

# Ripensare i dati come risorse digitali: un processo difficile?

Nicola Barbuti

Università degli Studi di Bari Aldo Moro  
nicola.barbuti@uniba.it

## Abstract

**English.** The Art. 2 of the *EU Council Conclusions of 21 May 2014 on cultural heritage as a strategic resource for a sustainable Europe (2014 / C 183/08)* recognizes the existence of the new digital cultural heritage (*born digital* and *digitized*). Starting from this assumption, we need to rethink digital and digitization as social and cultural expressions of the contemporary age. Digital resources shall record and represent both digitization processes and themselves in their life cycle, they are no longer mere gateway to improve the access to reality. So, we have to define clear and homogeneous criteria to validate and certify them as a memory and sources of knowledge for future generations. In this regard, the present paper outlines a first proposal for identification of digital cultural resources, based on the expansion of the *R: Reusable* of the *FAIR Principles for the management of scientific metadata* in *R<sup>5</sup>: Reusable, Readable, Relevant, Reliable and Resilient*.

**Italiano.** L'art. 2 delle *Conclusioni del Consiglio UE del 21 maggio 2014 sul patrimonio culturale come risorsa strategica per un'Europa sostenibile (2014 / C 183/08)* riconosce l'esistenza del nuovo patrimonio culturale digitale (*born digital* e *digitalizzato*). Partendo da questo assunto, si rende indispensabile ripensare il digitale e la digitalizzazione come espressioni sociali e culturali dell'età contemporanea. È necessario, perciò, riconsiderare le risorse digitali come registrazioni e rappresentazioni di processi, e non più come semplici mediatori atti a migliorare l'accesso alla realtà, definendo criteri chiari e omogenei per convalidarli e certificarli come memoria e fonti di conoscenza per le generazioni future. A riguardo, nel presente lavoro si delinea una prima proposta, basata sull'ampliamento della *R: Re-usable* dei *FAIR Principles* per la gestione dei metadati scientifici in *R<sup>5</sup>: Reusable, Readable, Relevant, Reliable and Resilient*.

## 1 Introduzione

L'art. 2 delle *Conclusioni del Consiglio dell'UE del 21 maggio 2014 sul patrimonio culturale come risorsa strategica per un'Europa sostenibile (2014 / C 183/08)* recita<sup>1</sup>: *Il patrimonio culturale è costituito dalle risorse ereditate dal passato in tutte le sue forme e aspetti – tangibile, intangibile e digitale (nato digitale e digitalizzato), inclusi monumenti, siti, paesaggi, abilità, pratiche, conoscenze ed espressioni della creatività umana, nonché raccolte conservate e gestite da enti pubblici e privati come musei, biblioteche e archivi. Ha origine dall'interazione tra persone e luoghi nel tempo ed è in continua evoluzione. Queste risorse sono di grande valore per la società dal punto di vista culturale, ambientale, sociale ed economico e quindi la loro gestione sostenibile costituisce una scelta strategica per il 21° secolo.*

Partendo da questo presupposto, dobbiamo necessariamente cambiare il nostro approccio al digitale e alla digitalizzazione iniziando a considerarli rappresentazioni qualificanti l'età contemporanea e la *digital transformation* che la connota. Ciò implica l'urgenza di individuare e classificare tra le risorse digitali prodotte fino a oggi e in produzione, siano esse singoli oggetti, o complesse *digital libraries*, o sistemi 3D, quelle che possono essere identificate come il nuovo *Digital Cultural Heritage* (DCH), distinguendole da quelle prodotte per mera semplificazione di processi gestionali o per la fruizione estemporanea e immediata di scadenti rappresentazioni relative a entità tangibili e intangibili.

---

<sup>1</sup><https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52014XG0614%2808%29>

Da diversi anni, la digitalizzazione e la qualità e preservazione delle risorse digitali sono riconosciute tra le principali emergenze da affrontare in tutto il mondo. Nel 2012 l'UNESCO ha tenuto la sua conferenza a Vancouver con il significativo titolo *The Memory of the World in the Digital Age: Digitization and Preservation* (Duranti and Shaffer [ed. by], 2012), nel cui ambito è stata redatta ed emanata la *Vancouver Declaration on Digitisation and Preservation*<sup>2</sup>, con l'IFLA e l'International Council of Archives (ICA) tra i principali responsabili.

Da allora, la situazione non sembra essere molto cambiata, nonostante gli sforzi intrapresi per accelerare l'elaborazione di soluzioni a criticità di una complessità che, forse, non hanno precedenti storici.

