

EOSC-A PID Policy and Implementation Task Force draft report:

# Community-specific (and stakeholder category-specific) perspectives on the EOSC PID architecture and the EOSC PID policy

1 September 2022

## Authors/Editors:

The following Task Force members comprised the editorial group responsible for producing this document: Matthew Buys, Britta Dreyer, Margareta Hellström, Tibor Kálmán, Gabriela Mejias, Tomasz Parkoła, Michal Růžička, Tommi Suominen.

## Members of the PID Policy and Implementation Task Force:

Gaelle Bequet (ISSN International Centre), Pablo de Castro (euroCRIS), Tom Demeranville (ORCID), Britta Dreyer (TIB), Nathalie Fargier (Center for Direct Scientific Communication), Beate Guba (TU Wien), Maggie Hellström (Lund University/ICOS/ENVRI-FAIR), Florian Krebs (German Aerospace Center), Paolo Lai (CNRS Inist), Gabriela Mejias (DataCite), Jozef Mišutka (LINDAT/CLARIAH-CZ), Tomasz Parkoła (PSNC), Claudio PISA (GARR), Michal Růžička (Masaryk University/e-INFRA CZ), Jochen Schirrwagen (OpenAIRE), Tommi Suominen (CSC/FAIRCORE4EOSC), Wojtek Sylwestrzak (University of Warsaw), Mario Valle (CSCS), Adam Vials Moore (Jisc). EC Observers: Christian Cuciniello (European Commission), Stephan Siemen (European Commission).

Chairs: Tibor Kalman (GWDG), Themis Zamani (GRNET).

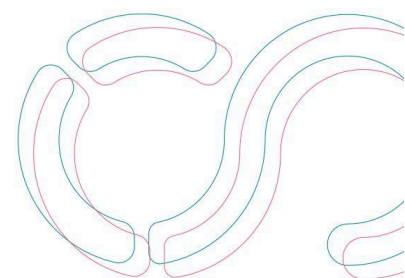
Task Force Support Officer: Brina Klemenčič (EOSC-A).

## Abstract

This document is the first comprehensive overview of the focus group established to “collect community-specific (and stakeholder category specific) perspectives on the EOSC PID architecture and the EOSC PID policy, in particular, priorities for addressing gaps, as well as use cases”.

## EOSC Association AISBL

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu  
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



### *Status of this document*

This document is an Internal Report that was used to be discussed with the community. This version was reviewed by the “Quality Review Committee (QRC)” of the EOSC Association, and it is being published as an official outcome of the “EOSC-A PID Policy and Implementation Task Force”.

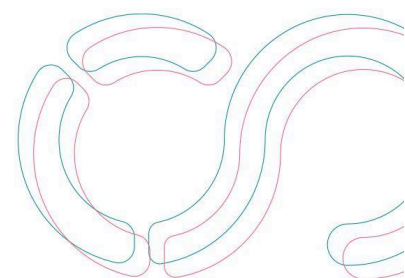
### *Current and previous versions:*

Current version: This version is the latest version of the document.

Previous version: Previous drafts are not available publicly.

### **EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
+32 2 537 73 18 | [info@eosc.eu](mailto:info@eosc.eu) | [www.eosc.eu](http://www.eosc.eu)  
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



## Report

EOSC-A PID Policy and Implementation Task Force draft report:  
Community-specific (and stakeholder category-specific) perspectives on  
the EOSC PID architecture and the EOSC PID policy

## Preamble

The EOSC-A PID Policy and Implementation Task Force established a focus group in early 2022 to “collect community-specific (and stakeholder category specific) perspectives on the EOSC PID architecture and the EOSC PID policy, in particular, priorities for addressing gaps, as well as use cases”.

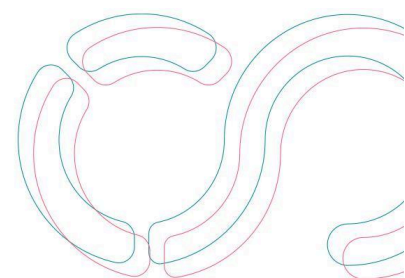
In order to gather feedback and input in a structured manner, the focus group collected survey responses. The survey questions are included in [Appendix 1](#).

This report consolidates and analyses the survey results for validation by the community.

## Process

The EOSC-A PID Policy and Implementation Task Force has four key aims in their [charter](#), namely;

- *Provide input to the EOSC Board of Directors starting from the gaps identified in the PID ecosystem mentioned in the SRIA.*
- *Ensure EOSC objectives are attained by working with [EOSC related] projects to develop standardised identifiers for resource types that have not as yet become standard practice, develop a ‘meta resolver’ that can deal with any type of relevant identifier, define specifications (schemata) for PID records/kernel information to support machine-actionable PIDs, produce type definitions for the most common data formats or building blocks, provide standardised interfaces and protocols for exchanging information on PIDs to support the creation and use of a PID graph and develop tools to support the certification of PID infrastructure against the EOSC PID policy.*
- *Implement and refine the EOSC PID policy and architecture by aligning with the principles of Open Scholarly Infra and by building consensus within the community.*
- *Report progress to the EOSC Association Board of Directors.*



In focusing on the task force activities, a focus group was established to gather feedback. The agreed process was to be iterative:

1. Develop a questionnaire to gather feedback
2. Distribute the questionnaire via available channels
3. Draft a report summarising feedback from the community
4. Disseminate the draft report for comments
5. Publish a final report with recommendations and/or conclusions

The questionnaire ([Appendix 1](#)) was distributed via various channels available to task force members. The survey was open for approximately 3 months (May through June 2022). In addition to the questionnaire, the focus group consulted the FAIRCORE4EOSC Work Package 2 members and the FAIR IMPACT WP3 members for additional insight in developing this draft report. The FAIRCORE4EOSC Work Package 2 members are working on an EOSC PID compliance framework, and the FAIR IMPACT WP3 members are working on an EOSC coordination mechanism for PID service providers.

