

Tài liệu xác định phạm vi của KE (Phiên bản 5.5)¹

Rủi ro và lòng tin khi theo đuổi hạ tầng mã nhận diện thường trực hoạt động tốt cho nghiên cứu

Dịch sang tiếng Việt: Lê Trung Nghĩa

Dịch xong: 27/05/2023

Bản gốc tiếng Anh: <https://doi.org/10.5281/zenodo.5018216>

KE Scoping Document (Version 5.5)

Risks and Trust in Pursuit of a Well-functioning Persistent Identifier Infrastructure for Research

¹ Tài liệu này có ý định:

- Phục vụ để xây dựng sự đồng thuận nội bộ KE cho lý do tồn tại của hoạt động.
- Để phác thảo và cấu trúc vấn đề, biện minh cho hoạt động.
- Để xác định phạm vi hoạt động được giao cho Chuyên gia tư vấn.

Mục lục

1 Tóm tắt và các từ khóa	4
1.1 Tóm tắt	4
1.2 Các từ khóa	4
2 Giới thiệu	4
2.1 Trụ cột thứ ba của nghiên cứu khoa học	5
2.2 Đồng hồ cát - các nhà cung cấp và người sử dụng PID.....	6
2.3 PID ở vòng eo của đồng hồ cát.....	6
2.4 Tác động chiến lược của việc không có hạ tầng PID hoạt động tốt	6
2.5 Một hạ tầng PID hoạt động tốt	7
3 Tầm nhìn và tham vọng	8
3.1 Tầm nhìn	8
3.2 Tham vọng	8
4 Bối cảnh PID - Hệ sinh thái và vai trò của các bên liên quan	9
4.1 Các hệ sinh thái PID	9
4.1.1 Các mã nhận diện được thiết lập tốt	9
4.1.2 Các mã nhận diện nội tại	10
4.1.3 Nhận diện tự quản và các mã nhận diện phi tập trung (DID)	11
4.2 Các bên liên quan và các vai trò trong hệ sinh thái PID	12
4.2.1 Cơ quan có thẩm quyền PID	12
4.2.2 Nhà cung cấp dịch vụ PID	13
4.2.3 Nhà quản lý PID	13
4.2.4 Chủ sở hữu PID	13
4.2.5 Người sử dụng đầu cuối PID	13
5 Phương pháp, chỉ đạo, ngân sách và kế hoạch thời gian	13
5.1 Các khái niệm về rủi ro và lòng tin trong ngữ cảnh PID	14
5.2 Các tiêu chí đánh giá rủi ro và lòng tin	15
5.2.1 Cơ quan có thẩm quyền PID.....	16
5.2.2 Nhà cung cấp dịch vụ PID	16
5.2.3 Nhà quản lý PID	16
5.2.4 Chủ sở hữu PID	17

5.2.5 Người sử dụng đầu cuối PID	17
5.3 Phương pháp - Sử dụng Khung Uyên thâm Mở	18
5.4 Chỉ đạo dự án	19
5.5 Kế hoạch ngân sách và thời gian	19
6 Nghiên cứu các câu chuyện trong các trường hợp PID của KE	19
6.1 Các chủ đề về lợi ích	20
6.2 Giới thiệu tổng quát về các nhà cung cấp và người sử dụng PID quốc gia của các thành viên KE	21
6.2.1 Phần Lan (CSC)	22
6.2.2 Pháp (Trung tâm Nghiên cứu Khoa học Quốc gia – CNRS)	23
6.2.3 Đan Mạch (DeiC)	26
6.2.4 Đức (DFG)	27
6.2.5 Vương quốc Anh (JISC)	29
6.2.6 Hà Lan (SURF)	30
7 Tóm tắt các trường hợp PID của KE	31
7.1 Rủi ro	31
7.2 Lòng tin	32
8 Kết luận các vấn đề về rủi ro và lòng tin của PID	32
9 Các khuyến nghị	32
10 Các tài liệu hỗ trợ	32
10.1 Các dự án và báo cáo liên quan	32
10.2 Các tài nguyên khác	34
10.3 Tham khảo	35

1 Tóm tắt và các từ khóa

1.1 Tóm tắt

Các mã nhận diện thường trực - PID (Persistent Identifiers) và hạ tầng của chúng được coi là có tầm quan trọng chiến lược đáng kể, đối với thực tế nghiên cứu hiện đại kỹ thuật số ngày một gia tăng.

Với nghiên cứu này, chúng tôi mong muốn hiểu biết tốt hơn những gì là cần thiết để xây dựng và khai thác một hạ tầng PID hoạt động tốt cho nghiên cứu. Tham vọng của chúng tôi là để xác định, thông qua nghiên cứu, phân tích và các khuyến nghị, những gì có thể là lộ trình chiến lược và hoạt động có thể tốt nhất để có được một hạ tầng PID hoạt động tốt cho các quốc gia thành viên trong và ngoài tổ chức Trao đổi Kiến thức - KE (Knowledge Exchange).

Tài liệu xác định phạm vi này phục vụ như một điểm khởi đầu, bằng việc cung cấp tổng quan về các hệ sinh thái PID hiện hành, tập trung để trở tới các vấn đề cần phải được giải quyết thông qua nghiên cứu. Chúng tôi đề xuất ánh xạ các vấn đề đó với **Khung Uyên thâm Mở (Open Scholarship Framework)** của KE, trợ giúp cho việc nhận diện và xây dựng các giải pháp tiềm tàng.

1.2 Các từ khóa

Mã nhận diện thường trực; các tổ chức nghiên cứu; châu Âu; đánh giá rủi ro; xây dựng lòng tin; hạ tầng Khoa học Mở; Khung Uyên thâm Mở; Trao đổi Kiến thức

2 Giới thiệu

Sự trao đổi lâu đời giữa lý thuyết khoa học và bằng chứng thực nghiệm vẫn đang phát triển mạnh vì đây vẫn là hai trụ cột truyền thống của nghiên cứu khoa học. Minh bạch về các phương pháp và kết quả đầu ra khoa học, nhận diện và tham chiếu tất cả đều là một phần của những gì tạo thành Khoa học Mở và Uyên thâm Mở hiện đại. Tuy nhiên, với sự hiện diện của điện toán và tác động của nó lên phương pháp khoa học, cũng như kho dữ liệu số gia tăng chưa từng thấy, tình hình đang tiến hóa.

Vài năm qua, nhiều người đã viện lý rằng điện toán và dữ liệu số đã tiến hóa để trở thành trụ cột thứ ba của nghiên cứu khoa học.

Chúng tôi viện lý rằng các con trỏ tin cậy tới các đối tượng số khác nhau, ví dụ, các mã nhận diện thường trực - PID (Persistent Identifiers), là quan trọng tối thượng để nói về sự thay đổi. Tầm quan trọng ngang bằng nhau là giả thiết rằng các PID chỉ có thể làm việc theo một cách thức hiệu quả nếu chúng trỏ tới các kho nơi các đối tượng số được lưu trữ và bảo tồn đúng cách về lâu dài. Trong tài liệu này, chúng tôi xác định một mã nhận diện như sự tuân tự các ký tự biểu thị một cách duy nhất đối tượng tham chiếu (Referent). Sự tuân tự này được coi là thường trực khi mã nhận diện đó cùng sự ràng buộc với đối tượng được tham chiếu và siêu dữ liệu có liên quan của nó sống sót qua thời gian và những tiến hóa công nghệ.

Chúng tôi muốn nghiên cứu các rủi ro chính là gì khi theo đuổi một hạ tầng PID hoạt động tốt cho nghiên cứu. Chúng tôi muốn hiểu tốt hơn các yếu tố quan trọng nhất về lòng tin trong tạo lập một hạ tầng như vậy.

2.1 Trụ cột thứ ba của nghiên cứu khoa học

Trụ cột thứ ba (Third Pillar) này là về cách để phần cứng máy tính, phần mềm và dữ liệu số cùng nhau tạo ra các phân tích, mô hình hóa và mô phỏng dữ liệu có thể ở mức độ phạm vi rộng lớn hơn và toàn cầu. Gần đây nhu cầu này đã trở nên rõ ràng để nghiêm túc phát triển các cách thức tốt hơn ở đó chúng ta xử lý dữ liệu số. Hệ quả là, ngôn ngữ chuyên nghiệp đã tiến hóa, nhằm vào việc xác định các yếu tố và phát triển trụ cột thứ ba đó. Các khái niệm như *Quản lý Dữ liệu*, *Quản trị Dữ liệu* và *các Nguyên tắc Dữ liệu FAIR (Findable, Accessible, Interoperable, Reusable)*, như một cách thức để cải thiện khả năng Tìm thấy được, Truy cập được, Tương hợp được, và Sử dụng lại được (dữ liệu số) tất cả đã giành được sự chú ý, như biện pháp quan trọng cho sự tiến bộ hơn nữa của khoa học, điều dựa vào hạ tầng và thực sự là toàn cầu. Ngoài ra, được thừa nhận rộng rãi rằng dữ liệu cần phải là máy truy cập được và hiểu được. Chỉ khi đó, toàn bộ tiềm năng của tài nguyên máy tính và dữ liệu phân tán trên toàn cầu mới có thể được khai thác, theo nhiều cách phân tán và đa dạng tương tự, để phân tích dữ liệu, mô hình hóa và mô phỏng.

Để điều này xảy ra, đã được gợi ý rằng một tập hợp ban đầu và rất tối thiểu các nguyên tắc hướng dẫn và các thực hành được cộng đồng đồng thuận là cần thiết.

