

TRAO ĐỔI KIẾN THỨC (KE)

KNOWLEDGE EXCHANGE

**Mã nhận diện thường trực cho các công cụ và tiện ích:
lĩnh vực PID cần thiết phối hợp đang nổi lên**

Tháng 2/2023

Dịch sang tiếng Việt: Lê Trung Nghĩa

Dịch xong: 16/06/2023

Bản gốc tiếng Anh: <https://zenodo.org/record/7330372>

**Persistent identifiers for research instruments and facilities:
an emerging PID domain in need of coordination**

February 2023

Trường hợp điển hình này là một phần của loạt trường hợp điển hình đã được sản xuất trong nghiên cứu về “Các rủi ro và lòng tin khi theo đuổi hạ tầng PID hoạt động tốt cho nghiên cứu” được Trao đổi Kiến thức - KE (Knowledge Exchange) ủy quyền vào tháng 7/2021. Kết quả chính của nghiên cứu này là một báo cáo xem xét bối cảnh PID hiện hành với sự nhấn mạnh vào các vấn đề có liên quan tới rủi ro và lòng tin của nó.

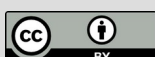
Loạt trường hợp điển hình bổ sung này nhằm cung cấp sự thấu hiểu sâu hơn trong các lĩnh vực hoạt động đặc thù, các tiến trình và các bên liên quan trong bối cảnh PID rộng lớn hơn.

Tiêu đề: Các mã nhận diện thường trực cho các công cụ và tiện ích nghiên cứu: lĩnh vực PID cần thiết phối hợp đang nổi lên

Các tác giả: Pablo de Castro, Ulrich Herb, Laura Rothfritz và Joachim Schöpfel (tư vấn khoa học scidecode).

Email: office@knowledge-exchange.info

DOI: 10.5281/zenodo.7330372



Tất cả các nội dung được xuất bản được chia sẻ theo giấy phép Creative Commons Ghi công (CC BY 4.0) creativecommons.org/licenses/by/4.0



All content published is shared under a Creative Commons Attribution licence (CC BY 4.0) creativecommons.org/licenses/by/4.0

Mục lục

1. Cơ sở lý luận	5
2. Trường hợp cho các PIDINST	6
2.1 Các cách tiếp cận kỹ thuật cho triển khai PIDINST	13
2.2 Vai trò của Hệ thống Thông tin Nghiên cứu Hiện hành (CRIS) để triển khai PID	18
3. Các vấn đề xung quanh các rủi ro và lòng tin liên quan tới triển khai các PIDINST ...	22
3.1 Các PID ‘kỹ thuật’ so với hướng quản trị viên	23
3.2 Sự tham gia của cộng đồng	23
3.3 Phân kỳ kỹ thuật	26
3.4 Quá tải các nhà cung cấp hạ tầng PID	27
4. Quyền tác giả	28
5. Tài liệu tham khảo	29

1. Cơ sở lý luận

Trường hợp điển hình này nhằm khai phá các thách thức phải đối mặt và các cơ hội do việc triển khai từng bước các PID đang nổi lên chào.

Trọng tâm chính của trường hợp điển hình này là các mã nhận diện thường trực cho các công cụ và tiện ích nghiên cứu (PIDINST từ bây giờ trở đi), nhưng phân tích này thực sự có mục đích đề cập tới bất kỳ hạ tầng PID đang nổi lên nào và vì thế có các đường liên kết tới các lĩnh vực PID khác như các mã nhận diện thường trực cho các hội nghị (ConfID) và - ở một mức độ nhất định - tới các PID được đề cập trong các trường hợp điển hình khác như IGSN cho các mẫu và ROR ID cho các mã nhận diện tổ chức.

Với Đăng ký Tổ chức Nghiên cứu - ROR (Research Organization Registry) đã được thiết lập và chạy, lĩnh vực sau của các OrgID rõ ràng tiên tiến hơn vào thời điểm viết bài này so với lớp PIDINST chỉ mới đang nổi lên, nhưng có những thách thức chung cho cả hai lĩnh vực để thành công. Trường hợp điển hình này xây dựng dựa trên nghiên cứu dành cho các OrgID.

Một khía cạnh chính về mặt này là thực tế rằng các trường hợp điển hình cho các OrgID đã được nhận diện từ lâu trong báo cáo “Rà soát lại các ID tổ chức được lựa chọn và phát triển các trường hợp điển hình cho Nhóm Làm việc về các Mã nhận diện Tổ chức của JISC CASRAI-UK” được phát hành vào năm 2015¹. Kể từ khi báo cáo này được phát hành, mục tiêu từng là để xác định cách tiếp cận phù hợp nhất để tạo ra một hạ tầng và cấu trúc điều hành có khả năng hỗ trợ ít nhất vài trường hợp điển hình như vậy. Mặt khác, phân tích trường hợp điển hình đó cho các PIDINST cũng được thực hiện, và điều này là thứ gì đó trường hợp điển hình này sẽ cố gắng giải quyết.

Hệt như trong trường hợp các OrgID và các PID đang nổi lên khác, rủi ro lớn nhất được thừa nhận hiện nay là sự phân mảnh và sau đó là thiếu hấp thu.

Trường hợp điển hình này vì thế sẽ tập trung vào việc xác định các sáng kiến khác nhau hiện có đang khai phá các cơ chế để triển khai các PIDINST và nhiều bên liên quan cùng đồng thời xem xét lĩnh vực này với ít sự phối hợp giữa họ. Sự nhấn mạnh sau đó sẽ nhằm vào nhu cầu đưa ra các cơ chế nâng cao nhận thức chung và có phối hợp cho cộng đồng để có khả năng cùng thúc đẩy trong lĩnh vực này.

2. Trường hợp cho các PIDINST

Như được nêu ở phần giới thiệu ở trên, có nhiều nỗ lực đồng thời diễn ra để khám phá cách thức tốt nhất để tạo ra lớp PIDINST, nhưng - trái ngược với cách Nhóm Làm việc về các Mã nhận diện Tổ chức (OrgID) của JISC CASRAI-UK giải quyết các nhu cầu được thừa nhận trong miền cụ thể này - không một khung rõ ràng nào liên các bên liên quan đã được cung cấp cho tới nay về các trường hợp điển hình mà các PIDINST cần phải giải quyết.

Nhiều bên liên quan đang làm việc về các yêu cầu quản lý cộng đồng và kỹ thuật xung quanh các PIDINST - các nhà nghiên cứu và các nhà cấp vốn là các bên nổi bật nhất - nhưng trường hợp cho các PIDINST còn chưa được đưa ra một cách rõ ràng. Điều này được thừa nhận như một rủi ro đáng kể cho việc hấp thu PID đang nổi lên này, vì thiếu nhận thức về các mục tiêu sáng kiến này nhằm tới để đạt được tác động rõ rệt tới thiện chí của các bên liên quan khác - các cơ sở và các nhà xuất bản - để tham gia vào nó.

Phân tích cơ bản về các trường hợp điển hình có liên quan cho các PIDINST (Persistent IDentifiers for research Instruments and facilities) được cố gắng nêu dưới đây, nhưng các gợi ý đó lý tưởng nên được thẩm định bởi một nhóm làm việc mà cũng có thể phục vụ như một diễn đàn phối hợp rất cần thiết xuyên khắp các bên liên quan. Trước tiên đáng để chỉ ra rằng các PIDINST đại diện cho nỗ lực phức tạp và tham vọng hơn nhiều so với các ID tác giả và tổ chức. Vài lý do cho sự phức tạp này được liệt kê bên dưới:

- Có thể triển khai ID tác giả, OrgID và các lớp ID đầu ra nghiên cứu mà không có sự can thiệp trực tiếp của các nhà nghiên cứu, nhưng điều này chỉ không khả thi cho PIDINST. Ngay cả nếu các tài sản đó có thể được thanh toán và được vận hành bởi các nhà cấp vốn và các cơ sở nghiên cứu, các nhà nghiên cứu là các bên liên quan thực sự có hiểu biết về khía cạnh các công cụ và tiện ích nghiên cứu. Đây là sự không trùng khớp mà Nhóm làm việc về PIDINST (PIDINST WG²) đang chạy trong Liên minh Dữ liệu Nghiên cứu - RDA (Research Data Alliance) là hoàn toàn do nhà nghiên cứu dẫn dắt với rất ít nếu có bất kỳ sự tham gia nào từ các tổ chức và các nhà cấp vốn nghiên cứuⁱ.