## Results

The summary survey results were analysed by task members, and additional graphs can be seen in [Appendix 2](#). It should be noted that whilst all care was taken in the analysis, the task force members welcome any feedback from the community.

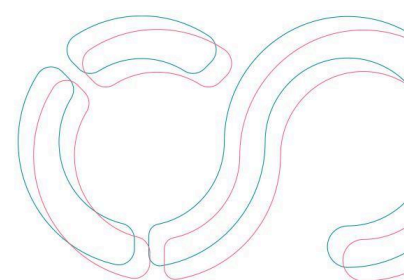
### About the respondents

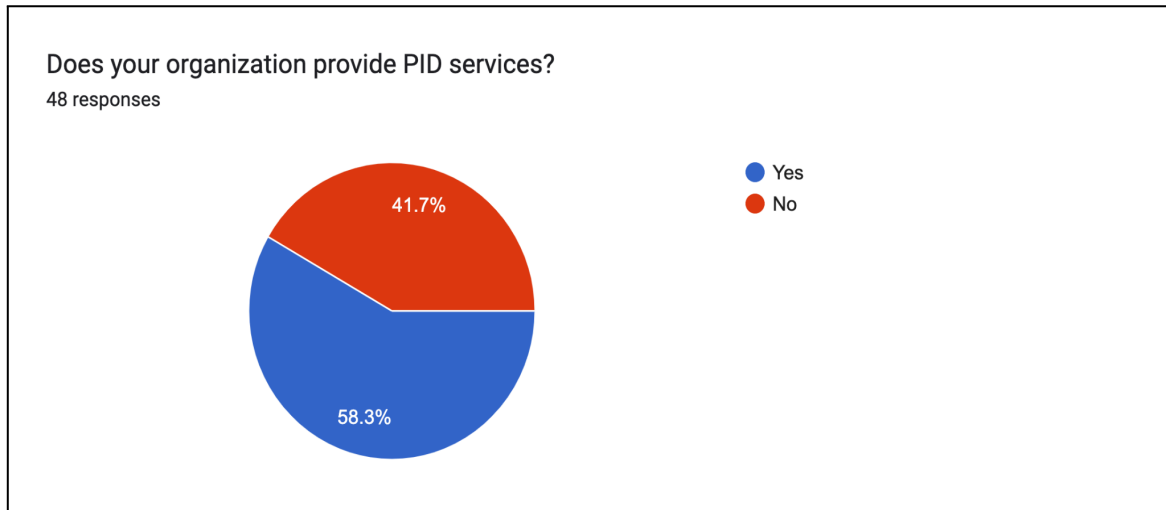
The survey had 49 respondents from 11 domains/disciplines, as defined in the survey.

**Table 1.** The distribution of disciplines who responded (*note that respondents were able to identify with more than one discipline*)

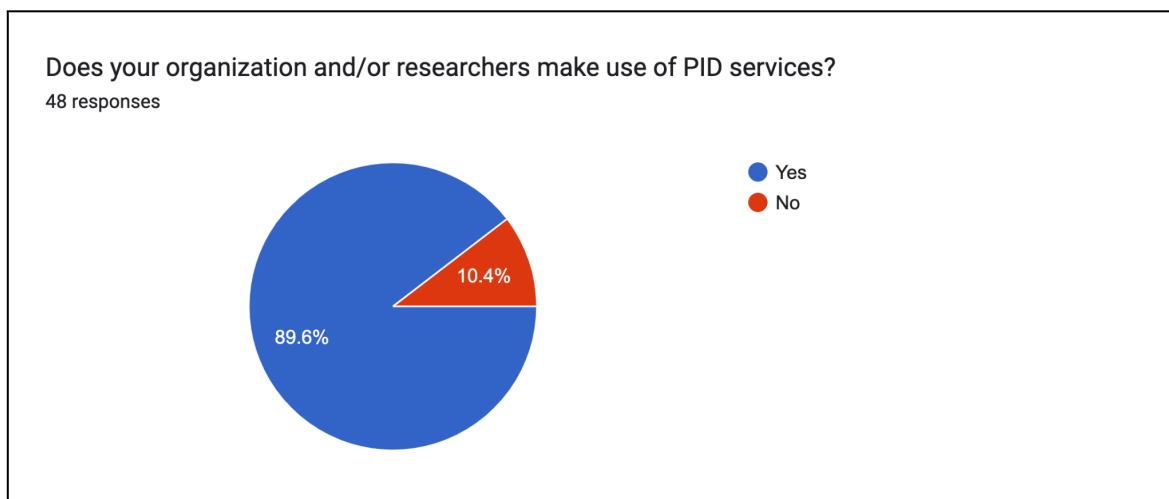
Natural Sciences	23
Engineering & Technology	19
Medical & Health Sciences	7
Agricultural Sciences	6
Social Sciences	9
Humanities	14
Other	7

58.3% (28) of the respondents said their organisation provides persistent identifier (PID) services; whereas 89.6% (43) of the respondents make use of PID services.





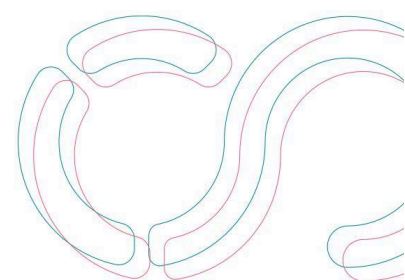
**Fig 1.** Percentage of respondents that provide persistent identifier services.

