

Open Science, dati FAIR e l'Osservatorio Virtuale

Marco Molinaro, Fabio Pasian

Istituto Nazionale di Astrofisica – Osservatorio Astronomico di Trieste

Abstract. Il ciclo di vita dei dati astrofisici non deve esaurirsi nel ciclo di progettazione, produzione, analisi e pubblicazione dei risultati. Vi sono varie ragioni a supporto: le pubblicazioni basate su dati d'archivio sono di volume comparabile a quelle del periodo proprietario; la ricerca è sempre più basata su dati multi-strumento, multi-banda, osservativo/numerici, multi-messenger; un'analisi esauriente delle attuali moli dati è fattibile solo da grosse collaborazioni o riutilizzando i dati; il finanziamento della ricerca astrofisica è spesso pubblico e tali devono essere i suoi risultati. Per permettere che le risorse dati astrofisiche siano efficacemente utilizzate occorre uno sforzo di omogeneizzazione dei metodi di ricerca e accesso alle risorse dati stesse. Esistono linee guida e tecnologie che si occupano di questo aspetto e sono sostenute dalla Comunità Europea. Qui spieghiamo come l'Osservatorio Virtuale sia un'implementazione di quelli che sono oggi noti come FAIR principles.

Keywords. Osservatorio Virtuale, Astrofisica, FAIR data, Open Science, Interoperabilità

Introduzione

La Comunità Europea sta spingendo affinché la ricerca scientifica costruisca una cultura del dato e della sua interoperabilità secondo il concetto dell'Open Science e i principi FAIR (Findable, Accessible, Interoperable, Re-usable, Wilkinson et al. 2016). A questo scopo ha iniziato il progetto EOSC (European Open Science Cloud) la cui fase iniziale è finanziata attraverso il progetto Horizon 2020 EOSCpilot. La Dichiarazione EOSC definisce in dettaglio gli intenti di tale iniziativa e fa chiari riferimenti ai principi FAIR per quanto riguarda le implementazioni.

Nel dominio dell'astrofisica quanto dettato sia dai principi FAIR che dalle prime analisi del progetto EOSC sono da tempo oggetto di discussione e implementazione all'interno della comunità dell'Osservatorio Virtuale (VO, Virtual Observatory) che si raccoglie attorno all'iniziativa globale dell'IVOA (International Virtual Observatory Alliance) di cui il progetto VObs.it è il rappresentante per l'Italia, mentre a livello comunitario le coordinazione avviene sotto il nome di EURO-VO e attualmente confluisce nel pacchetto DADI del progetto H2020 ASTERICS.

L'iniziativa VO, con gli standard tecnologici sviluppati dall'IVOA, rappresenta lo stato dell'arte per le buone abitudini nella gestione dei dati in astronomia e risponde ai requisiti di accessibilità e interoperabilità dei dati necessari per un efficace sfruttamento dei dati acquisiti delle osservazioni e prodotti delle simulazioni numeriche in campo astrofisico. Nel seguito di questo articolo verrà fornita una descrizione dell'architettura dell'IVOA, §1, della sua connessione diretta ai quattro fondamenti dei principi FAIR, §2, e infine si scenderà nei dettagli più specifico su come gli standard VO coprano lo scenario complessivo

della data FAIRness §3. Nelle conclusioni verranno discusse le differenze fra FAIR e VO.

1. L'architettura dell'IVOA

L'architettura dell'IVOA (Arviset et al. 2010) rappresenta, con tre livelli sovrapposti, lo schema di funzionamento del VO. I livelli, numerati 0, 1 e 2, vanno dal generale al particolare.

Il livello 0 (Figura 1) rappresenta una descrizione di alto livello delle componenti del VO. Vi si riconoscono due strati orizzontali principali, il Resource Layer (risorse disponibili) e lo User Layer (utenti): il primo a condividere le risorse (dati, servizi, ...), sharing, il secondo a utilizzarle, using.

