









C. Cipolloni, M. P. Congi





Un sistema integrato di validazione dei prodotti della ricerca: il caso del progetto GeoSciencesIR

Sessione organizzata dal Gruppo di Lavoro Open Science della CoPER nell'ambito della GenOA week 2025 - Genova 20-11-2025





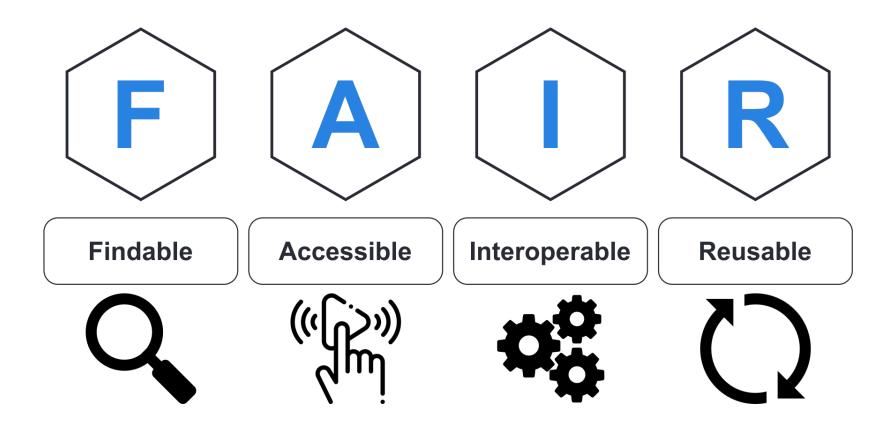








Principi FAIR come strumento di valutazione















Prodotti della ricerca FAIR?

I principi FAIR sono una serie di buone pratiche per la **gestione** e la **condivisione** dei prodotti della ricerca. I dati possono essere FAIR pur avendo restrizioni di accesso (e.g., per dati personali o segreto industriale), rispettando il principio:















Prodotti della ricerca FAIR?

Sebbene spesso si parli di "Dati FAIR", i principi si applicano a **tutti gli oggetti digitali prodotti dalla ricerca** (metadati, dati, vocabolari e ontologie, software, modelli, protocolli).

Per ciascun prodotto il processo di gestione per raggiungere i vari principi **FAIR** identificati per ognuno di essi, considerando che non tutti i principi saranno applicabili a specifici prodotti del progetto richiede la definizione di una buona **matrice di FAIRness**:

- ✓ Chiara (comprensibile nel suo scopo).
- ✓ Realistica (facile da implementare).
- ✓ Discriminante (misurare il successo e fornire indicazioni per migliorare).
- ✓ Misurabile (oggettiva e automatizzabile).
- ✓ Universale (applicabile a tutti i prodotti digitali).













Prodotti della ricerca FAIR?

Le catene di produzione di ciascun prodotto della ricerca evidenzia come ognuno di essi può essere in qualche modo connesso ad altri prodotti e che il processo di apertura di qualità è complesso e richiede un'azione integrata.

Il raggiungimento della **FAIRness** (adesione ai principi FAIR) di alcuni prodotti e/o sottoprodotti non può quindi prescindere dal raggiungimento di livelli **FAIR** anche nei prodotti collegati, in particolare nei **metadati**, che fungono da elemento fondamentale per la descrizione del prodotto e per migliorarne di conseguenza la sua **reperibilità** (**findability**) e **accessibilità** (**accessibility**).