Gli attuali approcci ai processi di creazione delle risorse digitali, infatti, sembrano non recepire l'evoluzione che negli ultimi anni ha riguardato la digitalizzazione, ancora oggi associata semplicemente alla riproduzione fotografica, mentre, invece, è diventata un processo complesso guidato da regole definite e condivise. Anche l'importanza della qualità dei dati digitali è del tutto sottovalutata nel relegarne la funzione a meri strumenti di mediazione per la fruizione del reale in forma virtuale, sebbene da più parti si riconosca che questi dovrebbero rispondere a requisiti di intellegibilità, affidabilità, pertinenza, persistenza, e registrare le trasformazioni delle funzioni legate al loro riutilizzo nel tempo. L'interpretazione strumentale, infatti, ancora orienta e condiziona negativamente soprattutto la strutturazione degli schemi di metadati con cui indicizzare gli oggetti digitali prodotti e la composizione delle descrizioni loro associate, formulate per essere meri codici funzionali esclusivamente alla ricerca e al recupero dei dati in rete.

Proprio i metadati e i contenuti descrittivi, invece, dovrebbero essere oggetto di particolare attenzione, in quanto sono la sola possibilità di registrare e rappresentare in modo intellegibile i processi di digitalizzazione, creazione e trasformazione che caratterizzano il ciclo di vita dei dati, e di conservare così le informazioni necessarie a conoscerli e a qualificarli come risorse digitali con funzioni anche culturali.

Il tema della funzione essenziale dei metadati nel management e nella fruizione dei dati digitali è il focus dei *FAIR Guiding Principles for Scientific Data Management and Stewardship*<sup>3</sup>, le linee guida per il management dei dati scientifici pubblicate nel 2016, da qualche tempo uno dei temi di maggior interesse nell'ambito del più ampio dibattito sulle possibilità di applicare le metodologie del *data science* alla creazione e gestione dei *data humanities*<sup>4</sup>.

A riguardo, nella CIDOC Conference 2018 si è tenuto un workshop sull'effettiva efficacia dei *FAIR Principles* rispetto agli scenari che oggi la digitalizzazione propone, e soprattutto rispetto a quelli che già si preannunciano imminenti<sup>5</sup>.

Il presente lavoro sintetizza alcune riflessioni maturate da quel proficuo confronto, relative alla necessità di provvedere a un ampliamento del requisito *R*: *Reusable* nei requisiti *R*<sup>5</sup>: *Reusable, Readable, Relevant, Reliable and Resilient*, finalizzato a facilitare l'applicabilità dei *FAIR Principles* ai *data humanities* e, conseguentemente, l'identificazione e certificazione come DCH dei dati rispondenti a tali requisiti nell'informe magma digitale in cui oggi fluttuiamo.

## 2 Verso un ampliamento dei *FAIR* da *R* a *R*<sup>5</sup>

L'assunto di partenza per avviare la riflessione è che i dati digitali non possono più essere creati finalizzandoli alla mera funzione di strumenti di mediazione per una fruizione della realtà alternativa a quella fisica: è necessario ripensarli quali risorse digitali che si qualificano come *record*, entità dinamiche e diacroniche che registrano e conservano nelle descrizioni i processi di digitalizzazione che li hanno creati e quelli che hanno caratterizzato il loro successivo ciclo di vita.

<sup>2</sup> [http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco\\_abc\\_vancouver\\_declaration\\_en.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_abc_vancouver_declaration_en.pdf)

<sup>3</sup> <https://www.go-fair.org/fair-principles/>

<sup>4</sup> <https://www.rd-alliance.org/open-consultation-fair-data-humanities-until-15th-july-2019>; <https://www.gofair.org/implementation-networks/overview/co-operas/>; <https://operas.hypotheses.org/>

<sup>5</sup> <http://www.cidoc2018.com/sites/default/files/CIDOC2018-BookOfAbstracts-Final-v-1-2.pdf>

I metadati descrittivi diventano perciò fondamentali e inscindibili dagli oggetti digitali, in quanto sono proprio l'accuratezza e la qualità delle descrizioni a qualificarli come *record* e, quindi, a renderli risorse digitali pensate e strutturate per essere fruite diacronicamente dagli utenti del futuro, che devono comprendere cosa il dato rappresenta alla pari degli utenti contemporanei.

