

The banner features a row of six icons: a globe, a book, a handshake, a money bag with a Euro symbol, a scale of justice, and a bicycle. Below the icons, the text 'AIUCD 2021' is prominently displayed. Underneath, it reads 'DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale' and '10° congresso annuale PISA 19-22 gennaio'. On the right side, a list of topics is shown in colored text: 'DIGITAL PUBLIC HUMANITIES' (red), 'OPEN CULTURE' (orange), 'RETI SOCIALI' (yellow), 'TECH ECONOMY' (green), 'E-PARTICIPATION' (blue), and 'TECNOLOGIE ASSISTIVE' (purple). The background includes binary code and a classical building facade.

AIUCD 2021

DH per la società: e-guaglianza, partecipazione, diritti e valori nell'era digitale

10° congresso annuale **PISA** 19-22 gennaio

DIGITAL PUBLIC HUMANITIES
OPEN CULTURE
RETI SOCIALI
TECH ECONOMY
E-PARTICIPATION
TECNOLOGIE ASSISTIVE

Versione PROVVISORIA del contributo presentato al Convegno Annuale

DISCLAIMER

Questa versione dell'abstract non è da considerarsi definitiva e viene pubblicata esclusivamente per facilitare la partecipazione del pubblico al convegno AIUCD 2021

Il Book of Abstract contenente le versioni definitive e dotato di ISBN sarà disponibile liberamente a partire dal 19 gennaio sul sito del convegno sotto licenza creative commons.

***Digital preservation “FAIRness” and “TRUSTworthiness”*: i principi FAIR e TRUST nei contesti di conservazione digitale**

Federica Marti¹

¹ Università di Macerata, Italia – f.marti@unimc.it

ABSTRACT

La presente relazione introduce i *FAIR Guiding Principles for Scientific Data Management and Stewardship* e i *TRUST Principles for Digital Repositories* e ha come scopo il vaglio di alcune esperienze di applicazione di tali principi in ambito internazionale, al fine di fornire una valutazione riguardo l'efficacia del loro impiego nel contesto della conservazione digitale e la coerenza rispetto all'adozione degli standard di conservazione.

PAROLE CHIAVE

FAIR, TRUST, conservazione digitale, accessibilità, *repositories*, contesto internazionale.

1. INTRODUZIONE: “Build TRUST to be FAIR”¹

Findable, Accessible, Interoperable, Reusable; Transparency, Responsibility, User Focus, Sustainability, Tecnology: queste le parole chiave in contesto internazionale in relazione alla creazione, diffusione e valorizzazione dei dati di ricerca e per l'implementazione di *repositories* in grado di conservare a lungo termine questi dati. Concepite come linee guida generali e applicabili a contesti diversificati², i *FAIR Guiding principles for scientific data management and stewardship* e i *TRUST Principles for digital repositories* auspicano la condivisione da parte degli enti di ricerca a livello internazionale, da un lato al fine di agevolare e incrementare la condivisione e la consultabilità di dati e informazioni, dall'altro di conservarli a lungo termine. Nello specifico, i *FAIR Guiding principles* forniscono indicazioni sulle modalità di predisposizione di dati, metadati e infrastrutture, perché siano, accessibili, interoperabili, ricercabili e riusabili³.

¹ Si cita, per il titolo di questo paragrafo, la tredicesima Sessione plenaria della Research Data Alliance (RDA) “RDA/WDS Certification of Digital Repositories: Build TRUST to be FAIR - Emerging Needs of Certification in Life Sciences, Geosciences and Humanities”: nel corso di questo evento, tenutosi a Philadelphia il 4 aprile 2019, i *TRUST Principles* vengono formalmente proposti e riconosciuti come *community discussion*.

² Ci si riferisce, ad esempio, a istituti di ricerca in ambito fisico-matematico, astronomico, biologico e sanitario; o, ancora, a enti preposti alla curatela del patrimonio culturale come biblioteche e archivi.