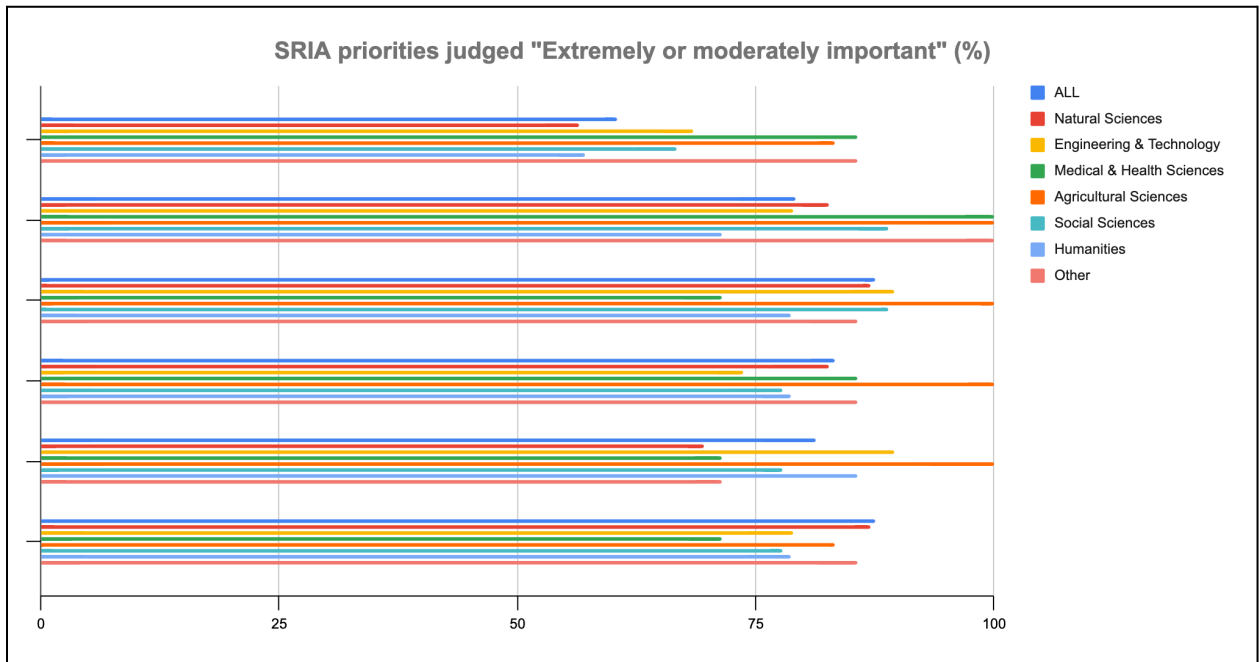


**Fig 2.** Percentage of respondents that make use of persistent identifier services.

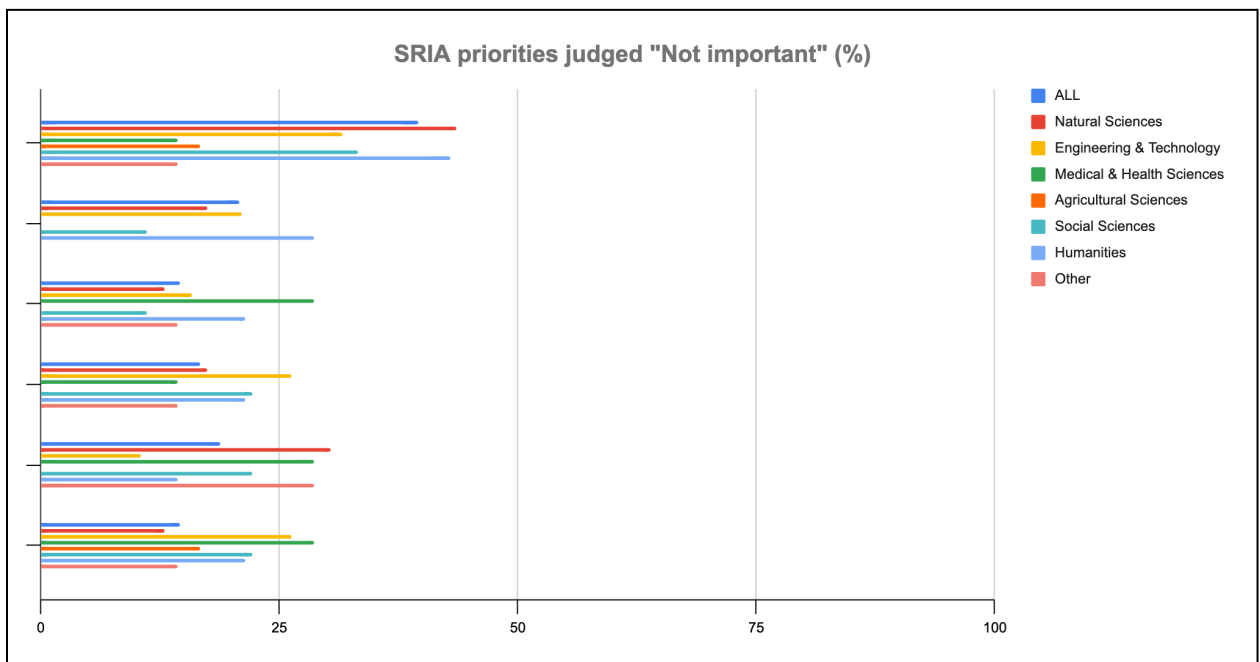
The survey responses cannot be seen as a complete representation of the EOSC community, although it is sufficient to provide insight into perspectives on the EOSC PID architecture and the EOSC PID policy. The survey responses have been consolidated to provide this report for further input and validation by the community.

As seen in the figures below, there is a general consensus that the SRIA priorities are considered moderately or extremely important, although there is no particular trend by domain or across domains. It may be that community stakeholders are unsure of the impact, and due to the technically complex nature of some of the priorities, the benefits remain unclear to community stakeholders.





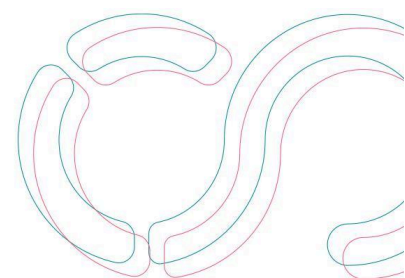
**Fig 3.** Percentage overview of respondents that indicated the SRIA priorities were considered extremely or moderately important by domain.



**Fig 4.** Percentage overview of respondents that indicated the SRIA priorities were considered not important by domain.

**EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
 +32 2 537 73 18 | info@eosc.eu | www.eosc.eu  
 Reg. number: 0755 723 931 | VAT number: BE0755 723 931



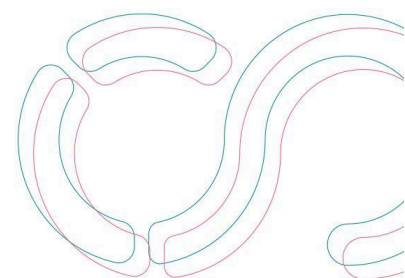
In reviewing the results of the questionnaire, it seems that there are mixed responses regarding the prioritisation of efforts and is somewhat inconclusive in nature. As such, the task force members gathered further insight from the open comments from respondents. In addition, the discussion with the FAIRCORE4EOSC and FAIR-IMPACT project members provided further insight as to their planned activities.

In reviewing the comments provided by respondents, there were several comments related to the emphasis on standardisation, supporting persistent identifiers for emerging resource types, minimal cost for services, and integration of widely used and adopted persistent identifier services.

Below are selected comments received from the respondents:

- (1)... standardised, preferably \*globally unique\*, regular expressions serving to identify a PID of a certain type.*
- (2) Promote the use of PIDs that are already common practice such as ROR, ORCID, DOI, Wikidata.*
- (3) Relations between PIDs need to be clarified and made visible in a graph.*
- (4) Produce type definitions for the most common data formats or building blocks: Definitions of types and their relations to metadata need to be clarified, as both data and metadata can be categorised in so many different ways.*
- (5) PIDs should be machine-actionable by using existing protocols as much as possible (e.g. HTTP Content-Type) This would be interoperable also with PIDs that do not support associated attributes.*
- (6)...the seamless integration of other, already existing, PID systems*
- (7) Affordable PID infrastructure and/or PID minting from PID providers.*

Further respondents alluded to the fact that perhaps the use of persistent identifiers is not yet common practice and that activities that support these efforts would be important. It is very common to use PIDs like DOI for publications, where it is a mature and established



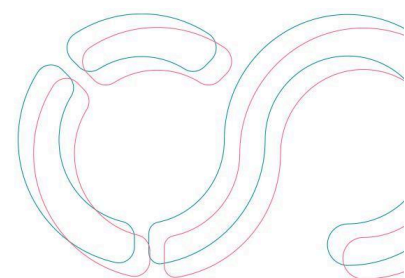
practice. On the other hand, PIDs for researchers like ORCID are still not integrated as common practices by all publication publishers, for example. There are still published documents with no PID of authors, editors etc. For 'more exotic' entities like datasets, scientific instruments, physical samples, institutions, projects, and consortia the maturity and the use of PIDs is even lower. As such, it is necessary to raise awareness of the possible PIDs and their use.

Respondents indicated that the EOSC-A should make sure that the broad range of persistent identifier services is interoperable and simple to use for the end user. Further, it is important to clarify the status of the various persistent identifier services, specifically regarding the adoption rate at scale across borders and domains.

## Conclusion

The task force has made the following conclusions in this draft report (to be shared for community feedback):

1. **In general, the SRIA priorities are considered important.** The survey results showed consensus on this.
2. **Need to gather further insight.** Due to the inconclusiveness of perspectives, it is critical that the community provides further insight to help understand the inconclusiveness drawn from the responses.
3. **Reach out to projects & activities regarding their work on priorities we clustered.** Specifically investigating interoperable services related to resolution and PID metadata creation, supporting emerging resource types.
4. **Development of metrics.** As the various EOSC-A task forces and projects develop services and make recommendations, it is critical within the context of persistent identifier services to develop metrics (beyond basic technology readiness level assessment). Specifically, persistent identifier services should be assessed based on adoption at scale (across borders and domains), ease of use, and interoperability across infrastructure services.
5. **Integration of widely used and adopted persistent identifier services.** The EOSC should leverage widely adopted open infrastructure that is used at scale across disciplines and domains globally.
6. **Interoperability.** There is a need to clearly define a pathway for interoperability within the EOSC framework and related services.
7. **Sustainability.** Critically, the (meta)data stored by the services should be community property and openly available. As such, it is important that the services have robust long-term sustainability plans and align with the EOSC plans for sustainability.



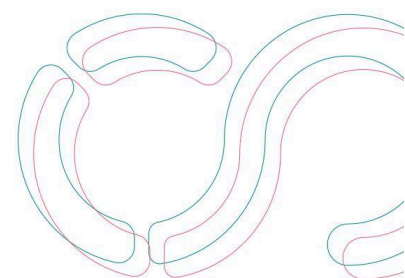


The EOSC-A seeks to serve a broad community of stakeholders. As such, **there is an increasing need to provide simple, easy-to-use technology within a complex technical infrastructure landscape.** PID service providers **should focus on reducing the barriers to entry, in particular, the technical barrier,** which is often overlooked.

Our task force will be coordinating closely with the FAIRCORE4EOSC and FAIR-IMPACT project members to track the impact of the services developed. Furthermore, the task force will take into consideration the report “Risks and Trust in Pursuit of a Well-functioning Persistent Identifier Infrastructure for Research” published later this year by the European Knowledge Exchange initiative. The report including PID recommendations and PID case studies will provide insight into the current PID landscape and its stakeholders (see <https://doi.org/10.5281/zenodo.5018216>).

## About the authors

The authors of this report contributed in the capacity of task force members, although it should be noted that authors are involved in various EOSC projects that are engaged in developing PID services for EOSC.



## Appendix 1: Survey questions

### EOSC-A PID Policy and Implementation Task Force survey on PID practices

As part of the TF efforts, we are seeking input and feedback regarding community-specific use cases. In this context, we are seeking to collect community-specific (and stakeholder category-specific) perspectives on the EOSC PID architecture and the EOSC PID policy, in particular, priorities for addressing gaps, as well as use cases.

The survey will be open until July 31, 2022.

Please complete the form below.