ⁱ The initial composition of the RDA PIDINST WG is available on pp. 12-13 in reference [3] below. It includes numerous representatives for data centres, supercomputing centres and other large-scale

- Các công cụ và tiện ích nghiên cứu có xu hướng sẽ được giải quyết chung trong thảo luận về các PIDINST, nhưng chúng là hai lĩnh vực khác nhau rõ ràng. Trong khi có thể là khả thi để đưa ra một cách tiếp cận thống nhất cho việc nhận diện thường trực của chúng, các phân tích sớm chỉ ra rằng các tiện ích là dễ dàng hơn nhiều để giải quyết so với các công cụ.
- Có lẽ lý do chính giải thích vì sao các công cụ nghiên cứu là khó hơn nhiều để giải quyết so với các tiện ích được tóm tắt trong hai câu hỏi đan xen nhau: “**Công cụ là gì?**” và “**Các công cụ nào là đủ thích hợp để được gắn một PID và các công cụ nào là không?**”
- Một trong những khác biệt rõ ràng nhất giữa các tiện ích và công cụ là tính đa dạng của bức tranh một sáng kiến PIDINST có thể cố gắng nhận diện duy nhất và thường trực. Trong khi các tiện ích nghiên cứu là các mẫu hạ tầng nghiên cứu khá ổn định, thì các công cụ mới đang được mua sắm và thiết lập trong hoạt động mỗi năm có thể cần phải được nhận diện lại một lần nữa. Nhiều công cụ khác đã ngừng hoạt động và điều này cũng phải được phản ánh trong hồ sơ PIDINST. Mức giám tuyển và duy trì các hồ sơ PIDINST cho các công cụ nghiên cứu sẽ yêu cầu và vì thế quan trọng hơn nhiều so với trường hợp các tiện ích.
- Tương tự như trong trường hợp của các OrgID, còn chưa được xác định rõ ai sẽ sở hữu các hồ sơ PIDINST vì các mục đích giám tuyển và quản lý. Các cơ sở có thể là các bên liên quan để đặt chỗ tốt nhất cho các công cụ, trong khi các nhà cấp vốn một cách tự nhiên có thể là lựa chọn mặc định cho việc chăm sóc các tiện ích, nhưng các tiến trình giám tuyển hồ sơ đơn giản là chưa sẵn sàng trong giai đoạn này.
- Trường hợp về tính hữu dụng của các PIDINST đối với các cộng đồng truyền thông học thuật rộng lớn hơn vào lúc này yếu hơn nhiều so với bất kỳ các PID nào khác đang được thảo luận. Thường có một mục tiêu chính rõ ràng cho từng lớp PID - đây là nhận diện duy nhất và sau đó là sự phân định rõ ràng của các nhà nghiên cứu đối với ID tác giả và nhận diện duy nhất liên kết cho các OrgID – nhưng trong trường hợp của PIDINST, điều này hiện chưa rõ ràng lắm.

research facilities but shows little presence of research funders, research-performing organisations or the wider scholarly communications community.

- Ngoài ra, các tiến trình của thông tin nghiên cứu nơi các mã nhận diện cho các công cụ và tiện ích nghiên cứu rõ ràng sẽ là một tài sản rất hữu ích cần có khác xa đáng kể so với các bên liên quan của truyền thông học thuật thông thường, nghĩa là các thư viện nghiên cứu. Điều này sẽ được chi tiết hóa trong phần về các trường hợp điển hình bên dưới, nhưng hiện tại hoạt động để thu thập thông tin về các công cụ và tiện ích nghiên cứu của các cơ sở thường nằm trong khuôn viên của các văn phòng nghiên cứu thay vì với các thư viện, trong khi các thảo luận về các tập hợp siêu dữ liệu có xu hướng thu hút các thư viện. Điều này đặt ra một thách thức phối hợp nội bộ bổ sung sẽ cần phải được giải quyết.

RDA PIDINST WG (Nhóm Làm việc về PIDINS của RDA) đã xuất bản một “tuyên bố trường hợp” vào tháng 12/2017 đưa ra trường hợp cho một “giải pháp do cộng đồng dẫn dắt để nhận diện rõ ràng và duy nhất trên toàn cầu các trường hợp công cụ đang hoạt động trong khoa học”³. Một loạt các trường hợp điển hình cho các PID cho các công cụ và tiện ích nghiên cứu đã được phác thảo vẫn là nỗ lực gần nhất sẵn sàng cho đến nay để trả lời cho câu hỏi “Vì sao chúng ta cần các PIDINST?”

Các trường hợp điển hình đó không có liên quan tới các bên liên quan nhất định, gồm:

- Các thước đo định lượng việc sử dụng các công cụ và cơ sở lý luận cho việc cấp vốn trong tương lai
- Liên kết dữ liệu tới các công cụ đã sinh ra chúng (nguồn gốc xuất xứ), cải thiện sự diễn dịch và thẩm định dữ liệu
- Trợ giúp lập kế hoạch cho sứ mệnh và hậu cần trang thiết bị
- Tạo thuận lợi cho tính tương hợp và chia sẻ dữ liệu mở, đặc biệt trong cải thiện các công nghệ thúc đẩy việc chia sẻ các công cụ
- Cải thiện khả năng phát hiện và tính trực quan của các công cụ và dữ liệu của chúng, được xuất bản trên web

Hai trường hợp điển hình mạnh nhất trong danh sách này là hai trường hợp đầu tiên, chúng chủ yếu phục vụ cho các nhu cầu của các nhà cấp vốn và các cơ sở nghiên cứu. Cả hai mong muốn thu thập dữ liệu sử dụng (bao gồm cho việc cộng tác giữa giới hàn lâm và giới công nghiệp) thường cho các tiện ích và công cụ nghiên cứu rất đắt giá và

các nỗ lực điều chỉnh phù hợp với Khoa học Mở để liên kết các tập hợp dữ liệu nghiên cứu với các công cụ cùng nguồn gốc xuất xứ của chúng có thể được coi là nằm trong lãnh địa của các nhà cấp vốn và các cơ sở.

Vài nhà cấp vốn ở các quốc gia của Trao đổi Kiến thức - KE (Knowledge Exchange) trên thực tế đang duy trì rồi các cơ sở dữ liệu các công cụ mức quốc gia sẵn sàng xuyên khắp các cơ sở ở quốc gia đó – xem, ví dụ, dữ liệu trang thiết bị của cơ sở dữ liệu được thúc đẩy bởi Hội đồng Nghiên cứu Khoa học Kỹ thuật và Vật lý (EPSRC) và JISC ở Vương quốc Anh với 17.538 hạng mục (các công cụ) từ 51 tổ chức của Vương quốc Anh (các trường đại học) vào thời điểm viết tài liệu này⁴. Các trường hợp điển hình chính cho các sáng kiến như vậy sẽ cải thiện tính trực quan và khả năng khám phá của các tài sản như vậy (hạng mục 5 trong danh sách liệt kê ở trên) và trên hết tất cả để thúc đẩy việc chia sẻ các trang thiết bị như vậy được tổ chức ở các cơ sở nhất định với các bên liên quan bên ngoài. Một loạt các bài đăng nhấn mạnh các mẫu hạ tầng nghiên cứu nhất định ở một loạt các trường đại học của Vương quốc Anh đã được xuất bản trên blog “Chia sẻ Trang thiết bị Nghiên cứu”⁵ của JISC cùng với cơ sở dữ liệu equipment.data.

FRIS: Metadata model for research infrastructure

► **Characteristics**
→ 25 metadata fields

[Links to other research objects](#)

- Identifier
- Federated identifier
- Name
- Acronym
- Description
- Keywords
- Type
- Location type
- Accessibility
- User modalities
- Starting date
- End date
- Location(s)
- Contact
- Website
- Technology classification (Fraunhofer-35)
- Research disciplines (FRDS)
- Data provider is consortiumcoordinator?
- Consortiumcoordinator
- Organisation(s) of consortiumpartners of infrastructure project
- Affiliations of consortiumpartners of the infrastructure project that provide data to FRIS
- [Link to funding project\(s\)](#)
- [Link to projects utilizing infrastructure](#)
- [Link to publications utilizing infrastructure](#)
- [Link to other infrastructure](#)

Slide từ bài trình bày “Các hạ tầng nghiên cứu: mô hình siêu dữ liệu & nắm bắt dữ liệu trong FRIS”⁶

Gần đây hơn, Bộ Kinh tế, Khoa học và Đổi mới (EWI) của Chính quyền Flemish đã bắt đầu sử dụng hệ thống Không gian Thông tin Nghiên cứu Flanders (FRIS) của họ để khai phá tính khả thi của việc thu thập thông tin về thiết bị và các tiện ích nghiên cứu có ở trong và ngoài các trường đại học Flemish bằng việc đề xuất một cách tiếp cận tập hợp siêu dữ liệu chung cho mô tả của chúng, xem slide từ trình bày gần đây ở trên.

Đáng lưu ý là dù yếu tố siêu dữ liệu đầu tiên trong tập hợp này là một mã nhận diện, không có lưu ý về ID thường trực, duy nhất mức quốc tế nào (điều tất nhiên không tồn tại ngay lúc này). Hai trường hợp đó trên thực tế nhấn mạnh rủi ro của bối cảnh phân mảnh qua các sáng kiến mức quốc gia. Mặt khác, như được thấy trong trường hợp điển hình đối với DAI của Hà Lan, các sáng kiến mức quốc gia có thể là cách tốt lành để bắt đầu phát triển trong khi một nỗ lực quốc tế được thiết lập tốt có khả năng để củng cố.

Cũng đáng nhớ trong đầu thực tế là các trường hợp điển hình về PIDINST đó không có “trường hợp điển hình nhà xuất bản” đặc biệt nào. Điều này ngụ ý sự khác biệt lớn với các trường hợp cho các ID tác giả và các OrgID được khám phá trong các trường hợp điển hình khác trong loạt này (và tất nhiên cho các Crossref DOI gốc ban đầu cho các xuất bản phẩm). Trong cả hai trường hợp như vậy, trường hợp điển hình nhà xuất bản từng là động lực rõ ràng nhất cho sự nổi lên của lớp quốc tế các PID - bất kể nó là vì phân định trong các xuất bản phẩm, hay cho việc nhận diện đúng các liên kết của họ. Còn tranh cãi rằng đó là các nhà xuất bản, được nhóm theo cách của Crossref, bên đã thúc đẩy nhiều nhất để triển khai lớp PID như vậy cho tới nay.