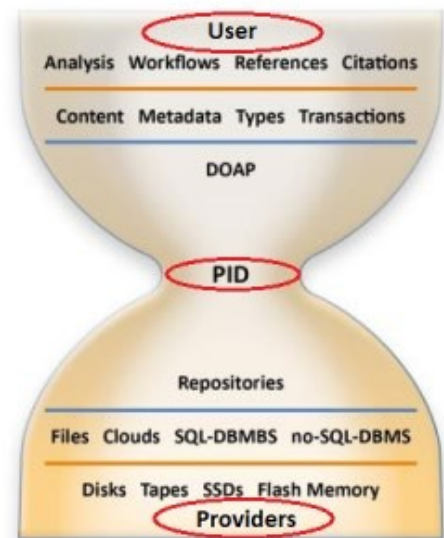
2.2 Đồng hồ cát - các nhà cung cấp và người sử dụng PID

Để xác định phạm vi một tập hợp tối thiểu các nguyên tắc hướng dẫn và thực hành được cộng đồng đồng thuận, có lập luận cho rằng các PID tạo thành vòng eo trong mô hình đồng hồ cát, với các nhà cung cấp và người dùng ở các đầu tương ứng:

2.3 PID ở vòng eo của đồng hồ cát

Tầm quan trọng của hạ tầng PID, được lập luận là cần phải khắc phục một điểm lỗi duy nhất hoặc mở rộng nút cổ chai:

“Các nguyên tắc và thực hành đơn giản đó nên xúc tác cho một dải rộng lớn các hành vi tích hợp và khám phá và hỗ trợ cho một dải rộng lớn các lựa chọn và triển khai công nghệ, hết như Giao thức Internet - IP (Internet Protocol) đã cung cấp một lớp tối thiểu - “cái eo” của đồng hồ cát - đã xúc tác cho việc tạo lập một ma trận khổng lồ về cung cấp, tiêu dùng, và các công cụ trực quan hóa dữ liệu trên Internet”.



Do đó, người ta lập luận rằng khả năng nhận diện duy nhất bất kỳ đối tượng số nào tạo thành điểm quan trọng nhất – tức là, vòng eo của đồng hồ cát. Vì thế, các mã nhận diện thường trực (PID) được giới thiệu như là một khái niệm hạ tầng rất quan trọng. Một hạ tầng PID hoạt động tốt đã trở thành một trong các yếu tố hạ tầng số khoa học chung cơ bản nhất để theo đuổi.

2.4 Tác động chiến lược của việc không có hạ tầng PID hoạt động tốt

Như là hệ quả của sự tiến hóa của trụ cột thứ ba của nghiên cứu khoa học, nhiều hạng mục có liên quan trực tiếp tới bản thân quy trình nghiên cứu hiện nay nên và có thể có các **mã nhận diện** được chấp nhận toàn cầu. Các hạng mục đó có nhiều, ví dụ, các xuất bản phẩm, trích dẫn, hội nghị, các nhà nghiên cứu, tổ chức, thử nghiệm lâm sàng, bằng sáng chế, tập hợp dữ liệu, kho dữ liệu, trợ cấp, dự án, mẫu, tổ chức cấp vốn.

Việc xây dựng dựa vào các hạ tầng PID hiện có và đặc biệt về mã nhận diện đối tượng số - DOI (Digital Object Identifier) như một mã nhận diện xuyên khắp các chủng loại của các đối tượng tham chiếu và ngữ cảnh, các công ty và các nhà xuất bản khác nhau đã phát triển một nhóm các dịch vụ cho các tổ chức và cơ quan nghiên cứu, ví dụ, các công cụ đánh giá tác động của tạp chí, các nghiên cứu đo lường khoa học, tác động của web đối với các bài báo học thuật, nơi mà giá trị gia tăng của chúng có nguồn gốc từ việc xử lý dữ liệu lớn bằng trí tuệ nhân tạo. Các dịch vụ dữ liệu lớn đó có mục đích để trở thành không thể thiếu trong các tiến trình khoa học và trong việc thông tin cho tất cả các dạng quyết định được các cơ sở nghiên cứu đưa ra, ví dụ, để tuyển dụng, cấp vốn nghiên cứu, và phổ biến các kết quả trong một môi trường cạnh tranh toàn cầu.

Là rất quan trọng cho các tổ chức nghiên cứu để hiểu các thách thức do việc sử dụng các PID đặt ra để biến các kết quả đầu ra của nghiên cứu thành các hạng mục FAIR mà không thiên vị cho bất kỳ khu vực và hệ thống cụ thể nào.

Chúng tôi vì thế kết luận rằng một hạ tầng PID hoạt động tốt phải được xây dựng và khả năng quản lý yếu kém trong việc đưa ra hạ tầng như vậy sẽ là một tổn thất lớn đối với khoa học.

2.5 Một hạ tầng PID hoạt động tốt

Một hạ tầng hoạt động tốt có nghĩa nó là:

- thân thiện với người sử dụng **về kỹ thuật** và có khả năng nhận diện thường trực và duy nhất bất kỳ đối tượng số nào, được cho là đáng bảo tồn.
- **được chấp nhận** toàn cầu (tương hợp được trong công nghệ và thiết kế lõi của nó) sao cho nó độc lập với công nghệ và địa lý, luôn trở đến tài khoản chủ sở hữu dữ liệu và siêu dữ liệu có liên quan (nghĩa là phân giải thành trang đích giải thích), nếu không phải là dữ liệu thực tế.
- **bền vững về mặt tổ chức và kinh tế**, nghĩa là, PID có thể vẫn phân giải ngay cả trong trường hợp thay đổi tổ chức hoặc hỗn loạn về kinh tế - về nguyên tắc là vĩnh viễn.
- đáng tin cậy **về chính trị** - theo đó có rủi ro tối thiểu về tính tương hợp ngẫu nhiên, các rào cản pháp lý hoặc khóa trói vào nhà cung cấp khai thác.

Chúng tôi lưu ý rằng các yếu tố ở trên là cần thiết cho bất kỳ hệ thống nhận diện nào cho kết quả đầu ra nghiên cứu bất kỳ dạng nào. Vì thế phạm vi PID của chúng tôi bao trùm tất cả các hệ thống phân giải, nghĩa là, tất cả các dạng PID.

3 Tầm nhìn và tham vọng

Khái niệm PID đã trở nên rất quan trọng để đạt được một cách thức đồng thuận toàn cầu trong việc tổ chức dữ liệu, sao cho người và máy đều có thể *tim thấy* bất kỳ đối tượng số nào được bảo tồn, ở bất cứ đâu, bằng bất kỳ phương tiện nào, bất cứ lúc nào; và nếu được coi là đáng giá, các vấn đề về đạo đức và pháp lý cho phép, cũng xử lý nó một cách khoa học và tính toán, theo các nguyên tắc dữ liệu FAIR.

3.1 Tầm nhìn

Tầm nhìn của chúng tôi là:

Một hạ tầng PID hoạt động tốt cho nghiên cứu đóng gói và triển khai khái niệm PID toàn cầu xuyên khắp tất cả các lĩnh vực khoa học.

Chúng tôi mong muốn không chỉ nghiên cứu những gì nó cần để đạt được trạng thái tốt như vậy cho các PID, mà còn cách để mọi người có thể đi với nó sao cho nó được áp dụng và hiện diện ở khắp mọi nơi. Trong khi bên hưởng lợi theo mục tiêu là các thành viên của KE, chúng tôi kỳ vọng các kết quả đầu ra từ hoạt động này cũng có thể là phù hợp cho các mối quan tâm rộng lớn hơn về hạ tầng mở.

3.2 Tham vọng

Tham vọng của chúng tôi là:

Lý tưởng, thông qua nghiên cứu, phân tích và kết luận, điều gì có thể là các khuyến nghị chiến lược và hoạt động được tốt nhất nhằm đạt được một hạ tầng PID hoạt động tốt cho các quốc gia thành viên của KE và hơn thế nữa.

Bằng cách làm như vậy, sự hiểu biết, đặc biệt, về 2 khái niệm “rủi ro” và “lòng tin” được tin tưởng sẽ là nền tảng cho việc áp dụng mặc định (de facto) một hạ tầng PID hoạt động tốt.

Kết quả nghiên cứu hy vọng có thể được sử dụng chung cho cộng đồng các tác nhân rộng lớn, làm việc về triển khai PID - toàn cầu và/hoặc địa phương.

4 Bối cảnh PID - Hệ sinh thái và vai trò của các bên liên quan

4.1 Các hệ sinh thái PID

Theo truyền thống, các PID được coi như là các mã nhận diện mà chúng là:

- duy nhất, nghĩa là, không bao giờ được sử dụng lại một lần nữa để đại diện cho một đối tượng khác,
- hoạt động được, nghĩa là, có các dịch vụ gắn với mã nhận diện đó,
- thường trực, nghĩa là, đối tượng y hệt không bao giờ được trao một mã nhận diện mới trong cùng một hệ thống.

Cần lưu ý là ý nghĩa “Thường trực” của các PID là một vùng xám: Vài bên liên quan kỳ vọng nội dung đằng sau các PID là bất biến ở mức bit, trong thực tế một yêu cầu mạnh như vậy tạo ra các vấn đề mới trong quản lý PID. Có nhu cầu đối với các chính sách rõ ràng về những gì được coi như là thường trực.

4.1.1 Các mã nhận diện được thiết lập tốt

Một số mã nhận diện đang được các tổ chức khắp trên thế giới sử dụng rồi, chẳng hạn, ARK, DOI, Handles, URN phân giải cho dữ liệu và các xuất bản phẩm, còn ISNI, ORCID thì trở tới thông tin xác thực về các cá nhân².

