Scottish Journal of Open Research
Article
10.36399/dme15z31

Expanding the research graph underpinning the implementation of Open Science: research instruments and facilities

Pablo De-Castro [1*]

[1]University of Strathclyde

*Corresponding author: pablo.de-castro@strath.ac.uk

CITE THIS ARTICLE
De-Castro P (2025) Expanding the research graph underpinning the implementation of Open Science: research instruments and facilities. *Scottish Journal of Open Research*, 1: 1. https://doi.org/10.36399/dme15z31

SUBMITTED
27 February 2025

ACCEPTED
06 June 2025

PUBLISHED
07 November 2025

The Scottish Journal of Open Research is an Open Access journal published by the University of Glasgow.

Keywords

Research information management; research instruments and facilities; information collection workflows; persistent identifiers; technicians; open research; research graph; PID graph

Abstract

Research instruments and facilities constitute an area of ever-growing relevance for research-performing organisations, research funders and research information managers alike. However, initiatives to gather the data on this equipment at an institutional or a national level are often working in isolation. A round table on emerging information collection workflows for research instruments and facilities, held during the Autumn 2024 euroCRIS membership meeting at the INRAE in Paris last November, offered an opportunity for an international discussion on the matter. The panellists were representatives of various national and regional research information portals across Europe. This paper summarises certain areas of the discussion and examines the way this additional research entity would fit into the steadily expanding research graph underpinning research information collection and its structure. Emphasis is made on the still early steps for the persistent identification of instruments and facilities and on the side-benefits that the consolidation of this additional area of research information might represent for institutions in the area of technician recognition.

1. Introduction

Emerging areas of practice in research information management often show a highly fragmented stakeholder landscape, making the coordination of ongoing initiatives particularly complex. This has traditionally been the case in the domain of persistent identifiers (PIDs) (De-Castro et al 2023). PIDs come in many shapes. For some of them – such as person IDs like ORCIDs – the record ownership and its end-users are very accurately defined. This is far less clear when we move onto RORs (Research Organization Registry) for organisations or DOI-based (Digital Object Identifier) grant IDs. Furthermore, it's also remarkably difficult to bring all involved stakeholders around a single table for an all-encompassing discussion.

This is why the round table on research equipment and facilities organised at the Autumn 2024 euroCRIS membership meeting in Paris (euroCRIS 2024, figure 1) started by asking the audience – some 50 attendees, most of them professionals in the domain of research information management – how many among them had ever attended a Research Data Alliance (RDA) plenary meeting. This is because the most prominent initiative to date on identifying research equipment and facilities – the PIDINST working group (Research Data Alliance 2017) – has mainly convened at RDA events. It was not surprising that only six people in the room (including the panellists in the round table) raised their hands in response to this question.

Fig 1. Round table on national-level data collection workflows for research equipment and facilities at the euroCRIS SMM2024. Left to right: Joonas Nikkanen (CSC-Research.fi, Finland), Ils de Bal (EWI-FRIS, Flanders), Ognjen Orel (SRCE-CroRIS, Croatia), Jan E. Garshol (Sikt-NVA, Norway) and Balviar Notay (Jisc-Equipment.Data, United Kingdom). Picture credit: euroCRIS President Jan Dvořák

The progress of the works undertaken by this PIDINST WG (in collaboration with DataCite) in the area of defining standards for the description of research instruments and facilities is significant (RDA

PIDINST 2022). However, the composition of the working group has traditionally been biased towards researchers, particularly in the area of Geosciences[1]. This is not necessarily a problem when the discussion is about the early steps in the definition of a metadata schema to adequately capture these additional pieces of research information. After all, researchers are the ones who know the most about the research equipment and facilities they use. However, the absence in the PIDINST WG of research funders and research information managers creates challenges when trying to expand the initial work so it can reach a wider community of users.

The panellists in the euroCRIS round table were all representatives from various European initiatives aiming to collect information on research equipment and facilities *at a national (or regional) level*. This included the Jisc-led Equipment.Data project in the UK, alongside similar efforts in Finland, Flanders/Belgium, Norway and Croatia. Critically, in all the latter cases, the data collection on research equipment and facilities is taking place *within the framework of a national Current Research Information System or CRIS*[2] (Research.fi in Finland, FRIS in Flanders, NVA in Norway and CroRIS in Croatia). The integration of research instruments and facilities into the network of interlinked research entities that constitutes the data model[3] underpinning these CRIS platforms makes a big difference. This enables, for instance the coupling of these records for research instruments and facilities to other entities in the data model (often with their own PIDs), such as the researchers and organisations using the research equipment or the outputs (both datasets and publications) arising from the use of such equipment and facilities. This linkage is of great value to research funders, who have usually invested significant resources in these facilities and expect some evidence on their widespread usage.

