

Observatoire Virtuel et FAIR, des principes fondamentaux à la pratique

Interopérabilité et pérennisation des données de la recherche : comment FAIR en pratique ? Retours d'expériences

Paris, 27 novembre 2018

André Schaaff (Observatoire Astronomique de Strasbourg), Laurent Bourgès (Observatoire des Sciences de l'Univers de Grenoble), Karin Dassas (Institut d'Astrophysique Spatiale), Jean-Michel Glorian (Institut de Recherche en Astrophysique et Planétologie de Toulouse), Pierre Le Sidaner (Observatoire de Paris), Jean-Charles Meunier (Laboratoire d'Astrophysique de Marseille), Michèle Sanguillon (Laboratoire Univers et Particules de Montpellier)
(Comité de programme du Semi-Hack-a-Thon annuel d'OV France)

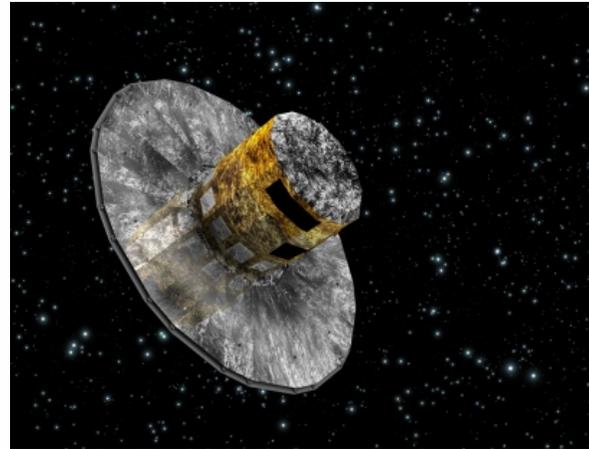


www.ivoa.net & www.france-ov.org

□ L'astronomie, une « grande » productrice de données, « Big Data » par nature...



Vue d'artiste de SKA
Crédit: SKA Organisation



Vue d'artiste de Gaia
Crédit: ESA



Vue d'oiseau du Very Large Telescope
Crédit: J.L. Dauvergne & G. Hüdepohl (atacamaphoto.com)/ESO



Hubble, le télescope spatial
Crédit: NASA, 2002



XMM-Newton
Crédit: Image courtesy of ESA
Interopérabilité et pérennisation des données de la recherche



Sous le charme des Nuages de Magellan, ALMA
Crédit: ESO/C. Malin

□ ... et avec une grande diversité

Observations des télescopes sol et spatiaux.

Très **grands relevés** du ciel (informations homogènes sur un grand nombre d'objets).

Bases de données à **valeur ajoutée** (CDS, NED).

Données **bibliographiques** (journaux académiques, base de données ADS maintenue par la NASA).

Données de **modélisation**.

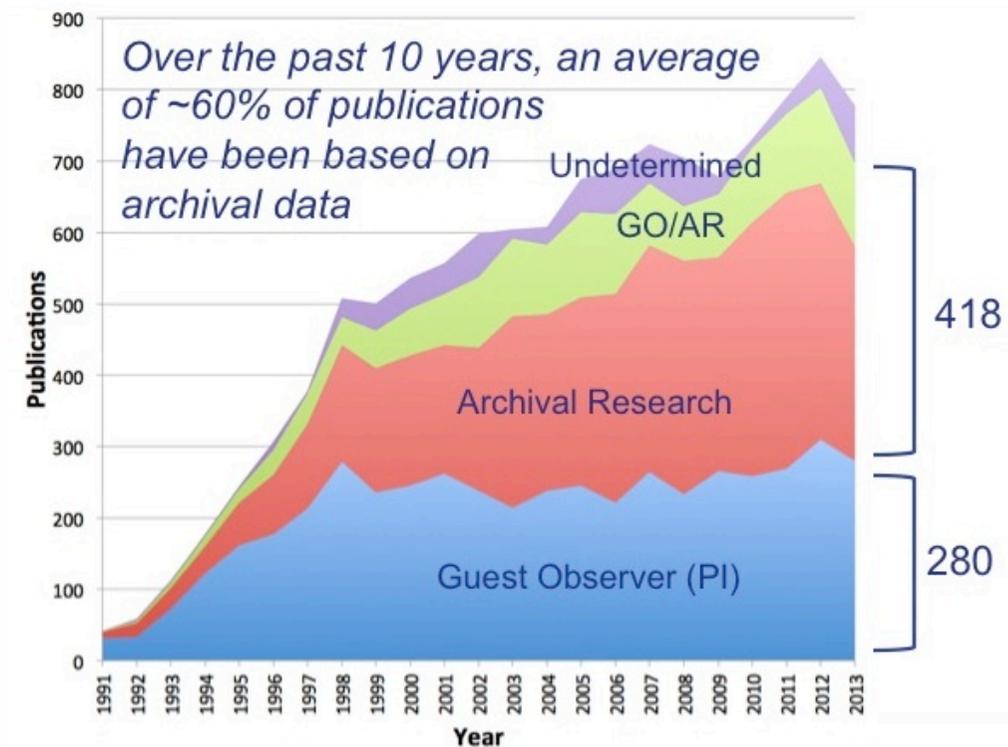
...