Từng trong số các mã nhận diện đó có hệ thống điều hành của riêng chúng thường được các cơ quan có thẩm quyền đăng ký ở cấp trung ương quản lý.

Trong khi các hệ thống mã nhận diện thường trực được thiết lập còn chưa giải quyết được tất cả các vấn đề xung quanh Rủi ro và Lòng tin, các hệ thống mới đang nổi lên, ví dụ như, các mã nhận diện nội tại (Intrinsic Identifier) nhận diện các đối tượng dựa vào nội dung mức bit của chúng. Khái niệm Nhận diện Tự quản (Self-Sovereign Identity) có ý định phi tập trung hóa việc đúc sẵn và xử lý các mã nhận diện của cơ quan đăng ký và

2 There are also PID ecosystems for citizens, like personal identification numbers in the Nordic Countries. Personal numbers are used for example in banking, for health and tax purposes.

các nhà cung cấp PID, cùng lúc giới thiệu các cơ chế để xây dựng lòng tin vào một cách tiếp cận phi tập trung hóa. Cả hai hệ thống mới đó được chi tiết hóa bên dưới đây.

4.1.2 Các mã nhận diện nội tại

Như đã được nêu, lưu ý quan trọng rằng các dạng PID khác nhau đang tồn tại. Vài dạng được các nhà cung cấp dịch vụ PID quản lý, vài dạng khác có thể được các chủ sở hữu PID quản lý. Điều này đáng chú ý là trường hợp đối với các hàm băm mạnh về mật mã có thể tính toán từ bất kỳ tệp nào một “chữ ký” ngắn, do đó cung cấp một mã nhận diện nội tại (Intrinsic Identifier) cho một đối tượng kỹ thuật số nhất định. Các đối tượng số đó, các mã nhận diện nội tại của chúng và siêu dữ liệu có liên quan có thể được chủ sở hữu PID đó quản lý và được truyền qua các chuỗi khối (Blockchain) và các hệ thống tệp phân tán, vì thế bỏ qua các nhà cung cấp dịch vụ PID. “Các mã nhận diện nội tại đó là rất mạnh, vì chúng cho phép không chỉ nhận diện duy nhất một đối tượng, mà còn xác thực rằng đối tượng được chỉ định còn chưa bị sửa đổi: nó đủ để tính toán lại mã nhận diện nội tại đó từ chính bản thân đối tượng đó để phát hiện ra bất kỳ sự thay đổi nào.” (Allen A. et al., 2020).

Kho lưu trữ Software Heritage - Di sản Phần mềm (<https://www.softwareheritage.org/>) đã phát triển và triển khai mã nhận diện thường trực Di sản Phần mềm, ví dụ, SWHID, các đặc tả của chúng là sẵn sàng miễn phí tại <https://docs.softwareheritage.org/devel/swh-model/persistent-identifiers.html>.

SWHID là mã nhận diện rất linh hoạt xúc tác cho người sử dụng để lựa chọn các dòng mã lệnh nhất định được gắn một mã nhận diện. Trong kho lưu trữ đó, hàm băm duy nhất này xúc tác cho bất kỳ ai để truy xuất mã lệnh được lựa chọn. Thông qua một trình duyệt, đường liên kết cố định đó bao gồm tiền tố phân giải và trở tới đúng kho lưu trữ và các dòng mã lệnh đó.

Một sáng kiến hay khác là mã nội dung ISCC sẵn sàng miễn phí (<https://iscc.codes/>) có thể được áp dụng cho bất kỳ tệp số các dạng đa phương tiện chung nào (văn bản, hình ảnh, âm thanh, video). Việc xử lý nội dung với các thuật toán được đặc tả ISCC xác định tạo ra một mã tổng hợp duy nhất, gồm 4 yếu tố chính, đó là, siêu ID (meta-ID), ID nội dung (content-ID), ID dữ liệu (data-ID) và ID trường hợp (instance-ID). ISCC hỗ trợ việc quản lý phiên bản nội dung, nhận diện sản phẩm có liên quan, dò tìm các phương án nội dung, thẩm định sở hữu dữ liệu. ISCC là sẵn sàng theo giấy phép CC BY-NC-SA 4.0.

4.1.3 Nhận diện tự quản và các mã nhận diện phi tập trung (DID)

Nhận diện tự quản (Self-sovereign identity) là “hệ thống quản lý nhận diện cho phép **các cá nhân** sở hữu và quản lý hoàn toàn nhận diện số của họ” [Mühle et al., 2018] mà tồn tại độc lập với các dịch vụ, ví dụ, Facebook, Google hoặc bất kỳ dịch vụ nào khác yêu cầu một cá nhân đăng nhập. Hệ thống quản lý nhận diện này là hướng vào người sử dụng và dựa vào sở cái phân tán của chuỗi khối để thay thế cơ quan đăng ký trong các hệ thống quản lý nhận diện kinh điển. Các nguyên tắc của nhận diện tự quản đã được [Allen, 2016] xác định. Vài dịch vụ tập trung vào nhận diện cá nhân đã bắt đầu hoạt động, ví dụ, Sovrin, uPort [Naik & Jenkins, 2020], Khung Nhận diện Tự quản của châu Âu - ESSIF (European Self-Sovereign Identity Framework).

Cả Sovrin và uPort đều sử dụng các mã nhận diện phi tập trung - DID (Decentralized IDentifiers) được W3C phát triển để xúc tác cho nhận diện số xác thực và phi tập trung. Tuy nhiên, một DID có thể nhận diện bất kỳ đối tượng được tham chiếu (Referent) nào (một “Đối tượng DID”, ví dụ, một người, tổ chức, một thứ gì đó, mô hình dữ liệu, thực thể trừu tượng, .v.v.) mà nhà kiểm soát DID quyết định rằng nó nhận diện. Các đối tượng của DID được xác định bằng việc sử dụng các tài liệu DID (DID Documents) có chứa siêu dữ liệu về đối tượng DID.

Ngược lại với các mã nhận diện thông thường, liên đoàn, các DID đã được thiết kế để chúng có thể được tách khỏi các đăng ký tập trung, các nhà cung cấp nhận diện, và các cơ quan chứng thực. Đặc biệt, trong khi các bên khác có thể được sử dụng để giúp xúc tác để phát hiện thông tin có liên quan tới một DID, thiết kế đó cho phép nhà kiểm soát một DID chứng minh sự kiểm soát đối với nó mà không yêu cầu sự cho phép từ bất kỳ bên nào khác. Các DID là các URI liên kết một đối tượng DID với một tài liệu DID cho phép các tương tác tin cậy liên quan tới đối tượng đó. Thông tin chi tiết là có sẵn tại <https://www.w3.org/TR/did-core/>. Các nghiên cứu về cách để sử dụng nhận diện tự quản và các DID để xác thực các tập hợp dữ liệu khoa học đang được phát triển [Barclay et al. 2020].

Vào thời điểm tài liệu này được viết (mùa xuân 2021) không tồn tại triển khai nào dựa vào DID đang hoạt động, tuy nhiên các sáng kiến như GAIA-X đang quảng bá mạnh việc

áp dụng chúng³. Các DID được thiết kế để máy đọc được và cung cấp các cơ chế tin cậy cho xác thực. Có thể là thú vị để nghiên cứu liệu các DID có tiềm năng bổ sung hoặc thay thế cho các hệ thống PID hiện có hay không.

4.2 Các bên liên quan và các vai trò trong hệ sinh thái PID

Đơn giản hóa, hệ sinh thái PID gồm những người sử dụng PID và các nhà cung cấp PID. Nghĩa là, ở một đầu, có những người sử dụng đầu cuối của hạ tầng PID, chẳng hạn như, các nhà nghiên cứu, các cộng đồng và các viện nghiên cứu. Ở đầu kia, chúng ta thấy các nhà cung cấp PID, ví dụ như, các nhà quản lý PID, như các kho lưu trữ, các nhà cung cấp hạ tầng PID (như DEIC, CSC, .v.v.), các tập đoàn PID (như những gì trong EPIC, DataCite).

Tuy nhiên, những người sử dụng cũng có thể là các nhà cung cấp, và ngược lại, nghĩa là, các nhà cung cấp hạ tầng nghiên cứu có thể hội tụ với những người sử dụng và các cộng đồng nghiên cứu, vì thế làm việc với nhiều, nếu không nói là tất cả, các yếu tố của đồng hồ cát.

Đi tiếp khi chúng tôi muốn hiểu tốt hơn cấu trúc của và các tác nhân trong bối cảnh hạ tầng nơi PID phải được nhúng vào, chúng tôi phải nghiên cứu các vai trò hiện hành của các bên liên quan khác nhau tham gia.

Các vai trò của các bên liên quan tới PID hiện hành (theo Hellström et al., 2020) có thể được phân tích như sau:

4.2.1 Cơ quan có thẩm quyền PID

Một nhà kiểm soát có trách nhiệm *duy trì các quy định để xác định tính toàn vẹn và tính duy nhất của các PID trong hệ thống PID*. Các quy định đó có thể bao gồm việc thiết lập các tiêu chuẩn cho các định dạng từ vựng, thuật toán và giao thức để đảm bảo tính duy nhất toàn cầu, cùng với việc thiết lập các điều kiện chất lượng dịch vụ để thực thi tuân thủ với các quy định đó. Các cơ quan PID có thể là các tổ chức thực thi kiểm soát đối với một hạ tầng PID. Nhưng cũng có thể là các cơ quan không có sự kiểm soát tập trung mà cung cấp một cơ chế tiêu chuẩn hóa của cộng đồng mà chỉ định sự tuân thủ của các PID đối với một hệ thống PID.