Prodotti della ricerca in GeoSciencesIR e loro paths



Prodotti GeoSciencesIR **Dataset**

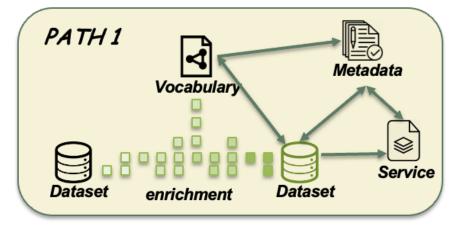


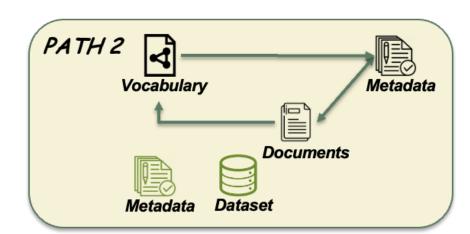


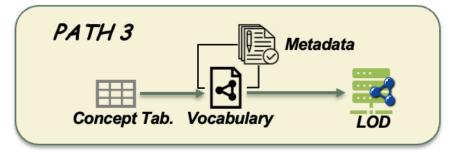


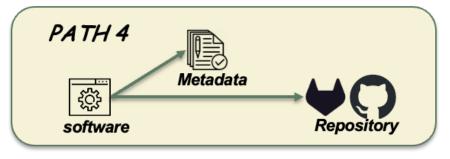


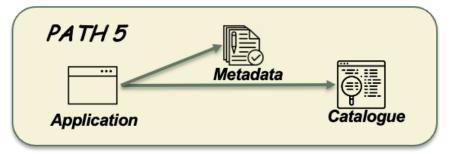




















F4-A1.1-1.2



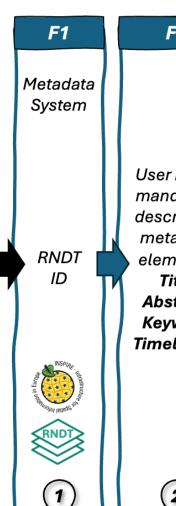
11-12-13



R1

Prodotto: Metadata

Per ogni prodotto sulla base dei principi FAIR applicabili è stata realizzata una catena per raggiungere la **FAIRness**, inserendo i riferimenti agli standard e ai processi necessari.



Metadata

F2 User insert mandatory descriptive metadata elements: Title **Abstract** Kevword **Timeliness** (2)

Metadata System User Metadata System is building in a Standard way 😃 DATASET in order to ID support machine request Document 3 (4)

F3

Metadata should have a strategy for maintain metadata User need to keep alive metadata after dataset dead (5)

A2

User has to follow metadata guideline for filling specific fields mainly using controlled vocabulary and thesaurus

6

User and system support the licence machine readable User needs to improve the quality of data filling element: Lineage Data Quality

(7)



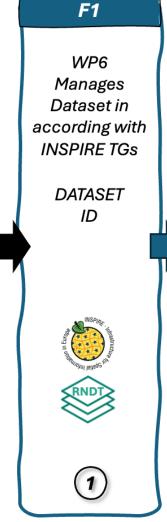








Prodotto: Dataset



Dataset

F2

User needs to documented well the dataset in metadata: Title Abstract **Keywords** Use limitation Access condition Data distribution link Distribution **Format** Lineage **Conformity** degree

(2)

F4-A1.1-1.2

Metadata
System
searching
+
A network
service check
will be
performed
following
INSPIRE TGs

A test on the INSPIRE performance critatia will be performed



3

11-12-13

The
interoperability is
based on the
INSPIRE TGs for
Dataset and
Network Service

User with WP6 harmonise syntactically and semantically the dataset





R1-R2

To improve the quality of data user fills metadata element but also the Network Service Capabilities To improve usability, correct distribution of datasets must be done in network services.









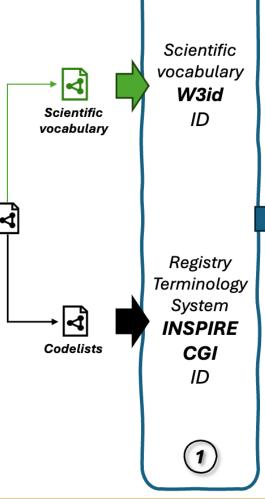
F1











F2

User insert mandatory and optional descriptive metadata elements:
Title
Licence
Editors
Language
Timeliness
Version
....