³ I *FAIR Guiding principles* sono stati formulati nell'ambito di un gruppo di lavoro comprendente ricercatori, editori accademici, professionisti, esperti commerciali ed enti di finanziamento e pubblicati nel 2016 sulla rivista *Scientific Data* (Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. «The FAIR Guiding Principles for Scientific Data Management and Stewardship». *Scientific Data* 3, no. 1 (March 15, 2016): 1–9. <https://doi.org/10.1038/sdata.2016.18>). Si appoggiano a diverse iniziative, che mirano sia al tracciamento di metriche di dettaglio per la *compliance* a FAIR (si vedano, a proposito, FAIRsFAIR al link <https://www.fairsfair.eu/> e Wilkinson, Mark D., Sansone, SA., Schultes, E. et al. «A design framework and exemplar metrics for FAIRness». *Scientific Data* 5, 180118 (2018). <https://doi.org/10.1038/sdata.2018.118>) sia all'utilizzo su larga scala di questi principi (si veda, ad esempio, GoFair (<https://www.go-fair.org/> - ultima consultazione: 12/09/2020).

The FAIR Guiding Principles	
To be Findable:	F1. (meta)data are assigned a globally unique and persistent identifier F2. data are described with rich metadata (defined by R1 below) F3. metadata clearly and explicitly include the identifier of the data it describes F4. (meta)data are registered or indexed in a searchable resource
To be Accessible:	A1. (meta)data are retrievable by their identifier using a standardized communications protocol A1.1 the protocol is open, free, and universally implementable A1.2 the protocol allows for an authentication and authorization procedure, where necessary A2. metadata are accessible, even when the data are no longer available
To be Interoperable:	I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2. (meta)data use vocabularies that follow FAIR principles I3. (meta)data include qualified references to other (meta)data
To be Reusable:	R1. meta(data) are richly described with a plurality of accurate and relevant attributes R1.1. (meta)data are released with a clear and accessible data usage license R1.2. (meta)data are associated with detailed provenance R1.3. (meta)data meet domain-relevant community standards

Fig. 1. I FAIR Guiding principles (Wilkinson et al., 2016)

I recentissimi TRUST Principles, invece, si propongono come linee guida per la costruzione di un repository, che consenta il persistere delle proprietà di affidabilità, accessibilità e utilizzabilità dei dati in essi riposti⁴.

The TRUST Principles	
Principle	Guidance for repositories
Transparency	To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence.
Responsibility	To be responsible for ensuring the authenticity and integrity of data holdings and for the reliability and persistence of its service.
User Focus	To ensure that the data management norms and expectations of target user communities are met.
Sustainability	To sustain services and preserve data holdings for the long-term.
Technology	To provide infrastructure and capabilities to support secure, persistent, and reliable services.

Fig. 2. I TRUST Principles (Dawei Lin et al., 2020)

Il presupposto della riflessione che segue è la considerazione dei principi TRUST come una “conseguente integrazione” di FAIR⁵; questi ultimi, infatti, non essendo elaborati esplicitamente ai fini della conservazione digitale, possono costituire sia il presupposto, sia la conseguenza dei TRUST: Lin Dawei (National Institute of Health – NIH), nel corso dell’evento “TRUST Principles Mini Symposium: The Future of Digital Repositories”⁶, ha illustrato come i principi di TRUST siano stati progettati anche allo scopo di influire positivamente sulla FAIRness⁷ dei dati all’interno dei repository. Infatti, gestire un repository in maniera conforme ai principi TRUST non solo permette di implementare la FAIRness nei dati per cui originariamente manca, ma consente anche di garantire il perdurare di questa caratteristica (ad esempio, nel caso in cui venga meno a causa di un cambiamento negli standard di metadattazione).

⁴ Rilasciati il 14 maggio 2020, anch’essi su *Scientific Data* e risultanti da mesi di consultazione e discussione della comunità di ricerca nell’ambito della Research Data Alliance (RDA), i principi TRUST sono stati sviluppati al fine di facilitare l’adozione di *best practices* per la gestione dei repository. Il primo giugno 2020 l’Open Preservation Foundation (OPF), organismo tra i più attivi in contesto europeo per l’elaborazione di strumenti e linee guida per la conservazione digitale, ha accolto questo set di principi, uno dei cui punti di forza risiede nell’aver acquisito l’endorsement di diverse istituzioni su scala mondiale (il cui elenco è consultabile al link <https://www.rd-alliance.org/trust-principles-rda-community-effort> - ultima consultazione 12/09/2020) e, dunque, nel rappresentare un documento ampiamente condiviso (Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, David Giaretta, Marisa De Giusti, et al. «The TRUST Principles for Digital Repositories» *Scientific Data* 7, no. 1 (14 May 2020): 1–5. <https://doi.org/10.1038/s41597-020-0486-7>).