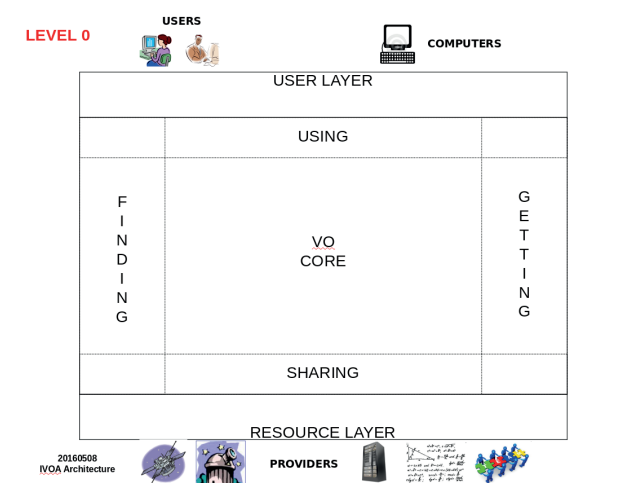


Fig. 1
Architettura dell'IVOA, livello 0,
descrizione generale dei concetti
e componenti di base

La connessione fra questi due strati è garantita dalle due sezioni verticali (getting e finding), che rappresentano, rispettivamente, i canali di accesso alle e di scoperta delle risorse disponibili da parte degli utenti. Il tutto è permesso da quello che è il cuore dell'architettura (VO core) che non è altro che la base dell'interoperabilità, parola chiave per la comunità VO che, proprio sotto il nome di Interoperability Meeting, si riunisce due volte all'anno per discutere e migliorare le specifiche tecniche.

Già al livello 0 troviamo un primo riferimento alle parole chiave dei principi FAIR, la F di findable, di cui parleremo meglio in seguito al §2, mentre abbiamo poco fa accennato alla I di interoperable.

Il livello 1 dell'architettura IVOA (Figura 2) aggiunge al livello precedente alcune definizioni nel gergo VO e descrive meglio le sue componenti; così il Resource Layer si definisce in funzione delle reali risorse che i providers mettono a disposizione: spazio dati, collezioni di dati e di metadati, risorse di calcolo e servizi di vario genere, mentre lo User Layer mostra che l'utilizzo delle risorse può avvenire tramite applicazioni di vario tipo, grafiche, a riga di comando e integrate nelle pagine web (questo avviene fondamentalmente perché il protocollo base di comunicazione del VO è l'HTTP).

I concetti di finding e getting vengono riscritti in funzione delle aree delle specifiche tecniche che se ne occupano: il Registry (finding) che è la parte del VO che si occupa di mantenere un elenco strutturato e annotato delle risorse disponibili attraverso le tecnologie VO, e i Data Access Protocols (getting) che definiscono le specifiche di accesso alle risorse (con chiaro accenno alla A di accessible dei principi FAIR).

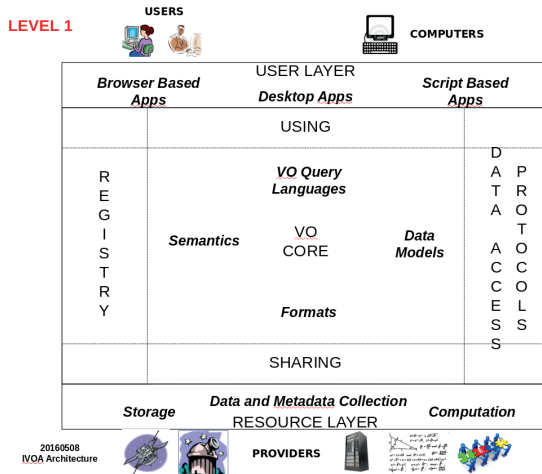
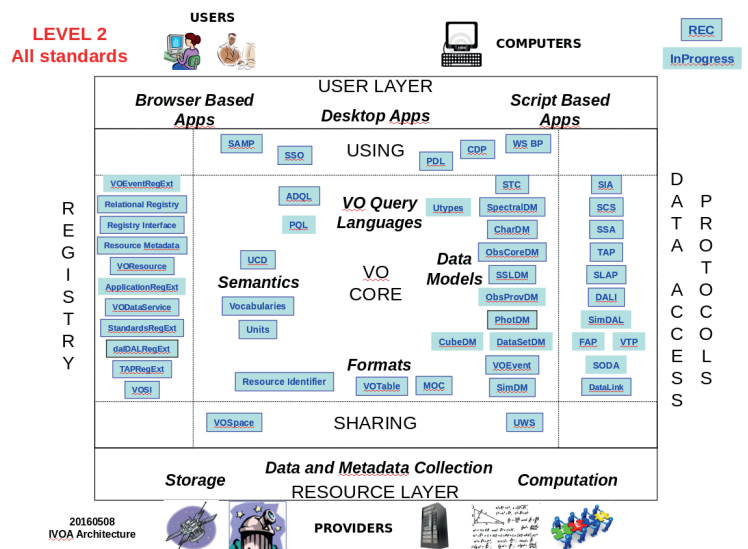