L'adozione dei *FAIR Principles* lascia aperte alcune questioni. Innanzitutto, non siamo del tutto persuasi che Ricercabilità (*Findable*), Accessibilità (*Accessible*, che assolutamente non è identificabile con *Open*) e Interoperabilità (*Interoperable*) siano requisiti idonei a qualificare i dati come *record* e risorse digitali, conferendogli funzioni nuove e più evolute da quelle strumentali attualmente riconosciute.

Un dato che sia ricercabile, accessibile e interoperabile con altri non fornisce alcuna garanzia di qualità, sufficienza e affidabilità dei contenuti informativi che contiene. Inoltre, la ricercabilità e, di conseguenza, l'accessibilità e interoperabilità che sono a essa vincolate hanno senso nella misura in cui un dato sia oggetto di interesse da parte dei fruitori. E l'interesse per un dato è legato strettamente non alla sua mera funzione di chiave d'accesso a un oggetto digitale semplice o complesso, ma all'essere risorsa informativa e cognitiva per la quantità e qualità dei contenuti descrittivi che mette a disposizione dell'utente già in fase di lettura dei suoi metadati.

Siamo perciò del parere che il requisito che conferisce significato e senso ai primi tre, e dal quale questi dipendono indissolubilmente, sia la Riutilizzabilità (*Reusable*). L'utilizzo e, soprattutto il riutilizzo dei dati, infatti, sono secondo noi i fattori che ne garantiscono la sostenibilità nel tempo e, quindi, la sopravvivenza, in quanto requisiti caratterizzati da dinamismo e diacronia che, quasi sempre, implicano trasformazioni nelle funzioni delle entità che ne sono oggetto: per avere un'idea utilizzando un paradigma analogico, si pensi al Colosseo e al suo ciclo di vita.

Le registrazioni descrittive dei metadati sono perciò fondamentali per garantire qualità e persistenza delle risorse digitali, qualora siano improntate a equilibrate soluzioni quantitative/qualitative e rispondano a ulteriori requisiti che, secondo noi, sono altrettanto essenziali quanto la riutilizzabilità.

Anche la *Reusability*, infatti, di per sé non costituisce una garanzia di qualità del dato e del suo valore quale risorsa informativa e cognitiva. Anzi: proprio le variabili cui una risorsa è soggetta perché riutilizzabile possono essere fonte di distorsione e difformità dei contenuti, il cui valore informativo e cognitivo può perciò non essere più certificabile come affidabile.

La R di Reusable andrebbe perciò, secondo noi, ampliata in R5 con i seguenti requisiti:

- *Readability*: da intendersi non nell'accezione semantica di leggibilità, ma in quella concettuale di *intelligibilità* della risorsa digitale per tutte i possibili target di utenti interessati a fruirne; è requisito fondamentale per conferire ai metadati la funzione informativa e cognitiva necessaria a qualificarli come risorsa culturale, e si basa sull'equilibrato rapporto quantitativo/qualitativo dei contenuti descrittivi e sull'accuratezza formale, stilistica e linguistica dei contenuti;
- *Relevance*: la *persistenza* nel tempo è legata all'interesse degli utenti per i contenuti informativi e cognitivi registrati nella risorsa; essa è strettamente legata al suo riutilizzo e alle possibili trasformazioni di funzione registrate nelle descrizioni; è, quindi, requisito indispensabile affinché la risorsa, di solito creata con funzioni e scopi non necessariamente culturali, possa essere identificabile e riconoscibile nella sua struttura formale e descrittiva anche se varia nel tempo le proprie funzioni evolvendosi in fonte di conoscenza sui processi che registra e, quindi, in risorsa culturale digitale;
- *Reliability*: l'*affidabilità* è la certificazione e validazione della qualità della risorsa digitale rilevabili dalle registrazioni delle sue descrizioni durante tutto il suo ciclo di vita, in relazione a tutte le possibili trasformazioni ed evoluzioni funzionali cui può essere stata soggetta; è, dunque, strettamente connessa alla capacità dell'entità digitale di registrare e preservare gli elementi qualificanti la qualità informativa e cognitiva dei suoi contenuti descrittivi, anche nell'evoluzione delle funzioni e nelle variazioni di forme e funzioni nel tempo;
- *Resilience*: come l'intelligibilità, anche la *resilienza* applicata ai dati e, soprattutto, ai metadati è requisito fondamentale per conferire alle risorse digitali la nuova dimensione culturale;