□ La Réutilisation des données

Pour des objectifs différents des objectifs initiaux.

Une **augmentation** significative du **retour scientifique** des gros investissements.

La combinaison d'observations par différents instruments permet de comprendre les phénomènes à l'œuvre et engendre une part significative et croissante des publications.



Crédit: Robert J. Hanisch, Space Telescope Science Institute

□ Les principes fondamentaux

Findable

Accessible

Interoperable

Re-usable

L'Observatoire Virtuel astronomique est né de la volonté de rendre les données de l'astronomie mondiale accessibles (A) à toute la communauté, de permettre de les retrouver (F) facilement, de faciliter les échanges entre services de données et outils (I), de documenter et de préciser les métadonnées (R).

« Mais il n'a pas vocation à promouvoir des technologies, libres ou propriétaires, dont le choix est du ressort des fournisseurs de données et des outils associés. »

Alors, comment faire ?

□ L'Observatoire Virtuel

Au niveau international, l'International Virtual Observatory Alliance (IVOA) est un consortium

- composé d'initiatives OV nationales, « continentales », d'agences intergouvernementales
- mettant en place des groupes de travail et d'intérêt se réunissant semestriellement depuis 2001
- élaborant des protocoles et des standards d'interopérabilité autour des données
- travaillant à la « à la W3C » (Notes, Working draft, RFC, Recommandations).

□ Les membres



- Argentine Virtual Observatory
- Armenian Virtual Observatory
- AstroGrid, United Kingdom
- Australian Virtual Observatory
- Brazilian Virtual Observatory
- Chinese Virtual Observatory
- Canadian Virtual Observatory
- Chilean Virtual Observatory
- European Space Agency
- European Virtual Observatory
- German Astrophysical Virtual Observatory
- Hungarian Virtual Observatory
- Japanese Virtual Observatory
- Observatoire Virtuel France
- Russian Virtual Observatory
- South African Astroinformatics Alliance
- Spanish Virtual Observatory
- Italian Virtual Observatory
- Ukrainian Virtual Observatory
- US Virtual Observatory Alliance
- Virtual Observatory India

L'Action Spécifique **OV France** du CNRS-INSU **soutient** la participation française à l'IVOA, en finançant des **missions** aux réunions semestrielles, en organisant des **journées annuelles** et un **Semi-Hack-a-Thon**.

□ Un projet organisé et bien avancé

Après 17 années d'existence et la définition des « briques » de base, leur intégration dans les services de données, la mise à jour ou la création d'outils « compatibles » OV, le projet se poursuit en étroite collaboration avec les grands projets du domaine, en partageant l'expertise acquise et en tenant compte des besoins nouveaux.

Il est référencé par  FAIRsharing.org
standards, databases, policies

□ Des chercheurs très impliqués

L'astronomie a une longue tradition de collaboration étroite entre les chercheurs et les ingénieurs dans le cadre de projets internationaux.

Les groupes de travail et d'intérêts sont présidés par des chercheurs ou des ingénieurs.

