



# MEETING

**Meeting Date** 25 April 2023

Ref MoMAAUG#10

ESA-ESDC-GEN-MOM-0032

**Meeting Place** WebEx

Chairperson Natalie Webb

**Minute's Date** 30 May 2023

## Participants

**AAUG Members:** Natalie Webb (AAUG Chair), Hervé Bouy, Annalisa De Cia, Ioannis Georgantopoulos, Søren Larsen, Christopher Conselice, Sandra Savaglio

**In attendance:** Guido de Marchi (ESDC Science Lead), Bruno Merín (Head of ESDC), Norbert. Schartel (XMM-Newton Project Scientist and Project Scientist's representative), Deborah Baines (ESASky Support Scientist), Rachana Bhatawdekar (ESDC Technical Lead), Bruno Altieri (Euclid Archive Scientist), Hector Canovas (Gaia Archive Support Scientist), Jos de Bruijne (ESDC Survey Science Lead & Gaia deputy Project Scientist), Javier Espinosa (ESDC), Monica Fernandez (ESDC Observatory Archives Technical Lead), Elena Jimenez Bailon (XMM-Newton Support Archive Scientist), Alvaro Labiano (JWST Support Archive Scientist), Ignacio Leon (ESDC), Sara Nieto (ESDC Survey Archives Technical Lead), Hector Perez (ESDC SCO-o8 Work Area Manager), Elena Puga (WinterWay), Maria Santos-Lleo (XMM-Newton Science Support Manager), Eva Verdugo (PLATO Archive Scientist).

## Absent:

**Subject** Minutes of the Astronomy Archives User Group (AAUG) meeting #10.

## Copy

Description	Action	Status
AAUG members to investigate one EU project each, investigate the type of data products from ESA missions that the project will be creating, advise ESDC what to ask the project PI, and if there is enough information at this point, recommend if the data should be added to any of the ESA science archives.	2019-02-13/13	IN PROGRESS
ESDC to begin contacting the PIs of the identified EU projects to gauge their interest in ESA preserving their project data.	2020-10-02/06	IN PROGRESS



Description	Action	Status
ESDC to also contact the PIs for the five recently identified EU H2020 and FP7 projects (action linked to Action 2020-10-02/06).	2021-04-14/03	IN PROGRESS
ESDC to find an 8 <sup>th</sup> member of the AAUG and ask for suggestions from stakeholders.	2023-04-25/01	
ESDC to contact ESA management regarding a question asking for gender in the ESA Astronomy Archives User Survey 2023.	2023-04-25/02	Q2 2023
D. Baines to update the 2023 archives user survey with the latest suggestions, circulate it to the AAUG and ESDC stakeholders for final validation and then publish and announce the survey.	2023-04-25/03	Before 10 <sup>th</sup> July 2023



## WELCOME:

N. Webb and G. de Marchi opened the meeting by welcoming everyone and summarising the main topics to be discussed during the meeting. The meeting was held via telecon during the afternoon of Tuesday 25<sup>th</sup> April 2023.

## ADOPTION OF THE AGENDA:

The agenda of the meeting was presented and adopted by the AAUG members.

## REVIEW OF ACTIONS:

The actions from the previous meeting were reviewed. Three actions remain open, all related to the EU H2020 and FP7 projects. The actions are the following: **Action 2019-02-13/13**: AAUG members to investigate one EU project each, investigate the type of data products from ESA missions that the project will be creating, advise ESDC what to ask the project PI, and if there is enough information at this point, recommend if the data should be added to any of the ESA science archives. **Action 2020-10-02/06**: ESDC to begin contacting the PIs of the identified EU projects to gauge their interest in ESA preserving their project data. **Action 2021-04-14/03**: ESDC to also contact the PIs for the five recently identified EU H2020 and FP7 projects (action linked to Action 2020-10-02/06). A new external contract covering the data curation and repatriation of ESA data started in July 2022. Work is progressing well and will be reported in a dedicated session in the next AAUG meeting.

