

La valutazione *FAIRness* di un archivio digitale certificato: tra principi teorici e azioni pratiche

«DigItalia» 1-2022
DOI: 10.36181/digitalia-00043

Cristiana Bettella — Yuri Carrer — Giulio Turetta

Sistema Bibliotecario di Ateneo, Università degli Studi di Padova

Principale obiettivo del presente contributo è quello di offrire un caso di buona pratica di attuazione dei principi FAIR, illustrando la metodologia e le azioni intraprese dal Sistema Bibliotecario di Ateneo dell'Università degli Studi di Padova al fine di rendere affidabile, e con caratteristiche FAIR, Phaidra (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets), l'archivio digitale sviluppato dall'Università di Vienna e, a partire dal 2010, adottato dall'Università di Padova con un accordo di collaborazione e di partecipazione a ulteriori sviluppi, ai fini della gestione, conservazione a lungo termine e disseminazione online di oggetti digitali culturali.

L'adesione ai principi FAIR di un archivio digitale non può prescindere dal riconoscimento formale e pubblico della sua affidabilità. La reputazione di affidabilità è per definizione intrinseca alla sua missione, la quale mira ad assicurare alla propria comunità di riferimento, ora e in futuro, l'accesso sicuro, robusto e a lungo termine delle risorse digitali gestite, con ciò garantendo la propria sostenibilità e durevolezza, sicurezza e preservazione dei dati archiviati nel tempo.

A partire dal processo di revisione critica a cui è stato sottoposto Phaidra allo scopo di ottenere la certificazione CoreTrustSeal (CTS), e sulla base dell'esperienza condotta come archivio pilota e di test nel contesto del progetto europeo FAIRsFAIR — Fostering Fair Data Practices in Europe, saranno illustrate le fasi fondamentali percorse e gli interventi resi necessari da realizzare, sia per il riconoscimento di affidabilità dell'archivio sia per attestare la valutazione di una misura della FAIRness dei suoi (meta)dati. Saranno altresì poste in evidenza le specificità proprie di un archivio caratterizzato dalla gestione di oggetti digitali culturali, definendo da un lato l'identità di un oggetto digitale culturale, dall'altro cosa implica misurare la conformità ai principi FAIR e la capacità di apertura dei suoi (meta)dati intesa nell'accezione più ampia di openness.

Introduzione e motivazioni

Nella sede di questa stessa rivista, veniva pubblicato nel giugno 2019 un ampio articolo dalla titolazione emblematica *Phaidra, un archivio digitale FAIR per la disseminazione e l'accesso integrato a testi, testimonianze e storie del patri-*

monio culturale¹, con l'intenzione di illustrare, *in primis* in termini di evoluzione infrastrutturale integrata rispetto al contesto di origine e sviluppo, le ragioni di una scelta tecnologica e di servizio poste in atto dal Sistema Bibliotecario dell'Università di Padova al fine di favorire l'esposizione distintiva e valorizzazione di collezioni digitali del patrimonio culturale di cui costituiscono un essenziale veicolo memoriale collettivo.

Phaidra², acronimo di *Permanent Hosting, Archiving and Indexing of Digital Resources and Assets*, è infatti la piattaforma del Sistema Bibliotecario dell'Università di Padova per l'archiviazione a lungo termine di oggetti e collezioni digitali³. La caratterizzazione distintiva dell'istanza patavina di Phaidra è stata fin dai suoi esordi da un lato la valorizzazione espositiva della ricchezza eterogenea delle collezioni digitali del patrimonio culturale di Ateneo, come tale proveniente da dipartimenti e centri di ricerca, archivi e musei, oltre che dai progetti di digitalizzazione delle biblioteche; dall'altro il fungere da polo attrattivo nei confronti di altre istituzioni culturali locali, innescando un processo virtuoso di osmosi infrastrutturale culturale e tecnologica che ha conferito a Phaidra, in aggiunta alla sua funzione primaria di sistema di *Digital Asset Management*, anche la funzione di aggregatore interistituzionale, ovvero di organizzazione che raccoglie e aggrega, modella e amministra i dati e i metadati provenienti da molteplici fornitori di contenuto⁴.

A quell'altezza temporale, e sempre in quella stessa sede, una evidenza soddisfacente della conformità di Phaidra ai principi ispiratori della *Findability / Rintracciabilità, Accessibility / Accessibilità, Interoperability / Interoperabilità, Reusability / Riutilizzabilità*, i principi FAIR⁵, era dimostrata attraverso un primo esempio di misurazione della cosiddetta *FAIRness*⁶, la cui valutazione, condotta

¹ Laura Tallandini — Lorisa Andreoli — Elena Bianchi — Linda Cappellato — Yuri Carrer — Gianluca Drago — Giulio Turetta — Antonella Zane, *Phaidra, un archivio digitale FAIR per la disseminazione e l'accesso integrato a testi, testimonianze, immagini e storie del patrimonio culturale*, «DigItalia. Rivista del digitale nei beni culturali», 14 (2019), n. 1, p. 147–157, <<http://digitalia.sbn.it/article/view/2281>>.

² <https://phaidra.cab.unipd.it/>.

³ 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, *Phaidra, un archivio digitale FAIR*, cit. e *Informazioni* in <<https://phaidra.cab.unipd.it/info/impressum>>.

⁴ Allo stato attuale Phaidra raccoglie un totale di circa 437.000 oggetti digitali dalle collezioni di biblioteche, musei, archivi, dipartimenti, centri e uffici dell'Università di Padova, delle Università Ca' Foscari e Iuav di Venezia, dell'Università di Bologna nonché di altre istituzioni cooperanti.

⁵ Mark D. Wilkinson — Michel Dumontier — IJsbrand Jan Aalbersberg, et al., *The FAIR Guiding Principles for scientific data management and stewardship*, «Scientific Data», 3 (2016), article n. 160018, <<https://doi.org/10.1038/sdata.2016.18>>.

⁶ L. Tallandini, *Phaidra, un archivio digitale FAIR*, cit., p. 155.

manualmente secondo il modello di liste di controllo tuttora disponibili online⁷, viene parimenti riprodotta nella Tabella 1.

F	FINDABLE (Rintracciabile)	Valutazione
F1	Ai (meta)dati viene assegnato un identificatore univoco e persistente	✓
F2	I dati sono descritti con ricchi metadati	✓
F3	I metadati includono in modo chiaro ed esplicito l'identificatore dei dati descritti	✓
F4	I (meta)dati sono registrati o indicizzati in una risorsa ricercabile	✓
A	ACCESSIBLE (Accessibile)	
A1	I (meta)dati sono recuperabili attraverso il loro identificatore utilizzando un protocollo di comunicazione standardizzato	✓
A1.1	Il protocollo è aperto, gratuito e universalmente implementabile	✓
A1.2	Il protocollo consente una procedura di autenticazione e autorizzazione, ove necessario	✓
A2	I metadati sono accessibili, anche quando i dati non sono più disponibili	X
I	INTEROPERABLE (Interoperabile)	
I1	I (meta)dati utilizzano un linguaggio formale, accessibile, condiviso e ampiamente applicabile per la rappresentazione della conoscenza	✓
I2	I (meta)dati utilizzano i vocabolari che seguono i principi FAIR	X
I3	I (meta)dati includono riferimenti qualificati ad altri (meta)dati	✓
R	REUSABLE (Riusabile)	
R1	I (meta)dati sono riccamente descritti con una pluralità di attributi accurati e rilevanti	✓
R1.1	I (meta)dati vengono rilasciati con una licenza di utilizzo chiara e accessibile	✓
R1.2	Ai (meta)dati è associata una specifica provenienza	✓
R1.3	I (meta)dati soddisfano gli standard rilevanti per specifiche comunità scientifiche	✓

Tabella 1. *Misurazione manuale della FAIRness di Phaidra nel 2019*

⁷ A titolo di esempio si veda: Sarah Jones — Marjan Grootveld, *How FAIR are your data?*, 2017, November 24, <<https://doi.org/10.5281/zenodo.1065991>> e *Assessing the FAIRness of data*, <<https://www.fosteropenscience.eu/learning/assessing-the-fairness-of-data/>>, corso di formazione online promosso nell'ambito del progetto europeo FOSTER Plus — *Fostering the practical implementation of Open Science in Horizon 2020 and beyond, 2017-2019*, <<https://www.fosteropenscience.eu/about>>. Per un quadro aggiornato sulla valutazione manuale si rinvia alla disamina sistematica in Anusuriya Devaraju — Robert Huber, *An automated solution for measuring the progress toward FAIR research data*, «Patterns», 2 (2021), n. 11, <<https://doi.org/10.1016/j.patter.2021.100370>>, p. 2-3.

