

«La goccia pronta per il mappamondo»: esperienze di Phaidra (e dintorni)

Lorisa Andreoli, Cristiana Bettella, Linda Cappellato, Yuri Carrer, Gianluca Drago, Giulio Turetta

Centro di Ateneo per le Biblioteche – Università degli Studi di Padova

{lorisa.andreoli, cristiana.bettella, linda.cappellato, yuri.carrer, gianluca.drago, giulio.turetta}@unipd.it

Introduzione

Phaidra¹, acronimo di *Permanent Hosting, Archiving and Indexing of Digital Resources and Assets*, è la piattaforma del Sistema Bibliotecario dell'Università di Padova per l'archiviazione a lungo termine di oggetti e collezioni digitali².

Il presente contributo intende illustrare Phaidra come caso studio secondo la convergenza di un triplice punto di vista, qui tematicamente declinato in tre paragrafi, *Rappresentare/descrivere – Esporre/guardare – Dimostrare/provare*, ciascuno dei quali è dedicato a temi di essenziale rilevanza e strutturante attualità come la rappresentazione digitale dei dati culturali, la loro manifestazione e garanzia di persistenza nel tempo.

L'approccio perseguito vorrebbe porre in evidenza l'empiria metodologica adottata unitamente all'insieme delle esperienze realizzate nell'arco di questo primo decennale Phaidra... quasi come una «goccia pronta per il mappamondo»³.

1 Rappresentare/descrivere: sul modellare i dati Phaidra

L'oggetto digitale è una entità complessa rappresentata da un *datastream*, suscettibile di citazionabilità e identificazione grazie alla dotazione di un identificatore persistente, la cui pregnanza identitaria è stabilita dall'insieme di proprietà descritte dai metadati. Stando a questa definizione, la connotazione dell'oggetto digitale Phaidra⁴ è rappresentata dalla aggregazione di contenuto o dato e formato di file, evidenza delle sue proprietà registrate nello schema di metadati, identificazione permanente⁵ ed esplicitazione dei termini di utilizzo, là dove la sua denotazione è espressa sia da un sistema di relazioni inter-entità⁶ sia dall'insieme aggregato di oggetti nell'oggetto «collezione»⁷.

¹<<https://phaidra.cab.unipd.it/>>.

² Phaidra nasce nel 2008 all'Università di Vienna. A partire dal 2010 l'Università di Padova ha avviato la collaborazione per il suo utilizzo e sviluppo per cui è approvata, nel febbraio dello stesso anno, la Convenzione con l'Università di Vienna da parte del Comitato Tecnico Scientifico del Centro di Ateneo per le Biblioteche. Si veda: L. Tallandini - L. Andreoli - E. Bianchi - L. Cappellato - Y. Carrer - G. Drago - G. Turetta - A. Zane, *Phaidra, un archivio digitale FAIR per la disseminazione e l'accesso integrato a testi, testimonianze, immagini e storie del patrimonio culturale*, «Digitalia», 1, 2019, pp. 147–157, <<http://digitalia.sbn.it/article/view/2281>> e *Informazioni* in <<https://phaidra.cab.unipd.it/info/impressum>>.

³ «La goccia pronta per il mappamondo» è capoverso di Milo De Angelis tratto dalla raccolta *Millimetri*, Milano, Il Saggiatore, 2013, p. 23.

⁴ <<https://wiki.lyrasis.org/display/FEDORA38/Fedora+Digital+Object+Model>> con riferimento alla comunità di sviluppo Fedora di cui Phaidra è adozione.

⁵ Con sistema di identificazione permanente Handle <<https://www.handle.net/>>.

⁶<<https://github.com/phaidra/phaidra-api/wiki/Relations>>.

⁷ Per la proprietà bidirezionale Phaidra hasPart/isPartOf. Una dimostrazione visuale dell'oggetto collezione è *infra* § 2. *Esporre/guardare: sul manifestare e raccontare Phaidra*.

