SF2A 2023
M. N'Diaye, A. Siebert, N. Lagarde, O. Venot, K. Baillié, M. Béthermin, E. Lagadec, J. Malzac, J. Richard (eds)

OPEN SCIENCE AND ASTRONOMY

F. Genova[1]

Abstract. Open Science is rapidly developing, with a strong political support in France and in more and more other countries, as well as at the European and international levels. Many disciplinary fields are joining the endeavour. Astronomy has been one of the early pioneers of Open Science. Astronomers from all around the world are using data and tools provided by others in their daily research work, thanks to the Virtual Observatory data sharing framework and to the dedicated work of the data and service providers. It is essential for us to track the ongoing developments and to fully participate in them, to ensure that astronomy continues to be recognized as one of the important components of the rapidly evolving Open Science landscape. This paper discusses several hot topics which have been tackled by the Research Data Alliance, and how they interface with the way astronomy implements Open Science in its field, namely, the Global Open Research Commons, the evaluation of FAIR (Findable, Accessible, Interoperable, Reusable), and the necessary evolution of research evaluation.

Keywords: Open Science, Astronomy, Evaluation of research, Global Open Research Commons, FAIR,

1 Introduction

Astronomy has been a pioneer of Open Science and FAIR, with the FITS standard, first described by Wells, Greisen & Harten (1981), which links data together with the metadata enabling to reuse it, and the standards developed and maintained by the *International Virtual Observatory Alliance* (IVOA, `https://www.ivoa.net/`), which started in 2002, enabling astronomers to Find, Access and Interoperate data. One can also remind that the *Centre de Données astronomiques de Strasbourg* (Strasbourg astronomical Data Centre - CDS) was created in 1972 to serve the international astronomical community, and that the data from the *International Ultraviolet Explorer* (IUE, which was active from 1978 to 1996) had been downloaded in average five times already in 1994 (Wamsteker & Griffin 1995).

Thanks to the hard work of the data and tool providers and of the IVOA community, most data produced in astronomy are available in the Virtual Observatory data sharing framework, which enables the world-wide community to access and use it for their research. To remain relevant, the astronomical disciplinary data sharing framework has to continue to evolve because the scientific, technical, and political contexts in which it operates evolve.

This paper summarizes the current evolution of the Open Science context and the way astronomy has been tackling the data aspects of Open Science over the years, long before Open Science gained its prominent position in the political scientific landscape. Three among the challenges astronomy has to face in the current landscape will then be discussed.

2 The fast-evolving Open Science context

Open Science, which includes open publication, open data and now open software, has been a hot topic on the political agenda in France and at the European and international levels for a while. In France the *Loi pour une République numérique** (Law for a digital Republic), was published in 2016 with the aim to regulate the circulation of data and knowledge. The ministry in charge of research, currently called *Ministère de l'Enseignement Supérieur et de la Recherche* (MESR - Ministry of Higher Education and Research) is engaged in the *Plan*

[1] Université de Strasbourg, CNRS, Observatoire astronomique de Strasbourg, UMR 7550, 67000 Strasbourg, France
*`https://www.gouvernement.fr/action/pour-une-republique-numerique`

national pour la Science Ouverte (PNSO - National Plan for Open Science). The initial version of the Plan[†], published in 2018, was built around three pillars: generalize open access to publications; structure research data and make it available through open access; and be part of a sustainable European and international open science dynamic. Its second version[‡], published in 2021, includes source codes produced by science and the necessary transformation of practices among its four paths: generalising open access to publications; structuring, sharing and opening up research data; opening up and promoting source code produced by research; and transforming practices to make open science the default principle.

The MESR implemented the *Comité pour la Science Ouverte* (CoSO - Committee for Open Science)[§] in 2018. The Recherche Data Gouv (`https://recherche.data.gouv.fr/en`) ecosystem was inaugurated in July 2022. It includes modules to support research teams on all issues linked to data, namely local data management clusters (*Ateliers de la Données*), thematic reference centres, resource centres and institutional reference centres, a repository to deposit and disseminate data (when no disciplinary repository is available) and a registry to search for data in Recherche Data Gouv and other repositories.