Các nhà xuất bản cũng được nhắc đến trong phân tích trường hợp điển hình mà PIDINST WG này đã tiến hành, nhưng chỉ nêu rằng các tham chiếu tới các công cụ và tiện ích nghiên cứu bao gồm trong các bản thảo được gửi cũng có thể được các nhà xuất bản xử lý dựa trên tính sẵn sàng rộng khắp của lớp PIDINST. Một yếu tố khác có thể trong thực tế được đưa vào trong các hệ thống đệ trình bản thảo để cho phép các tham chiếu đó sẽ được các nhà nghiên cứu cung cấp, nhưng điều này không phải là động lực đặc biệt vững chắc để thúc đẩy sự tham gia của nhà xuất bản trong một sáng kiến như vậy trừ phi có một yêu cầu rõ ràng cho điều này từ số đông các nhà cấp vốn nghiên cứu. DataCite tham gia rồi trong nỗ lực này xung quanh việc triển khai các PIDINST⁷ và có thể cung cấp vài đảm bảo như một lựa chọn thay thế bên tham gia do cộng đồng dẫn dắt, đặc biệt về sự liên quan tới trường hợp khẳng định nguồn gốc xuất

xử của tập hợp dữ liệu trong các DOI của DataCite cho các tập hợp dữ liệu. Vẫn còn phải xem liệu nỗ lực tạo ra lớp PID quốc tế có thể thành công mà không cần sự tham gia trực tiếp của các nhà xuất bản hay không.

Mặc dù các nhà xuất bản không hiện diện quá nhiều trong bức tranh tổng thể, nhưng là một bên liên quan cụ thể, có phần độc đáo được đề cập nhiều lần trong tài liệu của RDA PIDINST WG – nhà cung cấp cơ sở dữ liệu công cụ của cơ sở. Các nhà chế tạo công cụ, cũng được gợi ý là một tay chơi quan trọng tiềm tàng vì các mục đích nhận diện trang thiết bị là một bên liên quan đặc thù trường hợp khác mà công việc của PIDINST WG gợi ý làm việc cùng. Các bên liên quan này không có sự tương đương trong phân tích bối cảnh đối với các PID mới nổi khác. Điều này minh họa các mức phức tạp cao hơn đáng kể trong lĩnh vực các PIDINST. Các nhà cung cấp cơ sở dữ liệu công cụ của cơ sở đặc biệt là các sáng kiến công hoặc tư làm việc trực tiếp với các trường đại học để cho phép họ cung cấp một hình chụp nhanh các công cụ và tiện ích nghiên cứu có sẵn của họ. Chúng không chỉ phục vụ cho khả năng phát hiện và trường hợp điển hình về tính trực quan được cải thiện cho các cơ sở, mà chúng cũng thường bao gồm một hệ thống đặt trước yêu cầu thời gian nghiên cứu trong các tiện ích hoặc công cụ. Dự án Danh mục Bộ công cụ (Kit-Catalogue) do Đại học Loughborough dẫn dắt ở Vương quốc Anh⁸ và sau này được các cơ sở giáo dục đại học khác áp dụng ở quốc gia này là ví dụ tốt về một sáng kiến như vậy, nhưng còn có nhiều sáng kiến khác^{ii 9}.

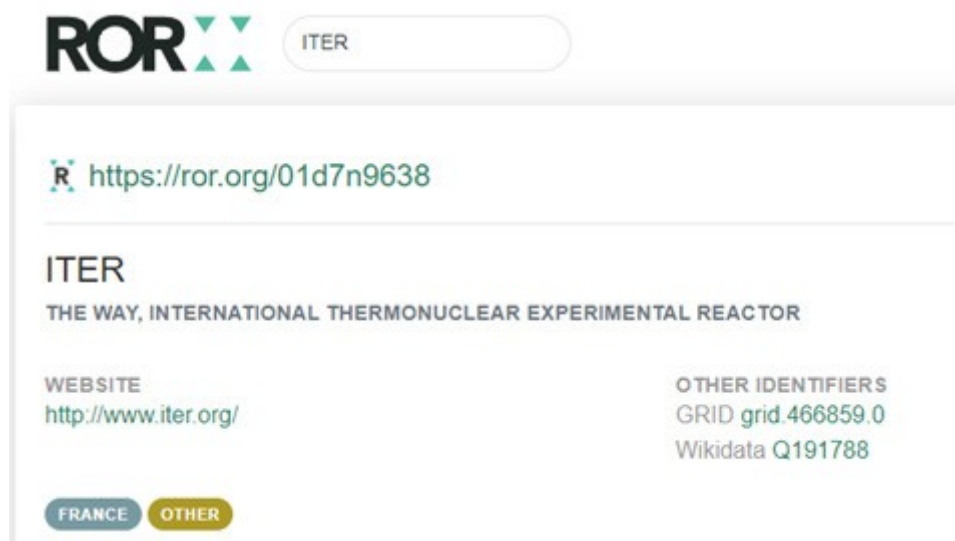
Một khía cạnh khác đặc thù trường hợp đáng chỉ ra trong bối cảnh PIDINST là vai trò chính các nhân viên kỹ thuật của các cơ sở trong hoạt động và mô tả các công cụ và tiện ích nghiên cứu. Các nhân viên kỹ thuật đó thường được thấy có trách nhiệm cho mẫu trang thiết bị cụ thể của cơ sở trong siêu dữ liệu có sẵn cho chúng. Trên thực tế, đây là các hồ sơ nghề nghiệp - thường bị bỏ qua vì mục đích đánh giá kết quả nghiên cứu của cơ sở nhưng ngày càng được đưa vào trong bức tranh của những nghiên cứu và báo cáo gần đây¹⁰ - những người cung cấp mô tả các tiện ích nghiên cứu của cơ sở đang được xuất bản trong các cửa hàng đặc thù ngành như *Tạp chí của các tiện ích nghiên*

ii Interestingly, the collection of institutional instrument database provider systems in [9] makes emphasis on the case for booking instrument/facility time at the appropriate equipment as a mechanism to promote equipment sharing and not so much on the (perhaps implicit) objective of increasing discoverability and visibility for the institutional equipment.

*cứu phạm vi rộng - JLSRF (Journal of large-scale research facilities)*¹¹. Trang chủ của tạp chí này nêu:

Tạp chí các tiện ích nghiên cứu phạm vi rộng (JLSRF) xuất bản các bài báo mô tả các trang thiết bị khoa học phạm vi rộng. Điều này bao trùm các trang thiết bị phạm vi rộng từ tất cả các ngành khoa học có ý định để các nhà khoa học sử dụng, những người không có liên kết tới việc vận hành các tiện ích đó của cơ sở (hoạt động của người sử dụng chuyên tâm). Các bài báo đó trên JLSRF cung cấp cho các nhà khoa học phương tiện đơn giản để tham chiếu các tiện ích phạm vi rộng trong các xuất bản phẩm của họ.

Tuyên bố tình huống được RDA PIDINST WG ban hành được trích dẫn trong thư mục trên thực tế gợi ý rằng các DOI được tạp chí phát hành cho các bài báo được nó xuất bản có thể được sử dụng như một dạng mã nhận diện thường trực để tham chiếu tới các tiện ích được mô tả. Cách tiếp cận này không phù hợp đầy đủ với cách thức các mã nhận diện thường trực thường trở tới một tập hợp siêu dữ liệu mô tả đối tượng đó (xem ví dụ về một OrgID bên dưới) và hé lộ giai đoạn đầu bị phân mảnh, do nhà nghiên cứu dẫn dắt mạnh mẽ, các nỗ lực từ dưới lên để xác định các tiện ích nghiên cứu.



Vài vấn đề được nhận diện trong phân tích công việc giai đoạn sớm được triển khai cho tới nay trong lĩnh vực các PIDINST được liệt kê bên dưới. Vài trong số đó sẽ được phân tích chi tiết hơn trong phần chuyên dành cho các rủi ro và lòng tin ở cuối của nghiên cứu này.

- Vài dạng tiện ích nghiên cứu và trang thiết bị và vài quốc gia được trình bày quá nhiều trong PIDINST WG, trong khi các bên liên quan khác như các nhà cấp vốn nghiên cứu và các trường đại học phần lớn bị bỏ quên khỏi các thảo luận. Điều này dẫn tới một phân tích hơi sai lệch về các trường hợp điển hình và lộ trình triển khai PIDINST.
- Một phần như là kết quả của sự thiếu vắng các trường đại học trong các thảo luận, các trường hợp hái quả ở chỗ thấp vì mục đích của việc nhận diện thường trực các trang thiết bị và tiện ích nghiên cứu (của cơ sở) được đề cập tới tràn trụi trong các báo cáo của WG (nhóm làm việc).
- Các bài báo nghiên cứu tóm tắt các kết quả công việc được nhóm làm việc này triển khai chỉ được xuất bản mới gần đây^{12, 13} và còn chưa có thời gian để gây ảnh hưởng tới những phát triển trong lĩnh vực này ngay cả nếu tham chiếu thứ cấp bao gồm rồi một lược đồ siêu dữ liệu được gợi ý để mô tả các công cụ và tiện ích.