**\* Required**

#### 1. Which categories align with your domain/discipline? \*

*Check all that apply.*

- Natural Sciences
- Engineering & Technology
- Medical & Health Sciences
- Agricultural Sciences
- Social Sciences
- Humanities
- Other: \_\_\_\_\_

#### 2. Please briefly describe your community or you as a stakeholder? \*

---

#### 3. Does your organisation provide PID services? \*

*Mark only one oval.*

- Yes
- No

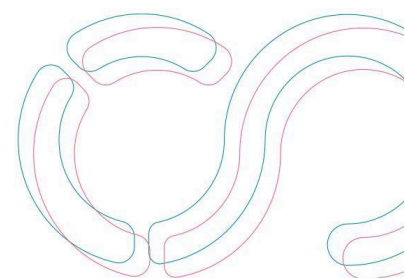
#### 4. Does your organisation and/or researchers make use of PID services? \*

*Mark only one oval.*

- Yes
- No

**EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
+32 2 537 73 18 | info@eosc.eu | www.eosc.eu  
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



**5. If possible, please could you provide your email address? By adding your contact information, you expressly opt-in to the below use of your contact information (exclusive use for this survey only).**

The contact details entered are for the exclusive use of this EOSC-A PID Policy and Implementation Task Force survey and are stored exclusively in the survey responses. Individuals may be contacted by EOSC-A PID Policy and Implementation Task Force members directly for additional information following the survey, although no individual's contact details will be transferred to any other database or mailing list whatsoever. This survey is administered by DataCite on behalf of the EOSC-A PID Policy and Implementation Task Force. For information about DataCite's general privacy practices and commitment to protecting your privacy, please review their Privacy Policy (<https://datacite.org/privacy.html>).

---

**6. Do you agree to the EOSC-A PID Policy and Implementation Task Force using your contact details as described above? \***

*Mark only one oval.*

- Yes
- No

### **EOSC PID Implementation**

Based on the Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC).

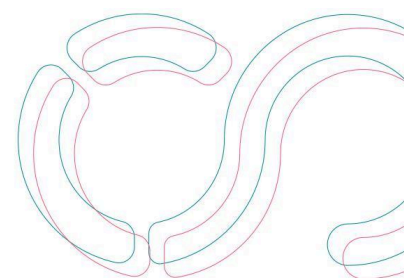
**7. How would you classify the following priorities outlined in the SRIA? \***

Based on the Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC). For detailed definitions of the priorities, please refer to the Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC), 2022, <https://data.europa.eu/doi/10.2777/935288>, p. 86.

*Check all that apply.*

### **EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
+32 2 537 73 18 | [info@eosc.eu](mailto:info@eosc.eu) | [www.eosc.eu](http://www.eosc.eu)  
Reg. number: 0755 723 931 | VAT number: BE0755 723 931



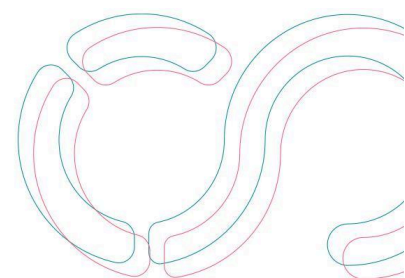
	Extremely important	Moderate importance	Less important	Not at all important
Develop standardised identifiers for resource types that have not as yet become standard practice. For general research use, EOSC would prioritise identifiers for instruments, services, organisations and software, although there is a need for particular domains to provide their own community standards.	0	0	0	0
Develop a 'meta resolver' that can deal with any type of relevant identifier.	0	0	0	0
Define specifications (schemata) for PID records/kernel information to support machine-actionable PIDs.	0	0	0	0
Produce type definitions for the most common data formats or building blocks.	0	0	0	0
Provide standardised interfaces and protocols for exchanging information on PIDs to support the creation and use of a PID graph.	0	0	0	0
Develop tools to support the certification of PID infrastructure against the EOSC PID policy.	0	0	0	0

**8. Are there additional priorities that you would like to share?**

---

**EOSC Association AISBL**

Rue du Luxembourg 3, BE-1000 Brussels, Belgium  
 +32 2 537 73 18 | info@eosc.eu | www.eosc.eu  
 Reg. number: 0755 723 931 | VAT number: BE0755 723 931



