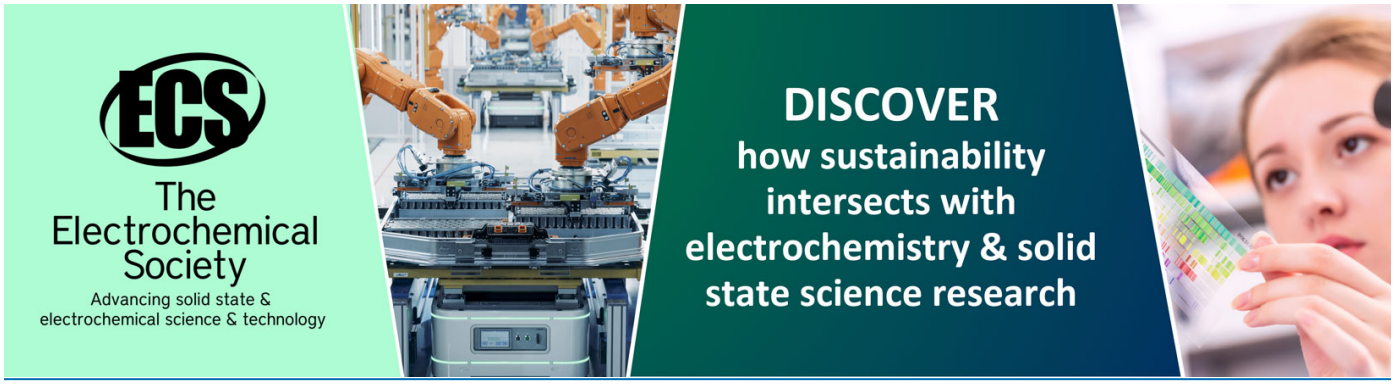
Open data policy and data sharing in Astroparticle Physics: the case for high-energy multi-messenger astronomy



Journal of Physics: Conference Series **718** (2016) 022016 doi:10.1088/1742-6596/718/2/022016

Open data policy and data sharing in Astroparticle Physics: the case for high-energy multi-messenger astronomy

M Spiro

CEA, Irfu, Saclay, France

michel.spiro@cea.fr

Abstract. This paper is a position paper on open data policy and data sharing in Astroparticle Physics, focusing on the case of high-energy multi-messenger astronomy. It is presented on behalf of ApPIC, Astroparticle Physics International Committee[1], IUPAP (International Union for Pure and Applied Physics) working group 10.

1. Introduction

The ApPIC (Astroparticle Physics International Committee[1]) terms of reference were discussed and decided in July 2013 by IUPAP (The International Union for Pure and Applied Physics) after consultation of APIF (the Astroparticle International Forum of funding agencies). They are the following:

- Review on a regular basis the scientific status of the field of Astroparticle Physics;
- Engage in a continuous dialogue with "The Astroparticle Physics International Forum (APIF)" of the Global Science Forum (GSF) and provide scientific advice to APIF, whose members are appointed by funding agencies;
- Comment on and liaise with similar national and international organizations on assessment and road-mapping activities as the need may arise, e.g. for promoting the global coherence of plans, priorities and projects in Astroparticle Physics.

Here the term « astroparticle physics » is defined in a broad sense to include investigations related to the properties of the high-energy universe as well as the dark universe and issues with cosmic relevance – at the interface of astrophysics, nuclear physics, particle physics and cosmology. It also pursues the relevant research in theory and technology development.

This paper is given on behalf of the present members of APPIC[1].

2. High-energy multi-messenger astronomy

High-energy Astrophysics has moved to a level of new quality during the last decade:

- High-energy gamma-ray astronomy has paved the way, with thousands of sources detected from satellites and around 150 from ground, providing fantastic new insights but also leaving many unsolved questions.
- High-energy neutrino astronomy has provided evidence for extra-terrestrial neutrinos at the TeV to PeV scale, with unknown origin.
- Cosmic-ray detectors confirmed the cut-off in the energy spectrum around 1020 eV, but its origin is still debated. No clear evidence for individual sources has been obtained.
- Gravitational wave detectors have increased their visible volume by a factor of 1000 and may open a new observational window in the next two years.

With many effects appearing first just above the level of background, significant results will require a multi-messenger approach, in order to mutually confirm low-significance results, and to trigger follow-up observations. This also includes data from optical astronomy. These considerations also favour an open data policy and virtual observatories combining all data.