Questa complessità aggregativa⁸ ascrive l'oggetto digitale Phaidra nella più ampia accezione di Oggetto Culturale, o *Cultural Heritage Object*, di cui la eterogeneità delle provenienze e tipologia di contenuti immessi (*ingestion*)⁹ sono la manifestazione del patrimonio culturale dei fornitori di contenuto Phaidra (*content provider*)¹⁰.

La duplice valenza fisica e digitale dell'oggetto culturale determina rispettivamente il rappresentato e il rappresentante, stabilendo una sorta di grammatura fusionale che realizza nella effettualità della rappresentazione digitale la ri-scoperta e restituzione dell'*hic et nunc* memorizzato del cosa vi è rappresentato, attestando al contempo la sua autenticità. Da questo punto di vista il rappresentante è dato dal tipo di risorsa informativa – testo, immagine, video, audio – e tipo di supporto o *media type*¹¹, veicolo del complesso artefatto culturale rappresentato quale può essere un codice miniatu, un libro antico, una pergamena, un ex libris, una dedica manoscritta, un carteggio, una mappa, una tavola parietale scientifica, una illustrazione botanica, un erbario, un modello matematico, un reperto archeologico, una fotografia in ogni sua tecnica di riproduzione, un video, un learning object¹².

L'iscrizione informativa delle componenti del binomio rappresentante/rappresentato costituisce le proprietà dell'oggetto culturale descritte secondo uno schema, o modello, di metadattazione la cui funzione primaria traduce in sostanza l'identità codificata dell'oggetto. Il modello di dati (*data model*)¹³ informa il *design* o struttura concettuale del dato dal duplice punto di vista della sintassi e semantica che esso assume rispetto al contesto informativo di riferimento, tuttavia potenzialmente applicabile anche a contesti informativi differenti e per differenti finalità informative (*application profile*). Il modello di dati fondativo di Phaidra, Universität Wien metadata o UWmetadata, ne è un esempio essendo il risultato della espansione dello standard Learning Object Metadata, LOM, (1484.12.1-2002 IEEE Standard for Learning Object Metadata) e della giuntura combinatoria di categorie ed elementi derivati o personalizzati di schemi di metadati diversi¹⁴.

La conformità all'essere modello, che ha esito normativo e certificabile di standard, espone il modello di dati a possibili processi di negoziazione concettuale, operati da linguaggi e procedure di trasformazione (*conversion*)¹⁵ sulla base di una tavola di corrispondenze identificate con relazioni inter-schemiche, nota anche con il termine di mappatura (*crosswalk*)¹⁶. Già altrove¹⁷ ci siamo icasticamente riferiti ai

⁸ Intesa dal punto di vista dell'oggetto come unità aggregativa primaria, della collezione come aggregazione di dati e per esteso di Phaidra come aggregatore, ovvero una organizzazione che raccoglie, aggredisce, formatta e amministra i (meta)dati provenienti da molteplici fornitori di contenuto, con duplice funzione di fornitore di servizi e fornitore di dati rispetto a fornitori di servizi o aggregatori come per esempio Europeana <<https://www.europeana.eu/portal/it>> e CulturaItalia <<http://www.culturaitalia.it/>>.

⁹ Indica il processo di immissione e assorbimento dei dati dei fornitori di contenuto in termini di acquisizione, mappatura e conversione se richiesto, e pubblicazione nel sito.

¹⁰ Biblioteche, musei, archivi, dipartimenti, centri e uffici dell'Università di Padova, le Università Ca' Foscari e IUAV di Venezia (dal 2014), il Sistema Museale dell'Università di Bologna (dal 2019), come altre istituzioni cooperanti.

¹¹ <<https://www.iana.org/assignments/media-types/media-types.xhtml>> e *Formati dei file raccomandati per l'archiviazione a lungo termine e per la disseminazione web in Phaidra* (maggio 2019), a cura di G. Drago, <<http://phaidra.cab.unipd.it/static/formati-file-compresso.pdf>>.

¹² Esemplicano le tipologie di oggetto culturale Phaidra sia esso digitale nativo o surrogato digitalizzato di originali analogici.