3 See Chapter 3.4 in the GAIA-X Technical Architecture document: <https://www.data-infrastructure.eu/GAIA-X/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf>

4.2.2 Nhà cung cấp dịch vụ PID

Một tổ chức cung cấp các dịch vụ PID tuân thủ với một hệ thống PID, tuân theo cơ quan PID của nó. *Các nhà cung cấp dịch vụ PID có trách nhiệm về cung cấp, tính toàn vẹn, độ tin cậy và khả năng mở rộng phạm vi các dịch vụ PID*, đặc biệt việc ban hành và phân giải các PID, mà còn cả các dịch vụ tra cứu và tìm kiếm nữa.

4.2.3 Nhà quản lý PID

Các nhà quản lý PID có trách nhiệm *duy trì tính toàn vẹn của mối quan hệ giữa các thực thể và các PID của họ, tuân thủ với một hệ thống PID được cơ quan PID xác định*. Một nhà quản lý PID thường sẽ thuê bao các dịch vụ PID để cung cấp chức năng cho chủ sở hữu PID trong các dịch vụ của nhà quản lý PID đó. Một ví dụ là nhà cung cấp dịch vụ mà sử dụng các dịch vụ PID như một phần của việc cung cấp dịch vụ của riêng nó. Ví dụ, các nhà quản lý PID có thể bao gồm một nhà cung cấp kho dữ liệu hoặc xuất bản phẩm, một catalog dữ liệu, hoặc một hệ thống tiến trình nghiên cứu.

4.2.4 Chủ sở hữu PID

Một tác nhân (một tổ chức hoặc cá nhân) có quyền *tạo lập một PID, chỉ định PID cho một thực thể, cung cấp và duy trì thông tin hạt nhân (Kernel Information) chính xác bao gồm vị trí cho PID đó*. Chủ sở hữu PID là chủ sở hữu của bản ghi PID cá nhân.

4.2.5 Người sử dụng đầu cuối PID

Người sử dụng đầu cuối các dịch vụ PID và các dịch vụ của người sử dụng PID. Chúng có thể, ví dụ là, các nhà nghiên cứu, hoặc phần mềm, hoặc các dịch vụ được sản xuất để hỗ trợ cho các nhà nghiên cứu. Những người sử dụng đầu cuối sẽ sử dụng các PID để trích dẫn và truy cập các tài nguyên hoặc thông tin hạt nhân về các tài nguyên đó.

5 Phương pháp, chỉ đạo, ngân sách và kế hoạch thời gian

Về phương pháp luận, các khái niệm Rủi ro và Lòng tin đang điều hành, ở đó chúng tôi muốn phân tích rủi ro vốn có đối với tầm nhìn được nêu, dựa vào sự hiểu biết nhận thức được về tình hình triển khai PID quốc gia, khu vực và toàn cầu.

Điểm xuất phát là một nghiên cứu về các trường hợp điển hình PID của KE. Lấy thành viên KE làm trọng tâm, chúng tôi tập hợp kiến thức từ:

- Phòng vấn các bên liên quan (các nhà nghiên cứu; các nhà cung cấp hạ tầng điện tử; các kho lưu trữ)
- Tài liệu (các báo cáo, các đặc tả, website, .v.v.)
- Các khuyến nghị về thực hành tốt, những điều nên và không nên làm
- Các rủi ro sẽ được xác định khi có liên quan tới các thực hành đó.

Chúng tôi sẽ xem xét mức độ ở đó lòng tin là cần thiết, ở đâu, vì sao và làm thế nào, đối với một hạ tầng PID hoạt động tốt để nổi lên.

5.1 Các khái niệm về rủi ro và lòng tin trong ngữ cảnh PID

Các tổ chức cung cấp các PID vốn dĩ tạo ra vài mức *khóa trói vào nhà cung cấp*, ví dụ, trong hệ thống PID được chọn. Ví dụ, ngay cả nếu cơ quan PID như *doi.org* có thể phân giải được các URN và cơ quan PID khác như *urn.fi* có thể phân giải được các DOI, các nhà cung cấp dịch vụ PID vẫn bị ràng buộc với các hệ thống quản lý tương ứng đặc thù cho một hệ thống được chọn. Các DOI với các trường siêu dữ liệu không thể dễ dàng được chuyển đổi sang các URN.

Các nhà nghiên cứu và các tổ chức sử dụng các PID, vì thế, cần tin tưởng sự điều hành của tổ chức đứng đằng sau một hệ thống PID, bất kể tổ chức đó có thể là một quỹ, tập đoàn thương mại, công chúng, tổ chức phi lợi nhuận (ví dụ như, ePIC, DataCite, .v.v.). Những người sử dụng cần đòi hỏi liệu hệ thống PID đó có bền vững hay không; liệu sự điều hành, chiến lược và chiến thuật là minh bạch và tin cậy hay không. Những người sử dụng phải tin tưởng rằng các nhà cung cấp PID trên thực tế phân giải nhất quán và lâu dài tới các URL đúng và siêu dữ liệu được tham chiếu là cập nhật.

Những người sử dụng cần đánh giá các rủi ro tài chính của họ cũng như rủi ro tiềm tàng cho tương lai chương trình nghiên cứu của họ, liệu hạ tầng PID đó có thất bại hoặc trở nên quá nặng nề hay không. Hệ quả là, các nhà cung cấp dịch vụ PID cần được điều chỉnh tốt để đảm bảo sự minh bạch, điều hành tốt, kinh tế khá và công nghệ vững, bằng cách đó đảm bảo tính bền vững của PID và chương trình nghiên cứu phụ thuộc đó. Điều này đúng đối với các nhà cung cấp thương mại cũng như được nhà nước cấp vốn, đối với các nhà cung cấp quốc gia cũng như các nhà cung cấp quốc tế.

Nhu cầu về quy định và điều hành nói trên rõ ràng là khá khó thực hiện, vì hai lý do rất khác nhau:

- Ở các cơ sở nghiên cứu, quả thực, bản thân các nhà khoa học hưởng thụ một mức độ tự chủ lớn trong việc tổ chức khoa học, lựa chọn các công cụ và hạ tầng. Có rất ít cấu trúc điều hành hỗ trợ thỏa thuận về tiêu chuẩn hóa hạ tầng và công cụ khoa học. Trong khi các nỗ lực chung về hạ tầng khoa học quốc tế là có thể, chúng hầu hết nằm trong các lĩnh vực cụ thể theo các ngành.

Rủi ro ở đây là các trường đại học thiếu quyết tâm và năng lực đáng tin cậy để hợp tác trong việc xây dựng hạ tầng PID chung hoạt động tốt.

Các nhà cung cấp thương mại có thể là tốt hơn trong việc điều phối trọng tâm nhằm vào việc thực thi hoặc thương thảo về các tiêu chuẩn cũng như việc phát triển hạ tầng/phần mềm cụ thể. Tuy nhiên, vì các lý do rõ ràng, họ chỉ ở đó vì lợi nhuận dài hạn. Việc sử dụng đa dạng rộng rãi các hệ thống PID cùng một lúc cho cùng một mục đích có thể gây lúng túng cho những người sử dụng và, tương tự, cho các nhà quản trị, nhưng các nhà cung cấp thương mại có xu hướng tìm kiếm sự thống trị hệ thống và tạo ra mối liên kết và sự phụ thuộc, điều này là có hại cho tiến bộ khoa học.

Rủi ro ở đây là một hạ tầng PID hoạt động tốt có chi phí rất cao, mà chúng ta không thể tin tưởng chỉ về khía cạnh tiền tệ.

Rủi ro cho những người sử dụng đầu cuối đang sử dụng bất kỳ hạ tầng PID cụ thể nào hay một PID được đúc sẵn nào sẽ là khiêm tốn hơn, miễn là bản thân dữ liệu có sẵn bằng cách nào đó. Các PID không phân giải chắc chắn là một mối phiền toái, nhưng không liên quan đến uy tín của người sử dụng PID. Nó cũng không gây gián đoạn nghiêm trọng cho nghiên cứu, giống như sự vắng mặt của quy định và điều hành của trường đại học/thương mại có thể xảy ra.

5.2 Các tiêu chí đánh giá rủi ro và lòng tin

Lòng tin được cho là không thể tránh khỏi, và lòng tin luôn kéo theo sự chấp nhận vài rủi ro. Đối với từng nhóm các bên liên quan, với một vai trò cụ thể, lòng tin vào một hạ tầng PID có thể dựa vào một số tiêu chí, từng tiêu chí kéo theo vài mức độ rủi ro, lớn hoặc nhỏ, tiềm tàng thảm họa hay không. Nghĩa là:

5.2.1 Cơ quan có thẩm quyền PID

- được đặt dưới sự kiểm soát của một cơ quan tiêu chuẩn hóa quốc tế thiết lập và thực thi các quy trình minh bạch cho việc tạo lập, phê chuẩn, duy trì và chấm dứt các tiêu chuẩn PID,
- đã ký một thỏa thuận hợp lệ với cơ quan tiêu chuẩn hóa quốc tế xác định các quyền và bổn phận của nó và cung cấp một kế hoạch kế vị,
- có tư cách pháp nhân, quy chế công khai, điều hành đại diện, các cơ quan điều hành với nguyên tắc biểu quyết rõ ràng.