⁵ A questo proposito, si sottolinea che, chiaramente, alcune raccomandazioni presenti in questi due documenti si sovrappongono: si veda il caso dei metadati, per descrivere i quali è raccomandato l’utilizzo di standard impiegati dalla vasta community di riferimento.

⁶ Nel corso di questo “minisimposio”, promosso da Research Data Canada, Research Data Alliance, and World Data Systems e svoltosi il 7 luglio 2020, sono stati discussi, secondo diverse prospettive e casi d’uso, i TRUST Principles. A proposito, si veda il resoconto di Robyn Nicholson (Research Data Alliance) al link <https://www.rdc-drc.ca/trust-principles-mini-symposium-the-future-of-digital-repositories/>.

⁷ Per FAIRness si vuole intendere, avvalendosi del gioco di parole tra il sostantivo “correttezza” e l’acronimo dei principi, la conformità dei dati ai principi FAIR.

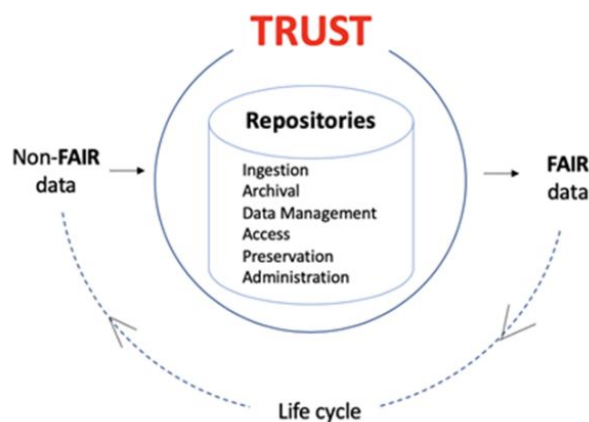


Fig. 3. Rappresentazione grafica della relazione tra i principi TRUST e FAIR (Dawei Lin, 2020)

2. ADAPTABILITY: I PRINCIPI FAIR E TRUST NEI CONTESTI DI CONSERVAZIONE DIGITALE

L'adattabilità di questi principi rispetto ai modelli e agli standard della conservazione digitale e, dunque, la loro efficacia in ambito applicativo è rilevabile attraverso dei contesti d'uso.

In primo luogo, un esempio di conferma della compatibilità dei principi FAIR con lo standard OAIS, riferimento internazionale per l'implementazione di sistemi di conservazione, si trova nel progetto ARCHIVER, iniziativa europea finalizzata all'implementazione di un servizio archiviazione e conservazione per i dati di ricerca, i cui *stakeholders* sono molti tra i maggiori istituti di ricerca internazionali⁸: si richiede, a bando, la presenza integrata di entrambi gli elementi tra i requisiti funzionali, al fine di poter disporre di strumenti software di base interoperabili entro cui si possano rintracciare, monitorare e riutilizzare dati predisposti dalla creazione come FAIR.

In secondo luogo, riguardo alla posizione dei principi FAIR rispetto alla certificazione CoreTrustSeal, la cui metrica di valutazione per certificare i *repository* si basa su ISO 16363 ed è ampiamente utilizzata in ambito internazionale, gli studiosi Mustapha Mokrane (Data Archiving and Networked Services - DANS) e Jonas Recker (GESIS Data Archive for the Social Sciences) hanno valutato che i requisiti CoreTrustSeal incorporano e riferiscono, di fatto, i focus dei principi FAIR, ma li completano visualizzando, appunto, gli aspetti connessi alla conservazione dei dati. Per esemplificare, se si considera il requisito FAIR "*F1. (meta)data are assigned a globally unique and persistent identifier*", si può notare che la concreta persistenza dell'identificativo consegue alla conservazione di questo dato all'interno di un *trusted digital repository*⁹.

Per quanto riguarda i principi TRUST, questi sono stati elaborati concependo come propria implementazione gli standard di conservazione e le certificazioni CoreTrustSeal e ISO 16363: David Giaretta (Primary Trustworthy Digital Repository Authorisation Body - PTAB) ha esaminato l'intersezione dei criteri di *auditing* ISO 16363 - sottolineando come questi rappresentino la componente "testabile" di OAIS - e ISO 16919 con i principi TRUST, definendo questi ultimi come un punto di partenza e affermando che sia, comunque, imprescindibile la verifica dell'effettiva *trustworthiness* dei *repositories* nei dettagli, attraverso i processi di audit¹⁰.