Regarding closed actions, G. de Marchi reported on **Action 2022-09-13/03**: G. de Marchi to contact CDS and ESO about the possibility of holding a special session with ESO and CDS at EAS 2023. He reported that he did contact CDS and ESO and the idea was received warmly however, it was intercepted by another request from the Euro-VO collaboration who have a full day session planned in which CDS and ESO will be strongly participating (and CDS are helping to organise). There will be a contribution from ESA during the Euro-VO session. Therefore, ESA have instead planned a lunch session again (see section 3), which won't clash with the Euro-VO session. Both ESO and ESA have decided this year to have lunch sessions. N. Webb asked if ESA had thought about having a slightly longer session, maybe half a day. G. de Marchi said ESA investigated this, the problem was the EAS organisers said there was already one full day devoted to the Euro-VO, and suggested merging with them. Therefore, ESA will have the lunch session this year, but could have a longer session next year.

Regarding closed actions, N. Webb sent an email after the Autumn AAUG meeting to poll all members for their continued membership. Eva Villaver recently communicated that she would step down due to other commitments. The ESDC and AAUG members thank her warmly for her contributions! An action was raised on the ESDC to find another member for the AAUG, possibly someone with similar expertise to Eva (exoplanets, star formation research in the UV, optical and near-IR; **Action 2023-04-25/01**). G. de Marchi suggested next week's presentation by N. Webb to the AWG could be an opportunity to ask the AWG members for suggestions.



## PRESENTATIONS:

The following presentations and discussion sessions were held on 25<sup>th</sup> April 2023:

1. Long-term preservation of archives (H. Perez)
2. Archival Research Visitor Programme status (G. de Marchi)
3. Lunch session at the EAS Annual Meeting (G. de Marchi)
4. Astronomy Archives User Survey 2023 (D. Baines)
5. Date and time of next meeting
6. AOB
7. Closed session

The slides of the presentations are available on the ESAC Science Data Centre (ESDC) public web site, under the heading Archives User Groups, Astronomy: <https://www.cosmos.esa.int/web/esdc/aaug/meeting10>

## DISCUSSIONS:

Discussions took place during and after each presentation. This section presents a summary of the discussions and is ordered by the scheduled agenda given above.

### 1. Long-term preservation of archives

H. Perez gave a presentation on the long-term preservation of the archives at ESAC. In summary, the situation at the ESDC is that there are normally 4 to 6 archives being developed for missions in development. For missions in operations, there are around 10 to 12 archives maintained and developed by the ESDC. For missions in legacy, these numbers are currently similar to archives in operations, but the number is continuously increasing, and in the future, there will be far more archives from missions in legacy compared to missions in development and operations. Currently, there is an independent archive per mission (except for the multi-mission Planetary Science Archive). When the missions move to legacy, each independent archive infrastructure needs to be maintained and updated to maintain homogenisation across the archives, technical improvements, and interoperability. All of this requires funding.

The solution proposed by the ESDC is a common infrastructure for the legacy mission archives, with the goal of providing a common access point for all ESDC legacy data. This will cover data from all areas: astronomy, planetary, heliophysics, and even from HRE (Human and Robotic Exploration). It will include common hardware and data storage infrastructure, a common layer of standards (IVOA, HAPI, etc.), and multiple user interfaces for different needs. The scientific advantages are foreseen to be unified access to multi-mission and multidisciplinary data and enabling cross matches among the mission's data. The technical advantages are foreseen to be that: new legacy archives can be added with little effort; maintenance costs will be reduced compared to maintaining individual archives; the accessibility of the data will be improved; and it should be



flexible enough that different types of clients can access it (e.g. mission-specific, multi-mission, ESASky , etc).

The ESDC has begun the initial work for the ESDC common legacy infrastructure by setting up the initial servers and working on the backend. The first legacy missions' data to add to the legacy infrastructure will be from Ulysses and EXOSAT, with a first version expected by the end of 2023. Meanwhile, work is ongoing to repatriate ESA legacy data not held by the ESDC and to make the datasets available via a new TAP, which can be accessed by ESASky and other TAP clients. Future work will then involve integrating the repatriated datasets into the common legacy infrastructure and providing one common ESDC TAP.