5.2.2 Nhà cung cấp dịch vụ PID

- có thỏa thuận với cơ quan có thẩm quyền PID,
- có một mô hình kinh doanh tin cậy được qua thời gian dài, các tài khoản được một cơ quan/công ty kiểm soát được thừa nhận phê chuẩn.
- xuất bản một báo cáo hoạt động thường niên,
- có kho người dùng ổn định và hợp tác trên cơ sở thường xuyên với các nhà quản lý PID triển khai PID bên trong các hệ thống thông tin của riêng họ,
- triển khai các biện pháp tài chính, kỹ thuật và con người để duy trì hệ thống thông tin PID chủ chốt và thực hiện các đầu tư cần thiết cho khía cạnh này,
- tạo lập một hệ thống thông tin theo đó PID được/bị ràng buộc vĩnh viễn với đối tượng được tham chiếu, đáng chú ý qua nhận diện & siêu dữ liệu mô tả cho các quy định và cơ chế tạo lập, sửa lỗi và cập nhật đã được xác định, triển khai và chia sẻ công khai.
- làm cho các PID sẵn sàng đúng lúc cho nhà quản lý PID,
- Các PID được nhà cung cấp dịch vụ PID không chứa ngữ nghĩa quản lý.

5.2.3 Nhà quản lý PID

- quản lý các quy trình để tương tác với nhà cung cấp dịch vụ PID để yêu cầu các PID cho hệ thống thông tin cục bộ địa phương,

- tải lên và cập nhật siêu dữ liệu của đối tượng được tham chiếu trong các hệ thống thông tin của nhà cung cấp dịch vụ PID,
- cung cấp hướng dẫn rõ ràng cho chủ sở hữu PID để tạo lập và thường xuyên cập nhật siêu dữ liệu của đối tượng được tham chiếu trong hệ thống thông tin cục bộ địa phương,
- điền vào các hệ thống thông tin cục bộ địa phương với các PID.

5.2.4 Chủ sở hữu PID

- tạo lập và cập nhật thường xuyên siêu dữ liệu đối tượng được tham chiếu của PID trong hệ thống cục bộ địa phương.

5.2.5 Người sử dụng đầu cuối PID

- mã nhận diện là duy nhất cho đối tượng được tham chiếu cụ thể và có thể được sử dụng như là sự ủy quyền hoặc như một đường liên kết tới đối tượng đó,
- mã nhận diện được thừa nhận như là một mã nhận diện thường trực và dạng đối tượng được đại diện là không mù mờ (siêu dữ liệu thuần túy, nhận diện mức bit, nhận diện nội dung thông tin, đối tượng hoặc bộ sưu tập động, .v.v.)
- mã nhận diện đủ chi tiết để nhận diện các đối tượng và các phần con của chúng theo các yêu cầu của người sử dụng đầu cuối,
- mã nhận diện được sử dụng ở mức quốc tế bởi các tổ chức sản xuất nội dung và tiêu dùng nội dung khác nhau,
- mã nhận diện phân giải trơn tru đối với đối tượng được tham chiếu hoặc thông tin về đối tượng này,
- siêu dữ liệu được kết nối với một mã nhận diện là sẵn sàng để các bên được ủy quyền sử dụng lại trong các định dạng đa dạng khác nhau,
- các mối quan hệ và nguồn gốc của siêu dữ liệu được ghi thành tài liệu tốt.
- PID được sử dụng rộng rãi khắp các lĩnh vực sản xuất và tiêu dùng nghiên cứu.

Các tiêu chí khác được kỳ vọng được xác định trong quá trình nghiên cứu này.

5.3 Phương pháp - Sử dụng Khung Uyên thâm Mở

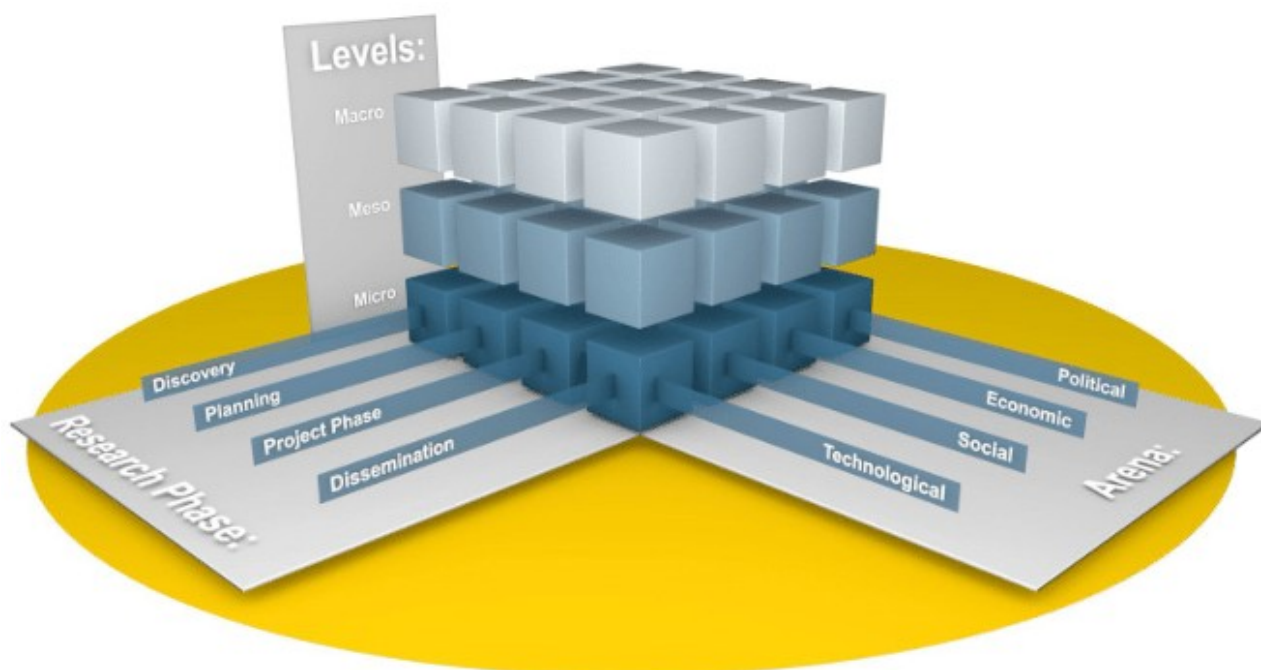
Khi các khái niệm về lòng tin và rủi ro có liên quan tới bất kỳ hạ tầng PID cụ thể nào được hiểu tốt, một phân tích các kịch bản triển khai và sử dụng khác nhau có thể được tiến hành từ 2 quan điểm: kịch bản những người sử dụng PID cũng như của các nhà cung cấp PID, theo một số vai trò. Phân tích như vậy được tiến hành bằng việc sử dụng các trường hợp của các thành viên KE, nơi từng thành viên KE trình bày các triển khai PID và nơi các mẫu sử dụng được phân tích có hệ thống:

- Nhà cung cấp PID trình bày
 - Các vấn đề về rủi ro
 - Các vấn đề về lòng tin
- Các mẫu sử dụng PID được tìm thấy qua suốt các ngành và tổ chức nghiên cứu
 - Các vấn đề về rủi ro
 - Các vấn đề về lòng tin

Khi xem xét và phân tích các mẫu triển khai và sử dụng PID khác nhau, xuyên khắp nhiều trường hợp của KE, là quan trọng để áp dụng một cách tiếp cận hệ thống để xác định các vấn đề rủi ro và lòng tin.

Còn tranh cãi, khung tốt nhất cho phân tích như vậy là **Khung Uyên thâm Mở (Open Scholarship Framework)**, đưa ra một số “điểm phân tích” theo 3 chiều - Pha nghiên cứu; Lĩnh vực; và Mức xã hội, như được minh họa bên dưới (Hình 1).

Nhiều vấn đề đang nổi lên được kỳ vọng sẽ là những thách thức đối với ‘hành động tập thể’. Với sự trợ giúp của Khung đó, chúng tôi có thể xác định và mô tả vấn đề đó, và các khuyến nghị cho từng tác nhân/bên liên quan. Nghiên cứu có thể xem xét các vấn đề và các thách thức đang nổi lên đối với một hạ tầng PID hoạt động tốt thông qua lăng kính của Khung Uyên thâm Mở đó.



Hình 1: Khung Uyên thâm Mở

5.4 Chỉ đạo dự án

Được đề xuất rằng một chuyên gia tư vấn ký hợp đồng để phác thảo một báo cáo - các ***Rủi ro và Lòng tin khi Theo đuổi một Hạ tầng Mã nhận diện Thường trực Hoạt động Tốt cho Nghiên cứu***. Báo cáo đó phải được phác thảo trong sự hợp tác với các nhóm KE thích hợp.

5.5 Kế hoạch ngân sách và thời gian

Tổng ngân sách bằng tiền EUR được phân bổ cho một chuyên gia tư vấn.

Hoạt động đó - phác thảo và phê duyệt báo cáo - phải hoàn thành vào mùa hè 2022. Sẽ có các báo cáo tiến độ theo từng quý được cung cấp cho Nhóm Chỉ đạo có kết nối với các cuộc họp để trình bày và thảo luận.

6 Nghiên cứu các câu chuyện trong các trường hợp PID của KE

Các trường hợp điển hình giữa các thành viên KE sẽ được tiến hành bằng việc phân tích các kịch bản triển khai và sử dụng PID khác nhau. Có ý thức lấy các quốc gia thành viên KE làm trung tâm. Đây không phải là rà soát lại ngang hàng các hệ thống PID khác nhau, mà là xem xét đặc biệt triển khai và sử dụng hiện hành, và xác định các vấn đề có

trong một hạ tầng PID hoạt động tốt. Khi làm như vậy, các vấn đề về rủi ro và lòng tin được xác định từ góc độ của người sử dụng cũng như nhà cung cấp.