2.1 Các cách tiếp cận kỹ thuật cho triển khai PIDINST

Bối cảnh PIDINST hiện hành chỉ ra một dải các cách tiếp cận hướng tới nhận diện thường trực các công cụ và tiện ích nghiên cứu, bao gồm các DataCite DOI, OrgID và vô số các mã nhận diện nội bộ ở các nơi đăng ký công cụ và các cơ sở dữ liệu hầu hết được nắm giữ và duy trì bởi các tổ chức thực hiện nghiên cứu nhưng cũng ngày một gia tăng bởi các nhà cấp vốn nghiên cứu ở các quốc gia khác nhau. Bức tranh phân mảnh này là đặc trưng của lĩnh vực PID đang nổi lên và không đại diện cho vấn đề đáng kể về triển khai một lớp PIDINST quốc tế một khi các thủ tục phù hợp và các tiến trình được xác định, mà nó nhấn mạnh sự cần thiết phải tìm ra các cơ chế phối hợp phù hợp xuyên khắp các bên liên quan khác nhau tham gia trong nỗ lực đó.

Bức tranh PIDINST bị phân mảnh được mô tả cho tới nay phần nào giống với nỗ lực triển khai ID tác giả cấp quốc gia, cụ thể ở Hà Lan trước khi các sáng kiến quốc tế như ISNI và ORCID xuất hiện trong lĩnh vực này. Trường hợp điển hình dành riêng cho DAI của Hà Lan trong loạt trường hợp điển hình này chỉ ra rằng sự thay thế và loại bỏ mã nhận diện giai đoạn sớm này bằng sự pha trộn của các ISNI và ORCID đã không đặt ra vấn đề chính nào trong đánh giá bối cảnh PID ở quốc gia đó.

Phần này sẽ phân tích và cung cấp các ví dụ về các cơ chế nhận diện khác nhau hiện đang được sử dụng để sinh ra các PIDINST. Một số đề xuất cũng sẽ được đưa ra để đạt được hiệu quả hơn cho việc triển khai PIDINST trên diện rộng hơn – có lẽ về mặt đạt được số lượng lớn tới hạn - triển khai PIDINST.

Các DOI là sự lựa chọn trực tiếp nhất cho việc nhận diện thường trực các công cụ và tiện ích nghiên cứu. Chúng không chỉ cung cấp một hạ tầng được thiết lập tốt, hoạt động tốt và lớp điều hành tốt, mà các DOI được DataCite tạo ra cũng phù hợp đầy đủ với trường hợp chính cho PIDINST, ấy là theo dõi nguồn gốc xuất xứ của các tập hợp dữ liệu nghiên cứu cũng được nhận diện bằng phương tiện của các DataCite DOI. Nhưng như chúng tôi đã thấy cho ví dụ của ITER-Cadarache ở trên, sự lựa chọn các DOI không phải là cách tiếp cận duy nhất, đặc biệt là đối với các cơ sở nghiên cứu quy mô lớn hoạt động như một thiết bị nghiên cứu độc lập bên ngoài các cơ sở.

Như một tổ chức thành viên phi lợi nhuận, do cộng đồng dẫn dắt, DataCite đang đóng vai trò nhà cung cấp hạ tầng PID cho các triển khai PIDINST giai đoạn sớm đó⁷. Điều này một phần là kết quả của các mạng lưới cộng tác được thiết lập trong quá trình của dự án FREYA-777523 được EU cấp vốn (“Kết nối các Mã nhận diện Mở để Khám phá, Truy cập và Sử dụng các Tài nguyên Nghiên cứu”, 2017-2020). Phần trên website dự án FREYA chuyên dành cho “các nguyên mẫu của các dịch vụ PID mới”¹⁴ bao gồm hai tiểu phần về các PID cho các công cụ khoa học và các PID cho các tiện ích nghiên cứu. Các đối tác của dự án PANGAEA và Hội đồng Tiện ích Khoa học và Công nghệ UKRI (STFC) được tham chiếu tới như là những người tiên phong trong triển khai các PID ở từng trong hai lĩnh vực.

Lớp ID thường trực dựa vào DOI cho công cụ nghiên cứu được nêu như là một ví dụ trong một bài đăng trên blog của DataCite ở [7] là cho một mạng địa chấn cụ thể - Mạng Shumagin-East Aleutian ở bán đảo Alaska, <https://doi.org/10.7914/SN/SH> – trong Liên đoàn Quốc tế các Mạng Địa chấn Số (FDSN), xem hình bên dưới. Trang đích cho DOI này chỉ ra bộ sưu tập siêu dữ liệu có cấu trúc lỏng lẻo bao gồm, trong số những điều khác, một mô tả mạng, vị trí địa lý của nó, một trích dẫn được gợi ý sẽ được bổ sung tới các bản thảo, cơ sở học thuật có trách nhiệm vận hành nó và một danh sách các trạm tạo nên mạng này. Hai yếu tố được nhấn mạnh trong hình bên dưới là bản

thân DOI và đường liên kết cho phép nhà vận hành mạng cập nhật thông tin trên trang ở những nơi thích hợp.

Trong khi điều này hoàn toàn là một ví dụ phức tạp về công cụ nghiên cứu, tính năng được FDSN xúc tác để cho phép vận hành cơ sở để giám tuyến và duy trì hồ sơ cho tiện ích đó là một ví dụ có giá trị về tiến trình giám tuyến cũng có thể áp dụng được cho các PID đang nổi lên khác.

Mặc dù quy trình này là không phức tạp và nhiều mạng còn chưa có một DOI, FDSN đã tạo ra các DOI cho các mạng địa chất trong liên đoàn như các DOI được liệt kê bên dưới. Lưu ý là tên các tác giả trong các trích dẫn là dành cho các nhà vận hành mạng đặc thù, bám theo thực hành được JLSRF thiết lập ở [11].

- Erin Pettit, Ted Scambos, & Martin Truffer (2015). RAPID: Việc quan sát sự phân kỳ của Scar Inlet Ice Shelf. Liên đoàn Quốc tế các Mạng Địa chấn Số. https://doi.org/10.7914/SN/1T_2015

- Heit, B., Yuan, X., Almendros, J., Abella, R., Carmona, E., Aguí, F., & Carrión, P. (2020). BRAVOSEIS Onshore Seismic Array. GFZ Data Services. <https://doi.org/10.14470/0Z7563857972>
- Alex Brisbourne, & Andy Smith (2016). Bed Access, Monitoring and Ice Sheet History (BEAMISH, 2016-2019) [Data set]. International Federation of Digital Seismograph Networks. https://doi.org/10.7914/SN/9B_2016

Như được nêu ở trên, PANGAEA từng là dự án của FREYA trong lĩnh vực các PID cho các công cụ và tiện ích nghiên cứu và vì thế là người tiên phong trong triển khai các PIDINST, xem, ví dụ, DOI cho tàu nghiên cứu Polarstern bắt nguồn từ một bài báo dành cho tiện ích nghiên cứu này được đăng trên JLSRF:

Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung (2017). Polar Research and Supply Vessel POLARSTERN Operated by the Alfred-Wegener-Institute. Journal of large-scale research facilities, 3, A119. <http://dx.doi.org/10.17815/jlsrf-3-163>

Ngoài ra, các tập hợp dữ liệu được PANGAEA quản lý đang sử dụng rồi tham chiếu tới cả tiện ích nghiên cứu và công cụ nghiên cứu từ đó dữ liệu đã có được, xem Witte, Hannelore (2018). *Processed 2 minutes-averaged continuous VM-ADCP (vessel-mounted Acoustic Doppler Current Profiler) profiles during Polarstern cruise PS82*. Alfred Wegener Institute, Helmholtz Centre for Polar and Marine Research, Bremerhaven, PANGAEA, <https://doi.org/10.1594/PANGAEA.887545>

Tập hợp siêu dữ liệu cho bộ dữ liệu PANGAEA được xác định liên tục này bao gồm các liên kết tới tàu nghiên cứu Polarstern được mô tả ở JLSRF, cho một chiến dịch đặc thù (PS82) qua một ID dựa vào DOI được TIB quản lý và cũng cho công cụ nghiên cứu đặc thù đã được sử dụng trên tàu, Acoustic Doppler Current Profiling (ADCP), TRDI Ocean Surveyor, 153.6 kHz, <https://doi.pangaea.de/10013/epic.47834.d001>. Đường liên kết này trao quyền truy cập tới mô tả công cụ dài 2 trang bởi nhà chế tạo công cụ nghiên cứu và phát triển Teledyne ở California.

Không phải ngẫu nhiên mà – giống như đối với PID mới nổi của IGSN được đề cập trong một nghiên cứu điển hình khác trong loạt bài này – cả hai ví dụ về việc triển khai PIDINST ở giai đoạn đầu đều thuộc lĩnh vực Khoa học Trái đất. Chính trong lĩnh vực này, với các kho dữ liệu lớn của nó¹⁵, nơi trường hợp cho việc liên kết các tập hợp dữ

liệu tới nguồn gốc xuất xứ của chúng được cảm thấy chính xác nhất vào lúc này. Cùng lúc, chúng là các ví dụ khá phức tạp về nhận diện thường trực cho các công cụ và tiện ích, và thật đáng để tự hỏi làm thế nào những thực hành ở giai đoạn sớm này có thể được tiếp quản bởi những trường hợp thực tế hơn như các trường đại học. Một gợi ý về khía cạnh này được thực hiện trong phần tiếp sau.

