



Project Name **FREYA**
Project Title **Connected Open Identifiers for Discovery, Access
and Use of Research Resources**
EC Grant Agreement No **777523**

D6.5 Final Report on PID Commons and Sustainability

Deliverable type Report
Dissemination level Public
Due date 30 November 2020
Authors Simon Lambert (STFC)
Josh Brown (consultant to British Library, for PID federation scoping study)
Abstract Mapping of the PID Commons through sustainability concerns and stakeholders including a PID Alliance for which an important study was undertaken in the final year of FREYA.
Status Submitted to EC 20 December 2020

The FREYA project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 777523.



FREYA project summary

The FREYA project iteratively extends a robust environment for Persistent Identifiers (PIDs) into a core component of European and global research e-infrastructures. The resulting FREYA services will cover a wide range of resources in the research and innovation landscape and enhance the links between them so that they can be exploited in many disciplines and research processes. This will provide an essential building block of the European Open Science Cloud (EOSC). Moreover, the FREYA project will establish an open, sustainable, and trusted framework for collaborative self-governance of PIDs and services built on them.

The vision of FREYA is built on three key ideas: the **PID Graph**, **PID Forum** and **PID Commons**. The PID Graph connects and integrates PID systems to create an information map of relationships across PIDs that provides a basis for new services. The PID Forum is a stakeholder community, whose members collectively oversee the development and deployment of new PID types; it will be strongly linked to the Research Data Alliance (RDA). The sustainability of the PID infrastructure resulting from FREYA beyond the lifetime of the project itself is the concern of the PID Commons, defining the roles, responsibilities and structures for good self-governance based on consensual decision-making.

The FREYA project builds on the success of the preceding THOR project and involves twelve partner organisations from across the globe, representing PID infrastructure providers and developers, users of PIDs in a wide range of research fields, and publishers.

For more information, visit www.project-freya.eu or email info@project-freya.eu.

Disclaimer

This document represents the views of the authors, and the European Commission is not responsible for any use that may be made of the information it contains.

Copyright Notice

Copyright © Members of the FREYA Consortium. This work is licensed under the Creative Commons CC-BY License: <https://creativecommons.org/licenses/by/4.0/>.

Executive summary

This deliverable on the PID Commons and sustainability reports on the important developments of this pillar of FREYA in the project's third year. A central aspect of the PID Commons and sustainability planning of FREYA was a study on the prospects for a PID federation (which the FREYA team prefers to call a PID Alliance), bringing together a wide range of stakeholders in the PID world for collective engagement and action. The idea of such a body came out of a FREYA project meeting in December 2019 and a broader meeting was held at the PIDapalooza conference in January 2020. As a result, the British Library (FREYA partner) engaged an in-house consultant to conduct a wide-ranging study. This deliverable presents that report, which concluded that there is a real appetite for some kind of PID Alliance, and that three primary focuses should be interoperability, inclusion and communication/advocacy. The deliverable goes on to present FREYA's response to the report, including recommendations.

The deliverable also considers the relationship between a PID Alliance and other aspects of sustainability, and produces a map of the PID Commons in which a variety of stakeholders are assigned particular concerns around sustainability—these stakeholders include the PID Alliance itself. The British Library is working with colleagues consulted in the study, to draft a plan for an open consultation beyond the end of the funded FREYA project to further understand the next steps for a PID Alliance as envisaged in this deliverable.

Contents

1	Introduction and developments in the third year of FREYA.....	5
2	The “PID federation” scoping study	8
3	The FREYA perspective: a PID Alliance	11
3.1	Overall view including scope and objectives.....	11
3.2	Priorities for action	12
3.3	Implementation of the PID Alliance	13
4	Sustainability and a PID Alliance.....	14
	Annex: Final report on PID Federation scoping study.....	18

1 Introduction and developments in the third year of FREYA

Work Package 6 of FREYA is devoted to the sustainability of the PID infrastructure, but it is not the only aspect of FREYA's work to be concerned with sustainability. The sustainability of the individual services and applications developed within FREYA, both the general foundations of the PID Graph and the pilot applications of it, are presented in the project's exploitation plan (not a public deliverable). Sustainability in the context of the European Open Science Cloud (EOSC) is an aspect of FREYA's work reported in deliverable D4.5 "Integration of the PID Graph with the EOSC", where relations with the future creation of EOSC, its policy and architecture are discussed. The concern of WP6 is with sustainability through the PID Commons, "defining the roles, responsibilities and structures for good self-governance based on consensual decision-making" as it was initially characterised.

The position at the end of the second year of FREYA was represented in deliverable D6.3, "Second Annual Report on PID Commons and Sustainability". During that year important work had been done on refining the understanding of the three facets of sustainability—namely maintainability, adaptability and desirability—in the context of the PID infrastructure specifically, and also on the "boundary conditions" of the PID Commons—that is, general attributes that any realisation of the PID Commons would need to satisfy in order to fulfil the expectations on it. The deliverable also reported on a backcasting exercise held at the project's plenary meeting in London in December 2019, which examined three complementary aspects of the PID Commons and generated a rich set of "raw data" for further analysis and planning.

Therefore at that stage the project was gearing up to embark on the crucial work of the third year, building on the analyses and surveys conducted earlier to produce a "clear and validated understanding of the nature of the Commons with roles of stakeholders." Mention was made of a dedicated meeting of partners, coinciding with the PIDapalooza conference in Lisbon in January 2020, to produce specific proposals for the structures and implementations of the PID Commons. What D6.3 did not make explicit was that this meeting was already planned to be a broader gathering of stakeholders going beyond the FREYA partners themselves to ascertain the general desire for a kind of "PID association", "foundation" or "federation" (as it was variously referred to at that time). The idea had been floating for some time in internal discussions of the PID Commons, as a way of giving form to the Commons through a definite body. Such a PID association, while it would not in itself *be* the Commons, would play a role in bringing elements of the Commons together and satisfy the boundary condition that "The PID Commons is not a single body or organisation (though there might be a central body that is the essence of the PID Commons), but it does have some existence and identity—it is not completely nebulous, and it will be possible to point to its structures and processes in action."

The backcasting exercise in London did in fact raise this possibility explicitly, and that was the motivation for the side meeting at PIDapalooza. Under "Scenario A: Community and Governance", the participants in the backcasting generated some thinking about options, captured in the following notes:

- Narrow vs. broad definitions of the Commons.
- A "PID association", perhaps with steering committee.
- Elected body representing the Commons.
- A single point for coordinating funding.
- Broad membership, not just PID service providers.¹

This, of course, was only a starting point; but the idea of a "PID association" was worth exploring, and that would need engagement with stakeholders. So the side meeting at PIDapalooza took shape: 22 individuals were invited, with a very brief appetiser for what would be covered:

¹ D6.3 p.12–13

On behalf of the FREYA project, I would like to invite you a breakfast meeting on 30 January to discuss a strawman for a new PID Foundation. The meeting will be an opportunity to discuss the idea of forming a new entity to advocate for and safeguard the future of persistent identifiers internationally.

At the meeting, Torsten Reimer of the British Library led the discussions, and the response was sufficiently positive that it led the FREYA project to conduct a more formal and comprehensive survey, whose contents, the implications of that, and the FREYA position on it form the bulk of this deliverable.

Although the conduct and follow-up of the study occupied the WP6 team for much of the third year of FREYA, they were not of course the only developments of relevance. Outside the project, important work was done within EOSC at the same time on sustainability. A succession of reports was produced by the Sustainability Working Group, with input from other WGs (Landscape, Architecture and FAIR) and external professional advice where appropriate: these were labelled “Strawman”, “Tinman” and “Iron Lady” (now rebadged as “FAIR Lady”).²

The “FAIR Lady” report (formally titled *Solutions for a Sustainable EOSC*³) explores possible means for sustaining the European Open Science Cloud beyond its initial phase which terminates at the end of 2020. Its approach is to consider the requirements for a “Minimum Viable EOSC” (MVE), then how that could be expanded through wider engagement, and finally examining governance of the EOSC. The report stresses the importance of federation of data and services in EOSC—the very first sentence of the introduction declares that “The European Open Science Cloud (EOSC) federates existing and emerging data infrastructures with the objective of offering a virtual environment in Europe to share and re-use research data across borders and disciplines.”

In the “FAIR Lady” report, PIDs are mentioned only in passing as one of components of EOSC Core: “Services to generate, resolve and validate persistent identifiers.” It is evident that the EOSC needs PIDs. A strong statement is made that “For EOSC to have the greatest impact and reach to external stakeholders it must establish itself as the Web of FAIR data as its primary Unique Selling Point”—and this of course implies the centrality of PIDs reflecting their primary role in FAIR.

It is interesting to examine how the “FAIR Lady” report relates to the three facets of sustainability identified in FREYA D6.1 and expanded in D6.3, namely maintainability, adaptability and desirability⁴. Desirability is certainly acknowledged: “the basic condition of success in ensuring EOSC sustainability is performance: how EOSC as an ecosystem operates and how the resources it federates are used and acknowledged by researchers.” Its performance is what will make it desirable to its users. There are some complications, such as the existence of established services: EOSC should “supplement rather than compete with activities of mature and well-structured research communities that provide thematic services”. Here one might observe that PID services, while not “thematic” in the sense meant here, are mature and well-structured and the same requirement applies. There is a vision of EOSC as a “super platform ... with the possibility of coordinating and connecting other existing platforms such as those operated by the service provider aggregators (GEANT, EGI, EUDAT and OpenAIRE) and the Research Infrastructures in a way that both they and EOSC would benefit by this strategy.”

Adaptability of EOSC is presumably handled by its Marketplace: allowing market forces (in some restricted sense) to respond to changing needs and expectations of users and new technological opportunities. It is envisaged that there will be a membership model for EOSC Core, but a transactional model for the Marketplace.

² A clear summary of the process is shown at https://www.esfri.eu/sites/default/files/BobJones_Session-1_ESFRI-EOSC_06.10.20_0.pdf

³ <https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-175468053>

⁴ These facets are, to quote D6.1, “those qualities that together define what we mean by the word in the context of (e-)infrastructures.”

The model adopted for establishment of EOSC starts with the Minimum Viable EOSC. It is accepted that it will first be necessary to ensure it appeals to its users, which will require funding subsidies at an early stage, and only later aim for financial sustainability. Regarding governance, a major development has been the establishment of the EOSC Association on 29 July 2020 as a not-for-profit international association based in Belgium (AISBL)⁵. This body will coordinate EOSC-related activities and have a role in definition of policies and adoption of standards.

There is a discussion of EOSC in a global setting, though not far advanced in this respect: “There is a clear willingness to collaborate and it is expected that the first agreements could be brokered by coordination fora during the first iteration of EOSC.” This again is a point of contact with the global PID infrastructure, though presumably the target is more “EOSC analogues” in other parts of the world rather than special-purpose infrastructures with global reach, such as PIDs.

The report also considers risk management, specifically in terms of Enterprise Risk Management which “includes the activation of an organised risk management structure, as well as a pool of actions and processes, rules and responsibilities through which decisions are taken and implemented in terms of risks.” Five specific risk categories were identified for the EOSC Association:

- operational;
- reputational;
- governance;
- financial;
- strategic.

These all have their correspondence in the sustainability of the PID infrastructure, and it is noteworthy that they emerged independently during the scoping study on the “PID Federation”, and have been taken into account when refining the facets of sustainability as described in section 4 of this report.

Taking account of this background (which is of course still evolving), this deliverable makes the final presentation of the PID Commons within the lifetime of FREYA. A summary is given of the study on the “PID Federation” (the full report being reproduced in the Annex), followed by FREYA’s response to it, which is to be taken as the final position of the project on such a body (which we prefer to call a PID Alliance). Other aspects of sustainability and their mapping to the stakeholders of the PID Commons are also presented.

⁵ <https://www.eosc-portal.eu/news/eosc-association-registered>

2 The “PID federation” scoping study

The meeting of stakeholders at PIDapalooza in Lisbon in January 2020⁶ had shown that there was interest in the prospect of some kind of PID federation; now it was up to FREYA to follow up, and the mechanism chosen was to conduct an in-depth study, reaching out beyond the parties present in Lisbon to wider communities with a stake in PIDs to obtain their views on the scope, priorities, representativeness and organisation of such a body. It was decided not to divert the effort of the FREYA team members themselves on to this task, but to engage an individual consultant to work within the Research Services section of the British Library. The reasons for this decision were the need for quite specific skills and experience in engagement, consultation and relationship management in the research and scholarly communications sector. While such capabilities could be found among the FREYA partners, there was also the desire for a degree of detached independence from FREYA, allowing an impartial assessment of the findings of the exercise. Therefore, an in-house consultant was involved who contacted the various stakeholders, conducted the analysis and presented their findings in a published report.