6.1 Các chủ đề về lợi ích

Các vấn đề quan trọng được xác định, có liên quan tới hạ tầng PID hoạt động tốt, được ánh xạ trong các chiều lợi ích - Pha nghiên cứu; Lĩnh vực; Mức xã hội (Hình 1) - và được mô tả và phân tích về các khía cạnh rủi ro và lòng tin. Một số câu hỏi rõ ràng là để trình bày những câu hỏi khác ít hơn, và vẫn còn những câu hỏi khác chưa được nêu ra. Một số câu hỏi thuyết phục có thể được thu thập theo các chủ đề:

- Mức phối hợp quốc gia (các chính sách, tổ chức, cấp vốn)
 - Các dịch vụ/hạ tầng PID được tổ chức như thế nào, và nó được phối hợp ở mức quốc gia/quốc tế như thế nào?
- Nhận thức/Bao quát
 - Các nhà cấp vốn có hay không yêu cầu sử dụng PID từ các nhà nghiên cứu?
 - Việc trích dẫn có sử dụng PID phổ biến không?
 - Các PID có được sử dụng trong hầu hết các tổ chức cung cấp dữ liệu lớn?
 - Bối cảnh các PID bị phân mảnh như thế nào?
 - Các hệ thống PID nào được sử dụng trong các tổ chức thành viên khác nhau của KE?
- Các triển khai
 - Công nghệ PID là tự duy trì hay thuê ngoài làm?
 - Siêu dữ liệu được duy trì và giám tuyển như thế nào?
 - Các mô hình điều hành và chính sách hiện diện. Nghĩa là, việc quản lý có là tập trung hay không; các trách nhiệm được phân bổ như thế nào (có liên quan tới các PID và các chiến lược thoát ra)?
 - Sử dụng các PID trong quản lý, đánh giá và/hoặc phân tích nghiên cứu (ví dụ, thu thập thông tin về con người hoặc dự án)

- Các nhà cung cấp của nhà nước và người sử dụng có mối quan hệ nào với các đối tác thương mại?
- Các mối quan hệ thương mại bất kỳ nào là tuân theo các điều khoản (rủi ro/lòng tin) nào?
- Các phát triển liên tục
 - Chiến lược hoặc kế hoạch của quốc gia và/hoặc cơ sở
 - Các tính năng PID cao cấp nào được trình bày và sử dụng, ví dụ, như các mã nhận diện một phần?
 - Các triển khai mới (trong, ví dụ, CRIS, dữ liệu, các nhà nghiên cứu, các kho lưu trữ, cấp vốn nghiên cứu, các dự án, .v.v.)
 - PID trong các hệ sinh thái Khoa học Mở, bằng việc xem xét cách để bất kỳ triển khai và sử dụng PID cụ thể nào được điều chỉnh phù hợp với các chính sách Truy cập Mở hiện hành của Liên minh châu Âu?
- Tính tương hợp
 - Các hệ thống PID hiện hành tương hợp được ở mức độ nào?
- Cấp vốn
 - Các hệ sinh thái PID quốc gia được cấp vốn như thế nào, sự thường trực của chúng được đảm bảo như thế nào?
 - Ngân sách nào (EUR và số lượng người theo từng tháng) được sử dụng, ở các mức quốc tế, quốc gia và/hoặc địa phương?

6.2 Giới thiệu tổng quát về các nhà cung cấp và người sử dụng PID quốc gia của các thành viên KE

Bên dưới chúng tôi mô tả một tập hợp cơ bản các thuộc tính có liên quan tới từng quốc gia thành viên của KE, dựa vào một mẫu template được tiêu chuẩn hóa. Rõ ràng, khi xem xét hơn nữa vào các kịch bản triển khai và sử dụng cụ thể đối với các thành viên của KE, mẫu template đó phải được coi như là điểm khởi đầu.

6.2.1 Phần Lan (CSC)

Ở mức độ chung, việc sử dụng PID khá chín muồi ở Phần Lan và cơ sở hạ tầng tương đối tập trung. Chúng tôi có một nhóm quốc gia về ORCID và DataCite. Đối với các xuất bản phẩm, *journals.fi* cung cấp các CrossRef DOI. Các URN được hỗ trợ khá tốt từ Thư viện Quốc gia và được sử dụng rộng rãi. Handles cũng được sử dụng, đáng chú ý nhất bởi các khuyến nghị quốc gia các dịch vụ của EUDAT cho các PID trong nghiên cứu có tồn tại. Dịch vụ *research.fi* quốc gia của CRIS sử dụng tích cực PID. Bên dưới chúng tôi tóm tắt các vấn đề và các lĩnh vực tiềm năng đáng khai phá.

Các rủi ro và lòng tin:

- Rủi ro: Các nguyên tắc FAIR bị sử dụng sai trong đánh giá
- Rủi ro: Sử dụng các URN trong ngữ cảnh FAIR chưa được xác định
- Lòng tin: Các vấn đề với thông tin động đăng sau các PID, nơi có thể thiếu lòng tin khi người sử dụng đầu cuối không chắc chắn các yếu tố nào nằm bên dưới → khía cạnh thường trực trở nên đáng ngờ. Chúng tôi cần xác định yếu tố thường trực nào thực sự bao gồm, và hệ quả là, chúng tôi cũng cần xác định các yếu tố nào có thể thay đổi.
- Rủi ro & lòng tin: Các vấn đề về tính bền vững khi các dự án sử dụng các PID tăng cường và/hoặc phát triển các cách tiếp cận mới cho PID: Ai duy trì chúng sau khi dự án kết thúc?
- Rủi ro: Tính tương hợp của PID (ở mức EOSC, ví dụ, các hồ sơ cốt lõi của PID so với của URN và DOI)
- Lòng tin: Các vấn đề chưa được giải quyết ở trên có các tác động tiêu cực nghiêm trọng lên sử dụng PID

Các vấn đề PID đáng để khai phá:

- Diễn đàn PID và toàn cảnh PID quốc gia
 - Nhóm DOI DataCite Phần Lan
 - CSC một thành viên của nhóm EPIC

- Thư viện Quốc gia Phần Lan tích cực cung cấp dịch vụ URN
- Các PID được sử dụng và có thể được sử dụng nhiều hơn trong tương lai trong hệ thống CRIS quốc gia *research.fi*
- Dịch vụ PID Micro ở CSC

6.2.2 Pháp (Trung tâm Nghiên cứu Khoa học Quốc gia – CNRS)

Các tổ chức khoa học và các cơ quan có thẩm quyền của Pháp bị thuyết phục rằng việc áp dụng các mã nhận diện quốc tế, ví dụ, các mã nhận diện cho các cá nhân, tổ chức, xuất bản phẩm, phần mềm, và các đối tượng số và dữ liệu nghiên cứu, chỉ có thể làm lợi cho tính trực quan và tác động của kết quả đầu ra khoa học của Pháp. Kế hoạch hành động quốc gia của Bộ Giáo dục Đại học sẽ tập trung vào các lĩnh vực sau:

- Tăng tốc áp dụng các mã nhận diện đối với các nhà nghiên cứu và các cơ sở nghiên cứu, và các tay chơi khác trong các hoạt động nghiên cứu (các nhà cấp vốn, các cơ quan đánh giá và thẩm định, dữ liệu thư mục, các kho lưu trữ mở, các nhà xuất bản, .v.v.)
- Xác định các mô hình kinh doanh khả thi để đảm bảo phát triển bền vững các mã nhận diện,
- Cải thiện tính tương hợp và tiêu chuẩn hóa các mã nhận diện trong khi đảm bảo cộng đồng khoa học giữ kiểm soát,
- Đóng góp cho việc giám sát và đánh giá các hệ thống nhận diện để đảm bảo tính mở và tính độc lập trong dài hạn.

Kế hoạch hành động chi tiết sẽ được phác thảo bởi Ủy ban Khoa học Mở đặc biệt, báo cáo cho Bộ Giáo dục Đại học. Kế hoạch này sẽ tập trung vào sự phối hợp của các tác nhân chính phủ, các tổ chức, các nhà nghiên cứu và các nhà vận hành dịch vụ. Nó sẽ thúc đẩy triển khai các hành động cụ thể để khuyến khích sử dụng các công cụ và các hệ thống nhận diện và duy trì sự kiểm tra cập nhật tình hình về mức áp dụng của chúng. Nó phải gắn với sự phát triển các dịch vụ hữu ích dựa vào các trường hợp sử dụng để thể hiện những lợi ích của các PID. Cuối cùng, nó phải phát triển các hành động truyền thông phù hợp hướng tới các nhà nghiên cứu để hướng dẫn họ trong sử dụng các mã nhận diện hàng ngày.

Bốn hành động riêng biệt sẽ triển khai về:

1. nhận diện các tổ chức,
2. nhận diện các cá nhân,
3. nhận diện các xuất bản phẩm,
4. nhận diện các dữ liệu và đối tượng số.

Hành động 1 - Các mã nhận diện cho tổ chức

Việc xác định các cấu trúc nghiên cứu và các liên kết của chúng vẫn đang được phát triển trên quy mô toàn cầu. Ngay cả nếu các tiêu chuẩn quốc tế còn chưa được xác định đầy đủ, là quan trọng để làm cho các hệ thống quốc gia tương thích với nhau và với phần còn lại của thế giới.