Fig. 2
Architettura dell'IVOA, livello 1,
descrizione più dettagliata
delle componenti principali

Anche il nucleo centrale di interoperabilità viene dettagliato meglio, proponendo la modellistica delle risorse, la semantica (nelle annotazioni delle risorse e nei dati stessi), i formati dei dati e i linguaggi di ricerca come nuclei fondamentali per garantire l'interoperabilità delle risorse.

Il livello 2 (Figura 3) non fa altro che posizionare nello spazio di competenza i vari standard definiti dall'IVOA per garantire l'interoperabilità delle risorse, così fornendo una panoramica, piuttosto dettagliata e densa, di formati, protocolli, modelli, sistemi di re-

Fig. 3
Architettura dell'IVOA, livello 2:
ovvero dove si posizionano tutti
gli standard definiti dall'Alliance



gistrazione delle risorse, vocabolari, linguaggi e quanto altro serva in funzione dei casi d'uso portati all'attenzione dell'IVOA. Si vedrà meglio, nel §3, a cosa corrispondano questi standard all'interno dei principi FAIR.

2. I principi FAIR e l'architettura IVOA

Descrivendo l'architettura dell'IVOA abbiamo già incontrato alcuni punti di connessione con i principi FAIR a cui la Comunità Europea fa riferimento nella sua volontà di costruzione di una European Open Science Cloud. L'unica lettera, se vogliamo, mancante, è la R di re-usable.

Tuttavia la comunità astrofisica si è riunita nel VO proprio per permettere un migliore e più efficiente utilizzo dei dati e per scambiarsi informazioni sulle migliori tecniche per descriverli e renderli accessibili e interoperabili. Quindi la riutilizzabilità dei principi FAIR è direttamente connessa al modo in cui il Resource Layer viene costruito a partire dalle specifiche del VO e investe un po' tutta l'architettura del VO, da chi distribuisce i dati a chi li riutilizza, passando per i metodi di ricerca e accesso agli stessi.

Un modo di descriverlo è rappresentato dalla Figura 4, in cui, unendo i primi due livelli dell'architettura VO, e tenendo conto di quanto descritto nel § 1, si mostra come il lavoro svolto dall'IVOA nel dominio dell'astrofisica nei suoi 15 anni di vita altro non sia che quello che oggi è stato definito dall'Open Science e dai concetti dei principi FAIR.

Non tutto quello che è identificato dai principi FAIR è tuttavia parte del VO, nè il VO si ferma alla sola data FAIRness, come si cercherà di mostrare in dettaglio e discutere nella prossima sezione e nelle conclusioni.

