016] *Archivi personali: la sfida del digitale. Una riflessione*, in “Allegrezza e Gorgolini”, 2016.

GUERRINI, Mauro, GIGLIO, Daniela [2006] *Conservazione del digitale*, Firenze, Università di Firenze.

HALBERT, Martin, [2010] *Overview of Distributed Digital Preservation Networks*, in *A Guide to distributed digital preservation*, Atalanta, Educopia Institute, 2010.

HEDSTROM, Margaret, MONTGOMERY, Sheon [1998] *Digital preservation needs and requirements in RLG member institution*, December 1998.

INGEBORG, Verheul, [2006] *Networking for Digital Preservation: Current Practice in 15 National Libraries*, Berlin- Munich, IFLA Publications 119, 2006.

KIRSCHENBAUM, Matthew [2008] *Mechanisms: New Media and the Forensic Imagination*, Londra, The Mit Press, 2008.

KUNZE, et al. [2019], *The BagIt File Packaging Format (V1.0)*, 2019.

LAVOIE, Brian [2016] *Il modello di riferimento per un Sistema informativo aperto per l'archiviazione*, a cura di Paul Gabriele Weston. Tradotto da Flavia Cancedda et al., Roma, Associazione Italiana Biblioteche, 2016.

MARZANO, Gilberto [2011] *Conservare il digitale. Metodi, norme, tecnologie*. Milano, Editrice bibliografica, 2011.

MC. GARRIGLE, Conor [2015] *Preserving born digital art: lessons from artists practice*. New review of information networking, Published online: 17 Dec 2015.

MICHETTI, Giovanni [2008] *Il modello OAIS*, in “DigItalia”, 1/2008, ICCU, 2008.

PIGLIAPOCO, Stefano [2007], *Le fasi del processo di conservazione digitale*, in “Archivi & Computer”, I, anno XVII, S. Miniato, Titivillus Edizioni, 2007.

PIGLIAPOCO, Stefano [2010] (a cura di) *Conservare il digitale*, Macerata, EUM, 2010.

PIGLIAPOCO, Stefano [2014] *La conservazione delle memorie digitali*, in GIUVA Linda-GUERCIO Maria, *Archivistica: teoria, metodi, pratiche*, Roma, Carocci, 2014.

- REDWINE, Gabriela [2015] Personal Digital Archiving, in “Digital Preservation Coalition”, UK, DPC Technology Report 15-01 December 2015.
- REDWINE, et.al. [2013] Born Digital: Guidance for Donors, Dealers, and Archival Repositories, 2013.
- RIDI, Riccardo [2007] La biblioteca come ipertesto, Milano, Editrice Bibliografica, 2007.
- RIDI, Riccardo [2010] Il mondo dei Documenti: cosa sono come valutarli e organizzarli, Roma, Laterza, 2010.
- ROBERT W. WOODRUFF LIBRARY [2007], MetaArchive, Distributed Digital Preservation Workshop, Emory University, Atlanta-Georgia, 2007.
- ROTHENBERG, Jeff [2000] An experiment in using emulation to preserve digital publication, RAND Europe, 2000.
- SANTORO, Michele [2001] Conservazione del digitale: definizioni, strategie, approcci, in “Biblioteche oggi”, XIX, 2001.
- SINIBALDI, Alessandro, BUONGIORNO, P. Bartolomeo [2012] Manuale di Conservazione digitale, Milano-Roma, Franco Angeli, 2012.
- STEFANI, Michele [2017] Gli archivi degli scrittori nativi digitali: Verso la conservazione e oltre, Venezia, Università Cà Foscari Venezia, 2017.
- VRANA, Radovan [2010] Public libraries in Croatia and the digitization challenge, in “Library Review”, 59 (2).
- WESTON, P. Gabriele [2013] (a cura di) Il libro, gli archivi e la memoria digitale, Milano, Edizioni Unicopli, 2013.
- ZHANG, Allison, GOURLEY, Don [2008] Creating digital collections: a practical guide, Oxford, Chandos, 2008.