2

DCAT

F4-A1.1-1.2

The user
must check
if other
vocabulary
and term exist
+
System make
concept
indexable
Accessible
with Content
negotiation



3

I1-I2-I3

User has to
prepare a
structure
vocabulary
following rules it's
important specify
hierarchy and
eventually
relation with
other resources

4

R1

User and system support the licence machine readable User needs to improve the quality of vocabulary with:

Description accessRights and Policy
GeneratedBy



OWL SKOS





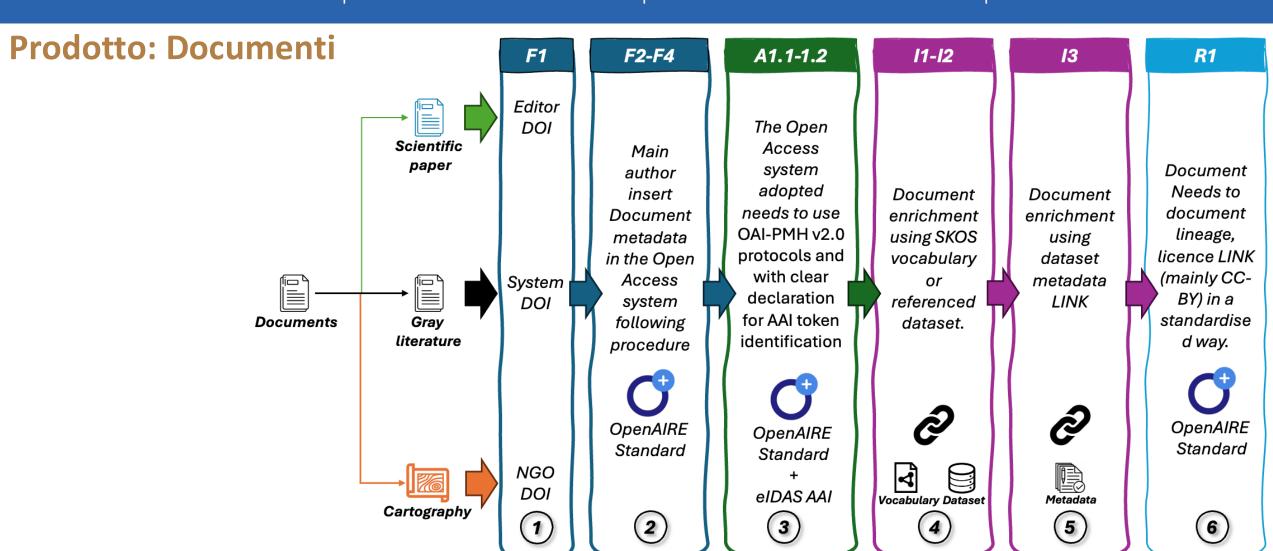
















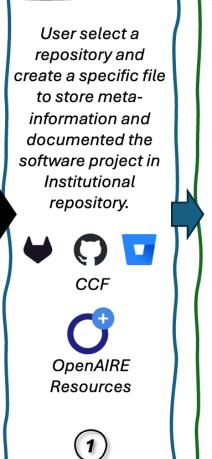
F1-F3





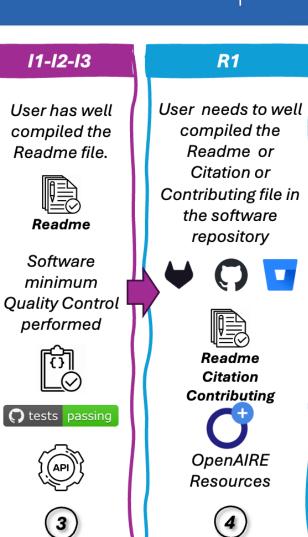


Prodotto: Software



software















Il sistema integrato di valutazione

L'interfaccia web GeoSFAIR è un'applicazione open source, modulare e flessibile, è stata realizzata per essere di facile utilizzo e coerente con gli standard, permettendo di testare tutte le risorse. Il sistema di monitoraggio integrato GeoSFAIR consente di scegliere i criteri di test per l'attribuzione del livello di FAIRness per ogni prodotto.