¹³ Su *data model* almeno: *DFT 2: Analysis & Synthesis*, in G. Berg-Cross - R. Ritz - P. Wittenburg, *Data Foundation and Terminology Work Group Products* (2015, August 3), <<https://dx.doi.org/10.15497/06825049-8CA4-40BD-BCAF-DE9F0EA2FADF>> e F. Jannidis - J. Flanders, *A gentle introduction to data modeling*, in *The shape of data in the digital humanities: modeling texts and text-based resources*, edited by Julia Flanders and Fotis Jannidis, New York, Routledge, 2019, pp. 26-95.

¹⁴ *University of Vienna Metadata (UWmetadata)*, 2006, <<http://phaidra.cab.unipd.it/static/phaidra-uwmetadata.pdf>>.

¹⁵ Realizzata da programmi, *parser*, scritti con linguaggi di programmazione o con linguaggio di trasformazione XSLT <<https://www.w3.org/TR/xslt/all/>>. MINT Services della National Technical University of Athens ne è un esempio come piattaforma di riferimento per la mappatura e conversione di metadati eterogenei conforme al modello di dati *Europeana Data Model*, utilizzata anche per la pubblicazione delle collezioni digitali Phaidra in Europeana nell'ambito del progetto europeo Linked Heritage 2011-2013.

¹⁶ Mappature e *parser* sono stati elaborati *ad hoc* per la conversione di metadati UNIMARC, ICCD e archivistici finalizzata all'importazione e creazione in Phaidra di collezioni digitali di libri antichi o rari e manoscritti, oggetti museali e archivi personali.

¹⁷ C. Bettella - A. Bellotto, *Metadata as Semantic Palimpsests: The Case of PHAIDRA@unipd*, in *Digital Libraries: Supporting Open Science. IRCDL 2019*, eds. P. Manghi - L. Candela - G. Silvello, Springer, Cham, 2019, <https://doi.org/10.1007/978-3-030-00000-0_10>.

metadati con l'espressione di «palinsesti semanticci», traslando in immagine la proprietà citazionale, insita in ogni processo di mappatura, della relazione inter-schematica, la quale si svela nella trasformazione ed esposizione degli elementi di destinazione prescelti per i corrispettivi elementi sorgenti. Esemplare, da questo punto di vista, è l'operazione di mappatura e conversione realizzata con la generazione del profilo Dublin Core Phaidra¹⁸ da UWmetadata, con la duplice finalità di rappresentare il contenuto informativo dell'oggetto culturale Phaidra nel sito web¹⁹ e di esporre i suoi metadati nel Data Provider OAI²⁰. Si è trattato di fatto di un'azione congiunta di *data modeling* e *data visualisation* che ha consentito sia di delineare con maggiore precisione l'identità bidimensionale rappresentante/rappresentato dell'oggetto Phaidra, sia di qualificare, anche sul piano formale, le classi di Persone, Opere, Spazio, Tempo, che determinano la specificità identitaria e ontologica di ogni oggetto culturale e l'autenticità della sua provenienza.

Gli effetti di tale qualificazione rappresentativa e descrittiva sono molteplici. Sottopone anzitutto ogni elemento costitutivo a un processo di discrezione semantica in grado di distinguere la granularità informativa dell'oggetto – in termini di supporto, materia e tecnica, tipo e genere, deissi spazio-temporale e autorialità – rispetto alla semantica del suo soggetto. Parallelamente, favorisce la modellazione e capacità adattativa del modello di dati nei confronti degli aggregatori di oggetti culturali di riferimento per Phaidra, come CulturaItalia e Europeana, per i quali lo scambio dei dati²¹ è governato da requisiti precisi sia dal punto di vista della qualità dell'oggetto digitale sia da quello dei suoi metadati²². Da ultimo, predispone lo stesso modello di dati alla riscrittura RDF (*Resource Description Framework*)²³ della sua struttura secondo un modello a grafo oggetto-centrico, pluridimensionale e relazionale, dove ogni elemento è il predicato dato da un'ontologia di riferimento²⁴, liberandone il potenziale connettivo e di espansione verso la nuvola dei dati aperti e collegati²⁵ e conseguentemente dischiudendosi a quella eterogenesi del riuso che può dare adito a forme di aggregazione inedite ed eterotopiche o alla sorpresa serendipitale del riconoscimento dell'oggetto perduto da parte del *flaneur*²⁶ nel suo viaggio atemporale del web.