---

---

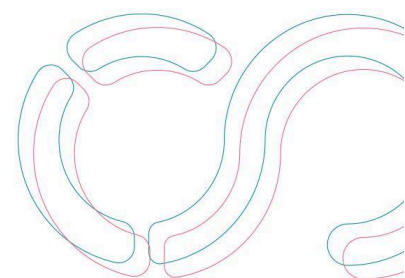
---

**9. Does your community have a specific PID use case/need/requirement that you would like to share?**

---

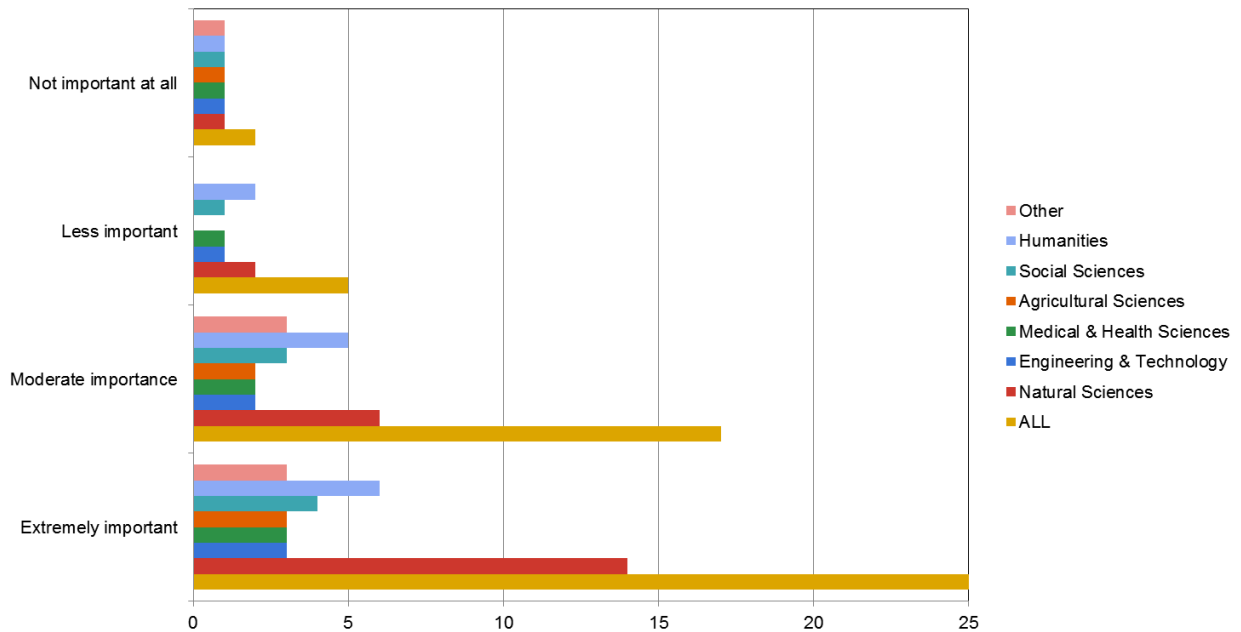
---

---

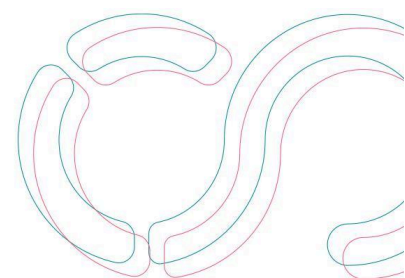


## Appendix 2: Additional graphs

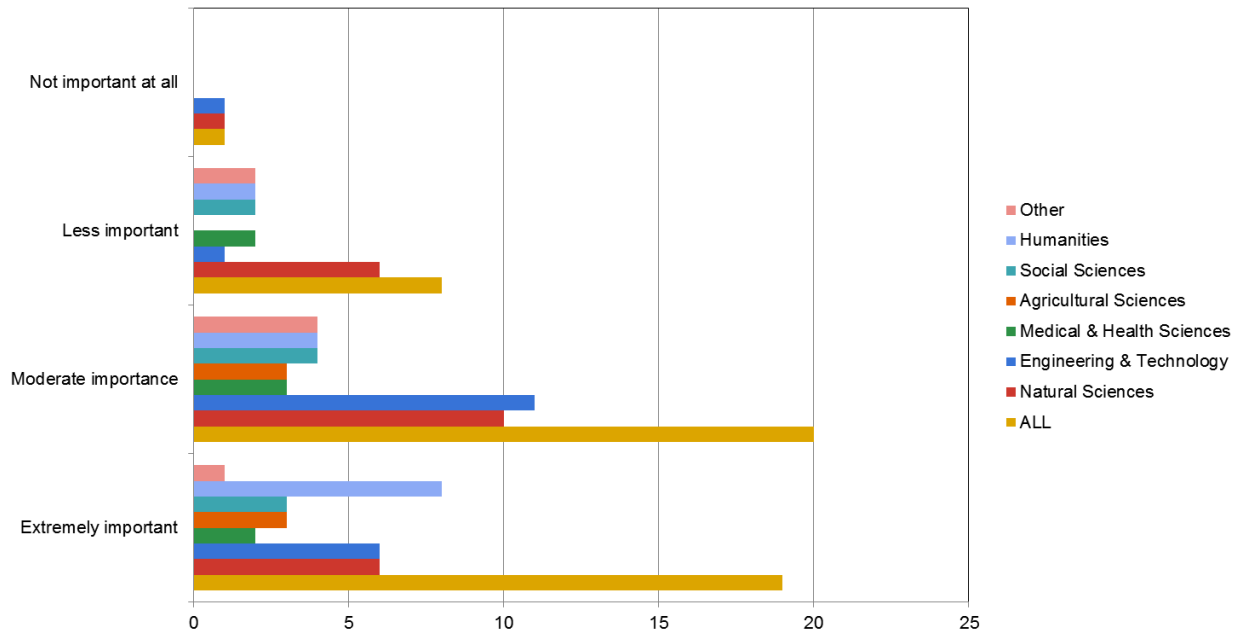
### SRIA P1: Develop standardised identifiers for resource types that have not as yet become standard practice.



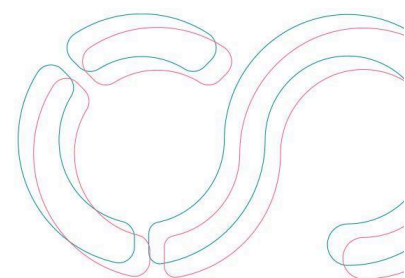
**Figure A2-1.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #1: Develop standardised identifiers for resource types that have not as yet become standard practice. The total number of responses, as well as responses broken down across disciplines, is indicated.