Một ngành nghiên cứu rất liên quan khác cho triển khai các PIDINST là ngành vật lý năng lượng cao, và đặc biệt cộng đồng liên châu Âu các tiện ích Neutron và Photon. Các mạng cộng tác được thiết lập tốt khắp các tiện ích và cơ sở này chính thức bắt đầu với dự án PaNdata ODI-283556 FP7 trong năm 2011 (“Dữ liệu Photon và Neutron - Hạ tầng Dữ liệu Mở”, 2011-2014). Mục tiêu chính của dự án là để thiết lập hạ tầng dữ liệu, lập thành liên đoàn giữa các tiện ích Neutron và Photon của châu Âu, để xúc tác cho các cộng đồng khoa học truy cập, phân tích và chia sẻ dữ liệu khoa học trong môi trường nghiên cứu cộng tác.

Xây dựng dựa vào mạng cộng tác này, liên đoàn các tiện ích này trong ngành vật lý năng lượng cao bây giờ đang quản lý hai dự án do H2020 cấp vốn dưới cái ô của Đám mây Khoa học Mở châu Âu - EOSC (European Open Science Cloud), ấy là PaNOSC-823852 (“Đám mây Khoa học Mở Photon và Neutron”, 2018-2022) và ExPaNDS-857641 (“Các Dịch vụ Dữ liệu Photon và Neutron của EOSC”, 2019-2023).

Như được chỉ ra trên hình bên dưới, lấy từ bài trình bày của dự án PaNOSC/ExPaNDS tháng 11/2020¹⁶, việc nhận diện thường trực các tiện ích và công cụ nghiên cứu được cộng đồng nghiên cứu này quan tâm rất nhiều theo khái niệm về “các hạ tầng nghiên cứu FAIR”. Mạng PaNOSC/ExPaNDS trên thực tế gần đây đã tổ chức các hội thảo đặc biệt về chủ đề các PID cho các tiện ích nghiên cứu¹⁷. Tương tự như trong trường hợp triển khai PIDINST giai đoạn sớm trong Khoa học Trái đất ở trên, mạng nghiên cứu HEP ‘ngách’ này được tách ra từ cộng đồng truyền thông học thuật dòng chủ lưu ở các trường đại học. Tuy nhiên, cách thức các nỗ lực đó đang được tiến hành gợi ý rằng EOSC và đặc biệt Đội đặc nhiệm Chính sách và Triển khai PID của nó có thể đóng vai trò đáng kể trong việc thúc đẩy các PIDINST qua các diễn giả khoa học khác nhau của EOSC trong nhiều ngành đã được triển khai theo dự án thí điểm EOSCpilot.

Các ví dụ về các PIDINST được khám phá cho tới nay tất cả đều dựa vào các DOI như là dạng mã nhận diện thường trực. Tuy nhiên, số lượng ngày một gia tăng các tiện ích

nghiên cứu cũng đang có được các OrgID mức đỉnh trong quy trình song song. ROR ID cho ITER-Cadarache đã được trình bày trước đó trong trường hợp điển hình này, và các tiện ích khác với các ROR ID có liên quan bao gồm:

- Nguồn Sáng Kim cương (Diamond Light Source) tại <https://ror.org/05etxs293> hoặc Đài quan sát SKA, <https://ror.org/01pwgd096>, ở Vương quốc Anh,
- German Electron Synchrotron (DESY) ở Đức, <https://ror.org/01js2sh04>,
- European Spallation Source ở Thụy Điển, <https://ror.org/01wv9cn34>,
- Trung tâm Siêu máy tính Barcelona (BSC) <https://ror.org/05sd8tv96> hoặc Kính viễn vọng Lớn Đảo Canary, <https://ror.org/03efc2j36> ở Tây Ban Nha.

Điều này có lẽ gợi ý các lựa chọn đa dạng cho việc nhận diện thường trực các công cụ và tiện ích nghiên cứu, đặc biệt nhớ trong đầu cách tiếp cận được mô tả trong trường hợp điển hình chuyên dành cho các OrgID nơi nhiều mức ROR ID có thể tiềm tàng được sử dụng để gắn các PID cho các công cụ được quản lý trong các tiện ích nghiên cứu phạm vi rộng.

FAIR Guidelines: Resource Identification

Persistent Identifier (PID) Services

- Purpose
- Scope
- Technology
- Governance
- Metadata
- Cost
- Uptake

What are the best choices for Facilities?

EOSC PID Policy

A survey of PID services is available in: FREYA project. D3.1 Survey of Current PID Services Landscape https://www.project-freya.eu/en/deliverables/freya_d3-1.pdf

These projects have received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 823652 and No. 957641

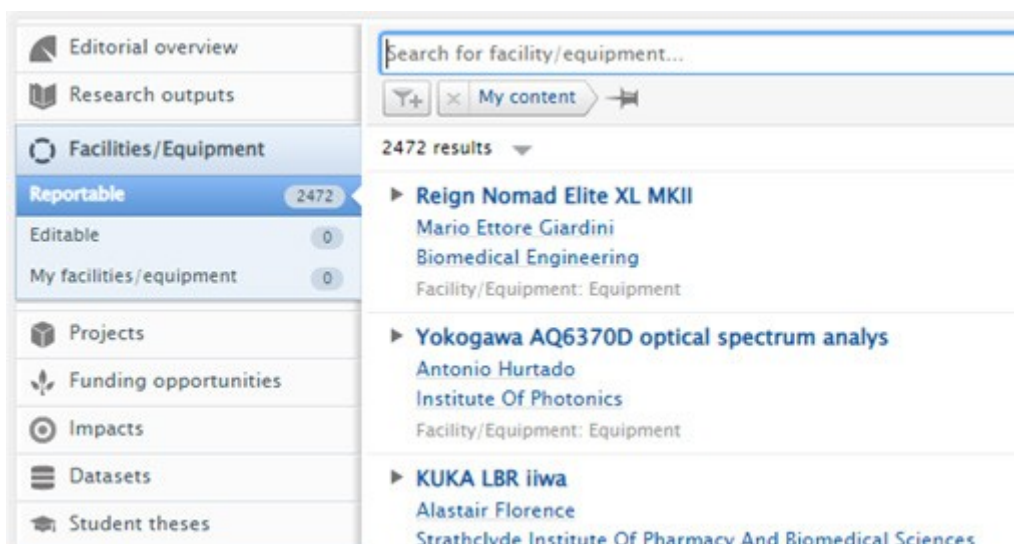
Publication	doi	PURL	Handle.Net Registry	
Data	ARK (Archival Resource Key)	Handle.Net Registry	PURL	doi
People	ID ORCID	RESEARCHERID	isni	
Organisation	ROR	GRID	isni	
Project	RAID			
Instrument	RDA	Persistent Identification of Instruments WG		
Sample	InChI TRUST	RRID	IGSN	
Software	doi	RDA	RDA/FORCE11 Software Source Code Identification WG	panosc

2.2 Vai trò của Hệ thống Thông tin Nghiên cứu Hiện hành (CRIS) để triển khai PID

Như được trình bày trong Thư mục các Hệ thống Thông tin Nghiên cứu - DRIS (Directory of Research Information Systems) được hiệp hội thành viên phi lợi nhuận euroCRIS¹⁸ duy trì, có hơn một ngàn Hệ thống Thông tin Nghiên cứu Hiện hành đang

hoạt động khắp trên thế giới chuyên nắm bắt tất cả các khía cạnh của hoạt động nghiên cứu được triển khai ở các tổ chức thực thi nghiên cứu, hầu hết là các trường đại học. Các hệ thống đó là rộng khắp ở hầu hết các quốc gia châu Âu và đặc biệt ở 6 quốc gia thành viên của Trao đổi Kiến thức - KE (Knowledge Exchange).

Các hệ thống CRIS, còn được biết tới là các Hệ thống Quản lý Thông tin Nghiên cứu - RIMS (Research Information Management Systems), thường vận hành trên cơ sở của tiêu chuẩn Định dạng Thông tin Nghiên cứu Chung châu Âu - CERIF (Common European Research Information Format) vì các mục đích mô tả siêu dữ liệu¹⁹. CERIF gồm một mạng các thực thể được kết nối lẫn nhau, với con người (các nhà nghiên cứu), các tổ chức và dự án trong hạt nhân của nó. Trang thiết bị nghiên cứu là thực thể mức 2 trong mô hình dữ liệu CERIF, và nó được liên kết với các thực thể cốt lõi và nhiều thực thể khác. Điều này về cơ bản là khái niệm tương tự như các đồ họa nghiên cứu để mô tả nghiên cứu và đặc biệt y hệt như các đồ họa PID sử dụng các mã nhận diện thường trực để thiết lập các kết nối lẫn nhau như vậy xuyên suốt các thực thể. Lược đồ Siêu dữ liệu [DataCite] được phát hành gần đây cho việc Nhận diện Thường trực các Công cụ¹³ trên thực tế được điều chỉnh phù hợp tốt với cấu trúc của thực thể CERIF cfEquipment.