This is the main reason why it's important to watch the developments around the collection of information on research equipment and facilities in the research information management domain. The groundwork conducted, for instance, to establish an appropriate metadata schema for the description of these entities is the cornerstone on which any further activity will build, but the link between the various areas of activity – for example instrument use versus data collection and exchange – is often hard to ensure.

Some presentations delivered at the euroCRIS meeting before the round table was held examined some of the concepts and workflows associated with this national-level data collection for research equipment. This was the case of the UK Equipment.Data project presentation by its manager, Balviar Notay from Jisc (Notay 2024), and particularly the explanation Joonas Nikkanen (CSC Espoo) delivered on the conceptual framework underpinning the data collection workflows for research equipment, facilities and services into the Research.fi national CRIS in Finland (Nikkanen 2024).

The discussion on research instruments and facilities was further enriched from the presence in the room of two additional initiatives from different geographies closely related to the topics under discussion. The PID-Network Germany project[4] is looking into all things PIDs in Germany – including PIDs for research equipment and facilities – while the Ohio Innovation Exchange (OIEx) portal[5]

[1] The initial membership of the PIDINST WG is shown on pages 12-13 at https://www.rd-alliance.org/wp-content/uploads/2024/10/rda-wg-pidinst-case-statement.pdf
[2] https://en.wikipedia.org/wiki/Current_research_information_system
[3] The data model underpinning Current Research Information Systems is usually the Common European Research Information Format (CERIF) maintained by euroCRIS, https://eurocris.org/services/main-features-cerif
[4] PID Network Germany, https://www.pid-network.de/en/
[5] Ohio Innovation Exchange (OIEx), https://www.ohioinnovationexchange.org/

collects research information from 10 different universities in the state of Ohio with the aim of facilitating collaborations between academia and industry.

2. Research instruments: what information to collect

Several criteria were raised by the panellists when asked what instruments and facilities their initiatives were aiming to collect. “Instruments and facilities in the country or region” was a frequent response. The geographic boundaries are unsurprisingly salient in these country-specific mapping efforts – even if when discussing joint use of facilities by international projects the relevance of the national borders often seems to fade somewhat.

Some economic criteria are also critical: it doesn’t make sense to add every tiny piece of equipment to the national database. This is particularly well defined for the Equipment.Data project in the UK, whose website states that "all new equipment purchased over £138,000 [is liable] to be registered on the equipment data national database. Institutions can also publish and share their research infrastructure asset records below the £138,000 threshold to support greater transparency and sharing of these resources" (Jisc 2024).

Several of the national CRIS representatives at the table mentioned legacy equipment databases that predated the consolidation of the national research information management platform[6]. This is because these expensive equipment and facilities are often funded by research funders external to the institutions that may host them, and it’s easier for these funders (who not only have covered the costs of the already existing equipment and facilities but are also constantly issuing calls to fund further research infrastructure) to provide the information on these objects. The workflows to allow research-performing organisations to provide their institutional infrastructure information to a central database are quickly consolidating but are also complex, and it’s difficult to ensure they are sufficiently comprehensive[7]. This is why a funder-maintained equipment database makes sense as a starting point.

3. How to collect the info on research equipment and facilities

The participants in the round table reported that it is typically the largest universities that are most proactive in providing their equipment and facilities information to central databases. This is because their resources, both technical (institutional CRIS) and human, allow them to devote a fraction of these to this purpose. National projects will provide the guidance on format and scope of the required research information to the data provider institutions. This is often part of a much wider set of guidelines on research information exchange when it’s a national or regional CRIS where this information is being centrally collected (see figure 2 below).

[6] The Croatian Šestar Information System for equipment funded by the Croatian Ministry of Science and Education has now been embedded into the CroRIS national CRIS where all new data is added nowadays, but the legacy system is still available as a read-only platform at https://sestar.irb.hr:8443/.