An Invitation to Tender was prepared, and the contract was awarded to Josh Brown⁷, formerly of ORCID and Crossref, then an independent consultant, with an excellent knowledge of the PID landscape and experience in stakeholder engagement. The tasks were specified as:

- To analyse the existing landscape of representation of persistent identifiers and PSPs to funding bodies, policymakers and other strategic organisations such as the European Commission and National Science Foundation
- To identify any gaps in representation amongst all of these organisations
- To document and analyse stakeholders’ requirements for a community body
- To assess and develop options for the composition and scope of such an organisation including identifying funding models
- To make recommendations with supporting evidence for the proposed organisation

Josh Brown undertook the work in two phases, as outlined below, with continuous interaction with the FREYA team, and delivered the final report on 29 September 2020.

The report is reproduced in the Annex of this deliverable. It must be emphasized that the report is not itself the “FREYA position” on a PID federation; it is the raw material, the findings of the survey of stakeholders plus recommendations that the author judged to follow from those findings. It is authoritative in its own right as a source document, but it nonetheless needed interpretation, qualification and refinement by the FREYA team in the context of WP6. That perspective is provided in the next section. Here the key points of the report are summarised.

The study took as its starting point a number of possible roles for a “PID federation” from the meeting at PIDapalooza:

Specifically, the new entity:

- *would produce a set of requirements which its members would have to meet including lodging sustainability and exit plans in the event of a PID provider ceasing to operate;*
- *would seek to provide and ensure best practice in the long-term administration of persistent identifier services, particularly in light of organisational and infrastructural change over time. For new types of persistent identifier, they can adhere to these guidelines as a mechanism of ensuring sustainability and building trust in their identifier;*
- *would advocate with funders and regulators on behalf of its members;*

⁶ See blog post at <https://www.project-freya.eu/en/blogs/blogs/an-intergalactic-federation-of-pid-providers>

⁷ <https://orcid.org/0000-0002-8689-4935>

- *could facilitate the safeguarding and preservation of the persistent identifier metadata in a number of national level memory institutions.*

Since the whole point of the study was to engage with a broad community, attention was paid to how this was defined, including organisations that provide PIDs, and those which are most engaged with PID adoption and governance, such as research infrastructures or funders and policy makers. It was envisaged that the community, suitably mobilised through a PID federation, could contribute to the sustainability of the PID infrastructure through a “collective response”—and determining the will for such collective action was the goal of the study.

The study was conducted in two phases, the first using interviews and questionnaires and the second assessing and refining the results of the first phase through focus groups. Seven interviews were conducted and 19 questionnaires were returned; interview candidates were selected for their leadership roles in the PID world. The interview questions and questionnaire are reproduced in the appendices of the study report (the Annex of this deliverable). The focus group attendees were initially selected from the list of participants in the interviews and questionnaires who had indicated that they would be willing to participate in further discussions of the initial research findings. These groups were extended with the addition of a small number of new participants, who were invited to participate specifically to bring perspectives from parts of the PID user community that were underrepresented in responses to phase one, such as infrastructure representatives from Africa and Latin America.

In phase 1, 73% of respondents indicated themselves in favour of the idea of some kind of PID federation, primarily either for greater coherence, interoperability, and coordination for the PID “network”, or to combat fragmentation. It was however clear that there was no single vision of a PID Federation that was held in mind. Yet, a number of potential roles for a PID federation did emerge and gained considerable support:

- communicating the value proposition for the global PID network;
- advocacy on behalf of the PID community;
- establishing and promoting best practice for PID providers;
- support in the development and delivery of sustainability plans to safeguard valuable PID services and associated data;
- helping to unite the community to tackle common problems or challenges.

It was also clear that inclusiveness (with geographical, financial and stakeholder dimensions) would be crucial. It would also be essential that any PID federation should have sufficient commitment and capability to achieve realistic goals in the short term.

The following recommendations were made from the first phase of the study:

1. *A PID Federation is generally welcome and its scope should be refined with the broader community*
2. *A PID Federation should drive adoption and support the sustainability of organisations within the PID network by articulating the value of the whole: PIDs are more than the sum of their parts*
3. *The federation should help to coordinate advocacy across the community without necessarily becoming a figurehead in its own right*
4. *The PID Federation should take on a role in ensuring technical resilience, operational continuity, and the sharing and recording of the community’s hard-won knowledge and expertise in these areas*
5. *The PID Federation should facilitate better and more inclusive communication across key actors, including PID users, and not limited to organisational members, to identify and articulate common problems and challenges*
6. *The Federation should support the widest possible range of existing PID providers in developing models of inclusion that bring truly global voices into the PID community. The PID Federation itself should be an exemplar of inclusion and ensure that underrepresented perspectives are brought to the fore*

7. *The Federation should explore ways that it could actively encourage and underwrite, including potentially via financial support, contributions from the global south*
8. *The PID Federation should consider a light-weight structure with a small secretariat to support activities and a range of ways organisations can contribute, financially, or in-kind, to avoid replicating the unintended barriers to participation created by existing membership models*
9. *The PID Federation should choose its targets wisely, and make sure that it is resourced appropriately to genuinely deliver visible change*
10. *The refined scope of the proposed PID Federation should be put out to a global, open consultation*

Phase 2, with the focus groups, confirmed the general interest in the PID federation with the same reservations about clarifying the scope and objectives. Derived from the findings of phase 1, three roles were proposed:

- advocacy and communication;
- improving interoperability;
- ensuring technical resilience.

All were considered important. The need for inclusiveness was reiterated; this was regarded as the most important feature, with independence and flexibility second.

In conclusion, the study report relates back to FREYA's PID Commons:

This project has shown that the Commons needs "commoners", and the PID Federation has the potential to be the structure that brings 'PID Commoners' together. There is no sustainability without the trust, collaboration and support of the full breadth of the global PID community at every level. An inclusive PID Federation at the heart of the Commons will help to sustain the PID network and make sure that our capacity to persist these infrastructures evolves alongside the community's needs for them.

3 The FREYA perspective: a PID Alliance

As has already been noted, the scoping study was conducted on behalf of but independently of FREYA, and this was a deliberate decision. The FREYA partners are well connected in the universe of PIDs, but the objective of the exercise was to allow a systematic and wide ranging study and report the results as found without direction from the project team. The results—the content of the report—is therefore “raw material” for the next steps, rather than being a FREYA position in itself.

The FREYA partners held a series of meetings, both with and without the attendance of the consultant Josh Brown, to discuss the findings of the study and come to a position. In this section the FREYA perspective is presented, reflecting the discussions and consensus of project partners’ own views and from the stance of the project as a whole (such as implications of the vision of the PID Graph). The aim is to interpret, qualify or extend the findings presented in the report, recognising where there are unresolved issues, highlighting opportunities and making recommendations for concrete next steps.

Two important points are worth making right at the start.

- The preferred name for the body or initiative is **PID Alliance**.
- There is further work needed to establish the role of such a body or initiative and the specific nature of such a body or initiative. The British Library is working with colleagues consulted in the study, to draft a plan for an **open consultation** beyond the end of the funded FREYA project (in which other FREYA partners could also be involved).

3.1 Overall view including scope and objectives

The report itself recommends a wider consultation to prepare the ground for a PID federation with wide acceptance. While recognising that there has to be a limit to the consultation process, this does seem a sensible next step. The study was widely publicised internationally, and has developed a proposition that can be openly discussed and refined, but it was based on consultation by invitation and this should be followed by a truly open process.

Recommendation: An open consultation with a wider range of stakeholders should be conducted (and indeed as noted just above is now planned) in order to clarify the role of such a body or initiative and the specific nature of such a body or initiative (e.g. initiative or standalone entity).

A particularly important question is about the name of whatever body emerges. In initial discussions in FREYA, the term “PID Foundation” was used, though this carries too many associations of an established and rather weighty body. Subsequently “PID Federation” (with or without a capital letter) has been employed. In any case the name should be short, descriptive, not confused with other organisations, and reflecting the scope and aims of the body.

Recommendation: Provisionally, the body should be referred to as the “PID Alliance”. This name highlights the equality and common interests of the participants, and does not imply a particular organisational structure.

Concerning the scope of the PID Federation/Alliance, the report finally settles on three major focusses.

Interoperability. The recognition that PIDs are more than the sum of their parts is of course entirely consonant with FREYA’s vision of the PID Graph as “networks of connected PIDs” with “new services operating over the Graph providing value to a wide range of users”⁸. The emphasis should therefore not be on PIDs themselves but on open PID infrastructure (services, metadata etc.), encouraging interoperability

⁸ M. Fenner, “The PID Graph in FREYA (additional project report)”, <https://doi.org/10.5281/zenodo.4028383>

through expansion of PID types and metadata. The RDA Interest Group “Open Science Graphs for FAIR Data”⁹, in which FREYA members are active, is relevant here.

Recommendation: The driving vision for the PID Alliance could be the expansion and enrichment of the PID Graph and local instances of it.

Inclusion. A very strong outcome of the report is the desire for greater inclusiveness, both geographically (the global south) and of stakeholders hitherto under-represented. FREYA, though an EC-funded project, has experience here with its programme of international ambassadors, but evidently there is much more to be done. It should also be recognised that existing PID service providers do engage with their users, and should be encouraged to extend that engagement accordingly.

Communication and advocacy. Given that it is admitted that the participation in a PID Alliance is not yet clearly defined, communication is taken to mean articulation of common challenges and advocating for the value of PIDs, both presenting a unified view from the PID “community” Advocacy of this kind seems uncontroversial, though (as noted in the report itself) any kind of lobbying should be avoided.

As with inclusion, the Alliance will need to recognise that existing PID service providers share many of the same visions and goals but have different priorities in terms of primary stakeholders, so their routes of communication may be different and these differences should be tolerated.

It is worth making a final remark on the question of certification of PID service providers. The report indicates that there is no desire for this as a role of the PID Alliance, as it might be counter to inclusiveness and reinforce existing positions. Certainly any certifying organisation would first have to gain the trust of the community. Nonetheless trustworthiness of PID services (in terms of accessibility, reliability and longevity) is essential for a PID infrastructure and the PID Alliance should somehow work towards that. EOSC is interested in certification of service providers and this might be a starting point to observe the progress made and how barriers could be overcome.

3.2 Priorities for action

There is a clear desire expressed in the report for the PID Alliance to identify and tackle common challenges, and to do so quickly: “it’s going to need to show tangible value early on, or support will drain away” as one respondent observed. It might be fruitful to think in terms of “quick wins” on a timescale of say 18 months that could be achieved with the minimum resources, followed by medium-term goals once the confidence of the community has been built up. Even establishing some sort of recognised and accepted “brand” for the PID Alliance in the inclusive way suggested by the report would be a valuable priority itself.

There is a clear role for the PID Alliance in knowledge exchange: experienced providers could assist those at an earlier stage. A kind of “incubation” role or bank of technical knowledge for new PIDs, dealing with issues such as robust systems behind the APIs and scalability (and thus helping to address technical resilience, though the PID Alliance could not be responsible for operational continuity itself).

Recommendation: The PID Alliance should have a rolling programme of short-term goals that can continue to be adapted, achieved and added to, and allow for continued review of the need for the Alliance.

Recommendation: The community forum *pidforum.org* (originating with FREYA, but taken over by the community under the lead of NISO¹⁰) should be taken as a building block of the PID Alliance: a ready-made, inclusive forum for communication and knowledge exchange.

⁹ <https://www.rd-alliance.org/groups/open-science-graphs-fair-data-ig>

¹⁰ News item on the future of the PID Forum: <https://www.project-freya.eu/en/news/newsitems/the-future-of-the-pid-forum-is-assured>

Recommendation: An initial role of the PID Alliance should be in supporting emerging or maturing infrastructures through sharing of knowledge. This has the advantage of addressing interoperability, inclusiveness and communication together.

Recommendation: The PID Alliance should align with general principles for PIDs like those of Bilder-Lin-Neylon for the Principles of Open Scholarly Infrastructure¹¹, aligned with the EOSC PID Policy, and covering aspects such as trustworthiness, longevity and governance, reflecting the inclusive hopes for the Alliance.

3.3 Implementation of the PID Alliance

A particularly pressing question is to what extent the PID Alliance should be a formal organisation, rather than an association of interested and committed parties. It is evident from the report that those consulted had different expectations and preferences—though something lightweight was the general preference. It must be noted that there will always be some cost to participation, whether met by a funding organisation, supported through some kind of subscription fee, or in staff time of participating organisations.