La modalità valutativa espressa nella tabella rappresenta un caso esemplare dello scarto concettuale che intercorre tra le linee ispiratrici promosse dai principi guida FAIR, la connotazione interpretativa a cui possono essere soggetti, e la definizione di una metodologia applicativa misurabile e ponderabile in grado di restituire la loro valutazione oggettiva⁸, come si cercherà anche di dimostrare nel corso di questo contributo coadiuvati dall'esperienza diretta di test pratici.

Inoltre, a partire dall'assunzione che «un repository affidabile per la conservazione e il riutilizzo dei dati favorisce lo sviluppo di servizi FAIR»⁹, già allora era stato intrapreso il processo di certificazione con l'ente certificatore CoreTrustSeal (CTS)¹⁰ al fine di ottenere il riconoscimento formale di Phaidra quale archivio digitale affidabile, *Trustworthy Digital Repository* (TDR), ovvero «one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future»¹¹. L'accREDITAMENTO di Phaidra di archivio digitale affidabile CoreTrustSeal¹², le azioni realizzate, e in corso di realizzazione, allo scopo di ac-

⁸ European Commission, Directorate-General for Research and Innovation, *Turning FAIR into reality: final report and action plan from the European Commission expert group on FAIR data*, Publications Office, 2018, <<https://data.europa.eu/doi/10.2777/54599>>; Mark D. Wilkinson — Michel Dumontier — Susanna Assunta Sansone — Luiz Olavo Bonino Da Silva Santos — Mario Prieto — Dominique Batista — Peter McQuilton — Tobias Kuhn — Philippe Rocca-Serra — Mercè Crosas — Erik Schultes, *Evaluating FAIR maturity through a scalable, automated, community-governed framework*, «Scientific Data», 6 (2019), article n. 174, <<https://doi.org/10.1038/s41597-019-0184-5>>; FAIR Data Maturity Model Working Group, *FAIR Data Maturity Model. Specification and Guidelines (1.0)*, 2020, <<https://doi.org/10.15497/rda00050>>; Anusuriya Devaraju — Patricia Herterich, *D4.1 Draft Recommendations on Requirements for FAIR Datasets in Certified Repositories*, 2020, <<https://doi.org/10.5281/zenodo.3678716>>; A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit.; Jan Magnus Aronsen — Rob W.W. Hooft — Ana Slavec — Sarah Jones — Pedro Principe — Andrés Holl — Natalie Harrower — Françoise Genova — Oya Beyan, *Recommendations on FAIR Metrics for EOSC (Report from the EOSC Executive Board FAIR Working Group (WG))*, European Commission, 2021, <<https://doi.org/10.2777/70791>>.

⁹ L. Tallandini, *Phaidra, un archivio digitale FAIR*, cit., p. 152.

¹⁰ <<https://www.coretrustseal.org/>>. L'ente certificatore CTS viene istituito nel 2018 dalla fusione dei sistemi di certificazione del World Data System of the International Science Council (WDS) e del Data Seal of Approval (DSA), anche in risposta all'urgenza di armonizzazione sollecitata nell'ambito della Research Data Alliance. Si veda: Hervé L'Hours — Mari Kleemola — Lisa de Leeuw, *CoreTrustSeal: From academic collaboration to sustainable services*, «IASSIST Quarterly», 43 (2019), n. 1, <<https://doi.org/10.29173/iq936>>.

¹¹ Secondo la definizione attribuita a "Trusted Digital Repository" da RLG/OCLC Working Group on Digital Archive Attributes in: *Trusted digital repositories: attributes and responsibilities. An RLG-OCLC report*, Mountain View, CA: Research Libraries Group, 2002, <<https://www.oclc.org/content/dam/research/activities/trustedrep/repositories.pdf>>, definizione ripresa anche in <http://sedataglossary.shoutwiki.com/wiki/Trusted_digital_repository> del *Science Europe Data Glossary*, edited by P. Doorn, R. Petrauskait, 24 October 2015 e, in versione circoscritta, dal glossario del manuale di riferimento della conservazione digitale *Digital Preservation Handbook*, 2. ed., Digital Preservation Coalition, 2015, <<https://www.dpconline.org/handbook>> (Trad. it.: *Manuale di Conservazione Digitale*, 2021, <<https://dx.doi.org/10.7207/ithb22-1>>, p. 302–303).

¹² <<https://www.coretrustseal.org/wp-content/uploads/2019/11/Phaidra-at-the-Library-System-of-the-University-of-Padova.pdf>>.

crescere e migliorare la valutazione oggettiva della adesione ai principi FAIR, sia a livello del *repository* Phaidra sia a livello dei dati e metadati *in e di* Phaidra, costituiscono il punto di partenza della discussione che si intende delineare nel presente contributo organizzato a tal fine secondo una struttura bipartita.

Nella prima parte, *Affidabilità, accreditamento e empiria della FAIRness*, si darà conto da un lato dell'esito della certificazione e delle azioni metodologiche e pratiche che questo processo ha implicato; dall'altro dell'esperienza sul campo ottenuta dalla valutazione della misurazione della *FAIRness* di Phaidra, in qualità di archivio digitale certificato, attraverso l'applicazione di uno strumento di valutazione automatica¹³, sviluppato e realizzato nel contesto del progetto europeo FAIRsFAIR — Fostering Fair Data Practices in Europe¹⁴ al quale Phaidra ha collaborato come archivio pilota e di test¹⁵, e ove ugualmente si pongono in luce le azioni pratiche e conseguenti che si sono intraprese.

Nella seconda parte, *Esemplificazione di alcune metriche FAIR e valutazione*, saranno prese in considerazione e discusse in modo puntuale alcune specifiche metriche FAIR, illustrando l'impatto migliorativo che l'esito oggettivo della loro valutazione ha reso possibile rispetto all'applicazione attuata per Phaidra, ed evidenziando, in particolare, le specificità proprie di un archivio digitale di oggetti culturali.

Da ultimo, nella sezione conclusiva, si discuteranno i potenziali margini di miglioramento e la lezione appresa che comprova come l'attualizzazione dei dati FAIR da un lato, e il riconoscimento di affidabilità degli archivi digitali dall'altro, non siano traducibili in assunzioni di principio indiscusse e non possano prescindere, per contro, dalla sollecitazione costante del processo di riesame sistematico.

Affidabilità, accreditamento ed empiria della FAIRness

Il neologismo inglese *FAIRness* definisce la misura in cui i dati possono essere interpretati e tradotti, correttamente e con coerenza, sia da macchine, o sistemi informatici, sia da persone in riferimento al significato e al contesto delle informazioni trasmesse, assicurandone al contempo l'autenticità e l'integrità informativa. Tale traducibilità dei dati realizza la cosiddetta interoperabilità semantica intesa, secondo la definizione proposta da Marcia Lei Zeng, come:

«the ability of different agents, services, and applications to communicate (in the form of transfer, exchange, transformation, mediation, migration, integration, etc.) data, information, and knowledge — while ensuring accuracy and preserving the meaning of that same data, information, and knowledge»¹⁶.

¹³ Si tratta di F-UJI Automated FAIR Data Assessment Tool
<<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>>.

¹⁴ <https://www.fairsfair.eu/>.

¹⁵ <https://www.fairsfair.eu/phaidra>.

¹⁶ Marcia Lei Zeng, *Interoperability*, «Knowledge Organization», 42 (2019), n. 2, p. 122–146. Disponibile anche in: *Encyclopedia of Knowledge Organization*, edited by B. Hjørland and C. Gnoli, <<https://www.isko.org/cyclo/interoperability>>.

L'interoperabilità semantica è indiscutibilmente tra le sfide interoperabili più complesse da raggiungere e rappresenta uno dei livelli costitutivi fondamentali, insieme ai livelli della interoperabilità tecnica, organizzativa e giuridica, che informano l'infrastruttura della European Interoperability Infrastructure (EIF), e come tale mutuata anche dalla European Open Science Cloud Interoperability Framework (EOSC IF) ove è assunta come requisito imprescindibile per l'attuazione dei dati FAIR¹⁷.

Conditio sine qua non, affinché sia espressa la *FAIRness* dei dati e i dati permangano nel tempo consentendo il più ampio riuso, consiste nel fare sì che essi siano *affidati* ad archivi digitali che rispondano a specifiche tali da permettere il riconoscimento pubblico da parte della comunità di riferimento della loro *affidabilità* in termini di Trasparenza, Responsabilità, Utilizzatori, Sostenibilità, Tecnologia¹⁸, ovvero che detti archivi possano essere accreditati e reputati *Trusted* e *TRUSTworthy Digital Repository* (TDR). Il compimento dell'affidabilità, *trustworthiness* in lingua inglese alla lettera l'«essere degno di fiducia», rivela la capacità ostensiva dell'*attendibilità* di un archivio digitale di dimostrare con oggettiva certezza di essere in grado di conservare, indelebilmente e permanentemente, il patrimonio digitale ospitato a garanzia del suo riuso costante nel tempo da parte di ogni attore coinvolto, che sia inteso come portatore di interessi (*stakeholder*), produttore e creatore (*producer*), o parimenti consumatore (*consumer*) e utilizzatore (*user*) di dati. La certificazione CoreTrustSeal (CTS), di cui Phaidra ha conseguito il sigillo nel novembre 2019¹⁹, è senza dubbio tra le attestazioni di accreditamento più auto-

¹⁷ European Commission — Directorate-General for Research and Innovation, Oscar Corcho, Magnus Eriksson, Krzysztof Kurowski, et al., *EOSC interoperability framework: report from the EOSC Executive Board Working Groups FAIR and Architecture*, 2021, <<https://data.europa.eu/doi/10.2777/620649>>, dove l'interoperabilità semantica viene così definita: «the ability of computer systems to transmit data with unambiguous, shared meaning. Semantic interoperability is a requirement to enable machine computable logic, inferencing, knowledge discovery, and data federation between information systems».