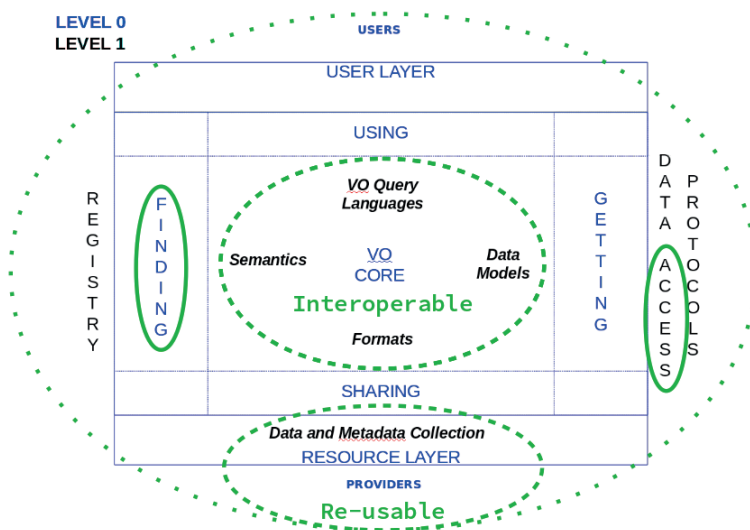


Fig. 4
Rimaneggiamento dei livelli 0 e 1 dell'architettura IVOA con sovrapposti i concetti chiave dei principi FAIR, schematicamente collocati sopra le componenti che più li rappresentano

3. Gli standard dell'IVOA a copertura dei principi FAIR

Andiamo ora in dettaglio, cercando di vedere quanto il VO e i principi FAIR siano simili. Per farlo partiamo riportando i principi FAIR così come riportati in Wilkinson et al. 2016 nella box 2. Li riportiamo in lingua inglese per comodità, anche perché riteniamo che

una traduzione, visti i termini tecnici, possa più nuocere che risultare utile. Nel seguito identificheremo i principi con l'etichetta lettera-numero (es: A1.2) con cui gli elementi dell'elenco puntato iniziano.

To be findable:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

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 re-usable:

- R1. meta(data) have 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 their provenance.
- R2. (meta)data meet domain-relevant community standards.

Quello che si nota subito è il ricorrere continuo del termine metadata, ovvero tutte le annotazioni ai dati che permettono a quest'ultimo di essere compreso da chi il dato deve utilizzarlo anziché apparirgli come un insieme di numeri o descrizioni testuali.

I metadati sono parte del cuore non solo dei principi FAIR, ma anche del VO, dove sono parte fondamentale del Data Model, per collegare i concetti astratti alle loro rappresentazioni fisiche (compito del Data Model Working Group dell'IVOA) e sono utilizzati dai protocolli e dalle descrizioni delle risorse per permettere un sistema di ricerca, accesso e utilizzo del dato che possa essere automatizzato quanto più possibile e, allo stesso tempo, fornisca un modo di interpretare correttamente il contenuto d'informazione ricercato o recuperato dal data provider.

Ma andando punto per punto nell'elenco possiamo stabilire meglio le connessioni con le specifiche tecniche dell'IVOA. Le specifiche, qui di seguito riportate virgolettate nel testo, sono tutte disponibili presso il Document Repository dell'IVOA all'indirizzo <http://www.ivoa.net/documents>.

3.1 Findable

F1 richiede l'uso di unique and persistent identifiers; a livello VO esiste una specifica, "I-

VOA Identifiers” che si applica a qualunque oggetto o risorsa (termine generico) si voglia sia identificabile e raggiungibile (collezione di dati, singolo dataset, organizzazione, ...). L'unicità globale di questi identifier è garantita dal modo in cui sono costruiti. Quello che manca, parzialmente è la persistenza indefinita, poiché è legata ai singoli provider essendo l'IVOA uno sforzo collaborativo. Parzialmente perché la specificità del VO per la descrizione di base delle sue risorse sta andando nella direzione di includere i DOI fra i possibili identifier.