One of the tensions in the report is between the need for capability of useful action with the desire for lightweight structures and (at least among some respondents) an aversion to a newly created body. Speaking already of a “secretariat” or “headquarters” is not right as it makes too many presuppositions. Admittedly there is a delicate relationship between understanding what are the key activities the community wants in the short to medium term and deciding on the structures or processes to implement them. There are two extremes, with a range of possibilities between: an established entity (whether new or leveraging some existing body), or a space where the conversations can continue.

Recommendation: The PID Alliance should take advantage of the relationship with existing forums/bodies such as RDA and FORCE.

Recommendation: Seek existing organisational models to follow or contrast—for example OASPA, RDA, PIDapalooza, ROR.

If the PID Alliance is not to be completely nebulous—which is required if it is to have credibility at all—it will need some notion of participation. Participation need not necessarily mean paying a subscription fee; it might just be signing up to some agreement (thereby lowering barriers). However, it must have some meaning, especially if advocacy is to be done—the Alliance would need some legitimacy and authority. Consequent on this, the size of the participating “membership” will be a major factor in determining its nature. There is a connection with geographical inclusiveness: it is clear from the report that this should be a broad initiative.

Recommendation: The wider consultation should cover implementation models; it is too early to propose a particular model now.

A final consideration is about sustainability of the PID Alliance itself. It should only last as long as it is needed, though there is a clear desire for a lifetime of at least a few years.

¹¹ <https://doi.org/10.24343/C34W2H>

4 Sustainability and a PID Alliance

The relationship between a PID Alliance and sustainability of the PID infrastructure is not straightforward. It is true that the idea of some kind of federation of PID stakeholders sprang from the work on sustainability in FREYA, and the possible roles for the Alliance refer to sustainability in terms of long-term administration, exit plans for providers, and preservation of PID metadata. However, the tasks specified for the conduct of the study were framed in terms of involvement of stakeholders, acquisition of their requirements and their desire for a federation. It was not the purpose of the study to quiz participants on their views on sustainability of PID infrastructure, but to determine their appetite for collective action, in which sustainability was one of the elements. It was necessary to be open-minded and to ask about challenges for PIDs for which a response might be appropriate, without focussing prematurely on sustainability *per se*.

It is apparent that the conclusions of the study do have a bearing on sustainability; in fact the report notes that sustainability/persistence was the most frequently cited challenge facing the PID infrastructure. However, other issues were also raised such as inclusiveness and communication—though it might be argued that these are aspects of the facet of *desirability*, helping to ensure the continued engagement between providers and users. In any case, it was never the expectation that the PID Commons would be identical with any PID Alliance; rather, their concerns overlap (Figure 1). The stakeholders might be the same, but their concerns are somewhat separate. As the scoping study report puts it, “this project has shown that the Commons needs ‘commoners’, and the PID Federation [Alliance] has the potential to be the structure that brings ‘PID Commoners’ together.” By “PID commoners” we understand those who come together to manage or influence aspects of the PID infrastructure for individual and collective benefit. A PID Alliance certainly can play a central role in the PID Commons—rather as its “conscience”, striving for what in D6.1 was called “an ideal, which could be summarised as ‘open collaborative governance structures with empowered stakeholders’.” It will be the PID Alliance that creates the trust and openness that are necessary for the proper functioning of the Commons.

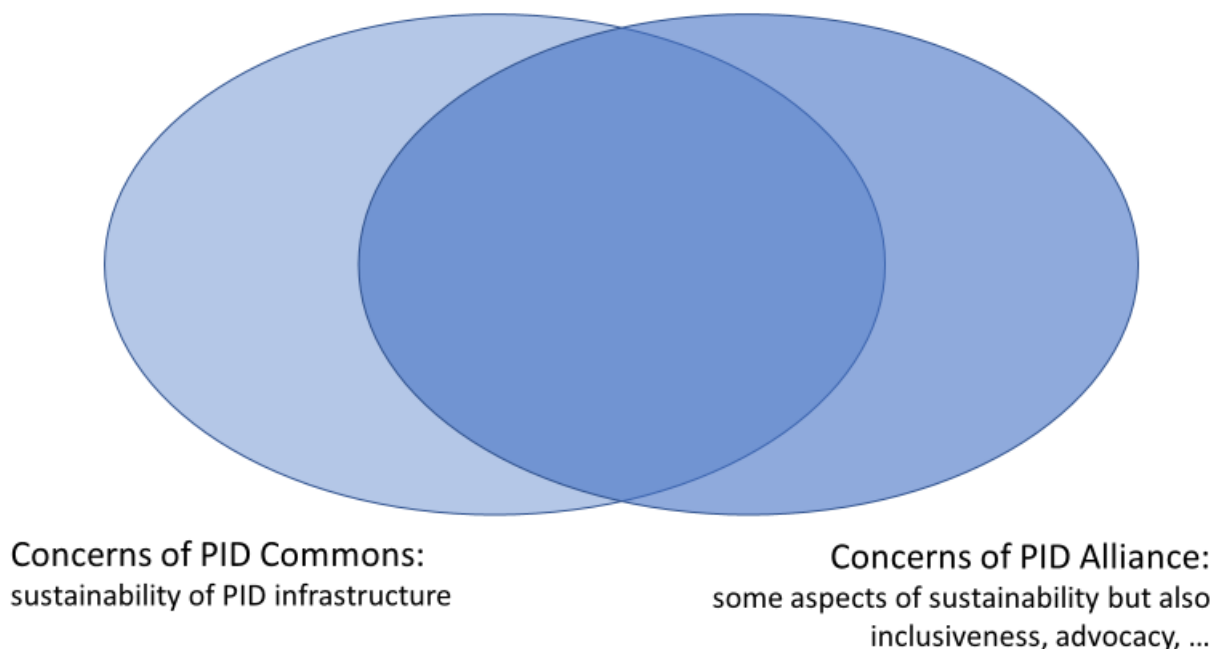


Figure 1 The overlap between PID Commons and PID Alliance

In order to conclude the sustainability work of FREYA it is therefore necessary to ask what aspects of sustainability are not covered by a PID Alliance and how they would be addressed. The starting point is the expansion of the three facets of sustainability that was presented in D6.3. By examining various sources

such as the discussion in D6.1, the material referred to therein (such as the well-known Bilder–Lin–Neylon paper), and the draft PID policy that FREYA helped to produce¹², a list of more specific concerns for each of the three facets was produced:

- **Maintainability:** service standards; funding; resilience of the infrastructure
- **Adaptability:** core and additional services; PID attributes; technology watch; supplier awareness
- **Desirability:** global developments; user awareness; community involvement; market characteristics; characteristics of PID Service Providers and PID Authorities

In third year of FREYA these have been reviewed, refined and extended; and one of the references for that work has been the scoping study itself. Examination of the findings of the study reveals a complex and subtle web of sustainability issues needing attention, covering all three of the facets of sustainability. The participants did not necessarily express their priorities in terms of sustainability, but the underlying theme can be seen: as one example, the requirements to manifest the value from PIDs is part of desirability, attracting stakeholders to the infrastructure and drawing in their commitment.

The result of this refinement of the facets of sustainability is shown in Table 1.

Facet	Component of sustainability
Maintainability	Resilience of the infrastructure including scalability Service standards Sunsetting/exit plans for providers and services Funding of providers and services Sharing of expertise
Adaptability	Core and additional services PID attributes Technical and organisational coherence/alignment Technology watch Supplier awareness Identification of common challenges Innovative use cases
Desirability	Global developments User awareness and community involvement Characteristics of PID Service Providers and PID Authorities Coordination between all stakeholders The whole PID “workflow” and evidence of value Core values (analogous to FAIR) Adaptation of services to meet global needs and constraints (such as limitations of funding or physical infrastructure)

¹² <https://op.europa.eu/en/publication-detail/-/publication/35c5ca10-1417-11eb-b57e-01aa75ed71a1/language-en>

Table 1 Refined components of sustainability

In order to map the PID Commons, it is necessary to be assured that these components are covered, in the sense that some stakeholder is concerned with them. This is not the same as being *responsible* for them, though in a particular case these might coincide. There is an analogy with the idea of “problem ownership” in an organisational context: the “owner” of a problem is not necessarily the person or unit that suffers from the problem, nor the one that has to fix it, but takes responsibility and works with others to ensure an appropriate response to the problem. *Concern* means recognition of importance, the current state and possible future, the risks and opportunities, and the capacity to propose courses of action. This is the essence of the “empowered stakeholders” referred to above: a recognition that some aspect of the environment is not beyond control, not something that has to be merely accepted, but is subject to influence to produce tangible benefits.

Now it would be possible to take each of these components of sustainability in turn and search for a stakeholder that holds it as a concern, but this would be a rather pedantic and unilluminating exercise. It is more worthwhile to group these concerns, not as the three facets of sustainability from which they arose, but according to the nature of the concern itself. In fact deliverable D6.1¹³ already proposed a simple three-way grouping in its Figure 2 (p.18), where it breaks down the “structures/mechanisms” of sustainability planning into “doing it” (meaning operating the infrastructure and providing services), “paying for it” and “taking a long-term view of it”. To these we could add—in the light of the importance of communication and inclusiveness highlighted in the study report—“engaging with it”¹⁴.

Table 2 shows the grouping of sustainability components according to this scheme—though the four short phrases have been further reduced to capture their essence. Furthermore, it identifies the primary stakeholders that will relate to those concerns, accept them as their own and act upon them. These stakeholders include the PID Alliance, but also other existing bodies playing a role on the PID stage and that will continue to play that role, or extend it, in the context of the PID Commons.

Type of concern	Component of sustainability	Primary concerned stakeholders
Engagement (“Engaging with it”)	Sharing of expertise Identification of common challenges Innovative use cases User awareness and community involvement Coordination between all stakeholders The whole PID “workflow” and evidence of value Core values (analogous to FAIR) Adaptation of services to meet global needs and constraints (such as limitations of funding or physical infrastructure)	PID Alliance

¹³ Deliverable D6.1 is available on Zenodo: <https://doi.org/10.5281/zenodo.3250702>

¹⁴ These four phrases have been deliberately chosen to be short, memorable and not exactly precise as to what the “it” refers to: “it” is understood as the PID infrastructure and everything needed to support and sustain it long-term.

Operation ("Doing it")	Core and additional services PID attributes Resilience of the infrastructure including scalability Service standards Technical and organisational coherence/alignment Characteristics of PID Service Providers and PID Authorities	PID service providers acting in concert and mediated by the PID Alliance
Assuring funding ("Paying for it")	Funding of providers and services	PID Alliance (meaning not that a PID Alliance would be <i>responsible</i> for funding of services, but that it would take a view on how services could and should be funded, considering among other factors the needs for inclusiveness) EOSC governance and other international bodies with an interest in the PID infrastructure
Long-term perspective ("Taking a long-term view of it")	Sunsetting/exit plans for providers and services Technology watch Supplier awareness Global developments	RDA through Interest Groups and Working Groups PID Alliance

Table 2 Grouped components of sustainability

Here we have a map of the PID Commons as it has emerged at the end of the FREYA project. It brings together the important work done on the prospects for a PID Alliance, and the FREYA team's response to that, with the analysis of sustainability elaborated in earlier deliverables. It makes explicit how the components of sustainability, arising from the three basic facets, can be grouped according to the type of concern they represent, and how they can then be mapped in quite a natural way to the organisations that will relate to those concerns.

There is, of course, an assumption that the organisations named will have a degree of sustainability themselves, but that is reasonable. There is also an implied need for some kind of overall coordination—for example, in proposing the creation of RDA groups for particular PID-related purposes. This is an aspect of engagement for which the PID Alliance will be particularly suited.

Annex: Final report on PID Federation scoping study

The final report produced by Josh Brown is reproduced below, preserving its original pagination. This report is also available as a standalone document at <https://doi.org/10.5281/zenodo.4059556>.



PID Federation scoping study: final report

Prepared by Josh Brown, research consultant for the FREYA project

September 2020.



Executive summary

The FREYA team has taken forward a conversation that has been developing over several years amongst persistent identifier (PID) users and providers: the idea that these critical research information infrastructures could be better supported by a coordinated community. A 'federation' of PID providers and users, spanning other research infrastructures, funders and policy makers, and the wide research community was mooted in January 2020. A project was commissioned to scope this potential 'PID Federation', undertaken between June and September 2020. This report outlines that project and summarises its key recommendations.