¹⁸ *Transparency, Responsibility, User focus, Sustainability e Technology* definiscono i principi guida TRUST identificati nel contesto del dibattito condotto all'interno della Research Digital Alliance (RDA) a seguito della sessione plenaria "RDA/WDS Certification of Digital Repositories: Build TRUST to be FAIR — Emerging Needs of Certification in Life Sciences, Geosciences and Humanities", tenutasi a Philadelphia il 4 aprile 2019. La discussione ha condotto dapprima al documento di indirizzo: RDA/WDS Certification of Digital Repositories IG, *The TRUST Principles for Trustworthy Data Repositories — An Update, Research Data Alliance (RDA)*, 2019, <<https://www.rd-alliance.org/trust-principles-trustworthy-data-repositories---update>>, su proposizione del tavolo di coordinamento CoreTrustSeal e, a seguire, alla pubblicazione manifesto a cura di Dawey Lin — Jonathan Crabtree — Ingrid Dillo, et al., *The TRUST Principles for digital repositories*, «Scientific Data», 7 (2020), article n. 144 <<https://doi.org/10.1038/s41597-020-0486-7>>, con l'obiettivo di offrire un quadro di riferimento comune e condiviso «as the essential components for assessing, developing, and sustaining the trustworthiness of data repositories, and thus supporting the scientist and other actors» <<https://www.rd-alliance.org/trust-principles-trustworthy-data-repositories-%E2%80%93-update>>.

¹⁹ La certificazione è pubblicata in evidenza nella home page di Phaidra con collegamenti di rinvio anche nelle pagine interne. Per una visione complessiva aggiornata circa gli archivi certificati CTS, 128 alla data corrente, si veda la mappa visualizzabile alla pagina

revoli e riconosciute sul piano internazionale²⁰ che comprovi l'affidabilità di un archivio digitale²¹. In accordo al quadro di riferimento delineato a partire dal 2018 nel *Memorandum of understanding to Create a European Framework for Audit and Certification of Digital Repositories*²², il quale identifica tre livelli di certificazione della affidabilità interdipendenti e dati in sequenza crescente — *basic, extended e formal* —, CTS si colloca infatti come certificazione di base (*core*) il cui conseguimento è considerato una tappa essenziale e di fatto obbligatoria per poter proseguire ai livelli successivi, rispettivamente basati sugli standard DIN 31644 e ISO 16363²³. La declinazione secondo sedici requisiti²⁴, distribuiti in quattro sezioni principali — *Background information, Organizational Infrastructure, Digital Object Management, Technology* — struttura e accompagna il processo di auto-valutazione da parte dell'archivio certificando il cui esito è affidato alla revisione di due membri valutatori CTS affinché ne attestino la sua conformità. Con riferimento diretto alla esperienza di

<<https://www.coretrustseal.org/why-certification/certified-repositories/>>. Con Phaidra rappresenta l'Italia l'archivio CLARIN-IT ILC-CNR <<https://dspace-clarin-it.ilc.cnr.it/repository/xmlui/>>, ospitato dall'Istituto di Linguistica Computazionale "A. Zampolli" (ILC) del Consiglio Nazionale delle Ricerche (CNR) con sede a Pisa.

²⁰ È indicata come componente di riferimento dell'ecosistema FAIR e raccomandazione esplicita nel documento programmatico *Turning FAIR into Reality*, cit., in particolare p. 44-45 e p. 67 dove si legga la raccomandazione 9: *Develop assessment frameworks to certify FAIR services*.

²¹ L'impulso decisivo per l'avvio dell'accreditamento degli archivi digitali dal punto di vista di *audit* e certificazione deriva dalla relazione finale della Task force of archiving of digital information, guidata da Donald Waters e John Garrett e commissionata nel 1994 dalla Commission on Preservation and Access e dal Research Libraries Group, dove nelle raccomandazioni conclusive si sottolinea come: «A process of certification for digital archives is needed to create an overall climate of trust about the prospects of preserving digital information» (Donald Water — John Garrett, *Preserving Digital Information. Report of the Task Force of Archiving of Digital Information*, 1996, pubb. 63, <<https://www.clir.org/pubs/reports/pub63/>>, p. 40). Per un inquadramento generale e storico sulle certificazioni esistenti, si rinvia al capitolo *Audit and certification*, in: *Digital Preservation Handbook*, cit., <<https://www.dpconline.org/handbook/institutional-strategies/audit-and-certification>>.

²² <http://www.trusteddigitalrepository.eu/Trusted%20Digital%20Repository.html>.

²³ DIN 31644:2012-04, *Information and documentation — Criteria for trustworthy digital archives*, <<https://www.din.de/en/getting-involved/standards-committees/nid/wdc-beuth:din21:147058907>>; ISO 16363:2012, *Space data and information transfer systems — Audit and certification of trustworthy digital repositories*, <<https://www.iso.org/standard/56510.html>>. La certificazione formale di terzo livello prevede la piena conformità ISO 16363 ed estende il processo di certificazione all'organizzazione cui fa capo l'archivio digitale nella sua integrità.

²⁴ *CoreTrustSeal Trustworthy Data Repositories Requirements 2020–2022* <<https://www.coretrustseal.org/why-certification/requirements/>>, ove sono illustrati i requisiti CTS validi per il triennio 2020-2022. La certificazione CTS di Phaidra ricade nel quadro dei requisiti 2017-2019 <<https://www.coretrustseal.org/why-certification/requirements/previous-requirements/>>, qui denominati *Core Trustworthy Data Repositories Requirements*. Per il tracciamento delle modifiche intercorse, si veda: CoreTrustSeal Standards and Certification Board, *Change file: CoreTrustSeal Trustworthy Data Repositories Requirements 2017-2019 to 2020–2022 (v01_00)*, 2020, <<https://doi.org/10.5281/zenodo.3828622>> e Id., *Change file: CoreTrustSeal Trustworthy Data Repositories Requirements: Extended Guidance 2017–2019 to 2020–2022 (v01_00)*, 2020, <<https://doi.org/10.5281/zenodo.3828636>>.

Phaidra²⁵, la dimostrazione formale della sua affidabilità segna il raggiungimento indubbio di una pietra miliare attraverso un percorso di sviluppo quasi decennale che ha inteso tracciare in chiave evolutiva la giustezza di una scelta organizzativa e tecnologica, la quale non solo si è rivelata essere sostenibile nel tempo ma che ha anche favorito la caratterizzazione distintiva di Phaidra, quale archivio deputato alla conservazione del patrimonio culturale digitale a servizio della comunità scientifica e civile, e suo fondamentale veicolo memoriale²⁶.

L'evidenza effettiva dell'impatto che l'acquisizione di una certificazione può avere presso la propria comunità e cittadinanza scientifica d'interesse, tuttavia, non sempre è facilmente determinabile e può dipendere da molteplici fattori, tra i quali campeggia anzitutto il gradiente di fiducia che riesce a ottenere dai suoi interlocutori come archivio *trustworthy*²⁷, non solo *trusted*, ponendosi con loro in una costante relazione dialogica²⁸. Per quanto riguarda Phaidra, sicuramente una prima manifestazione concreta in tal senso, pressoché contestuale all'accreditamento CTS, è stato l'esito positivo della sua candidatura come archivio digitale di test all'appello promosso dal progetto europeo FAIRsFAIR²⁹, teso al supporto nella progettazione di soluzioni prototipali concepite con l'intento precipuo di sollecitare il processo di *fairification*, migliorando e rafforzando «the FAIRness of data in repositories that enable FAIR data»³⁰.

Coordinato dall'ente olandese KNAW-DANS (Data Archiving and Networked Services)³¹, FAIRsFAIR ha previsto il coinvolgimento di 22 istituzioni partner³², provenienti da otto stati membri dell'Unione europea, con l'obiettivo primario di pro-

²⁵ Lorisa Andreoli — Cristiana Bettella — Linda Cappellato — Yuri Carrer — Gianluca Drago — Giulio Turetta, «La goccia pronta per il mappamondo»: esperienze di Phaidra (e dintorni), preprint, 2020, <<https://hdl.handle.net/11168/11.443591>> (in stampa, Milano: La Bibliografica, 2022), in particolare il paragrafo *Dimostrare/provare: sul certificare Phaidra*, p. 5-8.