F2 parla di metadata richness. Quasi ogni specifica del VO descrive quali metadati andrebbero associati alle collezioni, dati, servizi cui sono associati. Restando strettamente all'interno del Registry, ovvero del repository dove confluiscono e vengono mantenute tutte le descrizioni delle risorse del VO, esiste la specifica “VOResource” che descrive i metadati e la serializzazione di base che ogni risorsa documentata del Registry deve avere. Tale specifica è poi ampliata da estensioni specializzate in funzione del tipo di risorsa che debba essere descritta: “Standards Registry Extension”, “Simple Data Access Layer Registry Extension”, “VOResource Schema Extension for Describing Collections and Services”, “Table Access Protocol Registry Extension”. Questo vale per insiemi di dati e servizi, e comprende informazioni sulla creazione, cura, preservazione, localizzazione della risorsa, riferimenti bibliografici, estensione delle collezioni di dati in termini di coordinate spaziali, temporali, dello spettro elettromagnetico, disponibilità del servizio e policy di accesso. Tutti questi dettagli sono disseminati attraverso il registry, mentre altri metadati sono annotati alla risposta dei servizi per descrivere i dati stessi, oltre che i servizi. Per le annotazioni, molto lavoro rientra nelle specifiche del Semantics WG, dove gli UCD “An IVOA standard for Unified Content Descriptors” con relativo “Controlled Vocabulary” e lo standard sulle “Units in the VO” si uniscono ai vocabolari e al loro sistema di manutenzione per permettere l'annotazione di qualsiasi campo dati si voglia descrivere, senza contare le annotazioni specifiche dei Data Model.

F3 è coperto a livello VO dal Resource Registry parzialmente descritto nel paragrafo precedente. In più si può aggiungere che fra le sue interfacce ce n'è una che è basata sullo standard OAI-PMH (Open Archive Initiative – Protocol for Metadata Harvesting), rendendo le risorse astrofisiche interoperabili (fino a un certo livello) anche oltre il proprio dominio di ricerca. I metadati di base su cui l'OAI-PMH e lo standard “VOResource” si basano sono quelli noti come Dublin Core, utilizzati universalmente per la metadescrizione di risorse. Da notare che, oltre all'OAI-PMH, i Registry IVOA possono esporre un'interfaccia basata su un sistema relazionale (“Registry Relational Schema”) per una più efficace ricerca condizionale delle risorse desiderate.

F4 richiede che l'identifier sia parte dei metadati, cosa richiesta per ogni risorsa VO. Gli “IVOA Identifier” (IVOID) sono praticamente delle URI specializzate che possono essere risolte nelle rispettive risorse.

In generale, le capacità di ricerca delle risorse in campo VO copre in modo completo le richieste legate ai principi FAIR, forse perfino andando oltre i dettami in quanto si occupa di descrivere non solo i dati, ma i servizi stessi di accesso agli stessi e qualsiasi altra risorsa collegata alle collezioni di dati astrofisici.

3.2 Accessible

A1 ripercorre, per il VO (Registry in particolare), quanto già visto per F3 e F4. Esiste una specifica, “IVOA Registry Interfaces” che descrive i protocolli di accesso alle risorse identificate del rispettivo IVOID, ed esiste un protocollo specializzato (oltre all’OAI-PMH), “Registry Relational Schema”, per un accesso ottimizzato con capacità di ricerca filtrata a tutte le risorse del Registry. Oltre a questo, a partire da ciascun IVOID di risorsa dati, è possibile risalire agli eventuali servizi (standard o meno) di accesso alle risorse dati descritte. Questo è fatto attraverso i protocolli definiti dal Data Access Layer WG che descrivono l’accesso a cataloghi di sorgenti astrofisiche, immagini, dati spettrali e ogni altro tipo di risorsa disponibile, incluse serie temporali, sistemi di notifica di eventi transienti, e dati provenienti da simulazioni numeriche. Il punto A1.1 è automaticamente coperto in quanto tutte le specifiche VO sono open e pubbliche, mentre il punto A1.2 è solo parzialmente coperto perché sistemi di autenticazione e autorizzazione sono ammessi e descritti sulle singole risorse (si veda il “Single Sign-On Profile”), ma non esiste un reale sistema interoperabile fra tutti gli attori che richiedano data policy comuni. Esiste anche un protocollo, il “Credential Delegation Protocol”, che si occupa di descrivere come le credenziali possano essere passate da un servizio o risorsa ad un altro, permettendo così un’interoperabilità che arriva fino al livello delle risorse di calcolo.