L'OV est « science driven », les priorités scientifiques sont définies annuellement par un comité scientifique.

Des écoles sont organisées régulièrement pour sensibiliser les chercheurs aux outils et standards de l'OV.

□ Les difficultés

Au départ, des **temps** plutôt **long** pour trouver des **consensus** et valider les protocoles et standards.

Des **doutes**, un certain scepticisme d'une partie de la **communauté**.

Ne **pas** tout **réinventer** (ex.: OAI-PMH utilisé pour le Registre des ressources) et des standards existants. Mais des **risques d'obsolescence** et des choix à revoir (ex.: abandon de SOAP au profit de REST).

L'OV est très largement utilisé même si **les utilisateurs ne le voient pas !**

Faciliter l'**accès** aux **données** peut nécessiter un **redimensionnement** matériel.

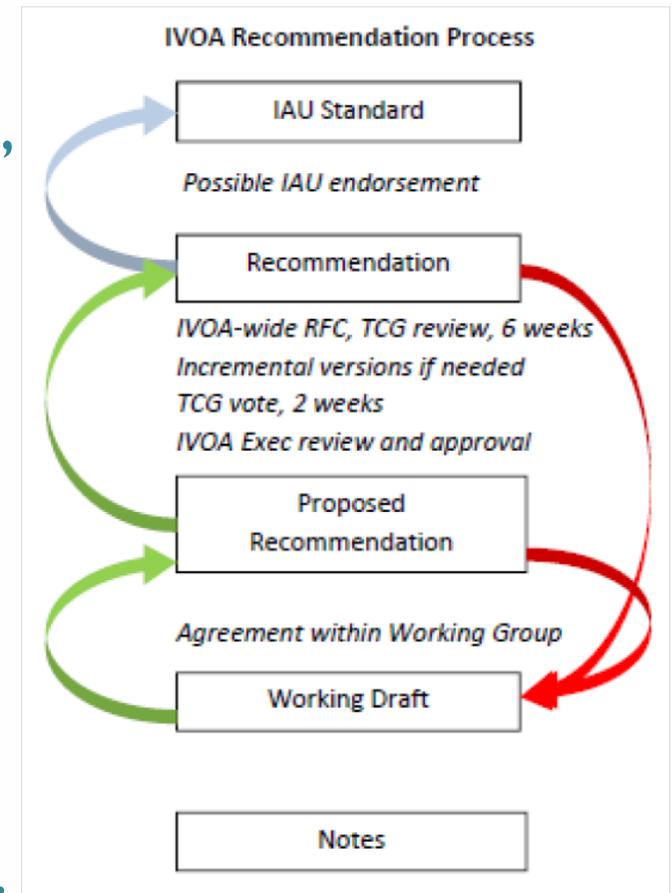
□ Les solutions

Une **définition** claire et **structurée** du **processus** jusqu'à la **validation** d'un **standard**, dans un **temps** raisonnable.

La définition de **priorités** scientifiques pour orienter les travaux.

Des **liens** étroits avec les **grands projets** du domaine.

Un **système** d'interopérabilité ouvert à tous.



□ Pour résumer

L'OV a notamment des « **pages jaunes** », le Registre.

L'interopérabilité, des données et outils associés, au travers de **standards** et de **protocoles** ouverts est la **raison d'être de l'OV**.

Findable

Accessible

Interoperable

Re-usable

L'implémentation des **standards** et **protocoles** rend les **données** immédiatement **accessibles**.

Des standards comme **Provenance** (pour tracer l'origine des données, d'un document ou d'un instrument), les **UCDs** (description) ainsi qu'un effort de **documentation** des **données**.

□ Conclusion

Près de 50 protocoles et standards que les grands projets intègrent maintenant en amont.

100 « autorités » déclarent au moins une ressource en ligne dans le Registre.

Un travail de longue haleine indispensable pour affronter (même pour l'astronomie) le changement d'échelle.

Une mise en pratique, souvent implicite, des principes FAIR, une architecture ouverte pour trouver, accéder, échanger, croiser, ..., les données.

Questions ?