²⁶ Ivi, in particolare i paragrafi *Rappresentare/descrivere: sul modellare i dati Phaidra e Esporre/guardare: sul manifestare e raccontare Phaidra*.

²⁷ «The notion of a 'Trustworthy Digital Repository' stems from the need to move beyond de facto trust in partner organisations to act as responsible stewards of data, towards de jure assertions of their trustworthiness», come spiegano H. L'Hours — M. Kleemola — L. de Leeuw, *CoreTrustSeal: From academic collaboration to sustainable services*, cit., p. 13.

²⁸ Devan Ray Donaldson, *Certification information on trustworthy digital repository websites: a content analysis*, «PLoS ONE», 15 (2020), 12: e0242525, <<https://doi.org/10.1371/journal.pone.0242525>> e D. Lin, *The TRUST Principles for digital repositories*, cit.

²⁹ FAIRsFAIR — Fostering Fair Data Practices in Europe (Project call H2020-INFRAEOSC-2018-2020 Grant agreement 831558, <<https://cordis.europa.eu/project/id/831558>>) è un progetto europeo Horizon 2020 di durata triennale (marzo 2019-febbraio 2022).

³⁰ <<https://www.fairsfair.eu/application-results-open-call-data-repositories>>. Si noti come la seconda *call* abbia riguardato la selezione di archivi impegnati, o interessati, nel perseguire la certificazione CTS, e ai temi della certificazione sono dedicati infatti i lavori del *working package* (WP) 4 «FAIR Certification (of Repositories)», <<https://www.fairsfair.eu/fair-certification>>.

³¹ <https://dans.knaw.nl/en/about/>.

³² <<https://www.fairsfair.eu/partners>>. Per l'Italia è coinvolto TRUST-IT, capofila del *work package*

muovere l'indirizzo di buone pratiche, nonché di realizzare soluzioni atte a informare una infrastruttura integrata della conoscenza, ove gestione dei dati della ricerca di qualità, procedure, standard e metriche si fondino essenzialmente sui principi FAIR. La declinazione progettuale FAIRsFAIR si è sviluppata lungo quattro assi tematici — *data practices, data policy, certification, e training, education and support* — facenti capo a sette *working package* (WP)³³, ciascuno dei quali incardinato su obiettivi specifici ma sempre in coerenza con una metodologia trasversale condivisa, incentrata sulla loro sostenibilità e applicabilità nei contesti di riferimento alla produzione, gestione e distribuzione espositiva dei dati della ricerca, e sempre con il fine ultimo di promuovere una autentica cultura dei dati FAIR³⁴.

In questa prospettiva, la partecipazione di Phaidra si è svolta secondo un ruolo e funzione duplici: da un lato, nel ruolo di archivio digitale certificato con qualificazione disciplinare specificamente rivolta al dominio *Cultural Heritage*; dall'altro, con funzione di *tester e pilot repository*, rispettivamente negli ambiti *Data practices* e *Certification* dei *working package* (WP) 2 “FAIR Practices: Semantics, Interoperability, and Services” (WP2, task 2.2)³⁵ e 4 “FAIR Certification (of Repository)” (WP4, task 4.5)³⁶.

“FAIR Semantics and Semantics in FAIR” traduce la sostanza dell’impegno assunto dal WP2, finalizzato a favorire l’interoperabilità semantica nell’ecosistema FAIR grazie alla definizione di linee guida di riferimento e realizzando dispositivi in grado di attuare funzionalità FAIR, come per esempio *FAIR Data Point* (FDP)³⁷ al quale Phaidra si è prestato come *tester*, la cui applicazione consente anche ad archivi non strettamente FAIR di condividere il proprio catalogo di collezioni di dati (*dataset*)³⁸, potenziando la loro esposizione, distribuzione e riusabilità, e aumentando la capacità *FAIRness* dei dati e metadati di interoperare semanticamente.

Per contro, tra i compiti affidati al WP4, dedicato ai temi della certificazione³⁹, vi è

(WP) 5 “Engagement, Communication and Uptake”, <<https://www.fairsfair.eu/engagement-communication-and-uptake>>.

³³ <https://www.fairsfair.eu/the-project>.

³⁴ <<https://zenodo.org/communities/fairsfair/>> in Zenodo dove è disponibile ad Accesso Aperto la documentazione prodotta nel corso del progetto.

³⁵ <<https://www.fairsfair.eu/fair-practices-semantics-interoperability-and-services>>, con i *repository* di dati della ricerca: Federated Research Data Repository (FRDR) — Canada, PANGAEA — Germania, Data SciencesPo — Francia, Dataverse Network Norway (DataverseNO) — Norvegia e National Biodiversity Network Atlas (NBN Atlas) — Regno Unito, dove emerge con nettezza la specificità dell’ambito *Cultural Heritage* di Phaidra.

³⁶ <<https://www.fairsfair.eu/fair-certification>>, in particolare con i *pilot repository*: PANGAEA — Germania, Commonwealth Scientific and Industrial Research Organisation (CSIRO) Data Portal — Australia, World Data Center for Climate (WDC) — Germania, <<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>>, dove ancora va evidenziata la specificità di Phaidra.

³⁷ <<https://www.fairdatapoint.org/>> e <<https://github.com/FAIRDataTeam/FAIRDataPoint>>.

³⁸ Basato su *Data Catalog Vocabulary (DCAT) — Version 2*, W3C Recommendation 04 February 2020, <<https://www.w3.org/TR/vocab-dcat/>>.

“FAIR Data Assessments: Pilots” (WP4.5) preposto alla formalizzazione di una metodologia di valutazione oggettiva, basata sull’evidenza empirica, del livello di *FAIRness* dei dati di archivi digitali certificati affidabili, attraverso l’identificazione di un modello di misurazione — *FAIRsFAIR Data Object Assessment Metrics*⁴⁰ — da applicare allo sviluppo di un dispositivo di test FAIR — *F-UJI (FAIR Research Data Object Assessment Tool)*⁴¹ — in grado di eseguirne la valutazione automatica. La sperimentazione pilota⁴², a cui Phaidra è stato invitato a collaborare su base volontaria, si è svolta secondo il duplice punto di vista del *valutato* e del *valutatore*, ovvero della valutazione del livello di *FAIRness* dei dati e metadati di Phaidra da un lato, e della valutazione degli indicatori metrici e del dispositivo di valutazione FAIRsFAIR

³⁹ In collaborazione con CoreTrustSeal, WP4 è stato artefice da un lato dell’allineamento dei requisiti CTS in relazione al modello e vocabolario FAIR in coerenza con le raccomandazioni dell’Unione europea; dall’altro ha promosso e guidato sia il percorso di certificazione CTS di dieci archivi digitali coadiuvandoli nella valutazione *FAIRness* dei loro dati, sia messo a punto la metodologia appropriata, ad uso di ogni attore d’interesse, per la corretta individuazione di archivi digitali affidabili e dei criteri di affidabilità, contribuendo in tal senso alla delineazione del profilo di *FAIR-enabling Trustworthy Digital Repository*. Si veda: Hervé L’Hours — Iona von Stein — Jerry deVries — Linas Cepinskas — Joy Davidson — Patricia Herterich — Robert Huber — Benjamin Jacob Mathers, *M4.3 CoreTrustSeal+FAIRenabling, Capability and Maturity (1.0)*, 2021, <<https://doi.org/10.5281/zenodo.5346822>>; Maaïke Verburg — Iona von Stein — Linas Cepinskas — Hervé L’Hours — Patricia Herterich — Joy Davidson — Kevin Ashley — Olivier Rouchon — Andrea Greco — Serenella Muradore Gallas — Sara Pittonet Gaiarin, *D4.3 Report on the certification support and guidance for repositories and reviewers (V1.0_DRAFT)*, 2021, <<https://doi.org/10.5281/zenodo.5137552>> e Anusuriya Devaraju — Robert Huber — Mustapha Mokrane — Patricia Herterich — Linas Cepinskas — Jerry de Vries — Herve L’Hours — Joy Davidson — White Angus, *FAIRsFAIR Data Object Assessment Metrics*, 2020, <<https://doi.org/10.5281/zenodo.4081213>> per l’allineamento dei requisiti CTS con i principi FAIR nel modello di misurazione FAIRsFAIR di cui si discute in seguito nel testo.