A2 è vero per il VO sebbene non sia auspicabile. La specifica “IVOA Support Interfaces” definisce, in particolare, il sistema grazie al quale una risorsa descrive la sua availability, ovvero il suo stato di attività al momento della richiesta. I metadati del Registry sono presenti e accessibili anche se le risorse specifiche sono temporaneamente non disponibili. Se una risorsa non è disponibile per lungo tempo è prassi etichettarla inactive mantenendo i metadati, ma senza distribuirli.

L’accessibilità di dati e metadati è uno dei punti chiave per l’IWOA, il cui lavoro parte sempre da requisiti, sotto forma di casi d’uso, che descrivono come una determinata risorsa possa essere trovata all’interno del mondo sfaccettato del VO e, una volta individuata, come possa essere acceduta e utilizzata, il tutto in maniera trasparente per l’utente, ovvero senza doverne conoscere ubicazione fisica reale e sistema di gestione specifico.

3.3 Interoperable

I1 descrive la necessità di un linguaggio comune per descrivere le informazioni contenute nei dati esposti. L’IWOA sta terminando lo sviluppo di un metalinguaggio, il “VO-DML - A consistent Modeling Language for IVOA Data Models”, proprio per uniformare non solo le annotazioni delle proprie risorse, ma il modo in cui tali annotazioni nascono, ovvero i modelli di dominio più specifico. Esiste inoltre uno di questi modelli, l’“Observational Data Model Core Components”, che definisce proprio gli elementi base comuni a qualsivoglia osservazione astrofisica si voglia esporre in rete (definendo inoltre le linee guida per l’uso di un protocollo per l’accesso a tale insieme di annotazioni).

I2 prevede che i vocabolari utilizzati siano a loro volta FAIR. I vocabolari del VO sono degli standard, in quanti tali open, descritti come risorse a loro volta all’interno del Registry, accessibili in vari formati convenzionali (non interni al VO, ma generici quali SKOS o

RDFa) quindi garantendone sia l'interoperabilità interna ed esterna al dominio di appartenenza che il riutilizzo.

La connessione interna, fra metadati e metadati e fra metadati e dati, richiesta da I3, è garantita proprio dall'ambiente complessivo descritto dagli standard dell'IVOA. Le annotazioni che una specifica impone nella descrizione di un servizio rimanda a namespace specifici che permettono di risalire all'insieme dei metadati utilizzati. Tutte le specifiche legate ai gruppi di Semantics e Data Model lavorano proprio in quest'ottica.

L'interoperabilità delle risorse, intese come generiche risorse dati o servizi standard per gli stessi, è al cuore delle iniziative dell'IVOA. Questo fa sì che quanto definito dai principi FAIR sia di fatto già attuato in campo VO, probabilmente anche oltre i dettami stessi.

3.4 Re-Usable

R1 richiede, nuovamente, ricchezza di metadati e ridondanza degli stessi. Ogni servizio che segua le specifiche del VO e sia registrato espone i propri metadati sia attraverso il Registry ("VOResource" e sue estensioni) che direttamente attraverso il sistema di capabilities previsto dalle specifiche "Data Access Layer Interface" e "IVOA Support Interfaces". Le risorse che rappresentano esplicitamente delle collezioni di dati hanno i loro metadati esposti in maniera globale attraverso il Registry e trasmettono i loro dati, con formato predefinito in forma di "VOTable Format Definition", accompagnati da un insieme di annotazioni così come richiesto dai vari protocolli e standard. Tutto questo usando i modelli, i descrittori e i vocabolari descritti, ad esempio, nella sezione §3.3 dedicata all'interoperabilità. R1.1 specifica che debba esistere un chiara espressione delle licenze applicabili ai dati. Una prima considerazione da fare qui è che le risorse del VO partono generalmente non dall'essere semplicemente open ma dall'essere solitamente pubbliche, fatto salvo un possibile iniziale periodo proprietario che ne garantisca un'indagine prioritaria da parte di chi ha richiesto l'osservazione o il run di simulazione. Fatto salvo questo, a livello di Registry c'è la possibilità di definire delle licenze per le risorse, possibilmente facendo riferimento (tramite URI globali esterni al mondo VO) alle licenze solitamente utilizzate, quali possono essere le Creative Commons o le GPL.