An initial consultation comprising interviews and questionnaires was conducted, to generate an initial vision for the PID Federation. The findings of this phase were then validated in a series of three focus groups, spanning the globe. The primary findings were:

1. A PID Federation is generally welcome and its scope should be refined with the broader community
2. A PID Federation should drive adoption and support the sustainability of organisations within the PID network by articulating the value of the whole: PIDs are more than the sum of their parts
3. The federation should help to coordinate advocacy across the community without necessarily becoming a figurehead in its own right
4. The PID Federation should take on a role in ensuring technical resilience, operational continuity, and the sharing and recording of the community's hard-won knowledge and expertise in these areas
5. The PID Federation should facilitate better and more inclusive communication across key actors, including PID users, and not limited to organisational members, to identify and articulate common problems and challenges
6. The Federation should support the widest possible range of existing PID providers in developing models of inclusion that bring truly global voices into the PID community. The PID Federation itself should be an exemplar of inclusion and ensure that underrepresented perspectives are brought to the fore
7. The Federation should explore ways that it could actively encourage and underwrite, including potentially via financial support, contributions from the global south
8. The PID Federation should consider a light-weight structure with a small secretariat to support activities and a range of ways organisations can contribute, financially, or in-kind, to avoid replicating the unintended barriers to participation created by existing membership models
9. The PID Federation should choose its targets wisely, and make sure that it is resourced appropriately to genuinely deliver visible change
10. The refined scope of the proposed PID Federation should be put out to a global, open consultation

The focus group analysis suggested that the themes of inclusion, interoperability, and communication resonated most strongly with the community. Each of these has a critical role to play in supporting the sustainability of the international PID network, and such an initiative



could help to address systemic imbalances in access to and benefits from large scale research information systems, and in levels of recognition for research contributions from less wealthy regions, communities, and disciplines.

Respondents felt that the PID Federation should be designed and resourced to be globally inclusive and, once established, it should focus on improving interoperability with and between PID systems.

This report suggests a path forward for the PID Federation, with the next crucial step being a much wider, open consultation on the scope developed and refined in the course of this research. Given that this work is taking place at the end of the FREYA project, the team should move urgently to identify a group of interested parties who can take this work forward.



Contents

1. Introduction	5
2. Community context	7
3. Research method.....	9
3.1. Phase one: Interviews and questionnaires	9
3.2. Phase two: focus groups.....	10
4. Findings of phase one.....	13
4.1. Responses to the idea	13
4.2. What should, or shouldn't, the PID Federation do?	14
4.2.1. Communicating the value proposition for the global PID network	15
4.2.2. Advocacy on behalf of the PID community	15
4.2.3. Establishing and promoting best practice for PID providers	16
4.2.4. Support in the development and delivery of sustainability plans to safeguard valuable PID services and associated data.....	17
4.2.5. Helping to unite the community to tackle common problems or challenges	18
4.3. Who should be involved?	19
4.4 How could a PID Federation be supported?	20
4.5. Observations	21
5. Phase one recommendations	23
6. Findings of phase two	25
6.1. Opening responses to the phase one report	25
6.2. Possible roles for the PID Federation	26
6.3. Overarching themes.....	29
6.3.1. Communication	29
6.3.2. Inclusion	30
6.4. Desired features of the PID Federation	31
6.5. Thoughts for the PID Federation.....	32
6.5.1. "What common challenges do you think a PID Federation should help to tackle first?"	32
6.5.2. Which communities or groups or regions should we be consulting next?	33
6.5.3. Are there issues that the PID Federation might face that have not been mentioned in our research?	34
7. Reflections on the focus groups.....	35
8. Conclusions.....	37
Appendices	40
Appendix 1: Interview questions.	40
Appendix 2: Questionnaire form.	41



1. Introduction

The idea of a ‘PID Federation’ as a way to bring together the global identifier community to address shared problems has been much discussed in recent years, and was seen by the FREYA project partners as a natural fit for their work on the ‘PID commons’. The PID Commons is one of the three pillars of the FREYA project, and is intended to address the post-project sustainability of PID infrastructures resulting from the FREYA partners’ work.¹

The discussions began in 2018, and continued in 2019, with workshops in Singapore, London,² and Portland, OR.³ These events were co-sponsored by FREYA partners (ARDC⁴, DANS⁵, ORCID⁶, STFC⁷) and external partners Jisc⁸ and California Digital Library⁹. During discussions of shared risks and values for PID providers the notion of some kind of ‘global union’ of PID enthusiasts emerged.

The promise of that discussion stems in part from its global nature: it was a conversation that began in Singapore with input from both global PID initiative representatives and experts from Malaysia, Singapore, Japan and other countries in the Asia-Pacific region, continued in London with a primarily European discussion group, and concluded in the United States with a discussion group dominated by representatives from the Americas. At each of these meetings there was a clearly articulated need for better coordination in critical areas for the sustainability of PID infrastructures: governance, interoperability and technical resilience.

At the annual PIDapalooza¹⁰ conference, the de facto annual meeting of the international PID community, the FREYA partners presented a proposal for a ‘PID Foundation’ to support the PID commons.¹¹ This proposal suggested a number of possible roles for the initiative to be explored during the meeting:

“Specifically, the new entity:

- *would produce a set of requirements which its members would have to meet including lodging sustainability and exit plans in the event of a PID provider ceasing to operate;*
- *would seek to provide and ensure best practice in the long-term administration of persistent identifier services, particularly in light of organisational and infrastructural change over time. For new types of persistent identifier, they can*

¹ <https://www.project-freya.eu/en/about/mission>

² <https://www.project-freya.eu/en/events/london-pids-workshop>

³ <https://www.project-freya.eu/en/events/pid-workshop-in-portland>

⁴ <https://ardc.edu.au/>

⁵ <https://dans.knaw.nl/en>

⁶ <https://orcid.org/>

⁷ <https://stfc.ukri.org/>

⁸ <https://www.jisc.ac.uk/>

⁹ <https://cdlib.org/>

¹⁰ <https://www.pidapalooza.org/>

¹¹ <https://www.project-freya.eu/en/blogs/blogs/an-intergalactic-federation-of-pid-providers>



adhere to these guidelines as a mechanism of ensuring sustainability and building trust in their identifier;

- *would advocate with funders and regulators on behalf of its members;*
- *could facilitate the safeguarding and preservation of the persistent identifier metadata in a number of national level memory institutions.”*

These potential goals were broadly accepted as areas for investigation, so the FREYA team commissioned a scoping study to evaluate the possible roles a ‘PID Federation’ could play. What would be desirable, or acceptable, to the PID community? What might the top priority challenges be? How could such an initiative operate and be sustained?

This document describes the initial results of that scoping exercise commissioned by the FREYA project.



2. Community context

Since this work is based on a preliminary consultation with a sample of the PID community, it is worth exploring how the 'PID Community' was defined at the outset of the project. For the purposes of this exercise, the PID community was broadly defined to include organisations that provide PIDs, and those which are most engaged with PID adoption and governance, such as research infrastructures or funders and policy makers.

While millions use PIDs as part of their daily work (clicking on DOI links or signing in with ORCID, for instance) there is a more limited pool of stakeholders that have the capacity and expertise to engage with PID provision, governance or technical development in the research space. This is compounded by the complexity of the PID community. For instance, some sectors of the PID community have something like a federal structure, with intersecting layers of activity and governance.

As an example of such a 'federal' structure, take the global Handle system.¹² It is overseen by the DONA foundation.¹³ DOIs are a kind of Handle and are overseen by the International DOI Foundation.¹⁴ DOIs in the academic space are provided by Registration Agencies which serve specific communities, such as Crossref for scholarly publishing.¹⁵ In this context, DONA represents the federal level, the DOI foundation the regional level, and Crossref the municipal level.

This type of structure has evolved as PIDs depend on global, multi-use infrastructure, but need to be adapted to specific contexts. This has obvious benefits, in that accountability is enforced from the bottom up by members, who tend to work most closely with the PID provider explicitly serving their needs, but at each step up the 'federal ladder' the distance from the PID user base increases. Another body which sits above or alongside the 'federal' level may increase the gap between the PID users and individual members and the level of ultimate accountability.

The closeness of many PID systems to their core community has fostered trust and responsiveness. As examples of this, think of the International Geo Sample Number which was developed by practitioners in the earth sciences, or identifiers.org which has evolved to serve the needs of many 'non-PIDified' databases in the life sciences. From the user perspective, diversity, community specificity and decentralisation can be highly beneficial and help to ensure that specific disciplinary use cases are addressed. For others, the multiple redundancies created are a source of resilience, and worth the added complexity they can bring.

That said, these 'Good Things' are not without drawbacks. They conspire to create a highly fragmented landscape. In the PID space there are many players, which can lead to healthy competition but also to poor coordination. Risks can go unseen and unmitigated. The wider

¹² <http://handle.net/>

¹³ <https://www.dona.net/index>

¹⁴ <https://www.doi.org/>

¹⁵ <https://www.crossref.org/>



research community is presented with many variations on core messaging about the value of PIDs. Valuable opportunities for collaboration may go unseized. Duplication of effort discourages adoption of PIDs by generating uncertainty over which PID to use. It also wastes scarce resources, a matter of increasing urgency in the current climate.

Funders want to see solutions to the challenges of research information management at internet scale, and may see the diverse PID landscape as disjointed. The complexity of the PID network may lead them to overlook the value of specific PIDs. As funders increasingly mandate specific PIDs in contexts such as Plan S¹⁶, or create guidance and policies around PIDs, such as the EOSC PID Policy¹⁷, clear PID community input and advice becomes even more important. New entrants could gain traction providing proprietary solutions, or new technologies could weaken or eliminate a coherent management layer for PID systems, making fragmentation even harder to mitigate. All of these issues, if left unaddressed, could hamper a collective response to the shared challenges of persistence.

A major feature of persistence is sustainability. The FREYA deliverable 6.3: “Second annual report on the PID Commons and sustainability” speaks of many components of sustainability including ‘empowered stakeholders’, ‘common interest’, ‘consensual decision-making’, ‘openness’, ‘representativeness’ and ‘trustworthiness’.¹⁸ These considerations show that sustainability is actually determined as much by social factors as by technical features. In light of this, some kind of collective contribution to the sustainability of the PID network is essential.

A PID Federation could provide the framework for such a collective response, but it can only do this with support and buy-in from a global, inclusive, and engaged set of stakeholders. This study sought to reach out to a small sample of this possible stakeholder community to evaluate the appetite for specific kinds of collective action.

¹⁶ <https://www.coalition-s.org/>

¹⁷ <https://doi.org/10.5281/zenodo.3574202>

¹⁸ https://www.project-freya.eu/en/deliverables/freya_d6-3.pdf, p15



3. Research method

Since the purpose of this exercise was to explore informed opinions about the possible scope of a PID Federation, the project team agreed that initial conversations should be with a balanced list of stakeholders with experience in PID provision, integration, or governance. There was clear guidance from the FREYA partners that the group should include a mix of:

- Geographical representation - covering the global south, and a good balance of the Americas, Europe, the Middle-East and Africa, and the Asia Pacific region.
- Practitioners and leaders - including those who operate or provide PID services, and those working in existing PID community bodies and governance boards.
- Perspectives on PIDs - incorporating the views of PID providing bodies, but also research infrastructure providers (who implement PIDs at scale and for a variety of use cases) and funders or policy makers who may mandate PIDs or rely on them for the delivery of concrete policy goals.

There were several rounds of discussion before a balanced list of respondents was agreed.

Participants took part under condition of anonymity, but agreed to be quoted under the Chatham House rule (i.e. without attribution). Their contributions are quoted directly, with only minor edits to preserve anonymity where necessary.

The research was split into two phases. Phase one comprised interviews and questionnaires. Phase two consisted of three focus groups examining the findings of phase one.

3.1. Phase one: Interviews and questionnaires

In all 51 individuals were invited to participate in this phase of the research. 12 were invited to interview, and 39 were asked to respond to a questionnaire. 7 interviews were conducted, and 19 questionnaires were returned, giving an overall response rate of 51%. Only one invited respondent declined to participate. All participants were asked if they would be willing to participate in a subsequent focus group.

Interview candidates were selected based on the nature of their experience of the PID community: they have all led a PID provider either at an executive or board level, or have leadership roles in PID or research information community initiatives. Interviews were conducted via the video conferencing platform Zoom, and lasted approximately 1 hour. The semi-structured interview script is included in Appendix 1. Interviews were recorded, and transcribed for accurate quotation and reference.

Questionnaires were shared as Word documents via email. The questionnaire outline is included in Appendix 2.

Recordings and documents were stored offline with local backups, and will be deleted when this project is completed.



The table below shows the numbers invited from each region, and the number who actually completed an interview or questionnaire.

Table 1: Participation in the study by global region

Region	Invited	Participated
Asia Pacific	7	6
Europe	16	9
Latin America	5	2
Middle East and Africa (MEA)	5	2
North America	7	3
Global bodies	11	4
Total	51	26

Of the 51 invited participants, 16 were from the policy/funding community, 19 were from the research infrastructure community, and 16 were from the PID provider community.

