



Open Science, dati FAIR e l'Osservatorio Virtuale

Marco Molinaro & Fabio Pasian
INAF – Osservatorio Astronomico di Trieste

Risorse in Astrofisica



Scenario attuale:

- Le pubblicazioni scientifiche basate su dati d'archivio sono di volume comparabile a quelle del periodo proprietario;
- La ricerca scientifica è basata su dati multi-strumento, multi-banda, osservativo/numerici, multi-messenger;
- Un'analisi esauriente delle attuali moli dati richiede grosse collaborazioni o riutilizzo dei dati;
- Il finanziamento alla base della ricerca scientifica è pubblico e tali devono essere i suoi risultati.

Per permettere che le risorse dati astrofisiche siano efficacemente utilizzate e ri-utilizzate dalla comunità scientifica, occorre uno sforzo di **omogeneizzazione dei metodi di ricerca e accesso alle risorse dati stesse.**

Principi e Standard

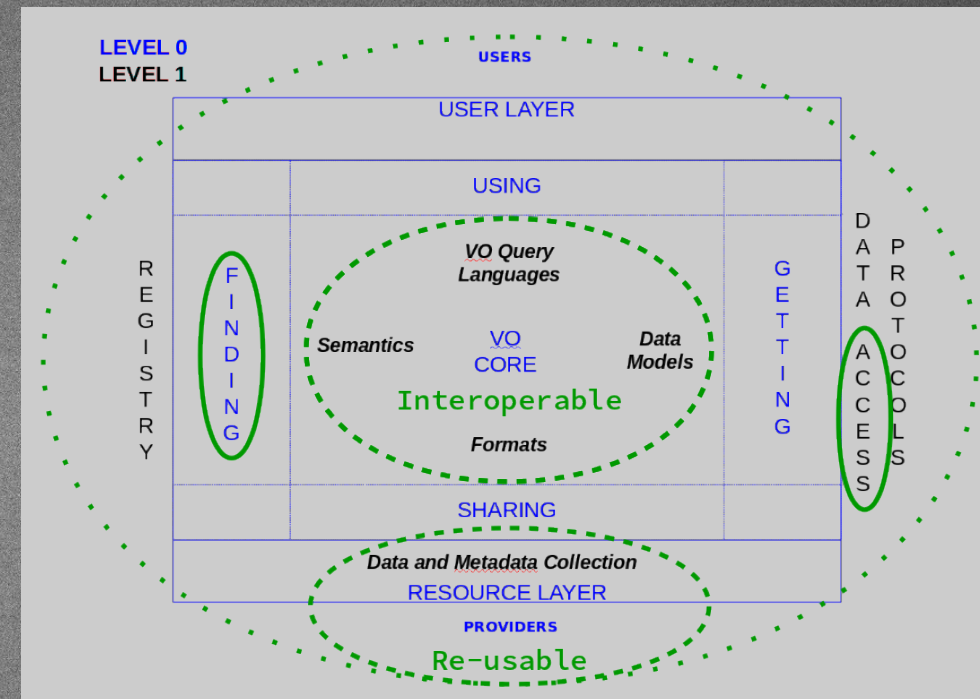


Open Science involves transitioning from a system in which it is difficult to access and locate the results of scientific research to one that openly distributes results to all kinds of end users [...]

(Open Science Conference, EU 2016)