R1.2 entra nel merito della tracciabilità del dato, dalla sua origine, passando per le possibili manipolazioni allo stesso, gli attori coinvolti e tutto quanto definisca la storia del dato come risorsa. Al di là delle informazioni di base reperibili attraverso l'uso del Dublin Core a livello di registrazione delle risorse VO, è ormai pronto per la standardizzazione il "Provenance Data Model", che si occuperà proprio degli aspetti sopra citati all'interno del VO, prendendo il via dall'analogo standard del World Wide Web Consortium (W3C) per garantirne l'interoperabilità verso risorse cross-domain fuori dal campo specifico dell'astrofisica.

R1.3, infine, richiede a chi espone risorse dati, nel nostro caso in astrofisica, nient'altro che quanto l'IVOA ha cercato di fare nei suoi 15 anni di vita e sta proseguendo a fare anche oggi.

Il riutilizzo dei dati e delle risorse, ovvero la possibilità di individuare le risorse più adatte al proprio lavoro di ricerca o indagine, parte della necessaria collaborazione fra più strumenti, sistemi di indagine e osservatori necessari alla ricerca astrofisica. L'infra-

struttura informativa definita dall'IVOA ne prende semplicemente atto e definisce le linee guida per poterla attuare.

Un punto base per capire quanto la ri-utilizzabilità dei dati sia importante nel VO lo definiscono i suoi standard di formato ("VOTable Format Description", "HEALPix Multi-Order Coverage Map", "Hierarchical Progressive Survey" e "Simple Application Messaging Protocol"), che uniscono in un sistema unico dati, metadati e applicazioni e permettono uno scambio fra le parti che garantisce la comprensione sia da parte di attori diversi che in diversi momenti, per esempio dopo un salvataggio su un sistema locale per un successivo riutilizzo.

4. Conclusioni

Nel capitolo §3 abbiamo descritto come le specifiche e l'architettura dell'IVOA non siano altro che un'implementazione che ha precorso i dettami dell'Open Science e dei principi FAIR per il dominio di ricerca dell'astrofisica.

Al di là di quanto l'IVOA ha fatto e sta facendo bisogna ricordare che il VO stesso prende spunto da una realtà, quella del formato FITS (Flexible Image Transport System, Pence et al. 2010) che già alla fine degli anni '70 preparava un formato di scambio dati completo di annotazioni e metadati, incluso un minimo modello interno per gli stessi. Questo formato ha giocato e gioca un ruolo attivo nell'accomunare la comunità astrofisica e l'IVOA ne ha sempre tenuto conto in fase di sviluppo dei propri standard, per non divergere da quanto la comunità già conosceva e utilizzava. Questo è vero al punto che i principi F2, I1, R1, R1.2 e R1.3 sono già rispettati proprio da questo formato, che definisce quindi già un piccolo sistema di scambio dati FAIR in sé stesso.

Restano ovviamente delle differenze fra quanto esiste nel mondo VO dell'astrofisica e i principi FAIR. Da un lato si può dire che tutti i dati esposti tramite il VO sono FAIR, dall'altro però bisogna specificare che non tutti i dati astrofisici sono raggiungibili attraverso il VO e che la conformità ai principi FAIR non è sufficiente per l'astrofisica.