The respondent pool was dominated by Europeans. Latin America and MEA were underrepresented in both invitations and responses. In part, this reflects the scale of activities and investment currently in play in Europe in such initiatives, via the European Open Science Cloud etc. but it also reflects the fact that the project team are European and have better connections in this region. This imbalance was raised with participants, and their responses on this topic are included in a special section on inclusion (below).

The second phase took some steps to address this imbalance, and it is strongly recommended that future stages of this investigation and any subsequent activities (consultations, outreach etc.) be designed and promoted to include additional representation from the global south.

3.2. Phase two: focus groups

An initial report (made up of sections 1-5 of this document) was shared with focus group attendees, and themes from the report were used as the basis for the focus group discussions. Focus group attendees were initially selected from the list of participants in the interviews and questionnaires who had indicated that they would be willing to participate in further discussions of the initial research findings.



These groups were extended with the addition of a small number of new participants, who were invited to participate specifically to bring perspectives from parts of the PID user community that were underrepresented in responses to phase one, such as infrastructure representatives from Africa and Latin America. Given that there was a strong emphasis on inclusion in the first set of recommendations, it was felt that participation from nations and groups that were described as underrepresented in the phase one findings should be increased. In total 10 invitees had not completed a questionnaire or interview in phase one. This brings the total number of research participants from both phases to 36 (with an original target of 50, it takes the participation rate to 72% of that envisaged at the outset).

Table 2: Breakdown of focus group invitees

Participant category/region	Americas	EMEA	APac	Total
Infrastructure	3	5	1	9
Funding/Policy	2	2	1	5
PID providers	2	2	4	8
Total	7	9	6	22

There was some attrition in participation. Some invitees did not respond to the invitation, and some were unable to attend the sessions.

Table 3: Breakdown of focus group attendance

Participant category/region	Americas	EMEA	APac	Total
Infrastructure	3	4	1	8
Funding/Policy	2	1	1	4
PID providers	1	1	2	4
Total	6	6	4	16

It is worth noting that of the 16 who attended one of the focus groups, two were only able to attend part of the session, and two had serious issues with their internet connection which limited their participation in discussions.

One participant who was unable to join the EMEA focus group chose to join the Americas focus group instead, and one respondent from North Africa who was unable to connect to the EMEA focus group gave their thoughts on the discussion in a post hoc phone call. These discussions took place across three focus groups, conducted online using the Zoom video conferencing service. Each focus group was timed to be in normal working hours in



three main global 'verticals': The Americas; Europe, the Middle East, and Africa (EMEA); and the Asia-Pacific region.

Additionally, each focus group was attended by a representative from the British Library on behalf of the FREYA project. These representatives were there to observe the discussions, and to answer any questions about the FREYA project in general and about the purpose of this research for the project in particular.

Focus groups were recorded for accuracy of note taking. The conversations were enriched with discussions in the chat, and by the use of online live polls, conducted using the Mentimeter platform. Responses to the live polls were used to direct discussions and to foster deeper discussions of key themes.



4. Findings of phase one

4.1. Responses to the idea

73% of respondents were in favour of the idea of some kind of PID Federation. 18% were ambivalent or undecided, and 9% were opposed.

Those in favour offered a consistent set of reasons for supporting this idea, primarily either the need for greater coherence, interoperability, and coordination for the PID 'network', or (and this is perhaps the same reason viewed from another angle) as a fix for the fragmented, competitive, and complex PID community:

"The lack of coordination makes the ecosystem more fragmented than it should be and stakeholders are missing value added and opportunities on the way. Coordination doesn't mean that each project should give its own identity away, it means making the whole system more efficient and scalable."

"Makes sense - a PID Federation can help to achieve necessary levels of technical and organizational coherence"

"With the potential of PIDs being the backbone of an open research information infrastructure, coordination, collaboration, and alignment about PIDs is essential."

"We believe the global academic community will be greatly benefited from an extended and integrated infrastructure of the PID collaborations."

Most respondents expressed excitement or enthusiasm for the idea succinctly ("a great idea") but on examining their detailed responses to specific aspects of the proposal, it is clear that they were offering support for one vision of the PID Federation as much as the idea per se.

Ambivalent responses ranged from those who felt that the purpose was unclear:

"The Federation should consider whether it is a community club, a lobbying organisation, a service provider, a consultancy or something else."

To those who felt that the nature of the PID community itself made success unlikely:

"I think that is a great idea, but I am absolutely skeptical about its success. The reason of my skepticism is the existence of so many international organizations doing the same thing, giving me the feeling that they are competing against each other instead of cooperating to achieve a common goal...Why do we have so many PID providers making the same thing and apparently believing they will reach better results comparing with their competitors, each one of them keeping their own infrastructures and access rules? ...I can't see the point of having so many non-



profitable organizations doing the same thing and apparently dreaming to reach different and better results. Seems to me we are dealing with bonfires of vanities.”

Of the respondents who opposed the idea of a PID Federation, their reasons include concerns around another layer of overhead:

“My overall take is that a PID Federation is not necessary and does not necessarily represent a “silver bullet” solution to the challenges it seeks to address and the goals it seeks to achieve. In my view, a Federation would add unnecessary overhead and Complexity.”

It was also compared to an organisational restructure in which another level of management is added, but nothing is gained.

Other responses focused on the sense that the benefits or purpose of a PID Federation are not clear (or perhaps not convincing): “What gap is it trying to fill?”.

It is worth noting that this objection was echoed by the individual who declined the invitation to take part in this research. They did so on the basis that they were opposed to the idea, and gave permission for their reason for declining to be quoted. It was:

“My gut feeling on this is that a PID Federation sounds a little unnecessary – and at a time when everyone is going to be looking to minimise expenditure setting up yet another organisation, sounds challenging. Given that we have ORCID, Crossref, DataCite etc – I’m not sure what problem this is trying to solve.”

Overall, there is a clear desire for better coordination in the PID community, and a strong sense that a unifying group or entity offers the best chance of delivering that coherence. This, however, is accompanied by the caveat that the ultimate level of support will depend on the priorities or role selected for the Federation.

4.2. What should, or shouldn't, the PID Federation do?

In discussions with the FREYA team, the potential roles of the PID Federation were broken down into a series of simpler 'value propositions' than those expressed before the PIDapalooza community meeting.

It is clear that attempting to deliver all of these is akin to 'boiling the ocean' and will require a level of organisational activity and size which is going to be hard to sustain without significant support from not just PID providers, but their members, funders and other bodies. Not one of the respondents supported all of the possible roles, but some gained significantly more support than others.



4.2.1. Communicating the value proposition for the global PID network

The need for better, more consistent and effective communication of the value proposition of PIDs as a network, rather than at the level of individual PIDs or services came up again and again in responses.

“Communication! Making sure that the end-users, such as funders, institutions, individuals, know what PIDs are and what value they can deliver. Focusing on a few key PIDs and communicating these along with applications, which really deliver value would be important, I think.”

“Each stakeholder needs to understand and see value in how PIDs are created / used upstream and downstream of their position in any particular workflow / sub-set of the research lifecycle(s). That’s what will make it an infrastructure rather than a set of components.”

“I would ask them to establish a real cooperative environment amongst them and with national partners. Celebrating agreements with partner or partners of a determined country, they could enlarge the number of institutions covered by their PIDs.”

“You know, I think the more we confuse the message, the more difficult it will be to evidence the value. So the more we can have a clear and consistent and coherent message, the better so that if we are talking to places like government, or major funders or whoever, you know, we’ve got that clarity around.”

There is a strong sense that the advancement of all PID systems is harmed by fragmentary, confusing, repetitive or contradictory messaging from the PID providing community. A PID Federation could drive adoption and support the sustainability of organisations within the PID network by articulating the value of the whole: PIDs are more than the sum of their parts.

“Community groups that are advancing the PID agenda as a whole rather than individual PID providers may be a better way to go.”

4.2.2. Advocacy on behalf of the PID community

A common refrain in responses is that there is a fine line between advocacy and lobbying. Many PID organisations have legal limitations on the extent to which they can lobby, and it is crucial that a PID Federation remains on the right side of that line. That said, advocacy can take many forms, and promotion would be acceptable to most:

“Most of the discussion on PIDs is within the community which is already very much aware and supportive to PIDs. The awareness should be raised beyond this community. A PID Federation can focus on this.”

In fact, the need for that promotion was noted in many responses, reinforcing the potential communications role endorsed in 4.1.1. above:



“Seems very difficult to rally public sector organisations into supporting and benefiting from PIDs in a coordinated way, in the way that publishers support and benefit from Crossref DOIs, support Crossref outreach, and accept Crossref guidance.”

Pragmatically, the best balance to the demands for better advocacy seems to be one of helping the community to coordinate its responses to specific advocacy challenges, without necessarily taking on the role of ‘representing’ the community:

“On an advocacy role: better to have a grassroots up approach so if a challenge emerges, the group can mobilise within the each organisation’s limits and generate a collective response that fits the bill, rather than having a full time advocacy role. The PID federation could help to facilitate those responses without being the thing to deliver them.”

4.2.3. Establishing and promoting best practice for PID providers

There is a clear need for best practice and consistency across PID systems. The challenges of overlapping metadata schemata were mentioned by several respondents as a real threat to the utility of the network.

Interestingly, there is a very strong steer from the community that this ‘best practice’ should be as lightweight as possible. Only one respondent felt that the PID Federation should “Develop and maintain a general trust and certification framework in which policies which ensure quality and persistency of PIDs are ensured”.

In contrast, numerous respondents felt that a certification or regulatory role should be avoided:

[There is a risk] “That it gets co-opted to become a regulator, issue seals of approval, or get in the way of the great bottom-up way ‘the big three’ have become the success they are, by listening to their members and solving their real-world problem.”

“I guess the risks that were perceived by some... of a change into regulatory or control of some kind, because, you know, this is much more about driving a shared agenda for mutual benefit.”

“It should avoid becoming a certification system”

“Would I want to sign up to a PID Federation that’s going to tell me “here’s how you certify yourself as a PID Federation certified PID provider”? I mean, what the hell does that even mean? We have the International DOI Federation about what it means to be a DOI provider or registration agency. We have that already... So I don’t necessarily think we need another group to tell us to be certified or to audit us.”

“Partners that come together willingly to try and make a difference is quite different than something that’s enforced from above or legislated or regulated or formed.”



There was a clear consensus that, alongside coordination on messaging and responses to emerging opportunities and threats, the Federation could drive greater consistency and predictability by means other than certification:

“Standardisation for interoperability”

“Maybe a simple core set of values about PIDs (akin to FAIR but not FAIR) that we sign onto?”

Throughout the responses, there was an emphasis on metadata exchange, more consistent practices around essential matters such as resolution or registration/creation of new PIDs and so on. There is a need to make PIDs predictable.

This raises a question for the proposers of the PID Federation: Can we achieve at least part of the desired outcome of ‘certification’ by adopting consistent, ubiquitous interoperability as a common strategic goal? Several respondents supported the idea of collaborative or shared ‘road mapping’. Interoperability is an unambiguous good for PID providers as well as the wider research community, and would serve broadly shared goals such as openness or efficiency.

Such coordination would also go a long way to addressing a threat that was mentioned in a majority of responses: the competitive or fractious nature of the PID provider community, variously referred to as the ‘bonfire of the vanities’ by a PID user and observer of the community across many years, or as a ‘competitive in-thing’ by a PID provider.

“The biggest challenge to PID providers, as I see, is to create an environment of real cooperation among them in the way they can improve the coverage and quality of the collections of digital objects they want to assign PIDs to.”

4.2.4. Support in the development and delivery of sustainability plans to safeguard valuable PID services and associated data

This possible role was an unarguable priority for the community, appearing as the most commonly cited challenge facing the PID community (both as users and providers) today, although it was manifested in several ways across the body of responses.

First, there was a common call to simplify the landscape and avoid overlaps or too many ‘non-profitable organisations’ competing for the same resources or users.

The second component is related to the call for less competition and fragmentation, and to the associated call for more cooperation. Sustainability for the whole network requires the sharing of expertise and support for struggling organisations. The PID Federation could collate practical skills and leverage the community’s experience in building up sustainable organisations to deliver infrastructure:



“It is a challenge to try to figure out longer term sustainability and governance alone, unless you have some significant backing or in many cases, business acumen to understand how to do that from a data perspective.”

This could also be seen as creating a degree of knowledge redundancy:

“Preservation guarantees that if one node loses, you have another node and so on.”

The third component is support for sunsetting for organisations and systems, such that the metadata and relationships created within them, and the links they support between external systems, can persist after any given provider has ceased to operate.

It is clear that sustainability in its broadest sense is a very high priority, understandably, for those building dependencies on PIDs into their systems and policies. The PID Federation should take on a role in ensuring technical resilience, operational continuity, and the sharing and recording of the community’s hard-won knowledge and expertise in these areas.