⁴⁰ A. Devaraju, *FAIRsFAIR Data Object Assessment Metrics*, cit. Formulato in 17 metriche secondo l’ultima versione 0.4, il modello metrico FAIRsFAIR si basa essenzialmente sugli indicatori proposti da FAIR Data Maturity Model Working Group, *FAIR Data Maturity Model, Specification and Guidelines (1.0)*, cit., con FAIRdat/FAIREnough, WDS/RDA Assessment of Data Fit Checklist e le indicazioni raccolte dai partner FAIRsFAIR. Si veda: Anusuriya Devaraju — Patricia Herterich, *D4.1 Draft Recommendations on Requirements for Fair Datasets in Certified Repositories (1.0)*, 2020, <<https://doi.org/10.5281/zenodo.5362222>> e Anusuriya Devaraju — Robert Huber — Mustapha Mokrane — Patricia Herterich — Linas Cepinskas — Jerry de Vries — Herve L’Hours — Joy Davidson — Vesa Akerman — Mivhael Diepenbrokek, *From conceptualization to implementation: FAIR assessment of research data objects*, «Data Science Journal», 20 (2021), n. 1, <<https://doi.org/10.5334/dsj-2021-004>>.

⁴¹ <<https://www.fairsfair.eu/f-uji-automated-fair-data-assessment-tool>>. Con le parole degli sviluppatori: «The ‘F’ stands for FAIR (of course) and ‘UJI’ means ‘Test’ in Malay. So F-UJI is a FAIR testing tool», in: <<https://github.com/pangaea-data-publisher/fuji>> ma si veda anche: A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit., p. 7-8.

⁴² Condotta prevalentemente nell’arco dei mesi di agosto-novembre 2020. Per l’illustrazione della metodologia applicata e la discussione critica dei risultati ottenuti dalla valutazione sperimentale condotta su Phaidra e i quattro *pilot repository* FAIRsFAIR, si rinvia a A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit., p. 9-10, in questa sede si osserva solo che la selezione del campione di oggetti digitali per la valutazione ha tenuto conto della specificità di Phaidra includendo ogni tipologia di risorsa.

dall'altro, assecondando l'impegno reciproco assunto da ambo le parti di dare un seguito fattivo, di buona pratica *FAIR enabler*, alle azioni mutuamente concordate.

È in questo quadro metodologico che si è sviluppato il percorso "formativo" della valutazione *FAIRness* di Phaidra, articolato, operativamente, nelle seguenti fasi: acquisizione informativa da parte dei referenti della valutazione sperimentale sul profilo complessivo di Phaidra quale archivio digitale certificato CTS⁴³ e, da parte dei referenti di Phaidra, acquisizione e analisi del modello metrico FAIRsFAIR; prima esecuzione del test pratico della valutazione automatica tramite dispositivo F-UJI con discussione dei risultati, definizione delle azioni FAIR e loro implementazione da entrambe le parti coinvolte; seconda esecuzione del test pratico della valutazione automatica tramite dispositivo F-UJI e discussione dei risultati; successive implementazioni FAIR e valutazione iterativa della *FAIRness* dei dati e metadati di Phaidra⁴⁴.

I risultati della valutazione *FAIRness*, emersi dalla esecuzione del primo test pratico⁴⁵, hanno dimostrato con evidenza la prova esplicita dello scarto interpretativo che si rivela nella valutazione dei principi FAIR nel momento in cui essi siano sottoposti all'interpretazione di un modello metrico computazionale. Se si eccettua l'esito positivo ottenuto esclusivamente per quanto riguarda le metriche relative al Principio F1 «(Meta)data are assigned globally unique and persistent identifiers»⁴⁶, infatti, la quasi totalità degli indicatori metrici del modello FAIRsFAIR è risultata non interpretabile da parte del *tool* di valutazione F-UJI, attestando il livello di *unFAIRness* dei dati e metadati di Phaidra in opposizione netta rispetto all'esito della misurazione manuale della *FAIRness* condotta nel 2019 (Tab. 1), nonché ponendo in discussione la stessa dimostrazione di affidabilità comprovata dalla certificazione CTS.

Naturalmente si è trattato di una valutazione *unfair*, non corretta dal punto di vista del profilo effettivo di Phaidra⁴⁷ ma che ha permesso, da un lato, di mettere a

⁴³ Nella forma di intervista esplorativa tesa a definire comprensivamente il profilo di Phaidra dal punto di vista organizzativo e tecnologico, del dominio di contenuto e del modello di rappresentazione dei dati, delle modalità di identificazione degli oggetti e collezioni di oggetti digitali, dei formati raccomandati, nonché da quello delle condizioni di accesso, recupero e ricercabilità dei dati e metadati. L'intervista si è rilevata un'occasione di confronto importante che ha consentito ai valutatori di assumere una lettura qualitativa dei risultati del primo test di valutazione, identificando quindi con maggiore attinenza la proposta delle azioni implementative FAIR.

⁴⁴ Su cui si veda *infra* il dettaglio in: *Esemplificazione di alcune metriche FAIR e valutazione*.

⁴⁵ Entrambi i test pratici previsti dalla sperimentazione sono stati eseguiti secondo la versione 0.3 del modello di misurazione FAIRsFAIR con 13 metriche implementate rispetto alle 17 della versione corrente. Si veda: A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit., p. 9-10 e A. Devaraju, *FAIRsFAIR Data Object Assessment Metrics*, cit., per la specificazione del modello metrico.

⁴⁶ *FsF-F1-01D — Data is assigned a globally unique identifier* e *FsF-F1-02D — Data is assigned a persistent identifier*.

⁴⁷ L. Andreoli, «La goccia pronta per il mappamondo»: esperienze di Phaidra (e dintorni), cit., p. 1-3.

fuoco il *gap* interpretativo che intercorre tra le azioni della *human* e *machine readability*, dall'altro, di identificare insieme ai valutatori alcune determinanti azioni implementative a favore del miglioramento della *FAIRness* di Phaidra. Nella fattispecie tali azioni hanno riguardato: la codifica dei metadati Dublin Core negli elementi *meta* del linguaggio HTML in cui è espressa la pagina web di ogni risorsa⁴⁸; l'adozione dei Typed Links di Signposting.org⁴⁹ negli elementi *link* HTML; l'esposizione dei dati strutturati Schema.org⁵⁰. La realizzazione di questi interventi ha implicato una immediata qualificazione espositiva e valutativa della *FAIRness* dei dati e metadati di Phaidra, positivamente dimostrata anche dai risultati ottenuti dalla esecuzione del secondo test pratico⁵¹, dove la misurazione della *FAIRness* supera il livello medio stabilito⁵² in *Findability / Rintracciabilità* e *Reusability / Riutilizzabilità* e ottiene il punteggio massimo in *Accessibility / Accessibilità*⁵³.

⁴⁸ L'operazione di incorporamento nell'intestazione HTML degli elementi Dublin Core ha consentito di potenziare il profilo Phaidra Dublin Core, qualificando dove possibile gli elementi secondo le proprietà del modello di dati agnostico Dublin Core Metadata Initiative (DCMI) Metadata Terms <<https://www.w3.org/TR/dwbp/#bib-DCTERMS>>. Tale qualificazione ha determinato la possibilità di utilizzo di vocabolari standard innalzando il livello semantico dell'esposizione dei (meta)dati Phaidra, tra tutti esemplare la qualificazione delle condizioni di accesso e termini d'uso con riferimento alle metriche *FsF-A1-01M — Metadata contains access level and access conditions of the data* (Principio A1) e *FsF-R1.1-01M — Metadata includes license information under which data can be reused* (Principio R1.1) su cui si veda *infra*: *Esemplificazione di alcune metriche FAIR e valutazione*, mentre <https://phaidra.cab.unipd.it/static/phaidra_dc-metadata-element-set.pdf> per la specificazione del profilo PHAIDRA_DC.

⁴⁹ <<https://signposting.org/>> e *FAIR Signposting Profile*, prepared by Herbert Van de Sompel — Martin Klein — Shawn Jones — Michael L. Nelson — Simeon Warner — Anusuriya Devaraju — Robert Huber — Wilko Steinhoff — Vyacheslav Tykhonov — Luc Boruta — Enno Meijers, 2021, <<https://signposting.org/FAIR/>>.

⁵⁰ W3C Schema.org Community Group, *Schema.org*, 2011, <<https://schema.org>>. L'implementazione Schema.org in Phaidra, sviluppato per il tipo CreativeWork, si è basata per la mappatura delle proprietà sull'operazione di *embedding* HTML degli elementi Dublin Core, anche qualificati secondo le proprietà del modello di dati DCMI Metadata Terms, e sempre con riferimento al profilo Dublin Core Phaidra. Si è inoltre tenuta in considerazione l'esperienza di applicazione Schema.org di aggregatori *Cultural Heritage*, come per esempio la Digital Public Library of America (DPLA) <<https://dp.la/>>, per cui si veda la discussione a partire da: Richard Wallis — Antoine Isaac — Valentine Charles — Hugo Manguinhas, *Recommendations for the Application of Schema.org to Aggregated Cultural Heritage Metadata to Increase Relevance and Visibility to Search Engines: The Case of Europeana*, «Code4Lib Journal», 36 (2017), <<http://journal.code4lib.org/articles/12330>>.