[7] In its programme for 2025, the Equipment Data Service in the UK states the intention to integrate the national database of research equipment and facilities with institutional PURE equipment modules in order to assign persistent identifiers (DOIs) to the equipment records held in the modules, https://research.jiscinvolve.org/wp/2025/01/28/equipment-data-service-development-update/

4.5.1 Overview of cfEquip and cfFacil elements in FRIS R4

For infrastructure 2 elements are used an equipment (cfEquip) and a facility (cfFacil):

- **Infrastructure**: global name for cfEquip and cfFacil.
- **CfEquip**: Instrument for scientific research, mostly off the shelf and located on one site. For example a bioreactor, telescope,....
- **CfFacil**: Virtual and/or distributed space for scientific research that has one or more equipments and/or e-resources. It has a service function and after an investment costs also has an operation cost to keep the facility running (maintenance, scientific personnel, ...) Note: in short a facility can contain equipments, but this is not necessary. Example: Flemish Super Computer (facility with e-resources), Elixir (facility without equipment), ...

4.5.1.1 cfEquip

	Id	Name	Type	FRISR4
1	cfEquipId	Equipment Identifier	Identifier (max 128 chars)	Yes
0-1	cfAcro	Acronym	string	Yes
0-1	cfURI	Uniform Resource Identifier	string	Yes
0-N	cfDescr	Description	Multi-lingual text field	Yes
0-N	cfKeyw	Keywords	Multi-lingual text field	Yes
0-N	cfName	Name	Multi-lingual text field	Yes
0-N	cfEquip_Class	Relationship with Classification		Yes
0-N	cfEquip_Fund	Relationship with Funding		No
0-N	cfOrgUnit_Equip	Relationship with Organisation Unit		No
0-N	cfPers_Equip	Relationship with Person		No
0-N	cfProj_Equip	Relationship with Project		No, please check 4.6.15
0-N	cfResPubl_Equip	Relationship with Result Publication		No, please check 4.7.29
0-N	cfEquip_Medium	Relationship with Medium		No

Fig 2. Section devoted to research instruments and facilities in the Flanders Research Information Space (FRIS) CERIF-based interoperability guidelines for data provider institutions. Source: https://www.ewi-vlaanderen.be/sites/default/files/integration_guide_fris_r4_version_2.12.pdf.

Panellists agreed it would be useful for these various international initiatives to have a communication channel to compare their data collection workflows and their effectiveness. The development of the appropriate interoperability mechanisms and research information exchange workflows to allow the information on research equipment to be directly exported to central databases from the most widely used institutional CRIS systems was also highlighted as a high-priority objective. This is a development that could moreover be shared across national-level initiatives.

Participants in the round table reported that all these central databases are allocating internal unique identifiers to their equipment and facility records (the first metadata element in figure 2 above, cfEquipId, is an example for such an internal identifier). This is seen as sufficient at this early stage, but several initiatives also reported their intention to explore the requirements to start issuing persistent identifiers for these entities. Again, this is an area where an exchange of best practices across initiatives may be very useful.

It's the PIDs for research instruments and facilities that will allow these objects to be referenced in datasets, in manuscripts for journal articles or in reports delivered to the research funders on their

usage (the use of very expensive equipment and facilities by bodies beyond the organisation that host them is something research funders are particularly interested in, with some emphasis on their use by industry). Likewise, the existence of PIDs will allow these cross-references to be surfaced so that – same as a personal profile for a researcher typically allows the user to check their affiliations, projects, publications and collaboration networks – the instrument/facility profile page in a national CRIS will show the persons and organisations working with it, the projects relying on the data that they are producing and the publications and datasets arising from its use.

4. What PID(s) to use to identify research equipment and facilities

Part of the landscape fragmentation alluded to at the start of this paper refers to the multiplicity of technical standards to persistently and uniquely identify specific objects. This is typically a severe issue at the early stages of the development of a PID that tends to gradually go away as a given solution consolidates. ORCID is universally seen as *the* person identification standard these days, but when it was launched, several national researcher identification systems coexisted with it and were eventually mapped to it (De-Castro et al 2023b). ROR is again seen as the default OrgID nowadays, but before it became mainstream Ringgold seemed to be an equally suitable alternative.