4.2.5. Helping to unite the community to tackle common problems or challenges

Successes such as PIDapalooza were lauded in many responses as an example of how well the PID community can come together, and held up as something respondents wanted to see much more of. Better and more inclusive communication across key actors “needs to continue and extend in a ‘shared listening’ mode to their overlapping memberships” to understand common problems and challenges.

There were several challenges which appeared in multiple responses, arranged with the most frequently cited first:

- Sustainability/persistence
- Scalability
- Interoperability
- Communication
- Metadata ownership and updates across organisations

Sustainability, interoperability and communication have all been covered in sections above, but scalability and metadata ownership (by which respondents meant the ability to re-use, extend or add to metadata as it is enriched ‘downstream’ from PID registration, and to share this valuable additional metadata openly through the primary PID network) are also critical challenges, and opportunities, which call for common action.

What the stakeholders in this study want is for these challenges to be reviewed, prioritised and chosen in an open, transparent community-led way. They represent an opportunity to act not just as a self-interested group, but as a wider partnership.



4.3. Who should be involved?

Inclusion is portrayed as a significant shortcoming of the PID community. While events like PIDapalooza and groups like the RDA PID Interest Group are praised as open, welcoming, and friendly to all comers, there are systemic issues at work which exclude many potential participants. As one respondent noted, “it is also worth paying attention to who and what are not part of these developments”.

In community terms, the Federation “should not be a continuation of the... model where it’s just Europeans talking to Europeans.” Whether this statement is fair or not, the feeling behind it is evident throughout responses to this research:

“PIDapalooza (as a proxy for the PID community) is Eurocentric and publishing-centric. Could this be more expansive? Yes. Should it be more expansive? I think this is a fundamental notion to consider in light of the proposal to develop a federation. We already have a sort of echo chamber. The federation might just might make this echo chamber more official. This is not necessarily a bad thing, but it depends on the goals the federation is meant to achieve.”

“The community needs to work on having more representation from different parts of the world – South American, Africa, and Asia. The organizations are often North America and Europe-centric.”

Risk: “The federation not being representative of the needs of the community world-wide. For example, the needs in Africa around PIDs are often different than in North America.”

“PIDs are not as well known or commonly used in East Asia as in the Western World. Providers have taken enormous efforts to engage or promote the PIDs while scholars seem still not perceiving immediate benefits from PIDs. Global academic networks will not be complete without universal participations from all over the world.”

The dominant model for business and governance in the academic PID space is membership, which for respondents from Latin America, Africa, and the Asia-Pacific is cited as a major barrier to participation. There is a concern that the PID Federation, if not explicitly and fundamentally designed to be inclusive, could reinforce and increase existing barriers:

“Every time you create let's say an instance that is global but is in many ways centralized, you generate a problem... exclusion So, the Federation of PIDs is a good idea if you are inside. If you are not inside, you generate a second level of exclusion.”

The third mode of exclusion which appeared time and again in responses is the perceived dominance of a small group of PID providers. This group is variously described as a “clique” or “cartel”, primarily made up of Crossref, DataCite and ORCID:



“The Community needs to engage better with the different system types – it is still too dominated by DOI and ORCID, and in danger of being too dominated by commercial vested interests. In the Life sciences the more prevalent and dominant identifiers are still accession numbers.”

The Federation should support the widest possible range of existing PID providers in developing models of inclusion that bring truly global voices into the PID community. The PID Federation itself should be an exemplar of inclusion and ensure that underrepresented perspectives are brought to the fore.

The call to arms here was well expressed by one respondent:

“Make open science inclusive, by showing clearly how PIDs can provide a level playing field for research from anywhere in the world, and across disciplines / fields. Take an informed view of the barriers to making that happen, and motivate Federation members to address those barriers together with those facing the barriers.”

4.4 How could a PID Federation be supported?

All but one respondent who expressed an opinion on whether an existing organisation could take on the role of running the PID Federation said ‘no’. The one respondent who said there was suggested that RDA could fill the role.

Several other respondents thought RDA was a potential candidate, as was FORCE11. However, a caveat emerges from the history of PIDapalooza which was established as a PID-centric gathering partially in response to the fact that “PIDs were always the last item on the agenda or the room no one went in” at these larger meetings.

Other possible groups were the DONA Foundation (on the understanding that its role would have to evolve considerably to deliver this) or the Digital Preservation Coalition, if advocacy for PIDs was ‘seen as part of preservation’.

A role for a national library or similar body was seen as inherently problematic by all those who addressed this specific possibility in their responses, as it raises concerns about purview and focus being too narrow or nationally determined, or rendering engagement politically difficult for organisations coming from certain countries.

“If it can be seen as related to any specific group or country, I am afraid that other Federations will be created and in some years to come we will be discussing the creation of a super federation to coordinate the federations.”

Respondents cited ‘neutrality’ and ‘independence’ as absolutely critical to the success of the proposed PID Federation. This would suggest that an independent organisation tasked with coordinating and managing the Federation would be required, but the idea of a new entity to deliver this received strong pushback across the board.



“Another coordinating body in a complicated landscape does not mean things are automatically more coordinated.”

“There’s a lot of evidence, a lot of good research that shows why creating entities creates a longer term problem because they’re bad at killing themselves off.”

“Yet another group.... So many groups are on the open science bandwagon, it’s hard to keep track even just those involved with PIDs. FREYA, Crossref working groups, EOSC working group, RDA interest group, newer initiatives like RoR and RAiD, NISO ID standards etc...”

“The risk is that it will be seen or perceived as “just another organization” competing with the existing ones for money and resources or that the different agendas of the existing organization won’t be able to align. In order for it to work it really needs to be broadly representative of all the different stakeholders and it should be seen as “needed” by its community. It should also immediately have a clear scope and deliverable, in addition to having the recognized authority needed to achieve the goals and its deliverables.”

While there is real opposition to the establishment of a dedicated entity or body to operate the PID Federation, there is a no-less-real appetite for collaborations to tackle defined challenges. This suggests that a lightweight, lower cost option would be the creation of a small core of support (a secretariat providing administrative resource and a network of contact points, for example), drawn from a range of organizations. This ‘core’ could contribute coordination efforts to facilitate the selection and prioritisation of challenges to be addressed, supported by a wider governance arrangement tasked with assuring inclusion and contributing to specific tasks as appropriate.

Contributions could be made financially, or in kind, and membership and participation could be decoupled from the financial or organisational sustainability of the Federation. This approach would also avoid replicating the unintended barriers to participation created by existing membership models (see 4.3 above).

4.5. Observations

Respondents in this project offered a great deal of insight and advice. While every effort has been made to include as much of this as possible, it is not possible to fully capture the richness of their responses in this summary. It is worth mentioning several suggestions here which did not have a clear ‘fit’ elsewhere in this report.

The PID Federation should choose its targets wisely, and make sure that it is resourced appropriately to genuinely deliver visible change. One respondent put it well, saying “As with anything, it’s going to need to show tangible value early on, or support will drain away quickly.” If the PID Federation launches with too little support or resourcing, the damage could extend beyond that to the Federation itself:



Risk: “That it will be set up with too little buy in and therefore will have hardly any power to do or achieve anything, which would be a waste of resources and could end up being counterproductive as observers of this project would be disappointed and might become somewhat disillusioned regarding PIDs.”

However, a broadly supported scope, and remit to tackle common challenges could help to mitigate this risk, not least by prompting stakeholders across the spectrum to contribute to a potential solution. Notable suggestions from respondents on a ‘problem set’ that might galvanise the widest possible swathe of the research community included leveraging the PID network and its associated metadata and relationships to evidence research impact or streamlining publishing workflows. If the Federation were to bring the wider community together to deliver on such a goal as a collective, it would unquestionably showcase the power and value of collaboration in the network.

The closed ‘by invitation only’ nature of this exercise, while based on pragmatic necessity and by no means intended to be the end of this consultation, raised concerns amongst respondents. It clashes with the principles of openness and inclusion which the PID Federation needs to embody if it is to succeed. As one respondent stated: “I strongly advise you to put this to open consultation.” The next phase in this investigation should address the gaps in consultation already identified, and in so doing should distil the proposal for a PID Federation down to something suitable for a truly global, open consultation.

Finally, the name ‘the PID Federation’ was regarded as variously ‘old fashioned’, ‘heavy’, ‘confusing’, and one respondent exhorted the project to ‘consider a better name than Federation’. The PID community contains many talented communicators. Perhaps once the proposed scope and model are clearer, they could be prevailed upon to provide this initiative with a suitably inspiring brand.



5. Phase one recommendations

The idea of a PID Federation is generally, cautiously, welcome. Clarifying the limits of the organisation will be as critical to its success as articulating a clear purpose and scope are. Creating another entity 'just' to manage the Federation will not be a popular move, but neither will a hosting solution from a single existing body, or too small a group of bodies, unless there is strong, transparent, and inclusive governance.

Based on the initial investigation of the potential role and scope of the PID Federation, this study makes the following recommendations:

1. A PID Federation is generally welcome and its scope should be refined with the broader community
2. A PID Federation should drive adoption and support the sustainability of organisations within the PID network by articulating the value of the whole: PIDs are more than the sum of their parts
3. The federation should help to coordinate advocacy across the community without necessarily becoming a figurehead in its own right
4. The PID Federation should take on a role in ensuring technical resilience, operational continuity, and the sharing and recording of the community's hard-won knowledge and expertise in these areas
5. The PID Federation should facilitate better and more inclusive communication across key actors, including PID users, and not limited to organisational members, to identify and articulate common problems and challenges
6. The Federation should support the widest possible range of existing PID providers in developing models of inclusion that bring truly global voices into the PID community. The PID Federation itself should be an exemplar of inclusion and ensure that underrepresented perspectives are brought to the fore
7. The Federation should explore ways that it could actively encourage and underwrite, including potentially via financial support, contributions from the global south
8. The PID Federation should consider a light-weight structure with a small secretariat to support activities and a range of ways organisations can contribute, financially, or in-kind, to avoid replicating the unintended barriers to participation created by existing membership models
9. The PID Federation should choose its targets wisely, and make sure that it is resourced appropriately to genuinely deliver visible change
10. The refined scope of the proposed PID Federation should be put out to a global, open consultation

The scale of the challenge the PID community (users, providers, policy makers, and supporters) face is daunting. However, the scale of the potential benefits to the wider research world are no less significant. A vision to inspire and energise this endeavour was put forward by one of the respondents:



“[F]ind a way [to] really reduce the time researchers spend in writing things that are not research related. That isn’t just a matter of reusing the same content many times, it’s much more than that. It’s about bringing content and opportunities to researchers without the need for them to actively look for them; it’s creating an ecosystem that allows any stakeholders to be able to access and modify content from any system and see that change propagated everywhere; it’s PIDs communicating to one another constructively; it’s agreeing on standards and workflows across the institutional spectrum and geography.”



6. Findings of phase two

The focus groups all began with a short recap of the background to the project, an overview of the process to date, and a summary of the recommendations and findings of phase one. Discussions opened with completely open questions about the group's preliminary responses to the findings.

The groups were then guided through a series of structured discussions around key themes using Menti polls to gather contributions and to bring out topics for more detailed examination. As well as high level themes from the recommendations (such as inclusion or advocacy), the participants were asked what they thought of the main roles suggested by respondents in phase one, and their priority concerns about the practical setup of the Federation.

The sessions finished with a series of open questions which were intended to surface any other issues with the recommendations or priorities of phase one which might not have been covered by the guided topic discussions, and to assess the extent to which participants' thinking had shifted as a result of the focus groups themselves.

6.1. Opening responses to the phase one report

The sense of 'cautious optimism' about the concept of a PID Federation which was shown in interviews and questionnaires responses was borne out in these discussions, and a clearer set of thematic priorities for any future work on the PID Federation are set out in this report.

One issue that several participants raised was the ambiguous purpose and scope of the PID Federation.

"In my analysis of the report that came out, it doesn't seem to me that there is a clear, consistent understanding of what the problem is that's trying to be solved"

"People are maybe just supportive of the idea, but they don't necessarily see how the Federation would really be the best way to help each of the individual PID providers or any or all of them advance in their individual goals."

One of those raising these concerns was amongst the more sceptical members of the original research cohort, but others spoke to a certain looseness in the definition of the role or purpose of the PID Federation, and a sense that some of the respondents quoted in the phase one report seemed to be responding to their particular interpretation of the idea. This is perhaps a product of the nature of this exercise. The study is designed to elicit various possibilities for a scope for the PID Federation, so there is necessarily no single, fixed statement of purpose at this stage. It does, however, suggest that a more limited set of possibilities should be presented in any future wider consultation to generate a more focused and unambiguous discussion.