Tới nay, Danh mục Quốc gia về Cơ cấu Nghiên cứu - RNSR (Répertoire National des Structures de Recherche) dường như là công cụ quốc gia thành công nhất để nhận diện các tổ chức nghiên cứu. Thư mục AureHAL được Trung tâm Truyền thông Khoa học trực tiếp – CCSD (Centre pour la communication scientifique directe) xây dựng là chi tiết hơn so với RNSR vì nó xác định các nhóm nghiên cứu. Kết nối các mã nhận diện quốc gia về các cấu trúc với các mã nhận diện quốc tế như ISNI hoặc ROR hoặc Q Wikidata là công việc đang tiến hóa được cơ quan biên mục giáo dục đại học (Abes) nắm giữ trong tổ chức đăng ký toàn cầu được giám tuyển tốt của nó *IdRef.fr*, và bởi Thư viện Quốc gia Pháp như một cơ quan cho ISNI, trong số các tay chơi khác. Bộ Giáo dục Đại học Pháp điều phối tất cả các bên liên quan của quốc gia.

Hành động 2 - Các mã nhận diện cho các cá nhân

Việc áp dụng ORCID của quốc tế ngày một gia tăng như một mã nhận diện cho các nhà nghiên cứu sẽ làm cho chính sách của Pháp hướng tới sử dụng có hệ thống các mã nhận diện tiêu chuẩn cho các tác giả và những người đóng góp cho khoa học khắp các ngành khoa học khác nhau.

Ở mức vận hành, điều này sẽ gồm việc:

- Xúc tác cho áp dụng rộng nhất có thể ORCID đối với các nhà nghiên cứu Pháp,

- Cung cấp bằng chứng về tính hữu dụng của ORCID qua các trường hợp sử dụng trong các trường đại học và các cơ sở và các hệ thống thông tin của họ,
- Cung cấp truyền thông và các công cụ và tư liệu đào tạo nhằm vào các nhà nghiên cứu,
- Đảm bảo rằng các nhà nghiên cứu của Pháp, cả sống và chết, có một IdRef, một ISNI, và một Q Wikidata, phù hợp với ORCID, IdHAL hoặc Researcher ID của họ khi sẵn sàng.

Vào năm 2019, Bộ Giáo dục Đại học, Nghiên cứu và Đổi mới Pháp, trong khung công việc của Ủy ban về Khoa học Mở, đã chỉ thị cho nhóm Couperin thiết lập một quan hệ đối tác chung với ORCID nhân danh các cơ sở nghiên cứu và giáo dục đại học. Cơ quan thư mục cho giáo dục đại học - ABES (Agence bibliographique de l'enseignement supérieur) có trách nhiệm về điều phối kỹ thuật của Nhóm ORCID Pháp.

Hành động 3 - Các mã nhận diện cho các xuất bản phẩm

Các bài báo khoa học là chủ yếu hiện nay - nhưng không chỉ - được nhận diện qua các DOI, cơ quan có trách nhiệm chính là Crossref. Crossref đã phát triển các dịch vụ bổ sung, ví dụ, liên kết chéo trích dẫn. Tuy nhiên, các DOI cũng có thể được các cơ quan khác gắn cho các xuất bản phẩm, ví dụ, DataCite, là các thành viên của Quỹ DOI Quốc tế có trụ sở ở Mỹ. Ngoài ra, các mã nhận diện khác được các cơ quan nghiên cứu của Pháp gắn cho các sử dụng cụ thể, như HAL ID (HAL/CCSD) hoặc Handle ID (Isidore/CNRS).

Hầu hết các xuất bản phẩm của Pháp đã được gắn rồi các mã nhận diện, mà nó sẽ giúp đánh giá tầm quan trọng của từng mã nhận diện để thông tin cho chiến lược quốc gia. Ngoài ra, sự điều chỉnh phù hợp giữa các mã nhận diện khác nhau đó chỉ được hiện thực hóa một phần và nên được khái quát hóa bởi sự hợp tác giữa các bên liên quan khác nhau (CCSD, Abes, ScanR, HumaNum). Cuối cùng nhưng không kém phần quan trọng, các trích dẫn mở là mong muốn cao để giám sát tiến bộ khoa học và một báo cáo về chủ đề này sẽ được các nhà chức trách Pháp ủy quyền sớm.

Bộ Giáo dục Đại học Pháp hỗ trợ Trung tâm Quốc tế ISSN, một tổ chức liên chính phủ là Cơ quan Đăng ký ISO để quản lý tiêu chuẩn ISO 3297:2020 - ISSN xác định các xuất bản phẩm trên các tạp chí và các tài nguyên liên tục. Tổ chức này nhận vốn cấp từ các

nhà chức trách Pháp và hơn 90 quốc gia thành viên. ISSN được sử dụng rộng rãi để nhận diện các xuất bản phẩm in và số trên tạp chí và các tài nguyên liên tục bởi các thư viện và các nhà xuất bản. Cổng ISSN (portal.issn.org) cung cấp dữ liệu nhận diện tự do không mất tiền và các URL cho các xuất bản phẩm trên tạp chí. Các dịch vụ bổ sung cung cấp dữ liệu về chuyển giao các đầu tạp chí và việc lưu trữ số. ISSN IC hiện đang nghiên cứu triển khai bộ phân giải URN (urn.issn.org).

Hành động 4 - Các mã nhận diện cho dữ liệu và các đối tượng nghiên cứu

Vì sự phức tạp và đa dạng của dữ liệu sẽ được quản lý và việc sử dụng của nó, việc nhận diện các dữ liệu nghiên cứu còn chưa chín muồi. Các hệ thống khác nhau tồn tại, chủ yếu dựa vào DOI và Handle, nhưng các nhu cầu của các cộng đồng và cách để đáp ứng chúng, còn chưa được xác định.

Sự phản ánh trong các nhóm khác nhau đã dẫn đến sự phát triển khái niệm Đối tượng Số (Digital Object) được nhóm chuyên gia quốc tế C2CAMP phát triển một phần. Phạm vi của các thực thể được xác định thực tế là vô hạn. Một sáng kiến quốc tế theo đó Pháp tham gia đã phát triển một hệ thống nhận diện phần mềm có hứa hẹn (<https://www.softwareheritage.org>). Một nút của Pháp trong Liên minh Dữ liệu Nghiên cứu tham gia vào trong dự án. Sự phản ánh được kích hoạt ở mức châu Âu với sự triển khai hiện hành của EOSC và các chỉ thị mới của Horizon Europe (<https://oaamu.hypotheses.org/2722?s=09> ở Pháp). Các kế hoạch hành động cụ thể dự kiến sẽ được đề xuất sớm.

6.2.3 Đan Mạch (DeiC)

Đan Mạch đã hình thành một nhóm quốc gia thuộc DataCite, nơi tất cả các chi phí của DOI được trang trải ở mức quốc gia bởi Tập đoàn hạ tầng điện tử Đan Mạch (DeiC), bên lãnh đạo nhóm này, đại diện cho tất cả các thành viên của nhóm ở mức quốc tế.

Ở Đan Mạch nhóm ORCID quốc gia đầu tiên (<http://orcid.dk/>) đã được thành lập từ 2014. Nhóm này do Đại học Aalborg lãnh đạo. Cách tiếp cận của nhóm quốc gia đối với ORCID và DataCite được các tổ chức nghiên cứu Đan Mạch chấp nhận rộng rãi.

Thư viện Hoàng gia quốc gia Đan Mạch sử dụng DOI từ Crossref trong nền tảng xuất bản của họ cho các tạp chí điện tử. Nó sử dụng nền tảng xuất bản được sử dụng rộng

rãi là Hệ thống Tạp chí Mở - OJS (Open Journal Systems), đây là một chương trình cho phép xuất bản trên trực tuyến nhanh chóng và dễ dàng.

Thư viện Hoàng gia cũng có một dịch vụ xuất bản các sách điện tử dựa vào một máy chủ của Nhà xuất bản Sách chuyên khảo Mở - Open Monograph Press server (<http://ebooks.au.dk/index.php/aul>), sử dụng DOI từ Crossref.

Đại học Aalborg cũng sử dụng OJS nhưng với DOI từ DataCite.

Đan Mạch sớm thành lập Danpid, một dịch vụ PID của Đan Mạch, nó được sử dụng để tạo ra các đường liên kết cho các đối tượng, ví dụ, trong các thư viện và viện bảo tàng quốc gia Đan Mạch hoặc các đối tượng được chính phủ, các tổ chức và các cơ sở giáo dục ở thành phố xuất bản. “Handle” là tên của hạ tầng theo đó Danpid dựa vào (hạ tầng y hệt cũng được DOI sử dụng).

Đan Mạch có lịch sử lâu đời trong các sáng kiến bảo tồn số. Một sáng kiến nổi bật là Bảo tồn Số (<https://digitalbevaring.dk>), một dự án hợp tác của Kho lưu trữ Quốc gia Đan Mạch cùng với Thư viện Hoàng gia. Nó bây giờ đã kết thúc, việc tái cấu trúc lưu trữ và các hạ tầng PID để có sự liên kết mạnh hơn với các trường đại học. Tuy nhiên, Kho Truy cập Mở của Thư viện - LOAR (Library Open Access Repository), trực thuộc Thư viện Hoàng gia, là một kho dữ liệu mở được thành lập năm 2016 như một dịch vụ lưu trữ và cung cấp truy cập tới dữ liệu nghiên cứu của Đan Mạch. Dịch vụ này đã được hình thành với các mục tiêu chính và động lực chính sách sau đây:

- Làm cho dữ liệu truy cập được cho việc rà soát lại các xuất bản phẩm
- Xúc tác cho các nhà nghiên cứu đáp ứng các yêu cầu đối với các trợ cấp của Đan Mạch và châu Âu
- Đảm bảo quyền riêng tư của dữ liệu và loại bỏ dữ liệu một cách phù hợp
- Xúc tác cho việc sử dụng lại dữ liệu khi thích hợp

Các nhà nghiên cứu tải dữ liệu lên được kỳ vọng chia sẻ dữ liệu đó bằng việc sử dụng các giấy phép Creative Commons.