As persistent identification for research equipment and facilities is at an early stage, it's not surprising that several competing, perhaps complementary IDs are currently being used in parallel for the purpose. The use of DataCite DOIs is very widespread in Australia and DOIs for instruments are gradually expanding to other countries like the United States and Germany. The recently launched PID-Monitor portal developed by the German Research Foundation (DFG)-funded PID Network Germany project is monitoring the DOI-based persistent identification of instruments (see figure 3 below). This project has also produced an animation showing the gradual adoption of DOIs for instruments and facilities across the world[8]. Given that most national initiatives at the round table expressed their intention to also use DOIs when they reach the PID-minting stage of their projects, it's fair to expect that the current relatively low uptake of DOIs for instruments will show a much more diverse geographical snapshot in the forthcoming months.

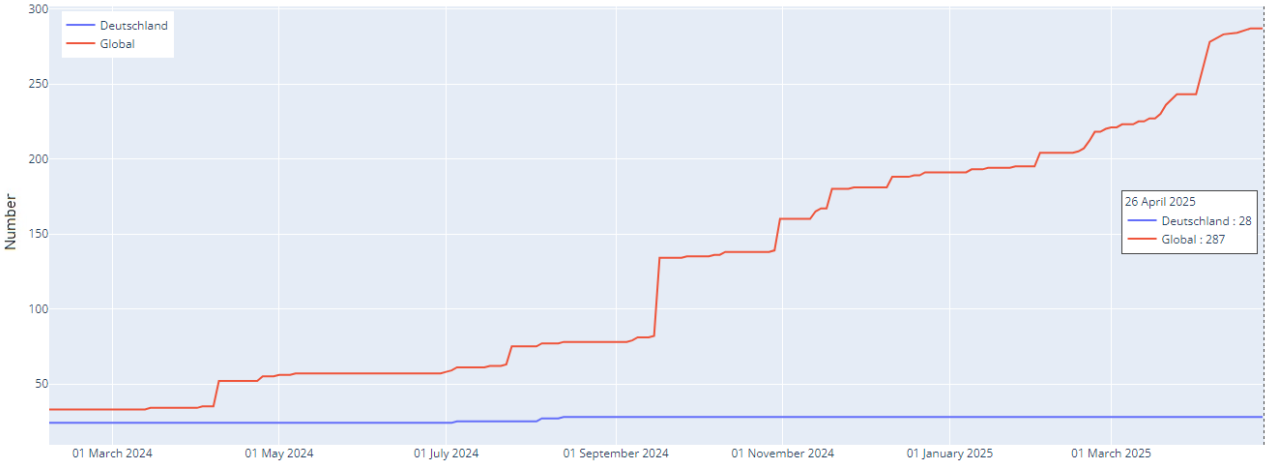

Fig 3. Number of DOI-based persistent identifiers for instruments and facilities in Germany and worldwide
Source: PID-Monitor, https://pid-monitor.org/Sparten/Instrumente/doi.html

[8] Development of the global registration of instruments at DataCite, https://pid-monitor.org/Sparten/Instrumente/worldwide.html

However, other solutions are being used for the persistent identification of instruments besides DOIs. The Research Resource Identifier (RRID) is a wider PID approach originally used in the United States to identify various kinds of objects, including instruments. This hints at a further geographical fragmentation on top of the stakeholder and technical fragmentation – it's very good news in this regard to see a section devoted to RRIDs in the latest white paper published by the PIDINST WG (PIDINST 2025).

RRIDs have a wider scope than just instruments and are, critically, already being referenced in journal articles. This uptake of RRIDs began for antibodies, cell lines or plasmids in articles in the biomedical/biosciences domains, where persistent identification was required, and grew bottom-up. This means that it's already possible to generate a partial research graph for a specific RRID-identified instrument that shows the people, the organisations and publications associated with it (see an example on figure 4).