The European Union has also been engaged for a while in supporting Open Science. Open Science[¶] is an item of the *Our digital future* chapter of its Research and Innovation Strategy 2020-2024: "Open science is a policy priority for the European Commission and the standard method of working under its research and innovation funding programmes as it improves the quality, efficiency and responsiveness of research." The ambitions of the European Union's Open Science policy are open data with a reference to FAIR; the European Open Science Cloud (EOSC); new generation metrics; mutual learning exercise on open science – altimetrics and rewards; future of scholarly communication; rewards; research integrity and reproducibility of research results; education and skills; and citizen science.

There are also notable activities at the international level in support to Open Science, with strong, repeated statements of the G7/G8 Ministers of Research[‖], for instance this recent one, included in the declaration from their meeting held in Sendai in May 2023: "The G7 will collaborate in expanding open science with equitable dissemination of scientific knowledge and publicly funded research outputs including research data and scholarly publications in line with the Findable, Accessible, Interoperable, and Reusable (FAIR) principles." The G7 Open Science Working Group was established in 2016. The OECD published in 2006 its pioneering *Recommendations on access to research data from public funding* and updated it in 2021 (`https://www.oecd.org/sti/inno/open-science.htm`). UNESCO also produced its *Recommendation for Open Science* in 2021 (`https://www.unesco.org/en/open-science`).

3 Astronomy and Open Science

Astronomy has been a pioneer in the data aspects of Open Science. Sharing astronomical data as widely as possible was required because astronomy relies on multi-wavelength, and now multi-messenger observations and performs statistical studies on large numbers of objects, and because astronomical objects can be variable on a wide range of time scales. It is also important to ensure the maximum possible return on the large investments in space and ground based telescopes by enabling their data to be reused as widely as possible.

Thanks to the IVOA standards and to the fact that they are used by the data and tool providers, astronomical data appears to its users as a single, global, multipolar data infrastructure. The infrastructure is open, anyone can publish a data resource in it or develop a tool using the data. The infrastructure, and thus the hard work required to make it happen, is hidden from the end users since it provides seamless access to data and the VO-enabled tools are interoperable.

Work is indeed required at the disciplinary level to make the most of Open Science: the knowledge of science needs and data resides there. This work has to be interfaced with the generic developments. For instance, the IVOA chose from the start to develop its Registry of Resources using the OAI-PMH standard and the Dublin Core with disciplinary extensions (Hanisch et al. 2007) as the basis for the resource metadata. These standards are widely used, and this enabled the IVOA Registry to be included in the EUDAT B2FIND

[†] `https://www.ouvrirlascience.fr/national-plan-for-open-science-4th-july-2018/`

[‡] `https://www.enseignementsup-recherche.gouv.fr/fr/le-plan-national-pour-la-science-ouverte-2021-2024-vers-une-generalisation-de-la-science-ouverte-en-48525`

[§] `https://www.ouvrirlascience.fr/the-committee-for-open-science/`

[¶] `https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science_en`

[‖] `https://www.ouvrirlascience.fr/les-ministres-de-la-recherche-et-de-la-technologie-du-g7-sengagent-en-faveur-de-la-science-ouverte/`

(https://b2find.eudat.eu/) registry with no major issue, a first step to ensure the visibility of VO-enabled resources in the EOSC.

Many disciplines are joining the Open Science endeavour, some with significant resources and in domains seen as societal priorities. Astronomy continues to update its disciplinary data sharing framework to respond to science needs and technical evolution. The SKA Organisation for instance joined the IVOA in 2022, and collaborations have been established to deal at best with multi-messenger astronomy. We have to make sure that we identify among the numerous developments linked to Open Science in the general context those that we have to follow and to tackle to ensure that we remain relevant. Three examples are discussed in the next section.