SITOGRAFIA (ultima verifica il 20/03/2021)

- FONTI:

<https://www.dpconline.org/docs/digital-preservation-handbook/299-digital-preservation-handbook/file>

https://en.wikipedia.org/wiki/Digital_preservation

<https://docs.italia.it/AgID/documenti-in-consultazione/lg-documenti-informatici-docs/it/bozza/conservazione.html#sistema-di-conservazione>

<https://www.dpconline.org/handbook/digital-preservation/preservation-issues>

<https://www.dpconline.org/handbook/technical-solutions-and-tools/cloud-services>

<https://www.oclc.org/content/dam/research/publications/library/2010/2010-11.pdf>

<https://www.dpconline.org>

http://www.soprintendenzaarchivistica Toscana.beniculturali.it/fileadmin/risorse/materiali/Guercio_-_Caratteri_documento_digitale.pdf

<https://www.archives.gov/preservation/electronic-records>

<https://www.ifla.org/preservation-and-conservation>

<http://www.loc.gov/standards/premis/v2/premis-report-2-1.pdf>

<https://www.clarin.eu/sites/default/files/preservation-CLARIN-ShortGuide.pdf>

<https://e->

l.unifi.it/pluginfile.php/227895/mod_resource/content/1/ConsevazioneDigitale_Guerrini_Giglio.pdf

<https://bitcurator.net/wp-content/uploads/sites/1099/2018/08/bitstreams-to-heritage.pdf>

<https://www2.archivists.org/standards/external/91>

<http://www.archivejournal.net/essays/disrespect-des-fonds-rethinking-arrangement-and-description-in-born-digital-archives/>

https://cdrh.unl.edu/articles/best_practices

https://www.iasa-web.org/sites/default/files/downloads/publications/TC03_Italian.pdf

<https://www.oclc.org/content/dam/research/activities/hiddencollections/borndigital.pdf>

http://www.aserl.org/wp-content/uploads/2013/04/Intro_DP_2013-3_Born_Dig_Docs.pdf

https://dcs.library.virginia.edu/files/2013/02/AIMS_final_text.pdf

<https://publications.arl.org/Managing-Born-Digital-Special-Collections-and-Archival-Materials-SPEC-Kit-329/>

https://www.agid.gov.it/sites/default/files/repository_files/manuale della conservazione_v6.4.pdf

https://www.digitalpreservation.gov/meetings/documents/ndiipp14/Developing_Born-Digital_Preservation-Workflow_20140723_Donovan.pdf

<https://phaidra.cab.unipd.it/static/formati-file-compresso.pdf>

<http://www.digitalpreservation.gov/>

<http://blogs.loc.gov/digitalpreservation/>

<http://www.clarin-it.it/it/content/archivio-vivo>

<http://www.soprintendenzaarchivistica toscana.beniculturali.it/index.php?id=304>

<https://www.aisoitalia.org/lassociazione/presentazione/>

https://www.academia.edu/37692678/Digitale_dautore_Un_archivio_possibile

https://convegni.unicatt.it/aiucd-Weston_et_al.pdf

www.francobuffoni.it

<http://eprints.rclis.org/13657/1/atti-testoFeliciati09.pdf>

https://www.iccu.sbn.it/export/sites/iccu/documenti/capitolato_tecnico_digitalizzazione.pdf

http://eprints.rclis.org/5216/1/metadati_torino05ott2004.pdf

<https://www.aib.it/aib/commiss/cnur/dltberga.htm>

<http://win.unsabbeniculturali.it/Bibliotecari%20%20Materiale%20didattico/Magliano.pdf>

<http://www.sa-toscana.beniculturali.it/index.php?id=172>

<http://www.ukoln.ac.uk/services/papers/bl/rdr6238/paper.html#foot>

<http://localwww.math.unipd.it/~derobbio/preserv-2003-10-28.htm>

<http://eprints.relis.org/10434/1/Web%26DigMemoryPreserv2007.pdf>

<http://files.archivists.org/conference/nola2013/materials/701-TimmsA.pdf>

<http://www.cdncouncilarchives.ca/archdesrules.html>

https://www2.archivists.org/sites/all/files/Hagley_Digital_Proposed_CMP.pdf

<https://www.tandfonline.com/doi/pdf/10.1080/23257962.2016.1144504?needAccess=true>