⁵¹ All'altezza del secondo test pratico (ottobre 2020) Schema.org, serializzato secondo il formato JSON-LD, non era ancora implementato. Da qui l'impatto negativo subito sulla metrica *FsF-I1-01M — Metadata is represented using a formal knowledge representation language* relativa al Principio I1 «(Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation».

⁵² Il punteggio FAIR è stabilito dai valutatori da 1 (massimo) a 0 (minimo) dove $\geq 0.5 < 0.7$ è considerato il livello medio di *FAIRness*. Si veda: A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit., p. 9.

⁵³ Come esempio di crescita incrementale *FAIRness* si veda la valutazione dell'oggetto Phaidra <<https://hdl.handle.net/11168/11.385937>> secondo i punteggi ottenuti dalle iterazioni: prima ite-

Si è trattato, nella sostanza, di una integrazione *FAIR enabling*, la quale ha posto in atto (*machine actionability*) una sorta di veicolo di lettura per il sistema di valutazione automatica dando adito alla corretta interpretazione e pesatura delle metriche FAIR applicate ai dati e metadati di Phaidra, per esempio, esplicitando le proprietà citazionali della conoscenza tacita espressa dal Principio F2 «Data are described with rich metadata»⁵⁴ aumentando la rintracciabilità e disvelamento del dato; definendo le relazioni della *network* dei dati ampliando la possibilità del riuso della risorsa che si espone (Principio I3. «(Meta)data include qualified references to other (meta)data»⁵⁵); dichiarando le condizioni di accesso (Principio A1. «(Meta)data are retrievable by their identifier using a standardized communication protocol») e l'estensione di utilizzo (Principio R1.1 «Metadata includes license information under which data can be reused»), o indicando il tipo di risorsa e il formato⁵⁶ (Principio R1.3. «(Meta)data meet domain-relevant community standards») a garanzia della sua conservazione nel tempo.

Esemplificazione di alcune metriche FAIR e valutazione

Sotto l'egida della Commissione Europea⁵⁷ sono emerse nel corso degli anni numerose iniziative e progetti⁵⁸ con lo scopo primario di raggiungere l'adozione completa dei principi FAIR nel nascente ecosistema EOSC di servizi federati per la ricerca, raccomandando, in particolare, lo sviluppo di metriche per facilitare la valutazione della *FAIRness* dei dati della ricerca⁵⁹.

La valutazione della *FAIRness* di un oggetto digitale può essere condotta in modalità automatica (o semiautomatica). In questa sezione sarà discusso e puntualmente illustrato un esempio di misurazione e valutazione basato sulle *FAIRsFAIR Data Object Assessment Metrics*⁶⁰ sviluppate nel contesto del progetto europeo FAIRsFAIR. Partendo dall'assunto che i principi FAIR «may be applied to any digital object»⁶¹, il modello di misurazione in esame, centrato sui

razione: 0.29 F_Principle, 0.0 A_Principle, 0.0 I_Principle, 0.0 R_Principle; seconda iterazione: 0.71 F_Principle, 1.0 A_Principle, 0.25 I_Principle, 0.56 R_Principle; iterazione corrente (post implementazione Schema.org): 0.86 F_Principle, 1.0 A_Principle, 0.5 I_Principle, 0.6 R_Principle.

⁵⁴ *FsF-F2-01M — Metadata includes descriptive core elements (creator, title, data identifier, publisher, publication date, summary and keywords) to support data findability, grazie all'embedding degli elementi Dublin Core e Schema.org.*

⁵⁵ *FsF-I3-01M — Metadata includes links between the data and its related entities, per la proprietà DCMI Terms DCTERMS.isPartOf.*

⁵⁶ *FsF-R1.3-02D — Data is available in a file format recommended by the target research community, per Typed link item.*

⁵⁷ *Turning FAIR into Reality, cit.*

⁵⁸ A titolo di esempio: EOSC-Nordic <<https://www.eosc-nordic.eu/>> e FAIRsFAIR, cit.

⁵⁹ F. Genova, *Recommendations on FAIR Metrics for EOSC*, cit.

⁶⁰ A. Devaraju, *FAIRsFAIR Data Object Assessment Metrics*, cit.

⁶¹ A. Devaraju, *From Conceptualization to Implementation: FAIR Assessment of Research Data Objects*, cit., p. 4.

dati della ricerca nella forma di oggetti digitali, può essere applicato anche agli oggetti digitali culturali.

La valutazione è effettuata utilizzando la versione online di F-UJI⁶², il *tool* progettato e sviluppato nell'ambito del progetto FAIRsFAIR per la valutazione automatica del livello di *FAIRness* di un oggetto digitale. F-UJI applica le *FAIRsFAIR Data Object Assessment Metrics* all'oggetto individuato tramite l'identificatore sottomesso — l'Handle nel caso in esame (Fig. 1) — analizzando il contenuto e i metadati dell'oggetto. Ogni metrica è associata a ciascuno dei principi FAIR e ogni principio può essere misurato tramite uno o più test pratici⁶³. L'oggetto digitale denominato "Pavia"⁶⁴, archiviato in Phaidra e identificato dall'Handle 11168/11.458406, rappresenta il caso studio valutato. L'esito della valutazione si riassume in un punteggio per ognuno dei principi FAIR e le rispettive metriche, e con un punteggio complessivo (Fig. 2).

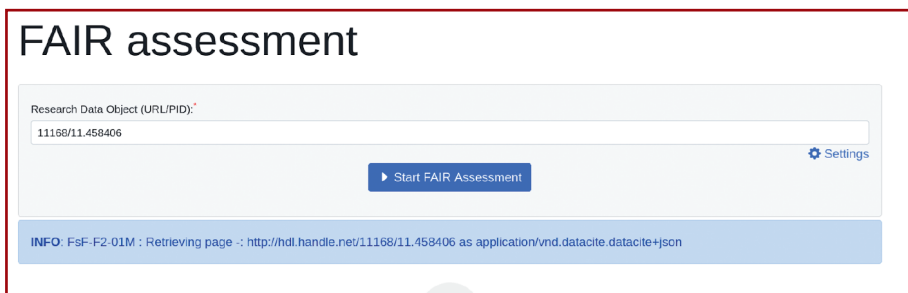


Figura 1. Valutazione metriche F-UJI, inserimento dell'Handle e avvio

Il *report* analitico restituito da F-UJI si struttura in quattro sezioni a cui è associata ciascuna delle metriche. Le metriche sono codificate secondo lo schema "FsF-<principio FAIR>-<numero della metrica>[D,M,MD]", dove "D" indica che la metrica si riferisce ai dati, "M" ai metadati e "MD" a entrambi. Per ogni metrica sono descritti i test effettuati e i risultati ottenuti. Un gruppo selezionato di metriche, i rispettivi test pratici effettuati dal *tool* sull'oggetto "Pavia" analizzato e i risultati ottenuti saranno descritti e discussi di seguito in dettaglio.

⁶² <https://www.f-uji.net/>.

⁶³ <https://www.f-uji.net/index.php?action=methods>.

⁶⁴ Rappresentazione digitale del documento conservato presso la Biblioteca dell'Orto botanico dell'Università di Padova: Giovanni Briosi, *La Stazione di Botanica Crittogamica (Laboratorio Crittogamico Italiano) in Pavia dalla sua fondazione (1871) sino all'anno 1910. Rapporto chiesto per l'esposizione di Bruxelles 1910 da S.E. il Ministro d'Agricoltura, Industria e Commercio*, estratto, «Atti dell'Istituto Botanico dell'Università di Pavia», 13 (1914), s. 2, <<https://hdl.handle.net/11168/11.458406>>.

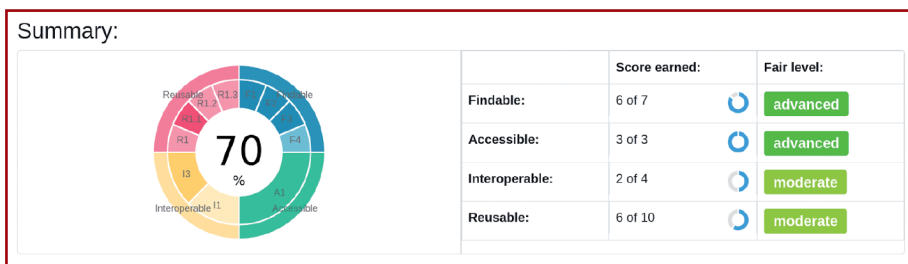


Figura 2. Riassunto del risultato della valutazione ottenuta

FsF-F1-01D — Data is assigned a globally unique identifier. La metrica valuta la presenza e la forma dell'identificatore associato all'oggetto digitale affinché la risorsa sia riferibile senza ambiguità sia da un essere umano sia da un sistema informatico. Alcuni esempi di standard di identificatori sono URL, URN, Digital Object Identifier (DOI), Handle System, w3id.org e Archival Resource Key (ARK). La verifica della metrica si declina nei test *FsF-F1-01D-1* e *FsF-F1-01D-2* (Tab. 2).