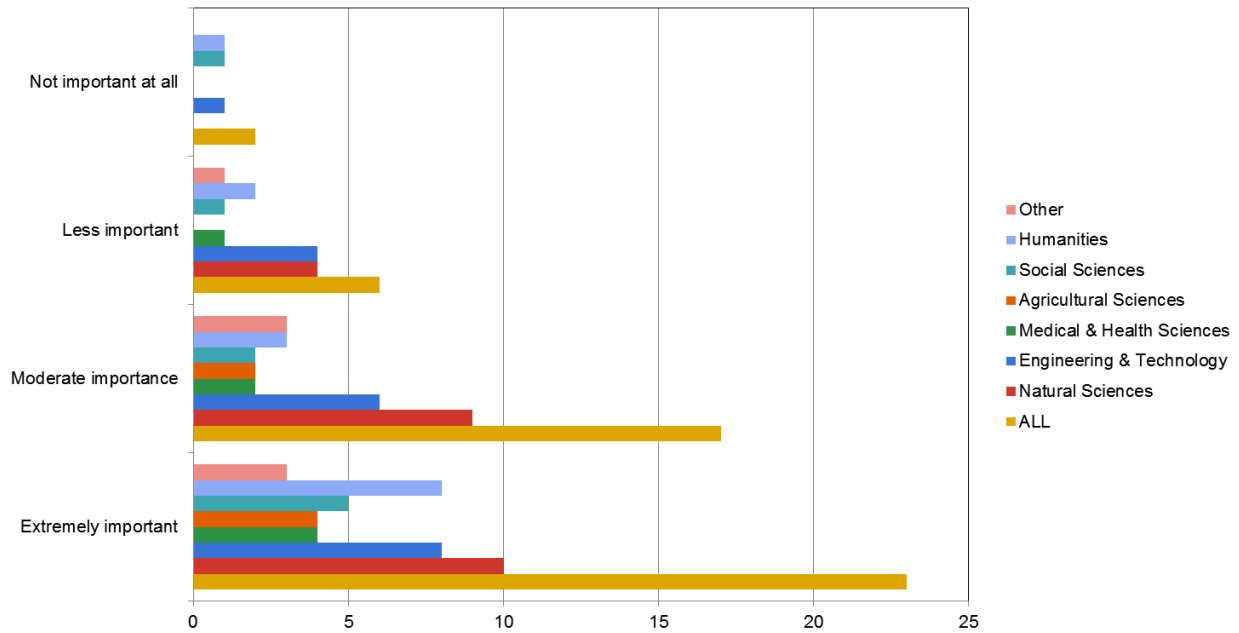
**SRIA P2: Develop a ‘meta resolver’ that can deal with any type of relevant identifier.**



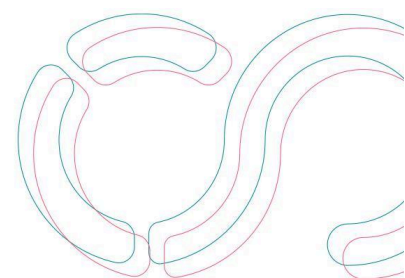
**Figure A2-2.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #2: Develop a ‘meta resolver’ that can deal with any type of relevant identifier. The total number of responses, as well as responses broken down across disciplines, is indicated.



**SRIA P3: Define specifications (schemata) for PID records / kernel information to support machine-actionable PIDs.**

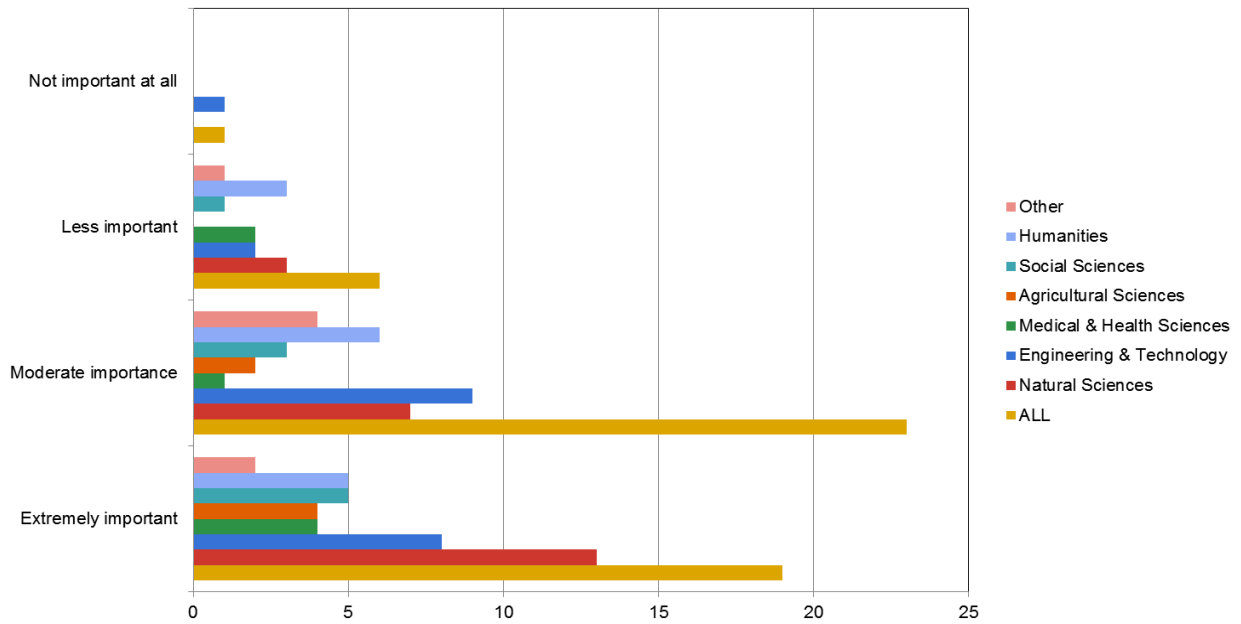


**Figure A2-3.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #3: Define specifications (schemata) for PID records / kernel information to support machine-actionable PIDs. The total number of responses, as well as responses broken down across disciplines, is indicated.