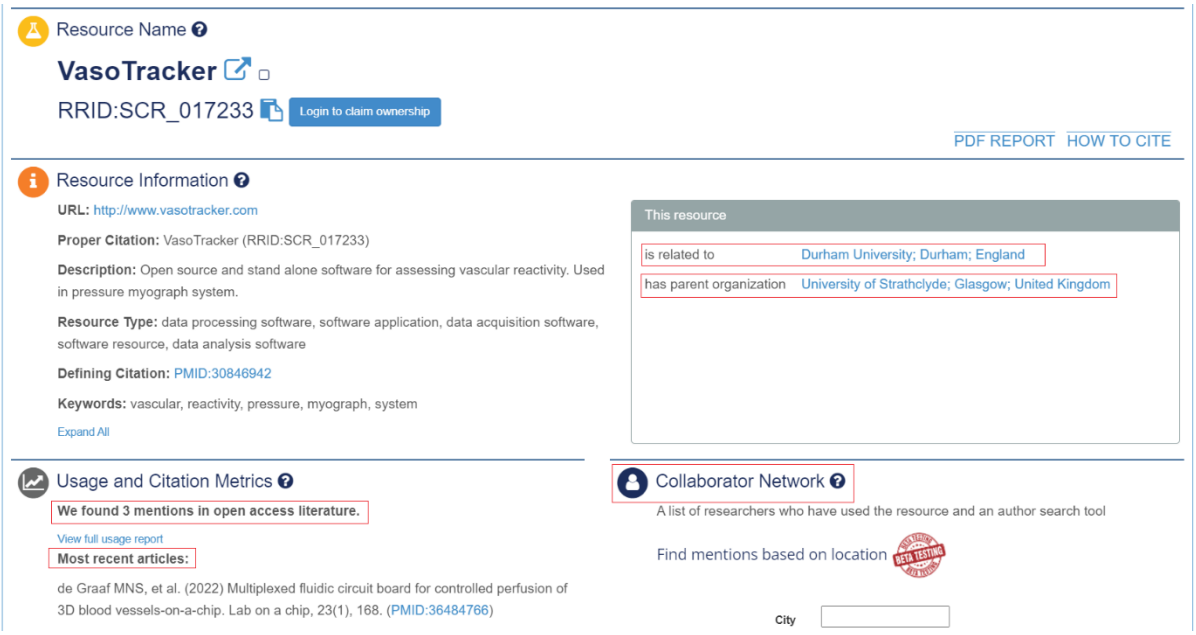

Fig 4. Example for University of Strathclyde research instrument persistently identified via a RRID. Links to other research entities like orgs, publications and persons (in beta) are highlighted in red colour
Source: https://rrid.site/data/record/nlx_144509-1/SCR_017233/resolver?q=vasotracker

The handle ID-based PIDs for instruments and facilities provided by the ePIC consortium[9] is yet another solution to provide these persistent identifiers. This standard is used by the B2INST instrument registration service offered by the EUDAT project (at https://b2inst.gwdg.de/), where external users registered with EUDAT may have handleID-based PIDs issued for their research instruments. At the time of writing (early May 2025), this EUDAT B2INST registry shows 853 instruments that have been persistently identified via this route[10].

Any attempt at a comprehensive monitoring of the uptake of PIDs for instruments should ideally try to cover at least these three sources of identifiers, but the multiplicity of sources makes this monitoring a challenging endeavour. It is good to see that the above-mentioned German PID Monitor

[9] ePIC consortium, https://www.pidconsortium.net/
[10] EUDAT B2INST instrument registry, https://b2inst.gwdg.de/records/

currently under development has included DOIs as the first category of instrument PIDs they aim to monitor, with additional ones probably awaiting their inclusion.

5. A possible knock-on effect: technicians and their areas of activity

In a recent parallel development to the gradual emergence of workflows for the collection of research information on research instruments and facilities, the technician commitment[11] was introduced in the UK in 2017 and currently has 120 signatory and supporter organisations. This initiative is all about making the work of this job family more visible and to identify mechanisms for a better recognition of their contribution to the research endeavour. Given that many of these technicians are often managers for research equipment and facilities, one good way to start increasing the visibility of their activity would be to highlight the instruments and facilities they are responsible for in their personal profiles in institutional CRIS systems and beyond.

This feature is already available in some cases (see figure 5 below), but it's far from being comprehensively implemented. Also, while it's possible to drill down on a specific research instrument by clicking on its entry under a personal profile, the subsequent links to other research staff using the equipment and to the publications and datasets resulting from its operation are usually not there (or not yet).

Given that the more standard research outputs (such as scholarly publications) recorded in the CRIS for this sort of technician profiles tend to be much lower than those for the average academic, the inclusion of research instruments and facilities in the wider research graph that a CRIS is able to display would be a good starting point to appropriately showcase the impact of a technician's work in the institutional research activity.