Các hội nghị và các cuộc họp thành viên CRIS được euroCRIS tổ chức trong quá khứ đã thấy những đóng góp trong lĩnh vực quản lý thông tin nghiên cứu qua các mã nhận diện thường trực và đặc biệt dựa vào bộ sưu tập thông tin về các công cụ và tiện ích nghiên cứu, xem, ví dụ, [20] và [21]. Vì các sự kiện đó thường có sự tham dự của các lãnh đạo CRIS cả CRIS cơ sở lẫn quốc gia/khu vực, chúng có thể đại diện cho cơ hội tốt

để đưa ra trường hợp triển khai các PID đang nổi lên như các PIDINST, để thảo luận về các tiến trình có liên quan và trao đổi các thực hành tốt trong lĩnh vực này.

Ngoài ra, các hệ thống CRIS cơ sở rất thường xuyên có các hồ sơ cho trang thiết bị nghiên cứu của cơ sở, xem ví dụ trên hình bên dưới về danh sách các công cụ và tiện ích sẵn sàng ở CRIS cơ sở tại Đại học Strathclyde ở Glasgow, Vương quốc Anh. Mô hình dữ liệu CERIF chống trụ cho các CRIS cơ sở đó sẵn sàng trong cơ sở dữ liệu của cơ sở - ngay cả nếu trang thiết bị gốc xuất xứ còn chưa có các mã nhận diện thường trực có liên quan tới nó. Cấu trúc UUID nằm bên dưới các thực thể CERID đó cho phép rồi các đường liên kết đó được thiết lập.

Như là kết quả của điều này, các PIDINST cho trang thiết bị sẵn có hoàn toàn có thể nhanh chóng được làm cho sẵn sàng bằng việc sử dụng các dịch vụ của DataCite cho việc gắn DOI mà các cơ sở thành viên đang sử dụng rồi cho việc gắn các DOI cho các tập hợp dữ liệu. Tuy nhiên, đây sẽ là một bài tập phần lớn vô nghĩa nếu lớp PIDINST mới không được chống trụ bởi trường hợp sử dụng mặc định để sử dụng các mã nhận diện thường trực như vậy để tham chiếu chính xác các công cụ và tiện ích trong - ví dụ - các bản thảo được gửi để xuất bản và tập hợp dữ liệu chống trụ của chúng.

Nhu cầu về các dịch vụ này sẽ được phát triển trên đỉnh của một lớp PIDINST cuối cùng gợi ý cho các nhà cấp vốn nghiên cứu có thể là động lực chính cho việc áp dụng chúng. Bằng việc yêu cầu tính năng bổ sung này - y hệt như các ORID đã được yêu cầu lặp đi lặp lại trong quá khứ của sự nổi lên dần dần bức tranh PID - các nhà cấp vốn có khả năng thúc đẩy các nhà nghiên cứu trích dẫn các công cụ và tiện ích trong các xuất bản phẩm và tập hợp dữ liệu nghiên cứu của họ. Các nhà cấp vốn cũng có khả năng giành được các thước đo sử dụng đối với các công cụ và tiện ích nghiên cứu họ đang cấp vốn ở các cơ sở. Các nhà xuất bản cũng có thể khám phá các cơ chế được yêu cầu để đưa các tham chiếu tới các PIDINST vào trong các bản thảo và các bài báo trên tạp chí. Điều này ngụ ý sự thay đổi văn hóa đáng kể, văn hóa sẽ yêu cầu các nhà nghiên cứu áp dụng một cách tiếp cận hệ thống để trích dẫn các PIDINST, mà điều này có thể tiềm tàng được hỗ trợ bằng việc từng bước đưa việc duy trì và vận hành các công cụ và tiện ích nghiên cứu của cơ sở vào trong các khung đánh giá cho các trường đại học như Khung Trao đổi Kiến thức - KEF (Knowledge Exchange Framework) ở Vương quốc Anh.

The screenshot shows a research profile page for the Kelvin Hydrodynamics Laboratory. On the left, there is a navigation menu with 'OVERVIEW' selected, and sub-items 'Relations' and 'Display'. The main content area has an 'Overview' tab. The profile information includes the laboratory name, manager Sandy Day, and the department 'Naval Architecture, Ocean And Marine Engineering'. Below this, there are search filters showing '12 results', 'Display: Short format', and 'Sort by: Publication year'. The results are filtered for the year '2021' and list two publications: 'Experimental and theoretical study of the effect of hull roughness on ship resistance' and 'Experimental investigation on the effect of heterogeneous hull roughness on ship resistance', both published in March 2021.

OVERVIEW
Relations
Display

Overview

Kelvin Hydrodynamics Laboratory
Sandy Day (Manager)
Naval Architecture, Ocean And Marine Engineering
Managing organisational unit: Naval Architecture, Ocean And Marine Engineering

12 results Display: Short format Sort by: ↓ Publication year

2021

Experimental and theoretical study of the effect of hull roughness on ship resistance
Song, S., Dai, S., Demirel, Y. K., Atlar, M., Day, S. & Turan, O., 17 Mar 2021, In: *Journal of Ship Research*. 65, 1, p. 62-71 10 p.
Research output: Contribution to journal > Article > peer-review

Experimental investigation on the effect of heterogeneous hull roughness on ship resistance
Song, S., Ravenna, R., Dai, S., DeMarco Muscat-Fenech, C., Tani, G., Demirel, Y. K., Atlar, M., Day, S. & Incecik, A., 1 Mar 2021, In: *Ocean Engineering*. 223, 1, 20 p., 108590.
Research output: Contribution to journal > Article > peer-review

3. Các vấn đề xung quanh các rủi ro và lòng tin liên quan tới triển khai các PIDINST

Các rủi ro nghiêm trọng liên quan tới triển khai các PID cho các công cụ và tiện ích nghiên cứu được thừa nhận là sự phân mảnh và thiếu phối hợp khắp các sáng kiến làm việc trong lĩnh vực này, điều có thể tới lượt nó dẫn tới sự hấp thu không đồng đều. Đây là một đặc điểm chung cho tất cả các PID mới nổi lên, nhưng trong trường hợp của các PIDINST, bức tranh phân mảnh không chỉ liên quan tới các khác biệt đáng kể về phương pháp luận khắp các ngành nghiên cứu, mà còn thực tế là toàn bộ các bên liên quan trong cộng đồng truyền thông học thuật hiện không nhận thức được về sự tiến bộ gần đây trong lĩnh vực này. Để củng cố cho các dạng PID này - ngụ ý không chỉ các PIDINST mà còn các PID 'kỹ thuật' bổ sung - đặt chúng vững chắc vào sự chú ý của các tác nhân như các nhà cấp vốn và các cơ sở nghiên cứu được coi là sự bắt buộc (a must).

Triển khai các PIDINST trong giai đoạn rất sớm ở thời điểm tài liệu này được viết làm dấy lên số lượng các rủi ro tiềm tàng và các vấn đề có liên quan tới lòng tin. Vài trong số đó - như rủi ro của việc phân mảnh và thiếu hấp thu bởi cộng đồng rộng lớn hơn - là phổ biến cho bất kỳ PID đang nổi lên nào. Đặc biệt hơn, các PIDINST là ví dụ đầu tiên trong loạt trường hợp điển hình này đối với những gì chúng tôi gọi là một 'PID kỹ thuật', ngụ ý rằng việc triển khai của nó cho tới nay chủ yếu đang được các nhà nghiên cứu và các nhà quản lý tiện ích nghiên cứu dẫn dắt với ít sự tham gia từ các bên liên quan khác như các cơ sở hoặc các thư viện nghiên cứu (các ISGN cho các mẫu địa lý là một ví dụ khác cho 'các PID kỹ thuật' mà các trường hợp điển hình của nó phục vụ rõ ràng hơn nhiều các mục tiêu của cộng đồng truyền thông học thuật rộng lớn hơn. Chúng không phải là các chủng loại đan xen rõ ràng, nhưng biết rằng các rủi ro là phổ biến cho các PID được đưa vào trong bất kỳ nhóm nào như vậy, điều đáng xem xét trong phân loại này chi tiết hơn một chút.

3.1 Các PID ‘kỹ thuật’ so với hướng quản trị viên

Có sự phân đôi trong bức tranh PID hiện hành và đang nổi lên giữa những gì chúng tôi có thể gọi là các mã nhận diện “kỹ thuật” và “hướng quản trị viên”. Các PID kỹ thuật được thúc đẩy như là các tiến trình từ dưới lên bởi các nhà nghiên cứu thừa nhận nhu cầu về việc nhận diện thường trực của các đối tượng họ thường xuyên làm việc với chúng, bất kể nó là các mẫu địa lý (IGSN), trang thiết bị và tiện ích nghiên cứu (PIDINST) hay các thử nghiệm lâm sàng (ISRCTN). Các PID hướng quản trị viên, mặt khác, (như - trong số những điều khác - các ORCID, OrgID và ConfID đang nổi lên) được triển khai theo cách thức từ trên xuống hơn (nếu không nói là hoàn toàn) bởi một loạt các bên liên quan không bao gồm các nhà nghiên cứu thường là các cơ sở, các nhà xuất bản và các nhà cấp vốn nghiên cứu - để giới thiệu vài sự tiêu chuẩn hóa rất cần thiết trong bức tranh truyền thông học thuật vì mục đích quản lý thông tin nghiên cứu.