Test	Descrizione
<i>FsF-F1-01D-1</i>	Identifier is resolvable and follows a defined unique identifier syntax (IRI, URL)
<i>FsF-F1-01D-2</i>	Identifier is not resolvable but follows an UUID or HASH type syntax
<i>FsF-F1-02D-1</i>	Identifier follows a defined persistent identifier syntax
<i>FsF-F1-02D-2</i>	Persistent identifier is resolvable

Tabella 2. Principio F1. Test pratici per la verifica delle metriche *FsF-F1-01D* e *FsF-F1-02D*

I test valutano che la sintassi dell'identificatore sia aderente a quella degli standard per gli identificatori permanenti. Nel caso in cui l'identificatore non sia risolvibile, viene verificato se l'identificatore alfanumerico è di tipo UUID o *hash*⁶⁵. Il primo test ha successo essendo l'oggetto in esame identificato da un Handle. Il secondo test non è applicabile poiché riguarda il caso di oggetti a cui è associato un identificatore non globale.

FsF-F1-02D — Data is assigned a persistent identifier. La metrica valuta la persistenza dell'identificatore associato all'oggetto. Anche se un URL è univoco, tale caratteristica non è sufficiente a garantire la sua persistenza. Identificatori come DOI, Handle, ARK sono sia globalmente univoci che persistenti, cioè non solo identificano una risorsa senza ambiguità ma sono anche gestiti da organizzazioni affidabili che ne garantiscono la stabilità e risolvibilità nel tempo. La metrica è ve-

⁶⁵ «A Universally Unique Identifier (UUID) URN Namespace», in: <<https://www.ietf.org/rfc/rfc4122.txt>>. Per la lista dei tipi *hash* utilizzati, si veda <<https://github.com/psypana/hashID/blob/master/doc/HASHINFO.xlsx>>.

rificata da F-UJI conducendo i test *FsF-F1-02D-1* e *FsF-F1-02D-2* descritti nella Tabella 2. Nell'oggetto analizzato F-UJI riconosce l'identificatore Handle come un identificatore permanente e ne verifica la risolvibilità. I due test hanno successo.

È possibile effettuare i test descritti fornendo a F-UJI l'URL dell'oggetto⁶⁶ in sostituzione dell'identificatore Handle. In tal caso il report evidenzia la non persistenza dell'identificatore fornito e, di conseguenza, viene attribuito un punteggio complessivo inferiore. Si noti che F-UJI, pur essendo in grado di determinare l'Handle dal contenuto dei dati strutturati Schema.org esposti nella pagina web che descrive l'oggetto, valuta il metodo di accesso all'oggetto premiando l'uso di un identificatore univoco, persistente e globale.

I test compiuti hanno quindi verificato che l'oggetto valutato soddisfa le metriche relative al principio F1 dei principi FAIR: «(Meta)data are assigned globally unique and persistent identifiers».

FsF-A1-01M — *Metadata contains access level and access conditions of the data.*

La metrica si applica al principio A1 e si concentra sulla presenza delle informazioni riguardanti la condizioni di accesso fornite all'utilizzatore dell'oggetto. La verifica della metrica si declina nei tre test pratici descritti nella Tabella 3.

Test	Descrizione
<i>FsF-A1-01M-1</i>	<i>Information about access restrictions or rights can be identified in metadata</i>
<i>FsF-A1-01M-2</i>	<i>Data access information is machine readable</i>
<i>FsF-A1-01M-3</i>	<i>Data access information is indicated by (not machine readable) standard terms</i>

Tabella 3. Principio A1. Test pratici per la verifica della metrica *FsF-A1-01M*

Nei metadati dell'oggetto valutato F-UJI riconosce la presenza della proprietà *DCTERMS.accessRights*⁶⁷ che descrive il livello di accesso, codificata in uno degli elementi *meta* del linguaggio HTML in cui è espressa la pagina web della risorsa. L'utilizzo di un termine tratto dal vocabolario controllato COAR⁶⁸ permette di assegnare una valenza semantica all'informazione veicolata dalla proprietà *accessRights*. Le condizioni di accesso dell'oggetto risultano quindi esplicitate e interpretabili anche da un sistema informatico (*machine readability*). Il terzo test non è applicabile essendo effettuato solo nel caso in cui fallisca la verifica dei due test che lo precedono.

⁶⁶ <https://phaidra.cab.unipd.it/o:458406>.

⁶⁷ <http://purl.org/dc/terms/accessRights>.

⁶⁸ Access Rights: open access <https://vocabularies.coar-repositories.org/access_rights/c_abf2/>. In: COAR — Controlled Vocabularies for Repositories <<https://vocabularies.coar-repositories.org/>>.

FsF-I1-01M — Metadata is represented using a formal knowledge representation language. La valutazione del principio I1 entra nel merito della rappresentazione formale dei metadati. La verifica della metrica si declina nei due test pratici descritti nella Tabella 4.

Test	Descrizione
<i>FsF-I1-01M-1</i>	<i>Parsable, structured metadata (JSON-LD, RDFa) is embedded in the landing page XHTML/HTML code</i>
<i>FsF-I1-01M-2</i>	<i>Parsable, graph data (RDF, JSON-LD) is accessible through content negotiation, typed links or sparql endpoint</i>

Tabella 4. Principio I1. Test pratici per la verifica della metrica *FsF-I1-01M*

In questo caso F-UJI identifica nella pagina web dell'oggetto una risorsa, serializzata nel formato JSON-LD⁶⁹, contenente i metadati dell'oggetto descritti secondo lo schema per i dati strutturati Schema.org. Inoltre F-UJI tenta di valutare la presenza dei metadati tramite *content negotiation*⁷⁰ o mediante l'interrogazione di un *endpoint* SPARQL⁷¹, con esito negativo. L'esposizione di *endpoint* è resa pubblica tramite i registri degli archivi digitali per i dati della ricerca come ad esempio re3data.org⁷². Se si osserva il risultato della valutazione si può notare che esistono margini di miglioramento della già buona interoperabilità raggiunta. Implementando meccanismi di *content negotiation*, o predisponendo un *endpoint* SPARQL, si può evitare di ricorrere al *parsing* della pagina web. L'adozione di tali meccanismi offre la possibilità di valutare risorse che potrebbero essere rese disponibili solo tramite l'interrogazione di API o altri *endpoint*.

FsF-R1-01MD — Metadata specifies the content of the data. F-UJI verifica la metrica del principio R1 ricercando nei metadati le informazioni di base sul contenuto dell'oggetto: tipo, dimensione, URL e la corrispondenza con l'effettivo contenuto delle risorse. La verifica si declina in quattro test, dei quali ci si sofferma nella descrizione dei primi due (Tab. 5).

⁶⁹ JSON for Linking Data <<https://json-ld.org/>>. La medesima risorsa viene utilizzata anche in altri punti della valutazione.

⁷⁰ *Best Practice 19: Use content negotiation for serving data available in multiple formats*, n: Data on the Web Best Practices, W3C Recommendation, edited by B. F. Lóscio — C. Burle — N. Calegari, 31 January 2017, <<https://www.w3.org/TR/dwbp/#Conneg>>.

⁷¹ SPARQL 1.1 Query Language <<https://www.w3.org/TR/sparql11-query/>>.

⁷² re3data.org — Registry of Research Data Repositories <<https://doi.org/10.17616/R3D>>.

Test	Descrizione
<i>FsF-R1-01MD-1</i>	<i>Minimal information about available data content is given in metadata:</i> A. <i>Resource type (e.g. dataset) is given in metadata</i> B. <i>Information about data content (e.g. links) is given in metadata</i>
<i>FsF-R1-01MD-2</i>	<i>Verifiable data descriptors (file info, measured variables or observation types) are specified in metadata:</i> A. <i>File size and type information are specified in metadata</i> B. <i>Measured variables or observation types are specified in metadata</i>
<i>FsF-R1.1-01M-1</i>	<i>Licence information is given in an appropriate metadata element</i>
<i>FsF-R1.1-01M-2</i>	<i>Recognized licence is valid and registered at SPDX</i>
<i>FsF-R1.2-01M-1</i>	<i>Metadata contains elements which hold provenance information and can be mapped to PROV</i>
<i>FsF-R1.2-01M-2</i>	<i>Metadata contains provenance information using formal provenance ontologies (PROV-O)</i>

Tabella 5. *Principio R1. Test pratici per la verifica delle metriche FsF-R1-01MD, FsF-R1.1-01M e FsF-R1.2-01M*

Nel primo test della metrica il *tool* esamina i metadati esposti secondo il modello Schema.org determinando il tipo di risorsa e le informazioni sul contenuto dell'oggetto, tra cui l'URL al contenuto. Nel secondo test F-UJI verifica che la dimensione e il tipo del file siano registrati nei metadati, mentre non è possibile verificare le variabili misurate⁷³ non essendo implementate nell'oggetto digitale culturale preso in esame.