3. Recommendations for High-energy multi-messenger astronomy open data policy and data sharing

ApPIC recommends an open data policy to enhance the discovery potential of this publicly funded field, in the spirit of sharing data broadly. This could be implemented at 5 subsequent levels:

- A data validation period, where the data are shared only within the collaboration formed around one experiment.
- First data release for joint analysis (combinations and cross-checks, complementary approaches). This would be implemented through agreements between collaborations.
- An open trigger, on- or off-line, for high-energy astronomy collaborations.
- Open access to data within the community of experts: this implies that the community gets prepared, following the virtual observatory models with help-desk for data and codes.
- Finally, data preservation and legacy have to be foreseen from the very beginning.

A few important considerations must be added:

- MoUs (bottom-up initiated and science driven) have to be signed by funding agencies, with attached resources which have to be planned from the very beginning.
- Competition (e. g. for funding opportunities, fame…) have to be flanked by cooperation and consensus on sharing of data and know-how.
- Within this framework of sharing data, possible false discovery claims are particularly delicate and regulations have to be implemented. Proper credit has to be given by quoting properly the source of data releases (collaboration data releases).

4. Benchmarking against funding agencies recommendations

These statements are in compliance with the section in the NSF (US) Award management guide related to data sharing: Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing. Privileged or confidential information should be released only in a form that protects the privacy of individuals and subjects involved. General

adjustments and, where essential, exceptions to this sharing expectation may be specified by the funding NSF Program or Division/Office for a particular field or discipline to safeguard the rights of individuals and subjects, the validity of results, or the integrity of collections or to accommodate the legitimate interest of investigators. A grantee or investigator also may request a particular adjustment or exception from the cognizant NSF Program Officer.
http://www.nsf.gov/bfa/dias/policy/dmp.jsp

The DOE Office of Science (US) has articulated similar principles guiding their digital data management policy:
The Office of Science affirms that the following principles related to the management of digital research data directly support fulfillment of its mission.
Effective data management has the potential to increase the pace of scientific discovery and promote more efficient and effective use of government funding and resources. Data management planning should be an integral part of research planning.
Sharing and preserving data are central to protecting the integrity of science by facilitating validation of results and to advancing science by broadening the value of research data to disciplines other than the originating one and to society at large. To the greatest extent and with the fewest constraints possible, and consistent with the requirements and other principles of this Statement, data sharing should make digital research data available to and useful for the scientific community, industry, and the public.
Not all data need to be shared or preserved. The costs and benefits of doing so should be considered in data management planning.
http://science.energy.gov/funding-opportunities/digital-data-management/

They also harmonize with the ASTERICS project of ESFRI (European Strategy Forum for Research Infrastructures) project (ASTERICS for Astronomy ESFRI & Research Infrastructure Cluster) proposal which aims to address the cross-cutting synergies and common challenges shared by the various Astronomy ESFRI facilities (SKA, CTA, KM3NeT and E-ELT). With the recent approval of the ASTERICS proposal, the implementation of the ESFRI telescopes will be supported and accelerated, their performance enhanced beyond the current state-of-the-art, and their interoperation as a multi-wavelength and multi-messenger facility realized. An important focal point is the management, processing and scientific exploitation of the huge datasets the ESFRI facilities will generate. In addition, ASTERICS will enable astronomers from across the member states to have broad access to the reduced data products of the ESFRI telescopes via a seamless interface to the Virtual Observatory framework.
https://www.asterics2020.eu/

References

[1] Pierre Binetruy (France), Roger Blandford (USA), Zhen Cao (China), Eugenio Coccia (Italy), Don Geesaman (USA), Kunio Inoue (Japan), Naba Mondal (India), Angela Olinto (USA), Natalie Roe (USA), Sheila Rowan (United Kingdom), Valery Rubakov (Russia), Bernard Sadoulet (USA), Subir Sarkar (United Kingdom, Denmark), Christian Spiering (Germany), Michel Spiro (France) Chair, Yoichiro Suzuki (Japan).
Karl-Heinz Kampert (Germany), C4 IUPAP chair is ex-officio member, Ani Aprahamim (USA), C12 IUPAP chair is observer.
http://iupap.org/working-groups/wg10-astroparticle-physics-international-committee-appic/