Các tiến trình để triển khai các PID hướng quản trị viên đó có xu hướng được phối hợp hơn nhiều xuyên khắp các quốc gia so với các nỗ lực thường bị phân mảnh để thúc đẩy các PID kỹ thuật. Chính sự phối hợp quốc tế này cho phép các nhà nghiên cứu rất cuộc cũng sử dụng và hưởng lợi từ sự áp dụng các PID hướng quản trị viên, nhưng vào lúc này không có lộ trình rõ ràng nào cho việc áp dụng các PID kỹ thuật bởi các bên liên quan trong lĩnh vực truyền thông học thuật. Vai trò quản lý cộng đồng mà các tác nhân như DataCite và Crossref có thể đóng ở khía cạnh này được coi là rất quan trọng. Tuy nhiên, vẫn có vài cách thức để đi về các khía cạnh tạo ra trường hợp về tính hữu dụng của các PID kỹ thuật sao cho các bên liên quan ngồi ở phía các PID hướng quản trị viên của bức tranh có thể bị/được thuyết phục để cũng thúc đẩy và hỗ trợ cho các PID kỹ thuật khác đó. Không có sự phối hợp như vậy giữa hai phía của bức tranh, rủi ro phân mảnh và thiếp hấp thu sẽ vẫn vô cùng lớn và có thể cản trở lớn sự nổi lên của lớp PID rộng khắp cho các PID kỹ thuật vì việc rất thiếu nhận thức.

3.2 Sự tham gia của cộng đồng

Việc triển khai mức quốc tế rất thành công của ORCID như là PID mặc định cho các nhà nghiên cứu đã biến nó thành một kế hoạch chi tiết cho các sáng kiến triển khai PID khác nơi theo. Điều này ngụ ý một cách tiếp cận tích hợp với việc nhận diện thường trực một dải đặc thù các đối tượng bởi nhiều tác nhân khác nhau tạo nên cộng đồng

truyền thông học thuật. Tuy nhiên, các PID khác nhau cũng có thể đi theo các chiến lược khác nhau để trở nên được áp dụng rộng rãi.

“Những điều tương tự về các tập hợp dữ liệu và các công cụ - đó là thông tin chúng tôi [các nhà cấp vốn] không nhất thiết nắm bắt như một phần của ứng dụng kỹ thuật số. Các chi tiết có thể được sản xuất như một phần của một trong những kết quả đầu ra và nơi chúng đang được lưu giữ. Thông tin đó có thể không nhất thiết quay lại chúng tôi, nó có thể được ký gửi với các tổ chức khác và sau đó chúng tôi [có thể sẽ] phải kết nối tới nó. Vì thế tôi nghĩ thậm chí có các rào cản kỹ thuật lớn hơn cho việc triển khai các PID nhỏ hơn đó. Chúng tôi sẽ không ở bất cứ đâu gần với việc trở thành chỗ lưu ký trung tâm cho tất cả các PID, ít nhất hiện nay. Tất nhiên, chúng tôi có thể không bao giờ là như vậy. Không có lộ trình chắc chắn nào cho điều đó. Và vì thế có thể một nhà cung cấp này nắm giữ tất cả các tập hợp dữ liệu đó, một nhà cung cấp khác nắm giữ tất cả trang thiết bị và công cụ, một nhà cung cấp khác nữa nắm các kết quả đầu ra chính sách, ví dụ thế, nếu điều đó trở thành một lĩnh vực bổ sung thêm. Và [như các nhà cấp vốn] chúng tôi phải tìm ra cách thức kết nối các thông tin đó cho mình, nhưng không phải bất kỳ ai cũng ở vào vị thế để có thể làm điều đó.”

Nói vậy, là rõ ràng ở thời điểm này rằng có rủi ro cao phân mảnh và thiếu hấp thu trong lĩnh vực các PIDINST. Có nhiều sáng kiến song hành đang được triển khai mà chông chéo với nhau với ít nhận thức về sự tiến bộ được các bên liên quan khác thực hiện, và có nhu cầu rõ ràng về các cơ chế để lấp đi khoảng trống giữa các cộng đồng nghiên cứu và các bên liên quan. Sáng kiến này của Trao đổi Kiến thức (KE) phân tích bức tranh PID hiện hành và khám phá các rủi ro và các vấn đề liên quan tới lòng tin của nó có thể giúp nhiều về khía cạnh này, nhưng các nỗ lực nhiều hơn nữa rõ ràng là cần thiết để đạt được sự áp dụng rộng khắp và sử dụng sau đó các PIDINST.

Một phần của công việc đang chờ xử lý này có thể được tiến hành bởi các tổ chức như DataCite, EOSC hoặc euroCRIS bằng việc tổ chức các sự kiện nơi hiện trạng của lĩnh vực PID được nêu có thể được khám phá bằng các trình bày từ các sáng kiến tiên phongⁱⁱⁱ.

iii An open webinar on persistent identification of instruments was held on May 19th, 2022 by the RDA PIDINST WG at https://www.rd-alliance.org/PID-instruments-May2022_webinar. This is very much along the lines of the dissemination initiatives suggested here.

Mục đích của các sự kiện đó có thể là để nâng cao nhận thức và lấp đi khoảng trống khắp các cộng đồng, thúc đẩy thảo luận rộng khắp nơi các tác nhân như các nhà cấp vốn và các cơ sở nghiên cứu có thể cân nhắc những lợi ích tiềm tàng của việc tham gia cùng với đoàn tàu PIDINST. Đáng để chỉ ra khía cạnh này và vài quốc gia có các cơ sở dữ liệu hạ tầng nghiên cứu giai đoạn phôi thai vì các mục đích quản trị có lẽ có thể được nâng cấp thành các tài sản được PIDINST hỗ trợ và biến thành cơ sở cho các thực hành được chấp nhận cho trích dẫn công cụ trong các tập hợp dữ liệu và các xuất bản phẩm.

Một khuyến nghị cụ thể về khía cạnh này có thể cho một nhà cấp vốn nghiên cứu ở quốc gia thành viên của KE nơi các sáng kiến áp dụng PIDINST giai đoạn sớm khá được củng cố (như Vương quốc Anh hoặc Đức) cân nhắc một lời kêu gọi cấp vốn cho dự án thu hút không chỉ các sáng kiến đi đầu tiên phong, đặc thù ngành, mà còn cả các tác nhân khác như các trường đại học, các nhà cung cấp cơ sở dữ liệu công cụ của cơ sở và cả các nhà chế tạo công cụ nữa. Mục đích chính của một dự án như vậy có thể là để tạo ra các mạng cộng tác mới khắp các bên liên quan và để khám phá các tiến trình để mở rộng các thực hành áp dụng PIDINST giai đoạn sớm trong các tác nhân dòng chủ lưu hơn như các trường đại học. Dạng hội đoàn dự án đa dạng này cũng sẽ cung cấp một mô tả toàn diện về các trường hợp điển hình khác nhau cho các PIDINST - một trường hợp hiện có phần lệch lạc do thiếu đầu vào trực tiếp từ các cơ sở giáo dục đại học và các nhà cấp vốn nghiên cứu.

“Chúng tôi cũng đã có các thảo luận khác nhau xung quanh các công cụ và chúng tôi hỗ trợ điều đó, ở mức rất cơ bản trong giai đoạn này. Vẫn còn điều gì đó phải cân nhắc. Nên tôi muốn nói, khuyến nghị của tôi, vì tôi không thể trao cho bạn một danh sách vết cạn, có thể phải đánh giá các trường hợp điển hình là cần thiết, hoặc các trường hợp điển hình nào bạn đang cố gắng giải quyết. Và sau đó làm một phân tích những gì có sẵn ở đó và sau đó phủ nó bằng lòng tin hoặc sự chín muồi của các dịch vụ đó và các cộng đồng có thể chào chúng. Và đó là những gì chúng tôi đang cố gắng làm khi đối tác và làm việc với các tổ chức khác nhau, để nâng cao lòng tin và độ chín của các dịch vụ đó và mở rộng phạm vi hiệu quả hơn.”

3.3 Phân kỳ kỹ thuật

Do bản chất khép kín các nỗ lực áp dụng PIDINST ở giai đoạn đầu, đặc thù ngành, nguy cơ về sự khác biệt kỹ thuật hiện đang xuất hiện trên toàn bộ lĩnh vực này. Vài ví dụ đã được cung cấp ở trên về cách làm thế nào các sáng kiến khác nhau đang sử dụng các tiêu chuẩn PID khác nhau như DataCite DOI, ROR DOI và các DOI cho các bài báo trong các tiện ích nghiên cứu được xuất bản trên các tạp chí nhất định như là cơ sở cho việc nhận diện thường trực các công cụ và tiện ích nghiên cứu. Một phần của vấn đề là do mục tiêu có lẽ quá tham vọng của việc giải quyết các tiện ích và công cụ như một lĩnh vực duy nhất, trong khi hai lĩnh vực này có thể được coi là các lĩnh vực khác nhau cần được giải quyết độc lập. Vấn đề này cũng đã nổi lên trong trường hợp điển hình về các OrgID, nơi các tiêu chuẩn cạnh tranh như ROR và Ringgold cùng tồn tại, và có thể không phải là vấn đề quan trọng miễn là có một chiến lược tổng thể để thu hẹp khoảng cách giữa các tiêu chuẩn kỹ thuật, nhưng các PIDINST thì hiện đang ở giai đoạn triển khai sớm hơn nhiều và rủi ro này được coi là nghiêm trọng hơn so với trường hợp của OrgID. Trên thực tế, đây là một vấn đề lớn hơn đối với các PID mới nổi lên cũng sẽ ảnh hưởng đến các grantID (mã nhận diện trợ cấp), một lĩnh vực nơi các cách tiếp cận giai đoạn sớm chủ yếu được Crossref hỗ trợ trong khi các RAID cũng đồng thời xuất hiện.