Per ciascun principio, un codice esecutivo basato su API (richiamabile anche in modo asincrono) per verifica la conformità raggiunta, ovvero assegnando un valore percentuale di FAIRness (0-100%).













Esempio di criterio di test

➤ The system check with an INSPIRE Validator embedded the conformance to «Common Requirements for ISO/TC 19139:2007 based INSPIRE metadata» if not pass stop the control and assign value = 0;

otherwise pass to check the presence of following element in the XML metadata:

gmd:title gmd:abstract gmd:keyword gmd:MD MaintenanceFrequencyCode

➤ If all the 4 element are empty stop the control and assign value = 0.3 (30%);

otherwise pass to check which element respect rules:

- Count value of gco: CharacterString in gmd: title if the value is between 10 and 100 get 1 point;
- Count value of gco: CharacterString in gmd: abstract if the value is between 250 and 600 get 1 point;
- Count first value of gmx: Anchor in gmd:descriptiveKeywords if the value is > 3 get 1 point if not count also value of gco:CharacterString in gmd:descriptiveKeywords and if the value is > 3 get 1 point;
- Select value of gmd:dateStamp@Date and compare with value of Test@Date if the number of day is < of number of gmd:MD_MaintenanceFrequencyCode@value [eg. Annual = 365, weekly = 7, ...] get 1 point;
- The control to assign value count total point ratio if (1/4) Value = 0.4, if (2/4) Value = 0.6, ...