<https://library.stanford.edu/research/digitization-services>

<https://clir.wordpress.clir.org/wp-content/uploads/sites/6/pub159.pdf>

<http://www.internetculturale.it/getFile.php?id=44411>

<https://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf>

https://www.webarchive.org.uk/wayback/archive/20140613220103/http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_p1finalreport.pdf

<https://public.ccsds.org/Pubs/651x0m1.pdf>

https://www.agid.gov.it/sites/default/files/repository_files/leggi_decreti_direttive/dpcm_3-12-2013_conservazione.pdf

http://www.jisc.ac.uk/publications/publications/pub_digipreservationbp.aspx

- SITI STANDARD:

OAIS: <https://public.ccsds.org/pubs/650x0m2.pdf>

PREMIS: <https://www.loc.gov/standards/premis/>

<https://www2.archivists.org/groups/standards-committee/preservation-metadata-implementation-strategies-premis>

INTERPARES:

http://www.interpares.org/display_file.cfm?doc=ip2_dissemination_ws_duranti_venice_2006b.pdf

<http://www.interpares.org/ip2/book.cfm>

http://www.interpares.org/book/interpares_book_f_part3.pdf

http://www.interpares.org/book/interpares_book_g_part4.pdf

[http://www.interpares.org/public_documents/ip2\(pub\)creator_guidelines_booklet.pdf](http://www.interpares.org/public_documents/ip2(pub)creator_guidelines_booklet.pdf)

<https://interparestrust.org>

MAG: <https://www.iccu.sbn.it/export/sites/iccu/documenti/manuale.html#standard>

<https://www.iccu.sbn.it/export/sites/iccu/documenti/Manuale.pdf>

OAI-ORE: <https://www2.archivists.org/groups/standards-committee/open-archives-initiative-object-re-use-and-exchange-oai-ore>

METS: <https://www.loc.gov/standards/mets/>

DUBLIN CORE: <http://www.internetculturale.it/getFile.php?id=44411>

- BUONE PRATICHE:

<https://www.minervaeurope.org/bestpractices/listgoodpract.htm>

https://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf

https://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf

<https://www.ninch.org/programs/practice/index.html>

<https://www.niso.org/sites/default/files/2017-08/framework3.pdf>

<http://born-digital-archives.blogspot.com/2012/01/aims-white-paper-now-available.html>

<http://www.internetculturale.it/getFile.php?id=44396>

<https://bibliothecae.unibo.it/article/view/7027>

[Buone pratiche di storia orale](#)

[Buone pratiche di storia orale. Questioni etiche, deontologiche, giuridiche](#)

- LINEE GUIDA:

<http://www.digitizationguidelines.gov/about/>

<https://www.nationalarchives.gov.uk/documents/information-management/digital-preservation-policies-guidance-draft-v4.2.pdf>

http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Intro_090915.pdf

http://www.digitizationguidelines.gov/guidelines/FADGI_BDV_p1_20141202.pdf

<http://www.internetculturale.it/getFile.php?id=44418>

<http://www.internetculturale.it/getFile.php?id=44398>

<http://www.internetculturale.it/getFile.php?id=44395>

<http://www.internetculturale.it/getFile.php?id=44397>

PARADIGM: <https://wayback.archive-it.org/org-467/20161101115538/http://www.paradigm.ac.uk/workbook/index.html>

<https://www.ifla.org/publications/guidelines-for-digitization-projects-for-collections-and-holdings-in-the-public-domain>

<https://phaidra.cab.unipd.it/static/guida-all-archiviazione.pdf>

- VIDEO DI APPROFONDIMENTO:

<https://www.youtube.com/watch?v=JmbJvuyO8UQ>

https://www.youtube.com/watch?v=X458qsSc6sM&list=PLd5bJJul8c5pvRcwmC3G5_mRhXQZfK7Sx

<https://www.youtube.com/watch?v=yJD9hY77nSE>

<https://archive.org/details/pda2017>