FsF-R1.1-01M — *Metadata includes license information under which data can be reused.* Il *tool* verifica che la licenza d'uso del contenuto dell'oggetto digitale sia esplicitata nei metadati e che la licenza applicata sia presente in un registro pubblico, ovvero che sia formalmente identificabile. La metrica si declina nei due test *FsF-R1.1-01M-1* e *FsF-R1.1-01M-2* descritti nella Tabella 5. La verifica avviene tramite l'individuazione, nei metadati dell'oggetto, dell'elemento che contiene l'indicazione della licenza — nel caso in esame la proprietà *DCTERMS.license*⁷⁴ — e che il valore in esso riportato esista nella lista di licenze delle specifiche *SPDX*⁷⁵, rendendo in tal modo espliciti i termini di utilizzo della risorsa.

FsF-R1.2-01M — *Metadata includes provenance information about data creation or generation.* F-UJI si occupa del reperimento delle informazioni sulla provenienza nei metadati dell'oggetto. La metrica comporta due test pratici, *FsF-R1.2-01M-1* e *FsF-R1.2-01M-2* (Tab. 5). Il primo test ha successo essendo presenti le informazioni sulla provenienza della risorsa. Nel caso in esame vengono rilevati,

⁷³ VariableMeasured, A Schema.org Property <<https://schema.org/variableMeasured>>.

⁷⁴ <http://purl.org/dc/terms/license>.

⁷⁵ *SPDX License List* <<https://spdx.org/licenses/>>.

tra gli altri, gli elementi *DC.creator*, *DC.publisher*, *DC.date*, oltre che i dati strutturati Schema.org, i quali vengono riconosciuti e mappati secondo i termini del vocabolario PROV⁷⁶. Al contrario, il secondo test fallisce in quanto nei metadati non sono presenti elementi specifici dell'ontologia PROV-O⁷⁷.

Confrontando l'esito della misurazione automatica della *FAIRness* con il risultato della misurazione manuale condotta nel 2019 (Tab. 1), si può riscontrare una sostanziale sovrapposizione dei risultati. La misurazione manuale, pur essendo limitata dalla determinante connotazione interpretativa dei principi FAIR, dall'assenza di metriche dedicate e specifiche indicazioni sull'interoperabilità con i sistemi informatici, evidenziava come la maggioranza dei principi guida erano soddisfatti dall'archivio digitale Phaidra.

La verifica della conformità ai principi A2 e I2, che aveva restituito un esito negativo nel 2019, è tuttora non valutabile automaticamente applicando le metriche e il *tool* che ne fa uso. La valutazione del principio A2 "I metadati sono accessibili anche quando i dati non sono più disponibili" non è attuabile in quanto coinvolge la verifica di pratiche di conservazione a lungo termine definite a livello di archivio digitale e non di singolo oggetto in esso depositato⁷⁸. Parimenti, il principio I2 "I metadati utilizzano i vocabolari che seguono i principi FAIR" non risulta valutabile in modo automatizzato poiché i criteri che definiscono le caratteristiche di un vocabolario FAIR non sono ancora stati esaustivamente sviluppati e formalizzati⁷⁹. La misurazione automatica, entrando nel dettaglio dell'implementazione tecnica dell'archivio, fornisce una lettura pratica di una realtà dove gli oggetti sono interconnessi e costruiscono una rete della conoscenza. Il principio I1 della valutazione manuale "I (meta)dati utilizzano un linguaggio formale, accessibile, condiviso e ampiamente applicabile per la rappresentazione della conoscenza" è tradotto nella verifica delle metriche *FsF-I1-01M* e *FsF-I1-02M*. In particolare, la metrica *FsF-I1-02M* "Namespaces of known semantic resources can be identified in metadata" evidenzia la necessità per un archivio digitale di dotarsi di strumenti dedicati all'identificazione semantica delle risorse mediante l'utilizzo di identificativi provenienti da vocabolari *machine readable*, globalmente riconosciuti e interoperabili, ad esempio facendo uso di un *Vocabulary Management System*^{80,81}.

⁷⁶ PROV Specification <<https://www.w3.org/2001/sw/wiki/PROV>>.

⁷⁷ PROV-O: The PROV Ontology <<https://www.w3.org/TR/prov-o/>>.

⁷⁸ A. Devaraju, *An automated solution for measuring the progress toward FAIR research data*, cit.

⁷⁹ Yann Le Franc — Jessica Parland-von Essen — Luiz Bonino — Heikki Lehtväslaiho — Gerard Coen — Christine Staiger, *D2.2 FAIR Semantics: First Recommendations*, 2020, <10.5281/zenodo.3707985>.

⁸⁰ iQvoc è un esempio di Vocabulary Management System per il Semantic Web <<https://iqvoc.net/>>.

⁸¹ Un possibile standard di riferimento è SKOS (SKOS Simple Knowledge Organization System) <<https://www.w3.org/TR/skos-reference/>>, esteso con SKOS-XL (SKOS eXtension for Labels) <<https://www.w3.org/TR/skos-reference/skos-xl.html>>.

Conclusioni

Nel presente contributo si è inteso illustrare un esempio di metodologia di valutazione del livello di *FAIRness* applicata a un archivio digitale affidabile certificato CoreTrustSeal. L'analisi dei risultati ha evidenziato come l'aderenza formale ai principi FAIR non sia espressamente determinabile dall'accreditamento di una certificazione. Ciononostante, il percorso di ottenimento della certificazione permette di formalizzare e di documentare la gestione dell'archivio, il dominio e la sua comunità di riferimento, ponendo le basi per il miglioramento della *FAIRness*, e consentendo di mettere a fuoco gli aspetti su cui può essere prioritario intervenire. Il processo di adeguamento e miglioramento della *FAIRness* sperimentato da Phaidra sembra seguire la successione delle lettere dell'acronimo FAIR: l'uso di standard consolidati, come gli identificativi univoci e persistenti, e l'utilizzo di metadati ricchi (F) si accompagna alla sistematica adozione di chiare politiche di accesso agli oggetti depositati (A) e all'impiego di modelli di riferimento che garantiscono l'interoperabilità (I), giungendo infine a dotare gli oggetti di quell'insieme di proprietà, dalla licenza alla caratterizzazione del contenuto dell'oggetto, che permettono il loro riuso (R).

Il contesto dei dati FAIR per la comunità di riferimento può essere in rapida evoluzione rafforzando, a maggior ragione, l'imprescindibilità del ruolo e delle funzioni di *FAIR-enabling Trustworthy Digital Repository* che, come tale, valuta "programmaticamente" la *FAIRness* dei dati. Nondimeno, l'attualizzazione del livello di *FAIRness* sottende un riesame qualitativo costante e implica come la valutazione non possa sottrarsi da un processo di revisione critica necessariamente aperto, iterativo e condiviso con la comunità e cittadinanza scientifica di riferimento, in stretta correlazione con gli sviluppi attivi nell'ecosistema della European Open Science Cloud (EOSC).

Un nostro ringraziamento speciale è rivolto a Anusuriya Devaraju e Robert Huber per la spinta determinante *FAIR*, alla professoressa Laura Tallandini per accordarci costantemente la sua *trustworthiness*, e ai nostri colleghi e compagni di strada *FAIR enabler* Loris Andreoli, Linda Cappellato e Gianluca Drago.

The primary aim of this contribution is to provide a case of good practice in the accomplishment of the FAIR Principles, illustrating the methodology and the actions undertaken by the Library System of the University of Padua in order to make trustworthy and FAIR Phaidra (Permanent Hosting, Archiving and Indexing of Digital Resources and Assets), the digital repository for the long-term management and preservation, and online dissemination of digital objects in place since 2010, whose system was created and developed by the University of Vienna.

The compliance of a digital archive with the FAIR principles cannot be separated from the formal and public recognition of its trustworthiness. The reputation of trustworthiness is by definition intrinsic to its mission, which aims to ensure its designated community, now and in the future, a secure, robust and long-term access to the digital resources it manages, while guaranteeing, at the same time, its sustainability and durability, security and preservation of the archived data over time.

Grounded on the process of critical reviewing unfolded for the acquisition of the CoreTrustSeal (CTS) certification, and based on the experience carried out as a pilot and test repository in the context of the European project FAIRsFAIR - Fostering Fair Data Practices in Europe, this paper outlines the main steps and actions needed to demonstrate the trustworthiness of the repository as well as the FAIRness of its data and metadata. Furthermore, the distinctive characterization of Phaidra as a cultural heritage digital repository will be highlighted, defining, on the one hand, the digital cultural heritage object identity, and on the other, what it implies to assess the degree of adherence to the FAIR Principles and the capacity of openness, meant in its broadest sense, of data and metadata.

L'ultima consultazione dei siti web è avvenuta nel mese di giugno 2022.