4 Challenges for Astronomy in the current Open Science landscape - Three examples

These three examples have been selected because the author has been involved both at the disciplinary level and with building the link with activities at the general level, in particular, but not only, in the framework of the Research Data Alliance (RDA, https://www.rd-alliance.org/). The RDA is a community-driven initiative, an international organisation with the mission to build the social and technical bridges that enable open sharing and re-use of data. Its activities are developed by the community organised in Interest Groups (IG), which are addressing a topic on the medium/long term, and Working Groups (WG) which have 18 months to prepare a Recommendation which will then be submitted to the community before endorsement by the RDA Council. In September 2023 there were more than 100 active RDA Groups tackling a large variety of subjects.

4.1 Astronomy among the Global Open Research Commons

One of the RDA Interest Groups is dealing with *Global Open Research Commons*"** (GORCs). "Commons" are defined as a global trusted ecosystem that provides seamless access to high quality interoperable research outputs and services. The IG is working to reach a shared understanding of what a "Commons" is in the research data space; what functionality, coverage and characteristics does such an initiative require and how this can be coordinated at a global level. The first IG output is a Typology of the essential elements in a Commons (Jones et al. 2023). The IG also set up a Working Group *GORC International Model*††, which recently produced a model or framework to describe Commons attributes. Its Version 1 (Leggott et al. 2023) was submitted for comments to the RDA community in September 2023.

The usual examples of GORCs are the European Open Science Cloud (EOSC), the Australian Research Data Commons or the African Open Science Platform. From the definition of a GORC, it seemed that the IVOA could apply to be recognized as one of them, but it is very different from the examples cited above, since it does not manage services per se, except its Registry of Resources which is provided by several of its partners, in implementations which harvest each other to ensure that their content is aligned. Participating in the WG activities enabled to test the idea and to propose the IVOA as a possible GORC. The WG organised a series of presentations of GORCs, to identify common elements which could appear in the model. The IVOA was the first to be presented, and the WG agreed that it was to be considered as a GORC. Its characteristics have been considered as examples in the definition of the model displayed in Version 1.

4.2 FAIR in Astronomy

By developing and maintaining the Virtual Observatory, astronomy has been providing its community with a rare global, operational framework enabling data sharing and interoperability. Our disciplinary data sharing framework has to be interfaced properly with the systems which are progressively being established at the general level. The key aspect to enable interoperability with the general level are the Registry of Resources and vocabularies.

The FAIR guiding principles encountered a remarkable success since their publication in 2016 (Wilkinson et al. 2016): they provided a simple way to define and check essential characteristics to be fulfilled to efficiently share data. This is evident when one compares them for instance to the properties identified in the OECD "Guiding Principles" in 2007: openness, flexibility, transparency, legal conformity, protection of intellectual

**https://www.rd-alliance.org/groups/global-open-research-commons-ig
††https://www.rd-alliance.org/groups/gorc-international-model-wg

property, formal responsibility, professionalism, interoperability, quality, security, efficiency, accountability, sustainability. This list is well thought and covers a large palette of key aspects of data sharing, but the FAIR guiding principles provide a subset of essential guidelines, as well as a powerful mnemonic.

The report *Six Recommendations for implementation of FAIR practice* produced by the FAIR in practice Task Force of the European Open Science Cloud FAIR working group (European Commission, Directorate-General for Research and Innovation 2020) offers six recommendations on how FAIR can be turned into practice: (1) Fund awareness-raising, training, education and community-specific support; (2) Fund development, adoption and maintenance of community standards, tools and infrastructure; (3) Incentivise development of community governance; (4) Translate FAIR guidelines for other digital objects; (5) Reward and recognise improvements of FAIR practice; (6) Develop and monitor adequate policies for FAIR data and research objects. The necessary evolution of evaluation of research cited in (5) will be discussed in the next Subsection. These recommendations strongly emphasise the importance of community standards, governance and support for FAIR to be turned into practice. This aligns perfectly with the way the astronomy community has been implementing FAIR, by setting up the IVOA for gathering the relevant community at the international level to ensure the governance, definition and maintenance of the standards.