After the presentation a discussion followed regarding these plans. E. Verdugo commented that for some missions, particularly Herschel, the way the data is stored can depend on the data model. Therefore, when moving the data, the ESDC will need to understand each archive and will need the expertise from the missions, archive scientists etc., whoever is still around. How does the ESDC plan to manage this? H. Perez confirmed that every single mission has a very specific data model and the ESDC can't move the data on their own. For every legacy archive that is moved, the ESDC will need scientific support (e.g. for EXOSAT, the ESDC are already talking with P. Kretschmar). E. Verdugo also asked if the ESDC plan in the future to remove the current legacy archive interfaces, or to keep the ones that are working well. H. Perez replied that the ESDC need to iterate with the scientists on whether the existing user interfaces provide an added value with respect to what will be in the new infrastructure, whether the features that are an added value can be implemented in the common infrastructure, or if the interfaces indeed need to be kept. This will have to be reviewed archive by archive.

H. Bouy commented that the cross matching between archives sounds like a very useful feature. He asked if the ESDC plan to coordinate this effort with any external agencies or observatories (such as STScI, ESO etc.) to at least use the same standards or same interface. H. Perez said it is in the scope of things to do but at a later phase in the process. B. Merin commented that the IVOA standards are used in the current archives and are proposed to future missions, helping to simplify the data modelling. H. Bouy asked if ESDC can be involved earlier in the mission development, so the work of the ESDC is more efficient. H. Perez confirmed the ESDC are involved at the development phase of the missions and do give advice on standards, formats etc., on what could be best for the long-term preservation of the data.

N. Schartel gave two comments, firstly that he was surprised there is not more technical commonality amongst all the ESDC archives. Secondly, that he did not think too many resources should be spent on preserving legacy data and that it's more important for young scientists to be using data from the latest missions. Various members of the AAUG noted how important legacy data can be for variability studies. B. Merin commented that it is part of the ESDC mandate to keep all legacy data available, regardless of the number of users accessing it. This new proposed infrastructure is designed to keep the resources as low as possible to preserve the legacy mission data. E. Puga pointed out that the type of data that is being repatriated are in general the high-level data products.

A discussion followed regarding the funding for the legacy infrastructure. N. Schartel argued to spend as little resources as possible for the preservation of legacy data and was concerned this would take resources away



from the archives in operations and development. R. Bhatawdekar added that the ESDC are creating a global technical road map, identifying the synergies between different projects and different missions, so that the archives development will be more effective technically, which will lead to a saving in resources. B. Merin confirmed that the ESDC receive funding from the ESA heritage space program specifically to work on the legacy mission archives and data. AAUG members emphasised the importance of dedicated funding for the legacy archives (see **Recommendation 2023-04-25/01**).

Finally, N. Webb asked how flexible the common infrastructure will be for it to integrate very different types of data. H. Perez replied that the existing archive data models can be put inside the infrastructure, and it should be flexible enough to integrate all types of data. G. de Marchi commented that this is an interesting topic to raise at the EAS during the discussion session of the ESA Archives lunch session.

## 2. Archival Research Visitor Programme

G. de Marchi presented a summary on the Archival Research Visitor Programme. The latest call is still open (deadline on 30<sup>th</sup> April 2023), therefore statistics for the call in Autumn 2022 were shown. This was the programme's fourth call. 19 applications were received, 15 of which were for astronomy archival research and 4 for planetary and heliophysics archival research. Most applications were to work on data from the large observatories/surveys and were evenly distributed between countries and nationalities. 6 visitors were accepted in this round and are listed in the programme's web page: <https://www.cosmos.esa.int/web/esdc/visitor-programme>.

In this call, like the previous call, the gender distribution of applicants was biased towards men. The numbers are small, so it is unknown if this is just small number statistics or if it is indeed a trend. Applications are now anonymised to limit unconscious biases, and this is advertised in the call. The call is advertised to mailing lists such as 'Women in Planetary Science', but G. de Marchi notes there must be more ways to reach out to women in the community to advertise the call. G. de Marchi pointed out, and AAUG members also noted, that women are underrepresented in the list of experts to contact at ESA in the visitor programme web page. The AAUG proposed to increase the representation of women and minorities in the table (see **Recommendation 2022-04-25/02**).