Per far sì che i dati astrofisici diventino tutti FAIR è necessario uno sforzo implementativo da parte dei centri dati e provider dell'astrofisica (istituzioni, organizzazioni, enti di ricerca, ...), ad esempio inserendo uno strato di traduzione per l'accesso ai propri dati che rispetti gli standard VO. Questo è possibile se gli attori in gioco e l'IVOA lavorano di concerto nello sviluppo e miglioramento degli standard.

Ma è anche necessario andare oltre la FAIRness dei soli dati. È importante che servizi di calcolo specifici possano essere agganciati alle risorse dati stesse. In astrofisica questo è necessario per fornire immagini ridotte, ritagliate sulle necessità dell'utenza, per generare cataloghi di sorgenti e oggetti, per permettere che le moderne tecniche di machine learning possano interagire direttamente coi dati. Questo è possibile se l'IVOA contribuisce a definire degli standard per la descrizione, reperibilità e utilizzo di queste risorse e se i provider dei dati si pongono nell'ottica di fornire capacità di calcolo sopra i dati stessi.

Riferimenti bibliografici

Christophe Arviset, Severin Gaudet and the IVOA TCG, "IVOA Architecture – version

1.0”, IVOA Note, 23 Novembre 2010

Pence, W. D.; Chiappetti, L.; Page, C. G.; Shaw, R. A.; Stobie, E., “Definition of the Flexible Image Transport System (FITS), version 3.0”, *Astronomy and Astrophysics*, Volume 524, id.A42, Dicembre 2010

Mark D. Wilkinson, Michel Dumontier, IJsbrand Jan Aalbersberg, Gabrielle Appleton, Myles Axton, Arie Baak, Niklas Blomberg, Jan-Willem Boiten, Luiz Bonino da Silva Santos, Philip E. Bourne, Jildau Bouwman, Anthony J. Brookes, Tim Clark, Mercè Crosas, Ingrid Dillo, Olivier Dumon, Scott Edmunds, Chris T. Evelo, Richard Finkers, Alejandra Gonzalez-Beltran, Alasdair J.G. Gray, Paul Groth, Carole Goble, Jeffrey S. Grethe, Jaap Heringa, Peter A.C 't Hoen, Rob Hooft, Tobias Kuhn, Ruben Kok, Joost Kok, Scott J. Lusher, Maryann E. Martone, Albert Mons, Abel L. Packer, Bengt Persson, Philippe Rocca-Serra, Marco Roos, Rene van Schaik, Susanna-Assunta Sansone, Erik Schultes, Thierry Sengstag, Ted Slater, George Strawn, Morris A. Swertz, Mark Thompson, Johan van der Lei, Erik van Mulligen, Jan Velterop, Andra Waagmeester, Peter Wittenburg, Katherine Wolstencroft, Jun Zhao & Barend Mons (2016), “The FAIR guiding principles for scientific data management and stewardship”, *Nature – Scientific Data*, 3, DOI: 10.1038/SDATA.2016.18

<https://ec.europa.eu/research/openscience/index.cfm>

https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf

<http://www.openarchives.org/OAI/openarchivesprotocol.html>

<http://dublincore.org/documents/dcmi-terms/>

Autori



Marco Molinaro marco.molinaro@inaf.it

Laureato in Fisica, con indirizzo in Astrofisica, si occupa ormai da oltre un decennio di data management e data science presso l'INAF-OATs e in connessione al centro Italiano Archivi Astronomici (IA2). È membro dell'IVOA e del suo Technical Coordination Group, dove attualmente ricopre il ruolo di vice-chair del Data Access Layer Working Group.

Fabio Pasian fabio.pasian@inaf.it

Astronomo Ordinario presso INAF-OATs, è stato a capo dell'archivio HST di ESO, PI degli archivi del TNG, manager del DPC di Planck e coinvolto attualmente nell'SGS di EUCLId. Da sempre coordinatore delle attività VO italiane e referente in tal senso a livello europeo (EURO-VO) e internazionale (come membro dell'Executive Committee dell'IVOA).