Vi è, infine, la complementarità tra gli stessi principi FAIR e TRUST: Ingrid Dillo (DANS) ha evidenziato l'importanza dell'ecosistema in cui i dati sono riposti (secondo la stessa logica esposta per CoreTrustSeal) e asserito come l'*endorsement* TRUST di un *repository* costituisca la base per l'affermazione della propria affidabilità e, di conseguenza, per la predisposizione FAIR dei dati.

⁸ Il *lead partner* del progetto è il CERN e il consorzio di *buyers* è formato dall'European Molecular Biology Laboratory (EMBL), dal Deutsches Elektronen-Synchrotron (DESY) e dal Port d'Informació Científica (PIC); tra gli *early adopters* si trovano National Research and Education Network in the United Kingdom, Friedrich Miescher Institute for Biomedical Research (Svizzera), Stockholm University Library (Svezia), European Institute of Oncology (Italia), National Research and Education Network in the Netherlands, Australia's Academic and Research Network (Australia). Si veda, a proposito, il sito web del progetto al link <https://www.archiver-project.eu/>.

⁹ Mokrane, Mustapha, Recker Jonas, «Coretrustseal-certified repositories. Enabling Findable, Accessible, Interoperable, and Reusable (FAIR) Data», 2019 https://ipres2019.org/static/pdf/ipres2019_paper_74.pdf (ultima consultazione 29/11/20).

¹⁰ Per questo caso di studio e per il seguente si fa riferimento al già citato "TRUST Principles Mini Symposium: The Future of Digital Repositories" (v. nota 6). In particolare, si fa riferimento agli interventi "Implementation Pathway: TRUST and ISO 16363/ISO 16919 Intersections" di David Giaretta e "TRUST and FAIR: Complementarity of the Principles" di Ingrid Dillo.

3. WHAT ABOUT THE FITNESS?

Emerge, dai casi presentati, che la duttilità che c'è alla base della formulazione dei principi FAIR e TRUST è particolarmente funzionale, in quanto questi consentono una scalabilità in contesti diversi che cerca di salvaguardare le diverse esigenze proprie della conservazione digitale. Pertanto, la genericità *by design* con cui sono stati concepiti non costituisce un limite, ma, al contrario, un vantaggio: il fatto che non vengano indicati standard e requisiti tecnologici specifici (si veda, per FAIR, “*I.I.: (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation*”) consente ai professionisti della conservazione digitale di adottare i più adeguati al proprio contesto di riferimento, mentre le definizioni piuttosto indeterminate (si veda, per TRUST, “*Transparency. To be transparent about specific repository services and data holdings that are verifiable by publicly accessible evidence*”) permettono ai *digital preservation practitioners* di accostarle a quelle del proprio ambito nazionale e, in particolare, del proprio quadro legislativo. Non si dimentichi che FAIR e TRUST sono stati elaborati in un ambito estremamente multidisciplinare, che comprende istituti di ricerca che vanno dall'ambito sanitario all'ambito archivistico, dall'ambito matematico a quello biblioteconomico, collocati in tutto il mondo: si tratta di *principi* scaturiti da un impegno collettivo, che riflettono necessità che possono essere molto distanti¹¹, il cui intento è giungere all'obiettivo comune di gestire *repositories* funzionali e durevoli nel tempo, senza contrastare con il contesto tecnologico e normativo specifico in cui si agisce¹².

Dunque, non si vede la necessità di enunciazioni teoriche o regole precise: queste linee guida non sono formulate a scopo introduttivo rispetto alla gestione dei dati e alla costruzione dei *repositories*, né si propongono come modelli circoscritti e “certificanti”, ma sono destinati a un pubblico di professionisti ed esperti che vi possano trovare un *framework* di riferimento per l'elaborazione delle proprie strategie di gestione e conservazione di dati e documenti e l'implementazione dei propri sistemi. A tal proposito, infatti, a completamento di FAIR sono state elaborate metriche e griglie di valutazione che consentano di misurare l'effettiva *FAIRness* delle soluzioni utilizzate¹³, mentre, riguardo TRUST, il processo può dirsi inverso: più che introdurre concetti innovativi, vogliono sistematizzare quelli già esistenti e, in parte, già delimitati da standard, sulla base dei quali si innesta la formulazione dei principi stessi. Su questa scia, l'elaborazione dei principi TRUST può essere considerata come espediente per sopperire alla mancanza, nei principi FAIR, di una prospettiva sul mantenimento a lungo termine della reperibilità, dell'accessibilità e della riusabilità dei dati, ricercata in precedenza con l'integrazione di questi principi con standard e certificazioni di conservazione¹⁴.