[11] Technician Commitment, https://www.techniciancommitment.org.uk/

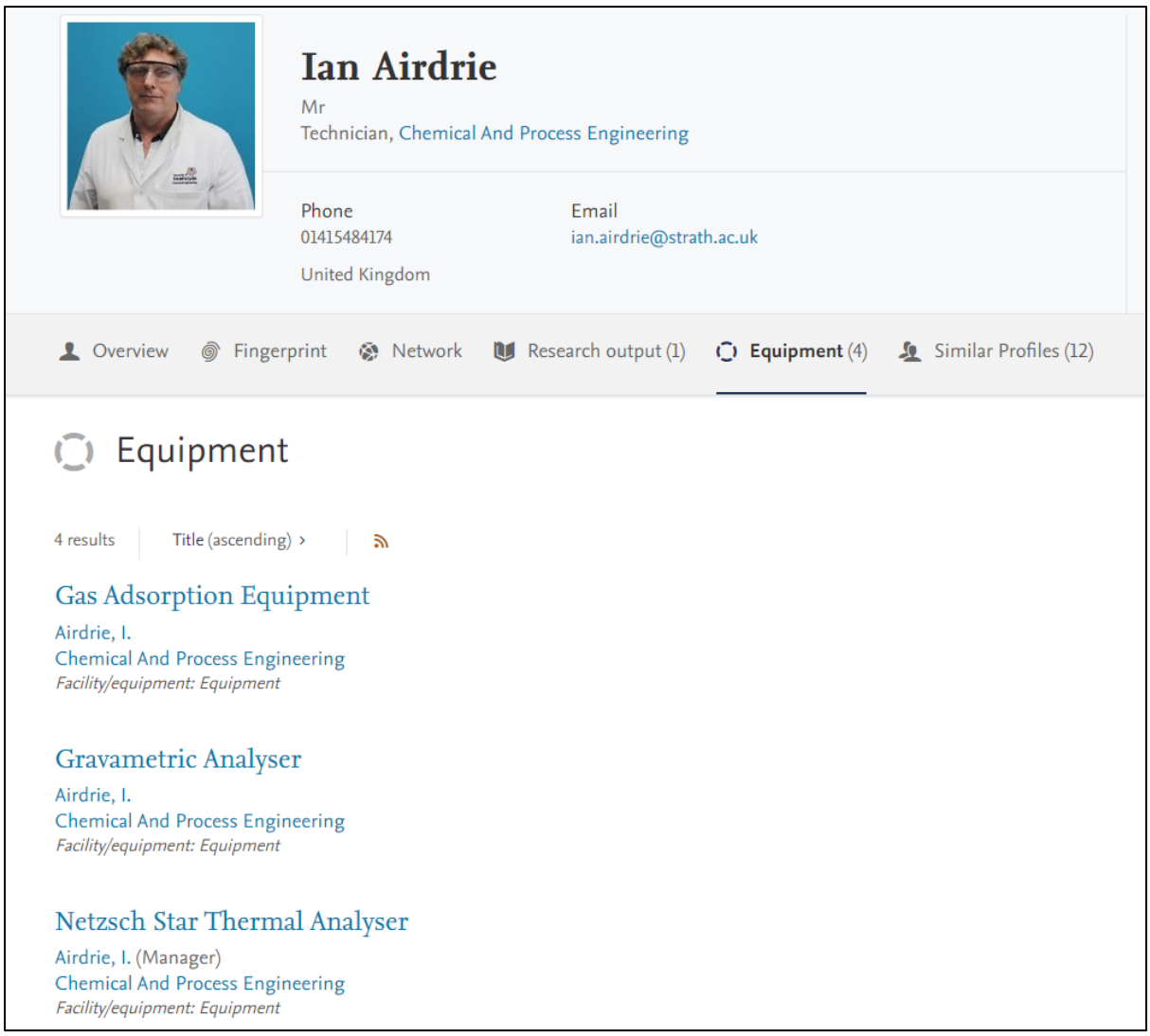

Fig 5. Personal profile for a technician on the University of Strathclyde institutional CRIS showing the research equipment he is responsible for.
Source: https://pureportal.strath.ac.uk/en/persons/ian-airdrie/equipments/

An eventual inclusion of a research equipment and facilities section in the ORCID personal profile would mean a big step forward in terms of making ORCID registration more attractive to technicians – with the subsequent increase in the visibility of their work. However, to avoid starting to build the house from the roof, it makes sense for this to wait until the PID infrastructure for these entities achieves a degree of consolidation it hasn't yet reached.