chiosando la definizione comunemente in uso in ambito informatico<sup>6</sup>, essa va intesa come *la capacità di una risorsa digitale di adattarsi alle condizioni di utilizzo e riutilizzo, di resistere all'usura, di essere duttile nelle trasformazioni e nell'evoluzione delle sue funzioni, al fine di garantire la disponibilità del proprio potenziale cognitivo e informativo nello spazio e nel tempo*; è, quindi, indispensabile per garantire la sostenibilità e il riutilizzo delle risorse digitali nel medio-lungo termine, provvedendo a preservare sia le informazioni utili a conoscere i processi della loro creazione, sia quelle sulla loro funzione originale, sia, infine, le registrazioni delle trasformazioni ed evoluzioni funzionali che ne hanno caratterizzato il ciclo di vita.

### 3 Conclusioni

Tirando le conclusioni su quanto sopra sinteticamente delineato, è nostra opinione che l'adozione dei requisiti FAIR con la  $R$  ampliata in  $R^5$  sia prerequisito indispensabile nel processo di creazione dei dati digitali e, soprattutto, dei metadati che li descrivono, in quanto gli conferirebbero le funzioni di potenziale DCH, rendendoli sostenibili, permanenti, affidabili e, nel contempo, storicizzandoli come fonti di conoscenza dei processi e delle complessità che caratterizzano la rapidissima evoluzione della *digital transformation*.

Non il dato in sé, infatti, ma l'interesse degli utenti presenti e futuri per la fruizione del dato in quanto risorsa informativa e cognitiva deve diventare il prerequisito su cui fondare l'intero processo di creazione, pubblicazione e preservazione di risorse digitali. L'applicazione dei requisiti  $R^5$ , dunque, deve diventare oggetto di attenzione fin dalla fase di analisi e progettazione dei processi sia di digitalizzazione che di creazione di qualsiasi schema di metadati con cui descrivere e gestire gli oggetti digitali in produzione.

Solo così si potrà dare un serio inizio, nel medio termine, all'individuazione di quanto possa essere identificato come DCH nella massa di dati che oggi sovrabbonda nel web e, nel contempo, si potranno definire linee guida omogenee e condivise che presiedano alla creazione di nuove risorse avendo chiaro fin dal principio se gli si voglia conferire il potenziale valore di entità culturali.

In questo modo, nel giro di pochi anni le *Conclusioni EU* del 2014 potranno finalmente essere sostanziate con un nuovo DCH ufficialmente riconosciuto. Diversamente, continueremo a considerare digitalizzazione e digitale solo come un modo diverso e accattivante di fruire il tangibile, perdendo di vista quanto invece tutto ciò sia già oggi l'*humus* identitario che, pur a livelli diversi, identifica l'era digitale contemporanea.

### Bibliografia

<https://www.go-fair.org/fair-principles/>

<https://www.go-fair.org/implementation-networks/overview/co-operas/>

<https://operas.hypotheses.org/>

[http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco\\_abc\\_vancouver\\_declaration\\_en.pdf](http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/CI/CI/pdf/mow/unesco_abc_vancouver_declaration_en.pdf)

<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52014XG0614%2808%29>

<http://www.interpares.org/>

Agenzia per l'Italia Digitale (AgID), Presidenza del Consiglio dei Ministri, *Linee guida sulla conservazione dei documenti informatici*, Versione 1.0 – dicembre 2015, pp. 45 ss. [http://www.agid.gov.it/sites/default/files/linee\\_guida/la\\_conservazione\\_dei\\_documenti\\_informatici\\_rev\\_def.pdf](http://www.agid.gov.it/sites/default/files/linee_guida/la_conservazione_dei_documenti_informatici_rev_def.pdf)