Tool di validazione









Vocabulary



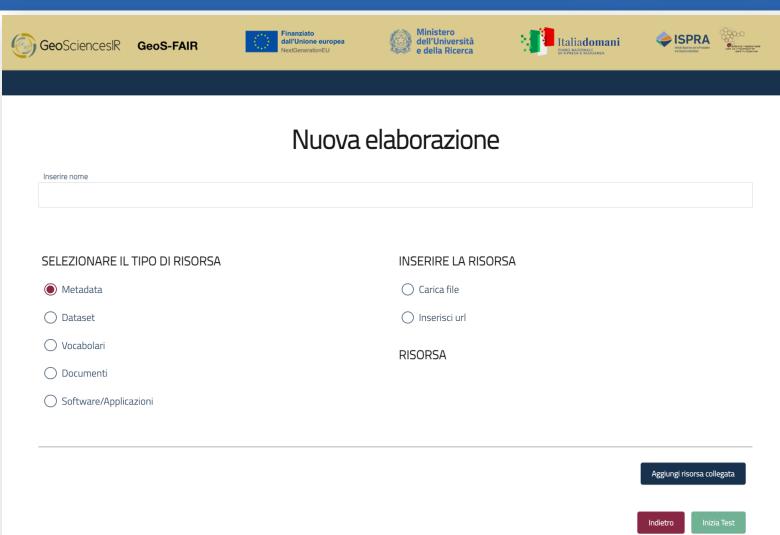
Documents





software Application

















Verificare Combinate tra prodotti

Risorsa collegata nº 1 **DATASET + METADATO + VOCABOLARIO(i)** SELEZIONARE IL TIPO DI RISORSA INSERIRE LA RISORSA Metadata Carica file Inserisci url Dataset Vocabolari **RISORSA** Documenti Inserire url Nuova elaborazione https://hub.geosciences-ir.it/MetaEditor/rest/document?id=ispra_rm%3AMeta Software/Applicazioni Dataset TEST Catena Dataset ID ispra_rm:001THP_DT Risorsa collegata n° 2 SELEZIONARE IL TIPO DI RISORSA INSERIRE LA RISORSA Metadata Carica file SELEZIONARE IL TIPO DI RISORSA INSERIRE LA RISORSA Dataset Inserisci url Metadata Carica file ○ Vocabolar **RISORSA** O Dataset Inserisci url O Documenti Vocabolari https://geosciencesir.isprambiente.it/geoserver/gsmle2/ITHACA-PS-GIR/ows?ser RISORSA O Software/Applicazioni O Documenti Inserire url https://hub.geosciences.cloud/git/geosciences/vocab/-/blob/main/paleosites/

O Software/Applicazioni











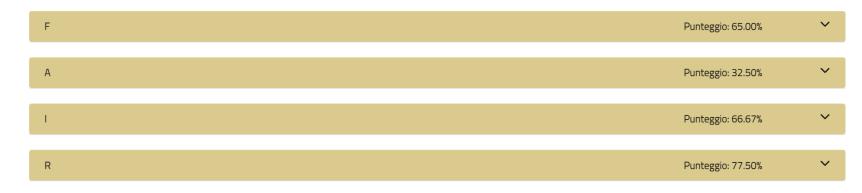


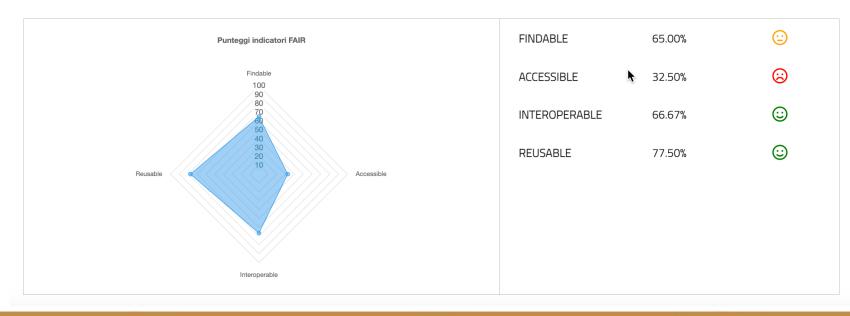




Risultato della validazione

TEST ESEGUITI

















Risultato della validazione

F		Punteggio: 65.00%	^
(i) F1.1	Metadata is identified by a persistent identifier	Punteggio: 50/100	
	★ Identifier URL present NOT_PASSED		
	★ Identifier URL resolvable NOT_PASSED		
	\checkmark Class 1: 'Basic metadata for datasets and datasets of classes 1.2 and	nd 1.3' PASSED	
© F1.2	Metadata is identified by a globally unique identifier	Punteggio: 100/100	
	★ Identifier defined in local system NOT_PASSED		
	✓ Identifier defined in global system PASSED		
	✓ Conformance Class 1: 'Basic Metadata for Datasets and Datasets P	PASSED	
⊕ F2	Resource are described with rich metadata	Punteggio: 100/100	
	✓ Optional elements PASSED		
	✓ Mandatory elements PASSED		
	✓ Class 1: 'Basic Metadata for Datasets and Datasets PASSED		
E F4	Metadata are registered or indexed in a searchable resource	Punteggio: 60/100	
	✓ Available in the system PASSED		
	➤ Dataset in external repository NOT_PASSED		
	➤ Dataset in national repository NOT_PASSED		
	✓ Namespace 'Appendix 1A' present in WFS, WMS or WCS PASSED		













PNRR "GeoSciences IR" - Missione 4 "Istruzione e Ricerca" - Componente 2 "Dalla ricerca all'impresa" - Linea di investimento 3.1 "Fondo per la realizzazione di un sistema integrato di infrastrutture di ricerca e innovazione" Finanziato dall'Unione Europea NextGenerationEU CUP: I53C22000800006



















Istituto Nazionale di Oceanografia e di Geofisica Sperimentale

