Le ragioni che possono addurre un modello datico a un processo di conversione verso un modello altro, se non a grafo, rispondono essenzialmente al principio della autonomia del significante che è proprio della interoperabilità, ovvero la capacità di un sistema di rendersi comunicabile, aperto, conformandosi alla lettura a distanza (*distant reading*) e all'apprendimento da parte di sistemi intelligenti (*machine readable* o *deep learning*). Proiettando un Phaidra del futuro potrà allora essere possibile rivelare l'evoluzione delle forme della comunicazione scientifica nel tempo attraverso i carteggi dei suoi protagonisti, visualizzare la matericità di *specimina*, ricostruire le interconnessioni profonde tra storia dell'arte e storia dell'illustrazione scientifica, ripercorrere a ritroso la genesi di una scoperta.

2 Esporre/guardare: sul manifestare e raccontare Phaidra

Biblioteche, archivi e musei compiono enormi fatiche nella creazione di biblioteche digitali. Si applicano

11226-4_14>.

¹⁸ C. Bettella, *PHAIDRA_DC Metadata Element Set*, version 1.1., 2018, <https://phaidra.cab.unipd.it/static/phaidra_dc-metadata-element-set.pdf>.

¹⁹ <<http://phaidra.cab.unipd.it/static/campi-di-phaidra.pdf>>.

²⁰ <<https://fc.cab.unipd.it/oaireader/?verb=Identify>>.

²¹ *Europeana Data Exchange Agreement (DEA) e Protocollo d'intesa per l'adesione al Portale della Cultura Italiana*.

²² Secondo il quadro illustrato in *Europeana Publishing Framework* <<https://pro.europeana.eu/post/publishing-framework>> e che CulturaItalia recepisce come aggregatore certificato di Europeana.

²³ <<http://www.w3.org/TR/2014/REC-rdf11-concepts-20140225/>> e, con riferimento a Phaidra, *Introduction to RDF*, a cura dell'Università di Vienna, <<https://github.com/phaidra/phaidra-ld/wiki/Introduction-to-RDF>>.

²⁴ <<https://github.com/phaidra/phaidra-ld/wiki/Metadatas-fields>>.

²⁵ <<https://lod-cloud.net/clouds/lod-cloud.svg>> e *infra* § 3. *Dimostrare/provare: sul certificare Phaidra*.

²⁶ *Infra* § 2. *Esporre/guardare*, cit.

alla digitalizzazione del patrimonio culturale, curando la qualità dei contenuti e fornendo metadati ricchi, esponendo generosamente in web le proprie collezioni per renderle accessibili a tutti, non solo allo studioso, al quale di norma viene dedicato lo sforzo progettuale, ma anche al visitatore occasionale. Oggi è anche parte della Terza Missione degli atenei²⁷.

A quest'ultima fase del processo, cioè la disseminazione online dell'informazione, deve essere dedicata la stessa cura già applicata alle fasi precedenti.

Dobbiamo attentamente figurarci il nostro visitatore. Quando approda alle nostre biblioteche digitali, proveniente dai percorsi di navigazione più disparati, gli si possono presentare due possibili esperienze tra loro contrastanti: l'emozione di esplorare un territorio nuovo che promette scoperte imprevedibili, oppure la frustrante sensazione di disorientamento di fronte alla soverchiante quantità di informazioni. Noi vogliamo che la prima delle due esperienze non sia annebbiata dalla seconda.

Non sappiamo a priori chi sia il nostro visitatore, ma vogliamo che Phaidra si apra a tutti, anche a quel particolare tipo di viaggiatore, intellettualmente curioso ma non necessariamente uno studioso – che è stato acutamente chiamato *information flaneur*²⁸ – che ama girovagare senza una meta predefinita tra le informazioni disponibili e farsi sorprendere da ciò che trova. Programmaticamente, figuriamoci che sia questo il nostro «utente di elezione»; a noi, progettisti del territorio che egli viene ad esplorare, spetta il compito di disegnargli lo spazio – con le vie di comunicazione principali e secondarie, i vicoli, i *passages* – in modo che egli possa cogliere chiaramente la topografia del luogo, invitandolo così all'esplorazione, e facilitandogli la comprensione dei singoli elementi costitutivi (suggerendo soltanto le possibili destinazioni, per non togliere respiro alla serendipità del suo vagabondare), per fornirgli poi gli strumenti di lettura che consentono la piena fruizione delle evidenze culturali, nelle loro differenti specificità.