Resourcing for the PID Federation was an issue that was raised in all three focus groups. Participants picked up on a potential contradiction in the initial findings, between the stated desire for a lightweight, minimal overhead collaboration, and the recommendation that it be resourced in proportion to the impact or value desired from the initiative.

“Enough resources should be made available for the PID Federation to be more than just viable.”

“It’s hard to think of doing something with no overhead but also one that is resourced.”

“If not a new organisation - who? Scope and remit of the coordinating org.”

“Resourcing - no one is going to do it for free”

Depending on the goals adopted by the PID Federation, this contradiction can also seem more or less significant. For example:

“We are just curious about how the PID Federation to become inclusive if it’s going to be lightweight in terms of organizational structure?”

Whatever scope is chosen for the PID Federation, there should be a clear alignment between the scale and nature of its activities and the resources or inputs demanded from the community and supporting organisations.

“In the scope of this organization, and perhaps giving it quite a focused scope to start with, is the way forward because as soon as you try and, you know, broaden the scope and become all things to all people is where you’re going to have a hiccup. But if you can have a very defined scope or you know, KPI that you want to get out of this, once that’s achieved, then you can incrementally improve or increase what the organization might be able to do.”

6.2. Possible roles for the PID Federation

Three possible roles were presented to the participants, derived from the original proposal and from the commonest suggestions from the phase one research.

These were:

- 1. Advocacy and communication:** This relates to sustainability in that high levels of adoption and engagement are essential if PID systems are going to reach sufficient coverage and visible value to retain users, integrators, and members. Also, as coverage and utility grow the value of PIDs becomes more evident and this should be communicated. Finally, for policy makers and others to understand the potential (and limitations) of PID systems in easing reporting, aiding policy compliance and so on,



they need to have sufficient understanding of their nature and abilities to drive the integrations and uses that best meet their, and their grantees, needs.

2. **Improving interoperability:** Enhanced interoperability bolsters the sustainability of PID systems in that it implies better use of standards and data exchange than siloed, incompatible systems and therefore suggests a PID network which is both simpler and more robust with the ability to ‘roll forward’ as technologies evolve. It also optimises the visible value of PIDs to end users (think of the ORCID-DOI auto-update system, or leveraging links between authors, articles and associated grant IDs to streamline and automate output reporting) which provides a strong business case for integrators and users.

3. **Ensuring technical resilience:** Technical resilience relates to interoperability in that it implies a reliable, predictable service upon which to build connections, but it also implies that the service will scale and be able to support unforeseen spikes in usage and innovative use cases that emerge from a global, evolving community of users. These are pre-conditions for longer term ‘persistence’ of the identifier, metadata, and resolution systems, and are critical aspects of genuine long-term sustainability.

The participants were asked to rate these according to the priority they accorded to each. Across the three focus groups, there was some variation in these priorities. The Asia-Pacific group assigned a similar priority to each of the three options, whereas in the Americas group there was much greater diversity of opinion on which was the highest priority.

Overall, improving interoperability was ranked highest, with advocacy and communication a close second, although there was no significant difference between their scores:

Table 4: Priority scores of potential roles for the PID Federation

Potential role	Average priority score (out of 5)
Improving interoperability	4.2
Advocacy and communication	4.0
Ensuring technical resilience	3.5

Improving interoperability was not only seen as critical to delivering on the promise of PIDs, but beyond this the lack of consistent interoperability is seen as a threat to the value and sustainability of the wider network of PID services and provision:

“Yes, there’s the value of individual PIDs but greater value in joining them up to improve visibility of research.”



“More and more PID systems come out and we found that there's no communication and no interoperative availability between these systems. This is not good for the PID development.”

“I think that the problem with having a lot of different providers is that no one's really in charge of coordinating the integration”

Advocacy was described as “one of the most important aspects” and “central to the idea”, especially with reference to the advantages that improved interoperability could bring to communication about the benefits or core value proposition of PIDs:

“It's about communication and it's about the added value of an integrated PID ecosystem”.

One participant felt that advocacy was the only plausible role for the PID Federation:

“Among these challenges it really only seems feasible that a hypothetical Federation could tackle the second point—and this has essentially already been attempted through initiatives like the PIDForum and the PID Workshops.”

It was noted that advocacy internationally could be challenging as policy or practical concerns vary from country to country, or indeed across disciplines. The need to engage with the context in which PID services are used was also raised:

“We keep on emphasizing that message and different communities, different stakeholders start getting the message and how it's aligned to other bigger issues of access, open science and etc. etc.”

The idea of technical resilience was seen as less of an issue for the wider community, but something that should be seen as a minimum requirement for a PID provider. It was described by several participants as ‘a matter for the individual PID providers’. One member of the funding community pointed out that without advocacy and subsequent uptake of PID services, technical resilience was not a visible issue:

“For us funders, the issue is not necessarily technical resilience, we probably don't even see those problems, because we're not technically versed enough. For us, it's more I think a lot of funders are not aware of what is out there or how they could use this. You basically have to almost use it intensely before you realize the problems of technical resilience. And I think from the funders' perspective, we need to first take that first step in many cases.”

It was also noted that consistent and reliable interoperability implies resilient and reliable connections between systems, so the pursuit of the former required the latter. Certainly, improvements to interoperability require technical enhancements that should be implemented in a way that reinforces PID services. A focus on information reuse and



improvements which could provide tangible efficiencies or other benefits could also underpin advocacy and communication campaigns.

6.3. Overarching themes

Despite the emphasis on interoperability as a priority for the PID Federation, two other themes dominated the discussions. These were themes which also emerged strongly in phase one of the research: communication and inclusion.

6.3.1. Communication

Participants overwhelmingly described a landscape in which communication around PIDs, their associated metadata and services, and the value of the whole is fragmentary or ineffective:

“Agree PID Fed is a good idea. There needs to be better coordination and cooperation between PID providers.”

“Value and use needs to be clearly articulated.”

“First of all, make it clear what the value of the PID network is.”

“Every time I have a conversation with someone who is not an expert in this area, they just kind of want a place to go for answers like this.”

In keeping with the findings of earlier discussions around advocacy, the focus groups felt that the PID Federation could be a driving force in crafting a unified ‘value proposition’ for PIDs that transcended the brand or self-interest of any individual PID provider, system, or service but would underpin communications to current and potential users of any PID system:

“Value statement is key as well as getting the buy-in of all. Owners of various persistent identifiers are key in working together and aligning.”

“You drive adoption because you explain the value of the ecosystem of a federation of an ecosystem of more than one PID and the reason why certain institutions might start using it is because they see an extra value of being part of something bigger than just one identifier and so they will also push for much stronger integrations and exchange of information among those identifiers. So I think that what's missing and what the Federation can actually help with is designing this framework of value.”

Designing and articulating a ‘framework of value’ for PIDs seems to be a crucial activity for the PID Federation. As expressed in the focus groups, it could not just boost adoption, but be used to showcase the beneficial impacts of existing infrastructures and be helpful in driving ongoing, community-guided enhancements. The focus groups clearly saw communication as a dialogue that will bring tangible value to all.



6.3.2. Inclusion

If we accept the idea of the PID Federation as a venue for ongoing conversations about the nature and evolution of the PID network, then inclusion, which emerged as a major theme in the interviews and questionnaire responses, gains even greater importance. Participants overall reinforced the emphasis on inclusion in the phase one recommendations:

“I think findings show the complexity of the matter. It is very important the inclusion component of the research.”

“Issues raised of global inclusion are unsurprising, as PIDs are driven by North America/Europe.”

Across the three focus groups, however, there was a much more practical focus on how inclusion could be achieved, with many suggestions and a rich discussion of what it meant to be truly inclusive.

It was noted that an effect of the North American and European dominance in PIDs was to reinforce existing imbalances between regions. As one north African participant described it: when researchers from the global south and their works are not included in the ‘big registries’, they are less visible. The major PID registries in this space provide a lot of data to other services, so when researchers are excluded from the PID services they are also missing from the downstream reuse and analysis of that data. This phenomenon serves to reinforce other exclusions.

A question, then, for the PID Federation is ‘how can it build inclusion in from the beginning?’. As one participant noted “you get the usual suspects volunteering to participate because they have the people and the resources to do so.”. This observation was echoed in other discussions: “Need other voices lifted up - Not just a conversation about top PID providers.”

A potential mechanism for achieving a globally inclusive organisation is actually implied in the concept of a ‘federation’ itself. Several participants suggested “making alliances with national institutions that can work as its national branch” or “finding local champions, who share the enthusiasm, can spread the word and can provide access locally...”

An added benefit of this approach would be the flow of intelligence on local priorities or regional issues inward to PID providers. These regional issues may be policy priorities, but they may also be more fundamental to the ability to interact with PIDs, metadata creation, services, APIs and so on.

“Through better understanding the challenges faced e.g. on the African continent. PIDs exist because of the Internet and working digital. Lack of infrastructure and expensive Internet are unfortunately a reality on the continent.”

Low bandwidth versions of services, simpler ways to provide metadata or register PIDs, and working with local partners to meet the needs of communities who lack the physical infrastructures necessary to access digital infrastructures could be an effective way to



extend the PID community and to increase the adoption and coverage of PID services globally.

Additionally, a flow of education, documentation and training materials outward to the 'local champions' would bolster these initiatives and enhance the dialogue even further:

"...the aspect of training. It is increased numbers. And if there's not so many people who are trained on the use of PIDs, then the communities can't grow - it will just be the usual experts, the usual perspectives, with no real growth."

The recommendation that some kind of subsidised access to the PID Federation, and to PID services more generally, was warmly welcomed in the focus groups:

"Make everything free, based on open source, and with free implementation support"

"Making affiliation taxes as low as possible so small organizations can be included."

"Low cost. Governance representative of the membership."

(The importance of reducing the costs of participation for those who need it is explored further in section 6.4 below.)

The discussion around inclusion also served to re-emphasize other themes, such as communication in the call to "promote internationally, [to] ensure membership is global and everyone feels they have ownership with an opportunity to input and guide its direction." It also drew on other recommendations, including the structure and nature of the operation coordinating the PID Federation:

"We can be more inclusive with operation with other organizations in the community if we can just maintain a lightweight organization."

The operationalisation of these overarching themes merged into many of the other discussions in each focus group. While this discussion focused on global inclusion, specific stakeholders who should be included or engaged with were also discussed, and this is covered in section 6.5.2 below.

6.4. Desired features of the PID Federation

Respondents in phase one cited a number of factors which they saw as critical to the success of the PID Federation. Focus group participants were asked to rank four of these features: cost, inclusivity, independence, and flexibility.

Inclusivity was clearly the most desirable feature: 90% of participants ranked this as their first or second choice.



Independence and flexibility came joint second. That said, there was much more of a clear split with flexibility with four participants ranking it as their number one priority, and eight placing it third. There was more of an even spread for independence, with ranking clustered around second or third choice. It is worth noting that the desire for independence was not as clear cut in the focus groups as it was in the phase one responses. It may be that this was influenced by the strong emphasis on discussions around inclusion: any initiative that was seen to be in thrall to the self-interest of a small group of organisations or countries could not be said to be truly inclusive. It may be worth exploring this issue more in any future consultation on the PID Federation.

Cost was consistently the lowest ranked: more than 50% of participants ranked it lowest. It is worth noting however, that some put it as their top choice, and issues of costs sparked significant discussions:

“So much of Scholarly Communications today is based on being able to pay to have a voice and to take part. And otherwise you are just picking up the dregs of the people that are building the system, because they can afford to be paying it and building it in their image.”

For those who ranked cost highly, it was seen as a corollary to inclusion:

“It seems like if you don't think costs are important to think about, I don't see how you can be thinking about inclusivity as that is one of the major...mechanisms for exclusion... cost is a barrier to entry.”

While there have been commendable efforts in this space, such as Crossref's fee waiver programme, one participant noted that “sometimes it has to be cheaper than free”. Waiving fees may not go far enough for those who need support for implementation and usage. We should not forget that integration and adoption can incur costs too.

6.5. Thoughts for the PID Federation

The final section of the focus groups posed a series of open questions to the participants. These were intended to elicit both new input to the scoping study, and to assess how discussions might have shifted participants' thinking during the course of the focus group.

6.5.1. “What common challenges do you think a PID Federation should help to tackle first?”

Response to this question tended to reinforce the sense of consensus around previous topics: Interoperability and communication of the value proposition of PIDs were the most frequently named. Boosting the adoption of PIDs and increasing the use of standards amongst PID systems were also mentioned. Interestingly, these suggestions are related to the two most popular: promoting the value of PIDs should increase adoption levels, and improving the standardisation of PID technologies and schemata would aid interoperability.