“Nỗ lực xung quanh các công cụ là nỗ lực thú vị, vì điều mà chúng tôi [các nhà xuất bản] tập trung rất nhiều vào là khả năng mở rộng phạm vi và tính nhất quán của các cách tiếp cận. Bây giờ, thiết bị đo đạc là rất, rất khác từ lĩnh vực này sang lĩnh vực khác và việc theo dõi chúng, tôi ngụ ý, chúng tôi xem xét trong thiên văn học, họ đã làm điều đó tốt không thể tin nổi, trong 20 năm có lẽ. Điều đó trao cho họ sự nhất quán lớn dữ liệu và sự hiểu biết. Nhưng sau đó nếu chúng tôi lật qua thành thứ gì đó như với khoa học hóa học, thì nó trở nên rất biến động trong các tiểu ngành hóa học. Và vì thế, nếu chúng tôi có thể có được nơi chốn đã có vài dạng mã nhận diện đã làm việc khắp các ngành, điều đó có thể rất có lợi, đặc biệt cho việc trừu tượng hóa, hiểu về tác động, xác định các dạng nghiên cứu và xu hướng đang diễn ra, v.v. Tôi nghĩ rằng khả năng điều đó xảy ra thấp hơn nhiều so với những điều như các cơ sở hoặc các đối tượng nghiên cứu, v.v., bởi vì khả năng thay đổi quá cao. Và, vì thế tôi nghĩ điều đó sẽ có lợi, nhưng nó khó xảy ra hơn.”

Theo một cách nào đó, những rủi ro về sự khác biệt kỹ thuật này là một phần của các vấn đề xung quanh sự tham gia của cộng đồng được đề cập trong phần trước và có thể được giải quyết như một phần của hoạt động nâng cao nhận thức chung hơn. Là quan trọng trong bất kỳ trường hợp nào phải nhớ chúng trong đầu để tránh phân mảnh.

3.4 Quá tải các nhà cung cấp hạ tầng PID

Các lĩnh vực PID đang nổi lên như OrgIDs, PIDINSTs, IGSNs, grantIDs và ConfIDs có khả năng được phát triển và chín muồi cùng một lúc. Thực tế là việc chống trụ cho hạ tầng PID và phối hợp cộng đồng cho tất cả các sáng kiến đó nằm lại với một tác nhân duy nhất trong cộng đồng - ấy là DataCite - cũng được coi là một rủi ro đáng kể. Điều này có thể là về sự quá tải cho các dịch vụ - đặc biệt trong lĩnh vực tham gia và quản lý cộng đồng - sẽ được một tổ chức duy nhất cung cấp. Rủi ro này có thể dẫn tới việc làm chậm lại sự tiến bộ trong triển khai các PID mới nổi lên đó và sự bất lực để tính đếm sự phân mảnh tự nhiên có xu hướng gia tăng từ các nỗ lực của giai đoạn sớm. Như được nêu ở trên, sự tham gia của các bên liên quan khác như EOSC trong các nỗ lực xác định lộ trình được tiêu chuẩn hóa, có phối hợp hướng tới lớp PID được triển khai rộng khắp có thể chống được rủi ro quá tải này. Các dự án được tiến hành dưới cái ô EOSC như FAIRsFAIR và được nói tới gần đây hơn FAIR-IMPACT và FAIRCORE4EOSC (“Phát triển một tập hợp các thành phần EOSC cốt lõi để xúc tác cho hệ sinh thái FAIR EOSC”, 2022-2025) được CSC ở Phần Lan dẫn dắt với DataCite trong số các đối tác của nó làm trong thực tế có nghĩa là một cơ hội để cùng giải quyết thách thức này của việc thúc đẩy vì sự triển khai và tăng cường cùng một lúc của nhiều PID khác nhau đang nổi lên.

4. Quyền tác giả

Trường hợp điển hình này chủ yếu được viết bởi Pablo de Castro (Đại học Strathclyde và euroCRIS, ORCID <https://orcid.org/0000-0001-6300-1033>) trong nhóm các nhà tư vấn bao gồm Ulrich Herb (Đại học Saarland, ORCID <https://orcid.org/0000-0002-3500-3119>), Laura Rothfritz (Đại học Humboldt Berlin, ORCID <https://orcid.org/0000-0001-7525-0635>) và Joachim Schöpfel (Đại học Lille và euroCRIS, ORCID <https://orcid.org/0000-0002-4000-807X>) dưới cái ô của tư vấn khoa học scidecode (ROR <https://ror.org/02c0bjd31>). Công việc được giám sát bởi Nhóm Nhiệm vụ & Hoàn thành của Trao đổi Kiến thức với thành phần của nó được liệt kê tại <https://www.knowledge-exchange.info/event/pids-risk-and-trust>.

5. Tài liệu tham khảo

1. Ferguson N, Moore R, Schmoller, S (2015) "Review of selected organisational IDs and development of use cases for the Jisc CASRAI-UK Organisational Identifiers Working Group". <https://repository.jisc.ac.uk/id/eprint/5853>
2. Research Data Alliance Persistent Identification of Instruments Working Group (PIDINST WG), <https://www.rd-alliance.org/groups/persistent-identification-instruments-wg>
3. RDA PIDINST WG (2017) Case Statement, https://www.rd-alliance.org/sites/default/files/case_statement/rda-wg-pidinst-case-statement.pdf
4. Jisc/EPSC database equipment.data, <https://equipment.data.ac.uk/status>
5. Ingram C (2018) Jisc "Sharing Research Equipment" blog, <https://equipment.jiscinvolve.org/wp/>
6. De Bal I, Poelmans H, Dengis P (2021) "Research infrastructures: metadata model & data capturing in FRIS". Autumn 2021 euroCRIS webinars, <http://hdl.handle.net/11366/1867>
7. Buys M, Dasler R, Fenner M (2020) "PIDs for instruments: a way forward". DataCite blog, <https://doi.org/10.5438/tdk2-2g94>
8. King M (2015) "Kit-Catalogue® - the open source equipment catalogue application". http://www.encyexchange.ac.uk/wp-content/uploads/KitCat_casestudy_ee_Feb15.pdf
9. Duca D (2017) "Five research equipment booking systems reviewed". Efficiency Exchange blog, <http://www.encyexchange.ac.uk/11219/managing-research-equipment-review-booking-systems/>
10. UKRI Research England (2021) "The Role of Technicians in KE: An explorative study", https://www.mitalent.ac.uk/write/MediaUploads/PDFs/TALENT_Technicians_and_Knowledge_Exchange_web.pdf
11. Journal of large-scale research facilities (JLSRF), <https://jlsrf.org/index.php/lfsf>
12. Stocker M et al (2020) "Persistent Identification of Instruments". Data Science Journal, 19(1): 18. DOI: <http://doi.org/10.5334/dsj-2020-018>

13. RDA PIDINST WG (2022) "Metadata Schema for the Persistent Identification of Instruments", <https://doi.org/10.15497/RDA00070>
14. FREYA project (2020) "Prototypes of new PID services", <https://www.project-freya.eu/en/pid-graph/prototypes-of-new-pids>
15. British Antarctic Survey UK Polar Data Centre (UK PDC) Discovery Metadata System, <https://data.bas.ac.uk/>
16. Matthews B (2020) "Achieving a Photon and Neutron community federated cloud in EOSC: Enabling our facilities to produce FAIR data". ExPaNDS-EGI Workshop on FAIR Data, <https://indico.egi.eu/event/5000/contributions/14511/attachments/13306/16179/ExPaNDS-EGI-workshop-FAIRData.pdf>
17. ExPaNDS project (2021) "Persistent Identifiers for Research Facilities Workshop", <https://www.panosc.eu/events/persistent-identifiers-for-research-facilities-workshop/>
18. euroCRIS Directory of Research Information Systems (DRIS), <https://dspacecris.eurocris.org/cris/explore/dris>
19. Joerg B. About Common European Research Information Format. European Commission Joinup, <https://joinup.ec.europa.eu/collection/eu-semantic-interoperability-catalogue/solution/common-european-research-information-format/about>
20. Cox A, Gutteridge C (2016) "equipment.data – Delivering a data autodiscovery infrastructure". CRIS2016: 13th International Conference on Current Research Information Systems (St Andrews, June 9-11, 2016), <http://hdl.handle.net/11366/500>
21. Tatum C, Brown J (2018) "Principles and Pragmatics of “as open as possible”: persistent identifiers as the interface between research information commons and closed systems". CRIS2018: 14th International Conference on Current Research Information Systems (Umeå, June 13-16, 2018), <http://hdl.handle.net/11366/658>

Knowledge Exchange

C/O Jisc

4 Portwall Lane,

Bristol, BS1 6NB

United Kingdom

T +44 203 697 5804

E office@knowledge-exchange.info