Il primo impatto con una biblioteca digitale è sempre visuale. È quindi importante prendersi cura delle immagini e degli altri contenuti multimediali: che siano di buona qualità, che siano rappresentativi, che descrivano esaustivamente la natura, la varietà e la complessità dei contenuti, presentati al visitatore per gradi, un passo alla volta, senza soverchiarlo di informazioni. In Phaidra partiamo dalle miniature in home page che rappresentano solo una selezione di collezioni, attraversiamo la pagina che raccoglie tutte le collezioni principali²⁹, ci soffermiamo nelle pagine che introducono alla singola collezione, per arrivare infine alla descrizione analitica dei singoli oggetti. Cerchiamo di invitare con garbo il nostro visitatore a intraprendere un proprio viaggio tra i contenuti, che costruirà da sé passo passo sulla base delle nostre suggestioni, spostandosi orizzontalmente tra collezioni e oggetti, calandosi in profondità nell'esplorazione di un soggetto specifico, oppure utilizzando la ricerca a faccette per trovare, filtrare, combinare³⁰.

Uno degli aspetti che caratterizzano Phaidra sono le pagine di introduzione alle collezioni più significative. Qui è possibile utilizzare tutta la gamma dell'armamentario web (formattazione, contenuti multimediali, collegamenti esterni) e si possono ideare piccole narrazioni che fanno da collante tra gli oggetti e ne evidenziano caratteristiche e peculiarità, estendendosi anche a documentare contesti culturali di ampio respiro, ad esempio la nascita della medicina moderna a Padova³¹, o la storia della botanica attraverso i suoi protagonisti³².

In una biblioteca digitale che include contenuti eterogenei – oggetti museali, antichi manoscritti, libri a stampa di epoca moderna, mappe, documenti nati digitali, fotografie – occorrono strumenti differenti

²⁷ M. Cassella, *La Terza missione dell'università: cos'è, come si valuta (con un'appendice sulle biblioteche)*, «Biblioteche oggi», 3, 2 (2017), <<http://dx.doi.org/10.3302/2421-3810-201702-059-1>>.

²⁸ M. Dörk - S. Carpendale - C. Williamson, *The information flaneur: a fresh look at information seeking*, «Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)», 2011, <<https://doi.org/10.1145/1978942.1979124>>.

²⁹ <<https://phaidra.cab.unipd.it/collections/featured>>.

³⁰ <<https://phaidra.cab.unipd.it/search>>.

³¹ <https://phaidra.cab.unipd.it/collections/scuola_medica_padovana>.

³² <https://phaidra.cab.unipd.it/collections/iconoteca_botanici>.

in grado di rappresentare nel mondo virtuale le specificità materiali degli oggetti originali.

In Phaidra, col Book Viewer, anche un manoscritto illustrato del XVII secolo può essere meglio compreso nella sua distintiva materialità: mentre lo si sfoglia si rivive virtualmente l'esperienza emozionante di aprire quelle pagine come se fossimo in una biblioteca storica³³.

Allo stesso tempo, con l'Image Viewer³⁴, uno strumento di visualizzazione che consente di ingrandire un'immagine fino ai dettagli più minimi, una mappa può essere fruibile come se l'avessimo davanti, sulla nostra scrivania³⁵.

Phaidra è una piccola goccia, ma si sente pronta per il mappamondo. Anzi, non si limita a occupare il suo spazio, ha la pretesa di volersi riverberare anche in territori altrove. Così succede che parte delle sue collezioni migrino in CulturaItalia e in Europeana³⁶, che si gettino nel gran calderone di Internet Archive³⁷ o che si allarghino in vere e proprie mostre virtuali in Movio³⁸, dove tanta ricchezza di contenuti diventa materia prima per approfondimenti, ampliamenti, narrazioni³⁹.