The last common challenge which was raised was the need for “better communication amongst PID providers”. This sparked a discussion, and the subsequent debate reinforced the perception that there is an ‘echo chamber effect’ amongst a group of PID providers who talk amongst themselves regularly and do not always see who is missing from their conversations.

6.5.2. Which communities or groups or regions should we be consulting next?

The proposed communities, groups, or regions for consultation comprise a mix of sectors or stakeholder types, and specific entities. It is worth noting that several of the named entities were approached during phase one, and some of those for a variety of reasons were not able to respond to the questionnaire in time to be included in the analysis.

Table 4: Proposed consultees, ranked by number of times proposed

Possible consultee	# of times proposed
Funders/policy makers	4
Researchers	3
Europe PMC ¹⁹	2
Research Data Alliance ²⁰	2
Research institutions	2
Codata ²¹	1
DORA ²²	1
Elixir ²³	1
FORCE11 ²⁴	1
Identifiers.org ²⁵	1
Library associations	1
National information systems	1

¹⁹ <https://europepmc.org/>

²⁰ <https://www.rd-alliance.org/>

²¹ <https://codata.org/>

²² <https://sfdora.org/>

²³ <https://elixir-europe.org/>

²⁴ <https://www.force11.org/>

²⁵ <http://identifiers.org/>



Pre-print servers	1
Submission systems	1

The Asia-Pacific region was cited as an area in which there was a significant readiness and need for PIDs but a low level of awareness.

6.5.3. Are there issues that the PID Federation might face that have not been mentioned in our research?

These suggestions were not discussed in detail during the focus groups. They were collected via the Menti platform as part of the session wrap-up, and are included here as issues that future work on the idea should consider and address as appropriate to the scope and remit of the PID Federation.

- Conflicts of interest between PID providers (NB: this relates to the 'independence' of the PID Federation)
- Think about scope and definitions: Which PIDs? Which community?
- Disconnect between self-described PID community and the broader community of research practice who don't think in terms of PIDs
- Map PID Federation goals to what already exists
- Incentives for using PIDs
- Where a registered HQ might be?
- Getting researchers, for example, to use PIDs especially if they feel it gets in the way of doing their research

These issues hint at several important factors which could help or hinder the PID Federation. The choice of PIDs and domains for inclusion in the Federation will have a major bearing on the scale of the operation, and the degree of focus that it can achieve. Using existing organisations and community groups may reduce overhead and could speed up progress in implementation. Ultimately, some of these questions may be eliminated by the final scope of the PID Federation, but they provide useful pointers for areas of investigation and future consultation.



7. Reflections on the focus groups

At the end of the focus groups, participants were asked to share a closing thought. These revealed that the 'cautious optimism' with which the sessions began had consolidated: eight out of ten responses were positive comments about the idea of the PID Federation or the scoping process.

One was a question for the consultant, and the remaining response was a suggestion:

"KISS²⁶ - focus on one major area of benefit. Get it right, build confidence and buy-in, expand."

The interviews, questionnaires and focus groups have yielded a set of priorities for the PID Federation that are interwoven in complex ways. They are inclusion, communication, and interoperability. As noted above, these aspects all have a critical role to play in supporting the sustainability of the global PID network.

The level of engagement in, and enthusiasm for the conversations which have made up this study are striking. To maintain and increase that momentum, it is clear that the PID Federation needs to be built from the ground up to be inclusive and open.

If we are to bring partners and communities into the PID Federation, there must be clear and accessible communication about the value of the enterprise. The process of refining the PID Federation and prioritizing its activities and agenda must be shared and understood as widely as possible. The results of its work, and the value of the global network of PID registries, metadata, and services should be known and appreciated at every level of the global research community.

Inclusion also relies on communication. Without clear and open channels to share messages from the PID Federation, and to welcome responses and feedback from the community, it will not break down the silos and barriers to participation that have accreted in parts of the PID network.

The first task of the PID Federation should be to consult widely, and to create a community around the initiative that is inclusive by design, and dedicated to eliminating barriers to participation. This will require both start-up funding, and ways to provide longer term subsidy or support to those communities or organizations which need it. In a time of global crisis, such an investment in global fairness and systemic resilience is vital, and the obvious place to start is by approaching funders, supranational bodies, and philanthropic foundations for support.

Once the broadest possible participation in the PID Federation has been established, and assuming that the findings of this study are borne out in wider discussions, the 'one major area of benefit' that it would make sense to focus on is improving interoperability between

²⁶KISS = Keep It Simple, Stupid



PID systems. The PID Federation partnership and community should be leveraged to identify the areas of most urgent need or greatest potential benefit for this work.

This should not be a matter of reinforcing the position (or dominance) of a small ‘clique’ of established PID providers. Instead, it should be about bringing disciplinary and regional identifier systems or initiatives into the global network, making sure that they use the same standards as the established players wherever appropriate, and guaranteeing that information flows benefit the established, the new, the large, and the small. These benefits could reach far and wide, helping the research community to engage with society and to share the fruits of their labour.

If this can be achieved, we will have not only a stronger, more resilient technical network, but we will have the social infrastructure in place to communicate the value proposition for PIDs and to ensure that value can be delivered wherever it might be needed.



8. Conclusions

There is an appetite for more engagement with and across the PID community. Stakeholders who have not been able to participate in current community mechanisms (whether collaborations like PIDapalooza or PID-specific like working groups, memberships or governance) want to be heard, and to shape the future of the PID network. FREYA has clearly tapped into a real need.

The stakeholders who participated in this project felt that communication of the value of PIDs has been fragmentary and confusing, and that this has slowed the adoption of PID services. At the same time, access to PID services is held back by technical and financial barriers. As a result, despite some excellent initiatives from PID providers, a significant portion of the global research community is excluded from the membership and governance of these critical research information infrastructures. A vision emerged from the questionnaires and interviews of a PID Federation that would work to eliminate barriers to PID adoption, to articulate the value proposition of the whole PID network, and strive to make the network more integrated and interoperable.

The recommendations from phase one resonated with focus group attendees. The priorities which emerged from the focus group discussions are well aligned with the three conceptual pillars of the FREYA project: the PID Forum, Graph, and Commons.

The PID Forum helps with intra-community communication, and these findings indicate that the Forum needs to be integrated with a wider community, and complemented with better coordinated communication. Participants referred to confusion about the value proposition of PIDs in general, and a perception of unhelpful competitiveness or self interest from some PID providers. The PID Federation could help to demonstrate cohesion and cooperation and strengthen trust in these critical infrastructures.

The PID Graph has shown the value of joining PIDs up, and provided real evidence of the benefits of bringing together the metadata and relationships between entities that PID services have captured. Extending the lessons of the PID Graph with a renewed emphasis on interoperability within real-world workflows could provide evidence of the benefits of investing in and working with PIDs. At the same time, swathes of the world are not creating PIDs or metadata. Enhancing the coverage of PID systems by eliminating barriers to adoption will boost the richness of metadata in PID systems. This enhanced coverage will also serve to underline the shared value proposition of the PID Graph.

The PID Federation could therefore enhance the PID Graph by improving access to PID services (and therefore increasing their adoption and coverage) and by enhancing their interoperability (and therefore increasing the number of connections and metadata flows between PID services). The need for more consistent and reliable interoperability between PID systems and other infrastructures was repeatedly expressed by the participants in this study.



The PID Commons as envisioned by FREYA supports the sustainability and governance of the whole PID network. This project has shown that the Commons needs ‘commoners’, and the PID Federation has the potential to be the structure that brings ‘PID Commoners’ together. There is no sustainability without the trust, collaboration and support of the full breadth of the global PID community at every level. An inclusive PID Federation at the heart of the Commons will help to sustain the PID network and make sure that our capacity to persist these infrastructures evolves alongside the community’s needs for them.

Of course, this raises an interesting factor: the size of the Federation. An organisation with hundreds of members drawn from a wide range of stakeholders is a very different prospect from one with tens of members drawn from a PID provider community. PID providers are adept at managing large communities, sometimes with thousands of members, whilst running very lean organisations, so there is expertise in managing precisely this challenge available to the Federation.

The desire for the PID Federation to reduce complexity and overhead costs is a related concern. The work of the PID Federation should not add friction. Similarly, costs should be as low as possible, and be managed in a way that does not pass costs on to PID users and members. Adding cost burdens would actively undermine the sustainability of the network. This re-emphasises the need for the lightest-weight, leanest Federation possible.

An organisation managed by a small team handling administration and logistics and supporting the work of a wider volunteer network could provide a hub for the Federation. This approach would fit with the findings of the focus groups in particular. Some funding to support participation around the world would seem to be necessary alongside voluntary or in-kind contributions if the goal of inclusion is to be met. Examples of inclusive organisations in this space that were mentioned throughout the research include FORCE11 and RDA. Understanding the successes and limitations of their approach to inclusion would be a good place to begin the design process. The FREYA team would need to identify interested parties to take this work forward urgently if it is to have an impact before the project formally ends.

If a wider, open consultation on the PID Federation produces evidence of sufficient demand for this endeavour, one way forward could be to leverage that demand by seeking medium-term funding from a consortium of sources, including national and supranational research funders, international organisations, or philanthropies. A useful analogy might be the Research4Life²⁷ initiative for access to research literature. Such an initiative for critical PID infrastructures would bring access to PID services and metadata to the entire global research community, whilst simultaneously providing a pathway for the contributions of researchers and institutions to be recognised in analytics and ranking systems that currently favour wealthier nations or disciplines.

The PID Federation could be built upon the expertise and effort of PID providers, with the active participation of a genuinely inclusive community, supported by global funders over

²⁷ <https://www.research4life.org/>



meaningful periods of time, maintained by a lightweight 'secretariat' or central office, all held together with memoranda of understanding. It could transform access to PID systems and the flow of open research information across the world. Building on inclusion to enhance interoperability and maximise the value of the PID network seems like a worthy contribution to the sustainability and resilience of these critical research information infrastructures.



Appendices

Appendix 1: Interview questions.

Preamble will include brief introductions, icebreaking discussions, and checking respondents are comfortable with the interview arrangements.

- 1) Before we begin, do you have any questions about the project or this interview?
- 2) How familiar are you with the FREYA project and its work?
- 3) You have been asked to participate in this interview because [customised rationale here]. Are you comfortable with this?
- 4) What was your first thought on hearing about the idea for the 'PID Federation'? [Follow up if needed]
- 5) Do you think there is an organisation or group which already plays a similar role for the research- or scholarly communications-related community?
 - a) If 'YES': How do you think they are helping to build and sustain the PID community in particular? [Follow up if needed]
 - b) If 'NO': Why do you think this role has not been filled already? [Follow up if needed]
- 6) How do you see the global PID infrastructure developing or changing in the next few years?
- 7) What do you see as the biggest risks or threats to the global PID infrastructure? [Follow up if needed]
- 8) What do you think the PID community does well, and what could it do better?
- 9) If you could ask PID providers, funders, and/or other research infrastructure providers to come together to tackle any problem, what would that problem be?
- 10) What possible roles do you think the 'PID Federation' could usefully play in the global community? [Follow up: What do you think the 'PID Federation' would need to do first to succeed?]
- 11) What do you see as the biggest risks or obstacles facing any potential 'PID Federation'?
- 12) Who would you like to see participating in the 'PID Federation'?
- 13) What do you think would be a natural home for the 'PID Federation'? How do you think it could be supported practically and financially?



- 14) Do you have any other thoughts or comments on the proposal for a 'PID Federation' that you would like to share with us?
- 15) Have your thoughts or opinions about a potential 'PID Federation' changed at all during this discussion?
- 16) Would you be willing to participate in a virtual focus group to discuss this idea further with a group of your peers?

Appendix 2: Questionnaire form.

Thank you for agreeing to participate in this research project. We have identified a short list of important topics, upon which we would value your thoughts. We invite your responses to the following questions. Please feel free to make your answers as long (or as brief!) as you feel is necessary. The more input you give us, the better the PID Federation will be able to serve you.

- 1) What do you think about the idea of a 'PID Federation'?
- 2) Do you think that there is an existing organisation that could take on that role?
- 3) What do you regard as the biggest challenges facing the PID community (including both users and providers), now and in the future?
- 4) What do you think the PID community does well, and what could it do better?
- 5) If you could ask PID providers, funders, and/or other research infrastructure providers to come together to tackle any problem, what would that problem be?
- 6) What possible roles do you think the 'PID Federation' could usefully play in the global community?
- 7) What do you see as the biggest risks or obstacles facing any potential 'PID Federation'?
- 8) Do you have any other thoughts or comments on the proposal for a 'PID Federation' that you would like to share with us?
- 9) Would you be willing to participate in a virtual focus group to discuss this idea further with a group of your peers?

Yes/No (Please delete as appropriate)

Thank you again for your responses. We will be in touch to share our findings with you soon.