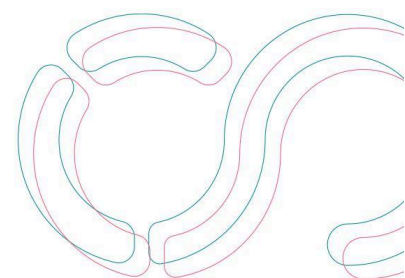




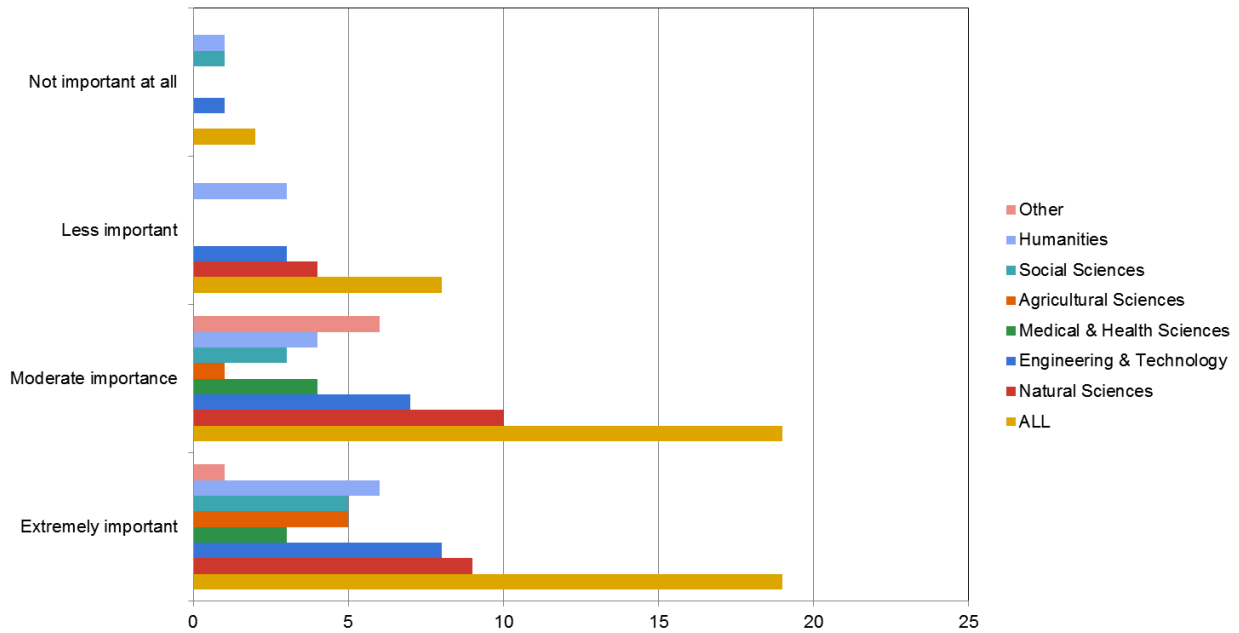
**SRIA P4: Produce type definitions for the most common data formats or building blocks.**



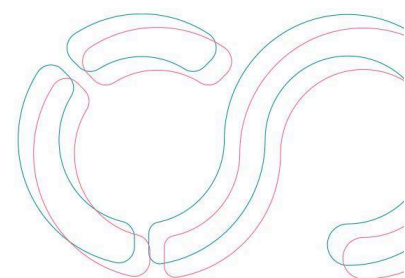
**Figure A2-4.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #4: Produce type definitions for the most common data formats or building blocks. The total number of responses, as well as responses broken down across disciplines, is indicated.



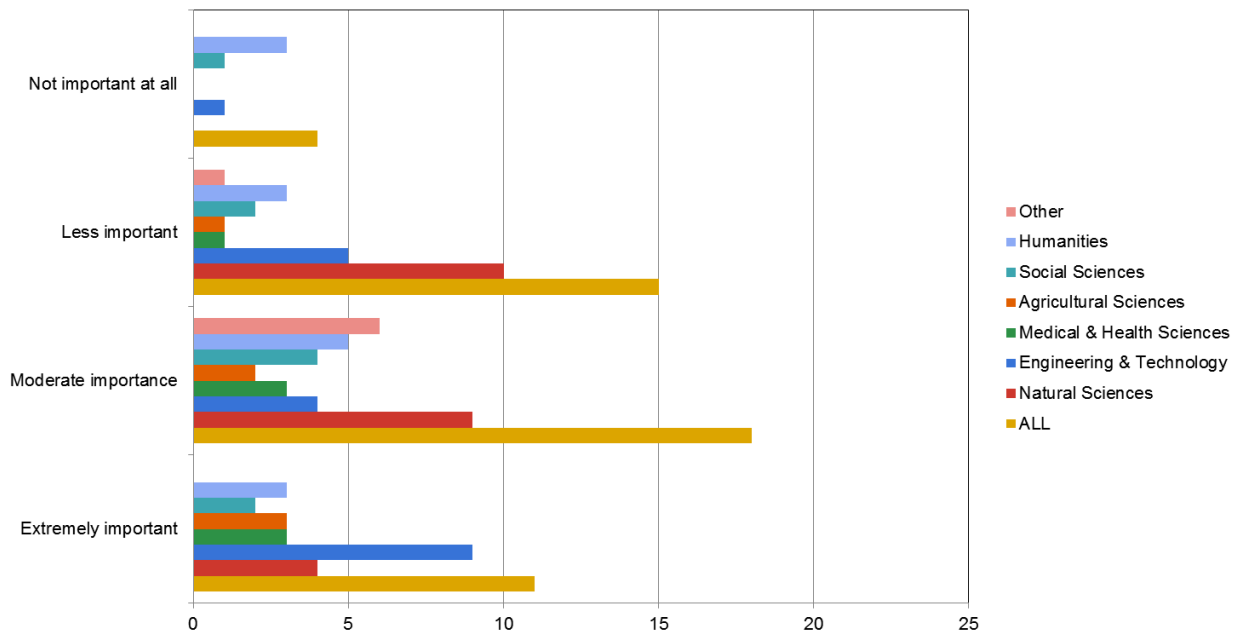
**SRIA P5: Provide standardised interfaces and protocols for exchanging information on PIDs to support PID graphs.**



**Figure A2-5.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #5: Provide standardised interfaces and protocols for exchanging information on PIDs to support the creation and use of a PID graph. The total number of responses, as well as responses broken down across disciplines, is indicated.



**SRIA P6: Develop tools to support the certification of PID infrastructure against the EOSC PID policy.**



**Figure A2-6.** Priority classification (“extremely important”, “moderate importance”, “less important” and “not important at all”) for SRIA priority #6: Develop tools to support the certification of PID infrastructure against the EOSC PID policy. The total number of responses, as well as responses broken down across disciplines, is indicated.