Organizzare, raggruppare, collegare tra loro le cose del mondo, indicare percorsi, allargarsi a promuovere nuove vie di accesso; perché, vuoi che siamo studiosi oppure *information flaneur*, noi siamo gli affezionati visitatori che ritornano, fiduciosi di poter ritrovare i luoghi già amati, «quel silenzio frontale dove eravamo già stati»⁴⁰, – la cui esistenza è garantita dall'apertura dei dati, da identificatori permanenti, da licenze chiare, dalla certificazione di Phaidra – per farci sorprendere dalle nuove prospettive di un territorio sempre in formazione.

3 Dimostrare/provare: sul certificare Phaidra

Il tema dell'affidabilità dei repository inizia ad essere affrontato dalla comunità internazionale negli anni '90, ma è con la pubblicazione nel 2010 del *Memorandum of Understanding to Create a European Framework for Audit and Certification of Digital Repositories*⁴¹ che vengono individuati tre livelli di certificazione – *basic, extended, formal* – raggiungibili tramite un percorso graduale: il livello base inizia con l'autovalutazione sulla base dei requisiti previsti da Data Seal of Approval (DSA), ora CoreTrustSeal (CTS); il livello esteso prosegue con un'autovalutazione più approfondita basata sugli standard ISO 16363 o DIN 31644; il livello formale si conclude con una valutazione effettuata con una verifica esterna sulla base degli standard ISO 16363 o DIN 31644.

Nel novembre 2018 la Commissione Europea ha pubblicato il rapporto *Turning FAIR into reality* che delinea un piano d'azione ricco di proposte tecniche e culturali sulla gestione dei dati aperti e descrive i cambiamenti necessari per l'attuazione dei principi FAIR (*Findable, Accessible, Interoperable, Reusable*)⁴². Tra le raccomandazioni del rapporto vi è lo sviluppo di sistemi di valutazione per certificare i servizi FAIR e il deposito degli oggetti digitali FAIR in repository affidabili per i quali si suggerisce la certificazione CoreTrustSeal (CTS).

³³ Come il manoscritto quattrocentesco *De herbarum virtutibus* in <<https://phaidra.cab.unipd.it/o:248201>>.

³⁴ Utilizza il software IIPIImage <<https://iipimage.sourceforge.io/>> che supporta il protocollo IIIF <<https://iiif.io/>>.

³⁵ Si veda il *Mappamondo Borgiano* <<https://phaidra.cab.unipd.it/imageserver/o:330988>>.

³⁶ In Europeana disponibili in <https://www.europeana.eu/portal/it/search?q=DATA_PROVIDER%3A%22Universit%C3%A0+degli+Studi+di+Padova+I+University+of+Padua%22&view=grid> a cui andranno ad aggiungersi le collezioni Phaidra di cui è in corso la pubblicazione in CulturaItalia.

³⁷ <<https://archive.org/details/universityofpadovalibrarysystem>>.

³⁸ <<https://www.movio.beniculturali.it/>>.

³⁹ Si veda per esempio la mostra dedicata al prefetto dell'Orto Botanico di Padova Giovanni Marsili in <<http://mostre.cab.unipd.it/marsili>> e <<http://mostre.cab.unipd.it/index>> per un elenco delle mostre realizzate dal Sistema Bibliotecario.

⁴⁰ Milo De Angelis, *op. cit.*, p. 23.

⁴¹ <<http://www.trusteddigitalrepository.eu/Memorandum%20of%20Understanding.html>>.

⁴² I principi FAIR sono uno degli otto pilastri su cui si fondano i principi dell'*Open Science* secondo quanto stabilito dalla Commissione europea.

CTS nasce nel 2018 dalla fusione di Data Seal of Approval (DSA) e World Data System (WDS) e fornisce la certificazione di base, cioè rappresenta il passaggio iniziale e obbligato senza il quale non si può giungere ai successivi livelli.