I dati devono essere **FAIR**

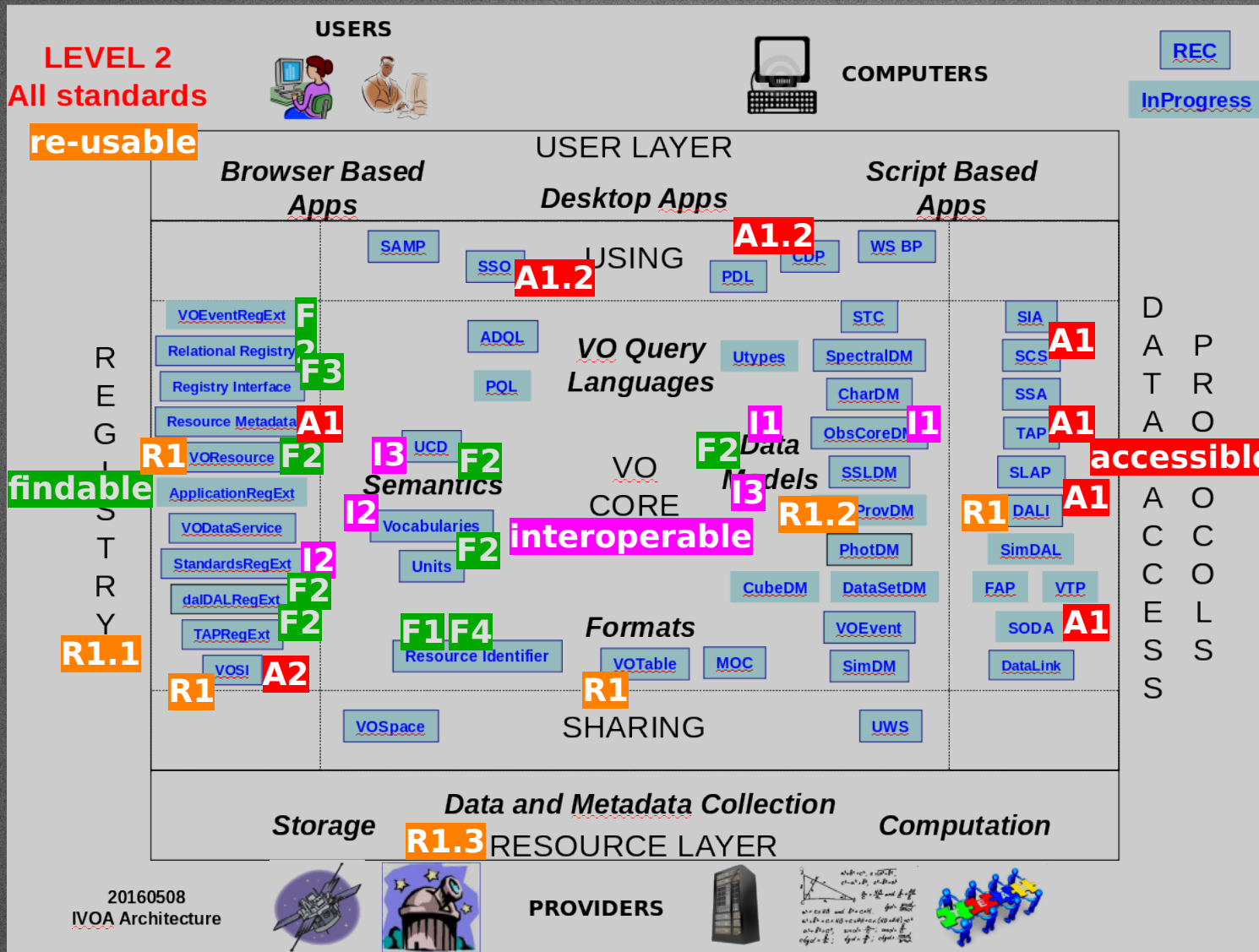
- **F** indable
 - metadata, persistent IDs, indexing
- **A** ccessible
 - standard communication protocols, accessible metadata
- **I** nteroperable
 - common metadata modelling, FAIR vocabularies, metadata referencing
- **R** e-usable
 - provenance, licensing, domain-relevant standards



(FAIR & IVOA Architecture)

(force11.org/fairprinciples)

In dettaglio



Formato FITS

- F2
- I1
- R1
- R1.2
- R1.3



In dettaglio



F1	(meta)data are assigned a globally unique and eternally persistent identifier	Resource Identifiers, Registry
F2	data are described with rich metadata	VOResource & extensions, Semantics, Data Model
F3	(meta)data are registered or indexed in a searchable resource	Registry Interfaces, Relational Registry, Resource Identifiers
F4	metadata specify the data identifier	VOResource, Resource Identifiers, Registry
A1	(meta)data are retrievable by their identifier using a standardized communications protocol	All the Data Access Layer Protocols
A1.1	the protocol is open, free, and universally implementable	VO standards and protocols are public, open and implementation agnostic
A1.2	the protocol allows for an authentication and authorization procedure, where necessary	Single Sign-On, Credential delegation Protocol
A2	metadata are accessible, even when the data are no longer available	VO Standard Interfaces, Registry
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation	VO-DML, Data Models
I2	(meta)data use vocabularies that follow FAIR principles	VO Vocabularies, Semantics, Standards Registry Extension
I3	(meta)data include qualified references to other (meta)data	Data Models, Semantics, Unified Content Descriptors
R1	meta(data) have a plurality of accurate and relevant attributes	VOResource, Data Access Layer Interface, VO Standard Interfaces, VOTable format
R1.1	(meta)data are released with a clear and accessible data usage license	Registry annotation, generally public
R1.2	(meta)data are associated with their provenance	Provenance Data Model (and protocols)
R1.3	(meta)data meet domain-relevant community standards	IVOA enabled data providers



Conclusioni



- Le risorse dati astrofisiche esposte tramite gli standard dell'IVOA sono FAIR
- Non tutte le risorse in astrofisica sono presenti nel VO
 - Serve uno sforzo per completare questo passo
 - Maggiore interazione fra data provider e IVOA
- La conformità dei dati ai principi FAIR non è sufficiente
 - Servono servizi di accesso e computazionali collocati sopra le risorse dati
 - IVOA e data provider devono lavorare assieme a questa standardizzazione

Grazie dell'attenzione!