6. An opportunity for international collaboration?

Broadly speaking, there are currently three hubs for emerging PID infrastructure for research instruments and facilities. Australia is the first one, thanks to the sustained efforts undertaken by the Australian Research Data Commons (ARDC)[12]. The second one is Europe, where most of the Research Data Alliance (RDA) plenary meetings have taken place (but there are very active RDA chapters

[12] Initiatives around research instruments and facilities tend to be driven by the wide-scoped work on research data management, as instruments and facilities are typically where research data comes from.

outside Europe too). And the third one is the United States, where the National Science Foundation (NSF) has funded programmes like FAIROS[13] (Findable Accessible Interoperable Reusable Open Science) under which the "FAIR Facilities and Instruments: Enabling transparency, reproducibility, and equity through persistent identifiers" project is currently unfolding[14].

However, as revealed by the show of hands at the euroCRIS meeting mentioned at the beginning of this paper, there is little, if any, participation by initiatives dealing with the nationwide collection of research information on instruments and facilities in these discussions on how to best address the design and development of the standards that will drive their description. It's in this research information management area where an organisation like euroCRIS can make a difference by pursuing its mission to promote collaboration across initiatives and provide opportunities to showcase best practice case studies.

Some of these case studies might include the initiatives undertaken by several European University Alliances to explore the development of databases of shared research infrastructures, laboratories and services that can be utilised by their member institutions. See for instance the resource-sharing platform built by the UNITA Alliance within their Horizon2020 SWAFS Re-UNITA project at https://www.research.univ-unita.eu/_resource/Documents/Flyer%20shared%20infrastructures.pdf. While these are very practical collaborative initiatives not concerned with persistent identification or metadata schemas, some degree of harmonisation in the provision of research information must be ensured for a coherent cross-institutional database to arise. Furthermore, any progress around the persistent identification of research infrastructure and its inclusion in the wider research graph will also benefit these efforts by international groups of universities to provide databases of their shared equipment and facilities to their member institutions.

Author statement

On top of his day job as Open Access Advocacy Librarian at the University of Strathclyde in Glasgow, the author has been a member of the euroCRIS Board since 2013 and serves as euroCRIS Technical Secretary since 2018. In this capacity, he organised and chaired the Nov 2024 round table on data collection for research instruments and facilities reported in this paper. The discussion held at the euroCRIS meeting was informed by previous consultancy work done by the author on the topic of persistent identifiers, references to which can be found in the bibliography.

References

De-Castro P, Herb U, Rothfritz L, Schöpfel J (2023) Persistent identifiers for research instruments and facilities: an emerging PID domain in need of coordination. *Knowledge Exchange* case study. https://doi.org/10.5281/zenodo.7330372

De-Castro P, Herb U, Rothfritz L, Schöpfel J (2023) Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI. *Knowledge Exchange* case study. https://doi.org/10.5281/zenodo.7327505

[13] Findable Accessible Interoperable Reusable Open Science (FAIROS), https://new.nsf.gov/funding/opportunities/fairos-findable-accessible-interoperable-reusable-open-science
[14] https://ncar.github.io/FAIR-Facilities-Instruments/#project-description

euroCRIS (2024) Autumn 2024 Strategic Membership Meeting (Nov 26-28, 2024; INRAE, Paris). https://dspacecris.eurocris.org/cris/events/events02510

Jisc (2024) Equipment data. https://www.jisc.ac.uk/equipment-data (accessed 26 February 2025)

Nikkanen J (2024) Handling research infrastructures within Research.fi. Autumn 2024 euroCRIS Strategic Membership Meeting (26-28 Nov, INRAE, Paris). http://hdl.handle.net/11366/2653

Notay B (2024) Equipment Data: A national approach to optimising equipment data management and discovery in the UK. Autumn 2024 euroCRIS Strategic Membership Meeting (26-28 Nov, INRAE, Paris). http://hdl.handle.net/11366/2677

PIDINST (2025) PIDINST White paper: Using other PIDs: RRIDs. https://docs.pidinst.org/en/latest/white-paper/metadata-schema-recommendations.html#rrids (accessed 26 February 2025)

RDA PIDINST (2022) Metadata Schema for the Persistent Identification of Instruments. https://www.rd-alliance.org/wp-content/uploads/2022/01/pidinst-schema-1.0_Final.pdf (accessed 26 February 2025)

Research Data Alliance (2017) Persistent Identification of Instruments (PIDINST) WG Case Statement. https://www.rd-alliance.org/groups/persistent-identification-instruments-wg/work-statement/ (accessed 26 February 2025)