I. Georgantopoulos asked if the gender distribution of applicants is similar to the distribution seen in proposals asking for observing time. N. ScharTEL commented that the gender distribution of applicants applying for observing time is far more even from younger scientists than the distribution from older scientists. He suggested, if we assume most people applying for the Archival Research Visitor Programme are young, then there could be an imbalance in applications for the programme. N. ScharTEL also suggested to change the list of contact people for the programme to female colleagues working outside of ESA, possibly within the AAUG, rather than the current list of colleagues working in ESA. S. Larsen commented if AAUG members become the contact people then perhaps they should not be part of the evaluation team and A. De Cia said the knowledge of who can support certain projects at ESA is only really known by ESA colleagues. G. de Marchi agreed it's interesting to consider this idea and may ask the other user groups for comments.

Regarding the gender distribution, AAUG members commented that there are two separate issues to address: firstly, the small number of women applying to the programme, and secondly, any biases during the evaluation process. A. De Cia gave two suggestions for the evaluation process: bias training could be given to the evaluators; and proposals could be passed through AI tools to even out the language and tone of the applications. S. Larsen commented that the larger problem currently appears to be the lack of female, and indeed any underrepresented groups, applying to the research programme. G. de Marchi agreed that the programme needs to increase the diversity in the pool of applicants.

The outcome of the latest call is planned to be announced on 1<sup>st</sup> July 2023, and the AAUG members agreed it is best to publicise the outcome as soon as possible and not to delay it until the EAS meeting (which starts on 10<sup>th</sup> July 2023). However, it was agreed that also announcing the outcome at the EAS is a nice way to advertise the programme. Further recommendations from the AAUG for the Archival Research Visitor Programme are given in ***Recommendation 2022-04-25/02***.

### **3. Lunch session at the EAS Annual Meeting**

G. de Marchi presented the plans for the ESA Space Science Archives lunch session to be held at the 2023 EAS annual meeting in Krakow on Wednesday 12<sup>th</sup> July. The structure this year will be similar to last year and will be 1.5 hours long. The EAS 2022 ESA Astronomy Archives lunch session page can be found here: [https://eas.unige.ch/EAS\\_meeting/session.jsp?id=LS5](https://eas.unige.ch/EAS_meeting/session.jsp?id=LS5). G. de Marchi proposed that the demo this year will only be from ESASky, allowing more time for the ESASky hands-on part and for the overall session discussion at the end. Any participants interested in Datalabs (which isn't fully open to the community yet), can ask the ESDC members for a demo at the ESA booth. Topics for the discussion will need to be prepared beforehand. G. de Marchi asked if a tool such as 'slido' would be useful during the discussion session and if some of the questions from the upcoming user survey are asked to the participants. He also said two to three topics for discussion should be prepared beforehand (in case questions are few) and suggests: legacy data preservation; and a poll about data and software becoming immediately public (e.g. no proprietary period, as has been proposed by NASA/ERC). And of course, there will be free questions from the audience.

### **4. Astronomy Archives User Survey**

During the last AAUG meeting, the AAUG recommended that another ESA Astronomy Archives User Survey is prepared and sent out to the community (given that the last survey was in 2019). D. Baines went through the latest test version of the survey. The following changes were suggested:

- In the Demographics question for the type of institution people work in, G. de Marchi suggested adding examples for the Research Institute (e.g. IRAP, and others). S. Savaglio suggested changing 'Research organisation' to 'International organisation (e.g. ESO, NASA, ESA)' and removing 'Space agency'.
- N. Webb asked if gender should be asked for in the Demographics questions. G. de Marchi suggests

ESA management are asked if a gender question can be added. The question can have the answers: 'Male', 'Female', 'Other', 'Prefer not to say'. A. De Cia said it is important to emphasise that if an organisation is committed to promote diversity, then studying the demographics is an important part of this commitment (**Action 2023-04-25/02**).