Any astronomer anywhere one the world – providing that they have access to the internet – can Find, Access, Reuse and Interoperate VO-enabled data and tools. For astronomers astronomical data are thus FAIR for their own needs. This is not necessarily true when considering the criteria defined by the RDA FAIR Data Maturity Model WG (FAIR Data Maturity Model Working Group 2020). A study performed during the ESCAPE Cluster project by F. Genova, F. Bonnarel and M. Molinaro showed that we can live with the criteria. The Virtual Observatory enables some of them "for free" once it is implemented on top of the data holdings. But we also found that we have issues with the priority levels assigned to criteria. There is for instance no "Essential" criterion among the Interoperability ones, whereas for us Interoperability has been one of the key IVOA driving forces – the IVOA six-monthly meetings are even called *the IVOA Interoperability Meetings*. The priority levels have been chosen by the RDA community, and the low ranking reflects the fact that very few disciplines have Interoperability as a priority or have been able to achieve it. Find is a dynamic process for us, with the persistent identifier, which appears prominently among the criteria, only one among many parameters. Three of the Reuse criteria refer to licenses: "metadata includes information about the license under which the data can be reused" is deemed essential, "metadata refers to a standard reuse license" and "metadata refers to a machine-understandable reuse license" are deemed important. Not all astronomical data found in the VO will have a reuse license attached. The data are however widely reused, usage being based on disciplinary ethics: cite what you use.

The astronomical community has thus to work in particular on defining and attaching licenses to the data and metadata it shares. Also the Interoperability criteria cite "FAIR-compliant vocabularies", to be used for data and metadata. We have to follow the activities on-going to define what a FAIR-compliant vocabulary is, and make sure that the IVOA vocabularies are compliant.

Even if it is time consuming for the individuals involved, we have to ensure that the discipline is actively represented in the committees and projects which tackle the evaluation of FAIR, to make sure that our point of view, which is the point of view of a legacy disciplinary data sharing framework which enables advanced functionalities and is widely used by its world-wide community, is considered.

4.3 Astronomy and the necessary evolution of evaluation of research

The evolution of research evaluation from being mostly bibliographic index-based is – at last! – being recognized as a priority at the international, European, national, and institutional levels. One can of course cite DORA (`https://sfdora.org/`), the Declaration on Research Assessment, which is a decade old in 2023. The fact that it has been well understood for some time now that the evolution of research evaluation is indispensable to enable Open Science has been a strong incentive to prioritize it, in particular in France and at the level of the European Commission. It is fully included in the Transforming Practices path of the Second French Plan for Open Science, which highlights the issue in two of the three measures it recommends: "Develop and value open science skills throughout the educational and career pathways of students and research staff" and "Value open science and the diversity of scientific productions in the assessment of researchers, of projects and of universities and research performing organizations." Rewards are also discussed among the ambitions of the European Union's open science policy, and evaluation, incentives and rewards are cited in a number of other international reference documents on Open Science.

Astronomy evaluation committees have had to face the issue somehow forever, since the recognition of the

work done by the astronomers involved in conceptualizing astronomical instruments has not always been an easy endeavour, to say the least, whereas these instruments are essential to the practice of astronomy. The scientists involved in the data and code aspects of the discipline encountered the same problem in their careers when these profiles emerged thanks to the early adoption of computing and digital data sharing by the discipline.

A recent evolution is the establishment of the Coalition for Advancing Research Assessment (CoARA – `https://coara.eu`) in December 2022. CoARA gathers research performing organisations which agree to work together to implement the changes described in the *Agreement for Reforming Research Assessment*. "The Agreement, Based on 10 commitments, establishes a common direction for research assessment reform, while respecting organisations' autonomy." As of 21 September 2023, 539 organisations have joined the Coalition, including more than 40 based in France, among which research organisations such as the CNRS, a number of Universities, the *Agence Nationale de la Recherche* (ANR) and the *Haut Conseil de l'évaluation de la recherche et de l'enseignement supérieur* (HCERES).