<sup>6</sup> <https://it.wikipedia.org/wiki/Resilienza>

- Lila Bailey. 2015. *Digital Orphans: The Massive Cultural Black Hole On Our Horizon*, Techdirt, Oct 13th 2015 <<https://www.techdirt.com/articles/20151009/17031332490/digitalorphans-massive-cultural-blackhole-our-horizon.shtml>>.
- Nicola Barbuti. 2016. *Le nuove entità culturali digitali tra Intangible Cultural Heritage e Patrimonio Culturale Immateriale*, in *The Creative Network – Conferenza GARR*, Firenze, 30 novembre-02 dicembre 2016 <<https://www.eventi.garr.it/it/conf16/home/materiali-conferenza-2016/paper/19-conf2016-paper-barbuti/file>>
- Nicola Barbuti. 2017. *Dalla Digital Culture al Digital Cultural Heritage: l'evoluzione impossibile?*, in *AIUCD 2017 Conference – Book of Abstract. Il telescopio inverso: big data e distant reading nelle discipline umanistiche*, p. 14-17, AIUCD <<http://aiucd2017.aiucd.it/wp-content/uploads/2017/01/book-of-abstract-AIUCD-2017.pdf>>
- Nicola Barbuti, and Ludovica Marinucci. 2018. *Dal Digital Cultural Heritage alla Digital Culture. Evoluzioni nelle Digital Humanities*, DH 2018 <<https://dh2018.adho.org/dal-digital-cultural-heritage-alla-digital-culture-evoluzioni-nelle-digital-humanities/>>
- Enrico Daga, and Leif Isaksen. 2016. *Proceedings of the 1st Workshop on Humanities in the Semantic Web*, co-located with 13th ESWC Conference 2016 (ESWC 2016), Anissaras, Greece, May 29th, 2016 <<http://ceur-ws.org/Vol-1608/paper-05.pdf>>.
- Luciana Duranti, and Elizabeth Shaeffer [ed. by]. 2012. *The Memory of the World in the Digital Age: Digitization and Preservation. An international conference on permanent access to digital documentary heritage*, UNESCO Conference Proceedings, 26-28 September 2012, Vancouver <[http://ciscra.org/docs/UNESCO\\_MOW2012\\_Proceedings\\_FINAL\\_ENG\\_Compressed.pdf](http://ciscra.org/docs/UNESCO_MOW2012_Proceedings_FINAL_ENG_Compressed.pdf)> .
- Vincenzo Gambetta. 2009. *La conservazione della memoria digitale*, [Rubano] : Siav.
- Pallab Ghosh. 2016. *Google's Vint Cerf warns of 'digital Dark Age'*, BBC News, Science & Environment, 13 February 2016 <<http://www.bbc.com/news/science-environment-31450389>>
- Maria Guercio. 2008. *Gli archivi come depositi di memorie digitali*, «Digitalia», Anno III, n. 2, Rom : ICCU, pp. 37-53.
- Maria Guercio. 2013. *Conservare il digitale. Principi, metodi e procedure per la conservazione a lungo termine di documenti digitali*, Roma-Bari : Laterza.
- Joint Steering Committee for Development of RDA. 2015. *Resource Description and Access (RDA)* <[http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/RDA\\_Traduzione\\_ICCU\\_5\\_Novembre\\_REV.pdf](http://www.iccu.sbn.it/opencms/export/sites/iccu/documenti/2015/RDA_Traduzione_ICCU_5_Novembre_REV.pdf)> .
- Wouter Kool, Brian Lavoie, and Titia van der Werf. 2014. *Preservation Health Check: Monitoring Threats to Digital Repository Content*, Dublin (Ohio) : OCLC Research <<http://www.oclc.org/content/dam/research/publications/library/2014/oclcresearch-preservation-health-check-2014.pdf>> .
- Brian Lavoie, and Richard Gartner. 2013. *Preservation Metadata (2nd edition), DPC Technology Watch Report*, 03 May 2013, DPC Technology Watch Series <<http://www.dpconline.org/docman/technology-watch-reports/894-dpctw13-03/file>> .
- Library of Congress. *PREMIS – Preservation Metadata: Implementation Strategies*, v. 3.0 <<http://www.loc.gov/standards/premis/v3/premis-3-0-final.pdf>> .
- Gilberto Marzano. 2011. *Conservare il digitale. Metodi, norme, tecnologie*, Milano : Editrice Bibliografica.
- Mellon Foundation and Digital Preservation Coalition Sponsor Formation of Task Force for Email Archives*, 1 November 2016 <<https://mellon.org/resources/news/articles/mellon-foundation-and-digital-preservation-coalition-sponsor-formation-task-force-email-archives/>> .
- OCLC. *PREMIS (PREservation Metadata: Implementation Strategies) Working Group*, 2005 <<http://www.oclc.org/research/projects/pmwg/>>.
- Sustainable Economics for a Digital Planet: Ensuring Long-term Access to Digital Information*, Final Report of the Blue Ribbon Task Force on Sustainable Digital Preservation and Access (F. Berman and B. Lavoie, co-chairs), La Jolla, February 2010 <[http://brtf.sdsc.edu/biblio/BRTF\\_Final\\_Report.pdf](http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf)>.