- D. Baines suggested adding some links to question 8 on machine interfaces and giving examples (e.g. to astroquery modules, TAPs etc).
- N. Webb suggested for questions 9 and 20 (satisfaction with the archives user support), that a 'not applicable' button should be added, for users that may never have used the helpdesk nor looked at the documentation.
- N. Webb suggested to include ESASky to the first page description, so that it is clear to participants that the survey is about the archives and about ESASky.
- S. Larsen suggested for question 12 'Would you be willing to register as an archive user and then have special privileges?', that it would be worth probing explicitly why a no answer has been chosen. It was agreed to add a yes / no check box. If no is selected another question appears asking the reasons why, and if yes is selected participants move on to the next section of questions.
- H. Bouy suggested adding a link in the text to ESASky in questions 14 and 15 (what users like the most and the least about the archives and ESASky).
- The AAUG were happy to see the progress made with the survey and suggest that the new legacy archives are mentioned in the survey to help publicise them (see end of **Recommendation 2023-04-25/01**).

It was agreed to make the survey public before the EAS meeting (before 10<sup>th</sup> July 2023). D. Baines took an action to update the survey with the above suggestions and circulate it to the AAUG and ESDC stakeholders for final validation before publishing and announcement (**Action 2023-04-25/03**).

## 5. Date and time of next meeting

A joint archives user group meeting is planned for 19 to 21 February 2024 at ESAC, Madrid. The agenda includes all three Archive User Groups (Astronomy, Planetary and Heliophysics) meeting for one whole day, and then splintering off for the individual User Group meetings on the other days.

## 6. AOB

No AOBs were raised.



## RECOMMENDATIONS

The following recommendations were formulated by the AAUG:

### *Legacy archives at ESA*

**Recommendation 2023-04-25/01:** The AAUG congratulates ESA on the development of the legacy archives and the inclusion of ESA satellite data that has been difficult to retrieve through ESA interfaces previously. This data is extremely useful to the scientific community, especially given the recent emergence of time domain astronomy, which often requires a long baseline to understand the variability of the source or the new developments in artificial intelligence, which often needs large volumes of data, but can also enhance a data set. This is also a good time to develop a legacy archive as there will be a number of missions that will have a definitive dataset to archive in the upcoming years, so the common legacy infrastructure will be well adapted to ingest the data and minimise the investment required. However, investment in the legacy archives must not be done at the cost of the investment in the archives of operational missions, as these require large investment to keep them up to date with current data and ever improving products. The AAUG recommends that a priority is given to the archives for operational missions and that ESA investigate the possibility of dedicated funding/support for the legacy archives. Finally, the AAUG was happy to see the progress made in putting together the next ESA Astronomy Archives User survey and suggest that the new legacy archives are mentioned in the survey to help publicise them.

### *The Archival Research Visitor Programme*

**Recommendation 2022-04-25/02:** The Archival Research Visitor Programme continues to go from strength to strength with a strong interest from the community, which is helping to raise the profile of the ESA archives and give maximum value to the data in them. The AAUG recommend following up the visits by asking for a short report written by the visitor, on the work accomplished thanks to the visit, possibly a year afterwards, to understand the impact of the visit. The AAUG would like to see the scientific output of the first laureates of the programme.

As noted previously, whilst participation of women and minorities was strong (compared to the average in astronomy) in earlier calls, this participation appears to have dwindled in recent calls. This may be due to small number statistics, but the AAUG recommend the continued monitoring of participation in the calls from women and minorities. The AAUG appreciate the efforts made to anonymise proposals as a measure to try to reduce biases in the evaluation process and recommend that this continues in future rounds. The AAUG also encourages ESA to use the sentence ‘We encourage applications from women and minorities.’, on the call for proposals, and the AAUG recommends adding ‘The evaluation process will be anonymised to ensure equal opportunities for all applicants’ and to follow with something like ‘ESA is actively committed to the diversity and inclusivity, and has recently obtained the EDGE (Economic Dividends for Gender Equality) Certification for gender equality’. This would demonstrate to the public that ESA are indeed taking action. Other action that could be taken by ESA is to increase the representation of women and minorities in the



table that lists the type of expertise that ESA mission scientists and archive scientists can offer, and given on the call for proposals. We understand that this is a long term effort, but it should be beneficial both to encouraging a more balanced pool of researchers to apply to the Archival Research Visitors Programme and to strengthening work at ESA.