The RDA also recently set up an Interest Group to identify its role and value in the domain, the RDA Evaluation of Research Interest Group (`https://www.rd-alliance.org/groups/evaluation-research-ig`). "The Interest Group (IG) will be a value-added contribution to the organisations, including research organisations and universities, funders and policy makers who use the RDA as an international, neutral forum: it will enable them to discuss their proposals on the evolution of research evaluation with the (RDA) community, whose comments will help improve these proposals and make them more relevant and applicable. It will also be a high value-added contribution to the community of RDA data experts, who will benefit from better recognition of their activities. The IG will bring their voice to the attention of the organisations, funders and policy makers who participate in the RDA."

The author's experience from years of participation in and chairpersonship of disciplinary evaluation committee shows that the evolution of evaluation of research from being mostly bibliographic index-based is extremely difficult to put into practice, even in a discipline like astronomy which relies on research infrastructures, and which has a significant fraction of its research staff belonging to the *Corps des Astronomes et Physiciens*, whose missions include the so-called *tâches de service*(functional duties). The diversity of **research** tasks, products and profiles is not always fully recognized by the *Conseil National des Astronomes et Physiciens* and the CNRS committee in charge of evaluating astronomy. Legitimate research activities are not only those which produce publications. One still has to fully base evaluation on impact and not on bibliometric indicators. The evolution of science evaluation in astronomy, not only on the aspects linked to Open Science, is still a challenge, even when CNRS is fully involved in the process and is producing very relevant guidelines, since its requires a deep cultural evolution of the whole community. This should be a topic for the next *Exercice de Prospective* (Long Range Planning Exercise) of the discipline.

5 Conclusion

The examples discussed in these proceedings underline only a few of the challenges faced by the discipline in the fast-evolving Open Science context. We have to follow and participate in the on-going developments at all levels, institutional (in CNRS and in our Universities), national, European and international. In France for instance, it is very important to participate in the activities driven by the Committee for Open Science and engage in the Recherche Data Gouv endeavour. At the international level we have to make sure that IVOA is known and recognized, and bring key new topics to the IVOA so that its data sharing framework is properly interfaced with the general framework which is progressively being set up. Participation in EOSC and evaluation and adoption of the new possibilities offered by Artificial Intelligence when they can fit our needs are also among the key challenges linked to the position of astronomy in the Open Science context.

References

European Commission, Directorate-General for Research and Innovation 2020, Six Recommendations for implementation of FAIR practice by the FAIR in practice task force of the European open science cloud FAIR working group, Publications Office. https://data.europa.eu/doi/10.2777/986252

FAIR Data Maturity Model Working Group 2020, FAIR Data Maturity Model: specification and guidelines, Research Data Alliance. `https://doi.org/10.15497/RDA00050`

Hanisch, R., the IVOA Resource Registry Working Group, the NVO Metadata Working Group 2007, Resource metadata for the Virtual Observatory Version 1.12, International Virtual Observatory Alliance. `https://doi.org/10.5479/ADS/bib/2007ivoa.spec.0302H`

Jones, S., Leggott, M., Lopez Albacete, et al. 2023, GORC IG: Typology and Definitions (Version 1.01), Research Data Alliance. https://doi.org/10.15497/RDA00087

Leggott, M., Treloar, A., Woodford, C. J., et al. 2023, The Global Open Research Commons International Model, Version 1, Research Data Alliance. https://doi.org/10.15497/RDA00097

Wamsteker, W., & Griffin, R.E.M. 1995, Ap&SS, 228, 385

Wells, D.C., Greisen, E.W., & Harten, R.H. 1981, A&AS, 44, 363

Wilkinson, M.D., Dumontier, M., IJsbrand, J.A., et al. 2016, Scientific Data, 3, 16001