CTS definisce i repository come «Organizations that preserve, manage, and provide access to many types of digital materials in a variety of formats. Materials in online repositories are curated to enable search, discovery, and reuse. There must be sufficient control for the digital material to be authentic, reliable, accessible and usable on a continuing basis»⁴³. CTS è di cruciale importanza perché, con criteri concordati a livello internazionale, attesta che i dati archiviati siano ben strutturati, sicuri, accessibili e conservabili nel lungo periodo, adattandosi alle iniziative emergenti relative a *Open Science, Research Data Management* e dati FAIR che operano sotto l'ombrelllo di *European Open Science Cloud* (EOSC).

Nel novembre 2019 Phaidra ha ottenuto la certificazione CoreTrustSeal⁴⁴, a seguito di un processo durato quasi diciotto mesi guidato da due revisori esterni CTS e condotto da una bibliotecaria e da un informatico con la collaborazione dello staff dell'Ufficio Biblioteca digitale del Sistema Bibliotecario⁴⁵. Il processo di autovalutazione del repository è più complicato di quel che sembra, perché le richieste dei revisori possono essere molto specifiche e soprattutto sono mirate alla ricerca di un quadro di coerenza generale rispetto ai requisiti.

La certificazione CTS⁴⁶ è formata da sedici requisiti obbligatori raggruppati in quattro sezioni: *Background information, Organizational Infrastructure, Digital Object Management e Technology*. Di seguito evidenziamo le specificità di ciascuna sezione e le criticità incontrate in corso d'opera.

Background information. Questo requisito è il presupposto di partenza per poter rispondere ai sedici requisiti successivi. Riguarda la descrizione del repository, richiedendo quindi una buona conoscenza dei contenuti, e la definizione della sua comunità di riferimento⁴⁷, ovvero i soggetti ai quali l'archivio è rivolto. Inoltre, a seconda della tipologia dei dati e dell'accordo di deposito, deve essere individuato il livello di cura dei dati fornito.

Organizational Infrastructure. Esamina la tenuta e la sostenibilità dell'infrastruttura organizzativa. La *mission* dell'organizzazione deve dichiarare pubblicamente che il repository è responsabile per la cura degli oggetti digitali depositati e che ne garantisce l'autenticità, la durata e l'accessibilità nel tempo. Questa sezione descrive le condizioni d'uso del repository e gli aspetti relativi all'accesso, disponibilità e tracciabilità dei dati in conformità con le questioni legali e le norme etiche e disciplinari (ad esempio, paternità degli oggetti digitali, privacy, licenze d'uso). L'infrastruttura organizzativa è sostenibile e implementabile se è governata da un piano di continuità di accesso e conservazione dei dati nel medio e lungo periodo e da uno staff di professionisti di cui vanno specificati i profili professionali, le competenze e i compiti (*expert guidance*). Va da sé che il piano di aggiornamento professionale deve essere dichiarato e che deve essere adeguato alle esigenze di mantenimento e sviluppo.

Digital Object Management. Illustra le caratteristiche degli oggetti digitali (integrità e autenticità

⁴³ CoreTrustSeal Standards and Certification Board (2019, November 20), *CoreTrustSeal Trustworthy Data Repositories Requirements: Glossary 2020–2022 (Version v02_00-2020-2022)*, s.v. *Repository*, Zenodo, <<http://doi.org/10.5281/zenodo.3632563>>.

⁴⁴ Si veda la certificazione CoreTrustSeal di Phaidra in

<<https://www.coretrustseal.org/wp-content/uploads/2019/11/Phaidra-at-the-Library-System-of-the-University-of-Padova.pdf>>.

⁴⁵ Per un caso studio sulla certificazione M. Lindlar - F. Schwab, *All That Work ... for What? Return on Investment for Trustworthy Archive Certification Processes – a Case Study*, OSF (20 June 2019), <<https://doi.org/10.17605/OSF.IO/8A3SC>>.

⁴⁶ CoreTrustSeal Trustworthy Data Repositories Requirements in <<https://www.coretrustseal.org/why-certification/requirements/>>.