In conclusione, è da sottolineare che la discussione riguardo alle potenzialità e alle criticità connesse all'utilizzo di queste *guidelines*, data anche la loro recente formulazione, è attuale e aperta nell'ambito internazionale della conservazione digitale e, di conseguenza, sensibile a sempre nuovi input e soggetta a prospettive diverse.

¹¹ Si pensi, ad esempio, all'utilizzo degli open data per la diffusione degli articoli scientifici di matrice divulgativa da un lato, e alla riservatezza dei dati gestiti in ambito sanitario dall'altro.

¹² Si consideri il già citato caso della certificazione dei *repositories*: si può scegliere di vincolarla a ISO 16363, a CoreTrustSeal o a griglie valutative customizzate su altri standard, come quelli rilasciati dall'European Telecommunications Standards Institute (ETSI).

¹³ Si veda, a proposito, la nota 3.

¹⁴ Questa “lacuna” è evidenziata nel rapporto della Commissione Europea «Turning FAIR Data into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data» (European Commission, and Directorate-General for Research and Innovation. «Turning FAIR Data into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data». 2018, p. 22. http://publications.europa.eu/publication/manifestation_identifier/PUB_KI0618206ENN – ultima consultazione 29/11/2020).

BIBLIOGRAFIA

- [1] Bartoli, Margherita, Fabiana Guernaccini, and Giovanni Michetti. «Preservation of Linked Open Data». *JLIS: Italian Journal of Library, Archives and Information Science = Rivista Italiana Di Biblioteconomia, Archivistica e Scienza Dell'informazione*: 11, 2, 2020, no. 2 (2020). <https://doi.org/10.4403/jlis.it-12633>
- [2] Biro, Timea, Harrower Natalie, Immenhauser Beat, Maryl Maciej «Sustainable and FAIR Data Sharing in the Humanities: Recommendations of the ALLEA Working Group E-Humanities Description», *ALLEA - All European Academies*, Berlin, February 2020.
- [3] European Commission, and Directorate-General for Research and Innovation. «Turning FAIR Data into Reality: Final Report and Action Plan from the European Commission Expert Group on FAIR Data». 2018. http://publications.europa.eu/publication/manifestation_identifier/PUB_KI0618206ENN
- [4] Lin, Dawei, Jonathan Crabtree, Ingrid Dillo, Robert R. Downs, Rorie Edmunds, David Giaretta, Marisa De Giusti, et al. «The TRUST Principles for Digital Repositories». *Scientific Data* 7, no. 1 (14 May 2020): 1–5. <https://doi.org/10.1038/s41597-020-0486-7>
- [5] McMurry, Julie A., Nick Juty, Niklas Blomberg, Tony Burdett, Tom Conlin, Nathalie Conte, Mélanie Courtot, et al. «Identifiers for the 21st Century: How to Design, Provision, and Reuse Persistent Identifiers to Maximize Utility and Impact of Life Science Data». *BioRxiv*, March 20, 2017, 117812. <https://doi.org/10.1101/117812>
- [6] Mokrane, Mustapha, Recker Jonas, «Coretrustseal-certified repositories. Enabling Findable, Accessible, Interoperable, and Reusable (FAIR) Data», 2019. https://ipres2019.org/static/pdf/iPres2019_paper_74.pdf
- [7] Wilkinson, Mark D., Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, et al. «The FAIR Guiding Principles for Scientific Data Management and Stewardship». *Scientific Data* 3, no. 1 (March 15, 2016): 1–9. <https://doi.org/10.1038/sdata.2016.18>
- [8] Wilkinson, Mark D., Sansone, SA., Schultes, E. et al. «A design framework and exemplar metrics for FAIRness». *Scientific Data* 5, 180118 (2018). <https://doi.org/10.1038/sdata.2018.118>