⁴⁷ «Designated Community: An identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities. A Designated Community is defined by the Archive and this definition may change over time», *CoreTrustSeal Trustworthy Data Repositories Requirements: Glossary 2020–2022*, s.v. *Designated Community*, cit.

dei dati, creazione e controllo di qualità di dati e metadati), i flussi di lavoro del repository (creazione, archiviazione e disseminazione dei dati), la rintracciabilità e il riuso degli oggetti digitali.

Il punto strategicamente più rilevante è il piano di conservazione digitale, la strada da percorrere per preservare i dati nel lungo periodo, il ponte tra l'oggi e il domani. Il piano deve descrivere tutte le fasi del processo di conservazione dei dati, in particolare:

- l'*ingesting* e la relazione tra le modifiche/aggiornamenti dei dati con un sistema di tracciamento
- i sistemi di backup, replica dei dati e gestione del rischio per quanto riguarda lo spazio di archiviazione (con eventuali partnership e collaborazioni)
- la necessaria sicurezza dei dati sia fisica che a protezione da intrusioni via rete
- la sicurezza dei locali e delle procedure adottate nella gestione dei dati
- eventuali note su procedure interne per la gestione dello spazio di archiviazione
- la strategia a lungo termine per la preservazione dei dati (sistemi tecnologici e informatici, gestione dei formati dei file e dei supporti, implementazione di procedure per la gestione sicura e robusta dei dati, efficacia nella continuità di accesso ai dati da parte della comunità di riferimento)
- il controllo della catena di accesso dei dati
- il controllo dell'integrità dei dati
- la continuità nella leggibilità dei dati nel tempo
- la collaborazione con il produttore dei dati sulle questioni etiche e legali
- i termini d'uso
- la gestione del piano stesso nel tempo
- un piano di migrazione basato su una strategia ben delineata e delle procedure di test

Technology. Descrive l'infrastruttura tecnologica coerente con i principi FAIR basata su tecnologie consolidate, software utilizzati e standard riconosciuti a garanzia dell'affidabilità del dato, la tracciabilità delle procedure, la capacità di prevenire potenziali problemi tramite l'analisi dei dati depositati, l'adozione di pratiche di sviluppo attuali e mature.

La crescente popolarità dei dati FAIR funge da volano per le questioni legate all'interoperabilità tra sistemi, alla semanticità dei contenuti e all'utilizzo di ontologie condivise. Partendo da una dotazione di base costituita da un set di API formalizzate e di formati di interscambio aperti e documentati, l'utilizzo di tecnologie di rappresentazione ed esposizione delle API⁴⁸ e di protocolli di disseminazione⁴⁹ consente la «scoperta» di un repository da parte di algoritmi e sistemi «intelligenti», nell'ottica della condivisione e dell'interoperabilità dei dati e dei sistemi.

Concludendo, sulla base di questa esperienza, possiamo affermare che il processo di certificazione è uno strumento prezioso per valutare i punti di forza e di debolezza di un repository, in quanto prende in esame i flussi di lavoro, individua i processi e le attività che necessitano di una migliore articolazione, osserva l'adeguatezza della documentazione nell'ottica della chiarezza, completezza e trasparenza,

⁴⁸ OpenAPI <<https://www.openapis.org/>>.

⁴⁹ Come OAI-PMH <<https://www.openarchives.org/pmh/>> o ResourceSync <<http://www.openarchives.org/rs/>>.

aumenta la visibilità dei dati conformi ai principi FAIR⁵⁰, contribuisce a riconoscere e migliorare la professionalità dello staff. Dimostrando all'esterno che viene rispettato uno standard di qualità e affidabilità, la certificazione crea fiducia negli *stakeholder* (depositanti, comunità di conservazione digitale, enti finanziatori, ecc.) e rafforza la reputazione del repository.

Ringraziamo Elena Canadelli (Università di Padova) per la lettura dedicata al manoscritto e i preziosi suggerimenti.

⁵⁰ Il progetto europeo FAIRsFAIR – *Fostering Fair Data Practices in Europe* <<https://www.fairsfair.eu/the-project>> ha selezionato dodici repository di dati, tra cui Phaidra, che parteciperanno alla progettazione di funzionalità finalizzate ad aumentare l'interoperabilità di un repository.