6.2.4 Đức (DFG)

Rủi ro chính: Áp dụng các PID và thỏa thuận về các tiêu chuẩn quốc tế

Bối cảnh hạ tầng nghiên cứu của Đức là rất phân mảnh liên quan tới các giải pháp hệ thống kho, các chính sách quản lý dữ liệu, các tiêu chuẩn siêu dữ liệu và các PID. Một thách thức khác là áp dụng các PID số, ví dụ, DOI, ORCID, ROR trong các quy trình đánh chỉ mục biên mục chính thức. Bằng cách này các URN, được Thư viện Quốc gia Đức quản trị và chỉ định, đóng vai trò quan trọng. Một yếu tố chính khác trong quy trình thu thập là Hồ sơ Cơ quan Tích hợp Đức - GND (German Integrated Authority File), nó mô tả các công việc, các thực thể theo địa lý, các hội nghị và con người. Nhóm ORCID Đức được thành lập năm 2016, do Thư viện thông tin Kỹ thuật (TIB) dẫn dắt và được dự án ORCID DE do DFG cấp vốn, hỗ trợ, đang cải thiện triển khai ORCID ở Đức. ORCID đang trong quá trình được tích hợp vào GND. Các PID được vài nhà hoạch định chính sách của Đức khuyến nghị nhưng không là bắt buộc.

Đức đang trong quá trình xây dựng hạ tầng dữ liệu nghiên cứu quốc gia (NFDI), nơi các nhóm xử lý các vấn đề đặc thù lĩnh vực bằng việc theo đuổi một chiến lược rộng khắp. Việc tùy chỉnh các PID dựa trên các tiêu chuẩn quốc tế là một trong những mục đích của nhóm (chủ đề xuyên suốt). Vì thế, có thể là trường hợp điển hình tốt để phân tích các vấn đề rủi ro và lòng tin phát sinh với 3 chiều đó.

Việc áp dụng Đăng ký Tổ chức Nghiên cứu (ROR) như một mã nhận diện của tổ chức ở Đức như một hệ thống PID mới sẽ là một trường hợp điển hình tốt khác và kết quả của một khảo sát được tiến hành gần đây khắp bối cảnh nghiên cứu của Đức có thể là điểm khởi đầu tốt. Đặc biệt thú vị là các vấn đề rủi ro và lòng tin có liên quan tới thách thức để vượt qua các cấu trúc khu vực tin cậy nhưng có giới hạn.

Các DOI DataCite đã được cung cấp miễn phí ở Đức từ 2012 cho tới cuối năm 2020, khi mô hình mới cho thành viên và phí DataCite đã được triển khai. TIB đã thành lập một nhóm và nhiều khách hàng trước đó đã trở thành các thành viên nhóm có đóng tiền, đã ra nhập DataCite trực tiếp hoặc chuyển sang Crossref. Nhưng đã có một số tổ chức không còn quan tâm tới DOI nữa. Là thú vị để phân tích các vấn đề rủi ro và lòng tin của các tổ chức dựa vào 3 chiều đó.

Thư viện Quốc gia Đức (DNB) gắn các URN vào các xuất bản phẩm số, hầu hết ký gửi hợp pháp của mình và cũng chào sử dụng hạ tầng URN của nó cho các đối tác được đăng ký (các cơ sở nghiên cứu, các thư viện và các nhà xuất bản). Dịch vụ đó là miễn phí cho các đối tác và tương tự, những người sử dụng đầu cuối. Các đối tác URN thường

quản trị và tự gắn các URN và cũng có trách nhiệm đảm bảo sự thường trực của họ. Đối với các URN được DNB gắn trực tiếp, thư viện lưu trữ một bản sao xuất bản phẩm trong kho lưu trữ số của riêng mình. Một phân tích các vấn đề rủi ro và lòng tin xung quanh các URN có thể có quan tâm để nhấn mạnh các khác biệt giữa các hạ tầng PID nhằm vào các dạng dữ liệu chong chéo nhau.

6.2.5 Vương quốc Anh (JISC)

<https://repository.jisc.ac.uk/8107/1/PIDs%20for%20OA%20project%20community%20survey%20report.pdf>

Các khuyến nghị chính dựa vào các phát hiện của khảo sát là:

- Các can thiệp là cần thiết để cải thiện mức độ phạm vi và chiều sâu tích hợp PID trong các tiến trình công việc hàng ngày
- Những người trả lời muốn thấy các PID đang được sử dụng tối ưu trong các hệ thống cấp vốn (cả cho đơn đề xuất/trao trợ cấp và cho việc báo cáo), các nền tảng nội dung mà tổ chức các kết quả đầu ra nghiên cứu (bao gồm dữ liệu và các sách điện tử cũng như các bài báo), và các công cụ quản lý thông tin nghiên cứu trong các cơ sở
- Siêu dữ liệu có liên quan tới các PID là sống còn. Nó cần phải hiện diện có thể dự đoán trước được, chứa các yếu tố nhất quán hơn, đồng thời được duy trì và cập nhật một cách đáng tin cậy
- Đối với các PID mới hoặc đang nổi lên, có các bài học kinh nghiệm từ các chiến lược tham gia của ROR, điều đã xúc tác cho nó để đạt được tiến bộ rõ rệt
- Các rào cản áp dụng cần phải được hạ thấp.

Một báo cáo khác, được Josh Brown viết sẽ được xuất bản cuối năm 2021, dựa vào công việc của nhóm các bên liên quan tới PID của Vương quốc Anh, nó đã được thiết lập với sự khuyến khích của UKRI.

Các PID đáng khai phá:
<https://scholarlycommunications.jiscinvolve.org/wp/author/manistaf/>

Dự án lộ trình Truy cập Mở/PID đã được chuyển cho một chuyên gia tại JISC. JISC đang dẫn đầu công việc tiếp theo sau trong cộng tác với MoreBrains Cooperative, LTD. Chuyên gia này cũng sẽ có khả năng cập nhật về kế hoạch chiến lược tài nguyên và hạ tầng nội bộ xung quanh các PID.

Địa điểm mới của diễn đàn PID với NISO sau khi kết thúc của dự án Freya: <https://www.project-freya.eu/en/blogs/blogs/a-new-home-for-the-pid-forum-at-niso>

Các mã nhận diện ORCID cho mọi người, Crossref và DataCite DOI cho các kết quả đầu ra, Crossref trao DOI, các mã nhận diện ROR cho các tổ chức, và các RAiD cho dự án.

Họ muốn thiết lập một chiến lược PID quốc gia dựa vào lộ trình 2019 (được thảo luận trong các bài đăng trên các blog đó)

6.2.6 Hà Lan (SURF)

Tại Hà Lan, có vài dịch vụ PID được nhúng vào các trường hợp sử dụng đặc thù. Từng dịch vụ (xem bên dưới) vận hành độc lập với các dịch vụ khác với mức phối hợp thấp giữa chúng. SURF hiện đang phát triển một khung quốc gia để gia tăng sự phối hợp PID và khai phá các cách thức để nâng cao sự tương tác.

Nhìn qua bức tranh tổng thể PID của Hà Lan:

ORCID	ID nhà nghiên cứu, người đóng góp (Nhóm ORCID Hà Lan)
ISNI	ID tác giả, các thư viện đại học (người đăng ký)
DOI: Dữ liệu HSS	EASY cho các tập hợp dữ liệu HSS (DANS)
DOI: Dữ liệu HSS	DataverseNL , đang nghiên cứu (DANS)
DOI: dữ liệu	Datacite (41 tài khoản kho qua 4TU)
ePIC: data	Handle cho các tập hợp dữ liệu trong nghiên cứu (SURF)
varius: đối tượng/dữ liệu	Di sản văn hóa số (hướng dẫn pid)
PURL: đối tượng/dữ liệu	Các đối tượng/các bộ sưu tập đa dạng sinh thái
URN: NBN	Các xuất bản phẩm, Thư viện Quốc gia Hà Lan

Rủi ro

Trong khi cộng đồng được các tổ chức hỗ trợ PID điều hành, và trong một vài trường hợp việc chấp nhận, phát triển hạ tầng nghiên cứu mở, những đóng góp từ các công ty xuất bản/dữ liệu vì lợi nhuận làm phức tạp nỗ lực này. Rõ ràng, đây không phải là sự căng thẳng giữa các tác nhân tốt và xấu. Thay vào đó, đây là vấn đề về các nguyên tắc.

Chính sự căng thẳng liên tục giữa các lợi ích của cộng đồng và các lợi ích của các bên liên quan, sự căng thẳng đó tạo ra sự không chắc chắn về tính bền vững dài hạn của các tiến trình PID phụ thuộc vào sự truy cập dài hạn tới siêu dữ liệu có liên quan.

Một mối quan tâm liên quan là gói siêu dữ liệu thương mại (bao gồm cả PID) với một bộ các dịch vụ được tích hợp chặt chẽ. Ví dụ, sự kết hợp của CRIS, thuê bao dữ liệu, và các phần mềm phân tích tạo ra một dạng khóa trói vào nhà cung cấp.

Giảm nhẹ các rủi ro

Nhận thức ở mức quốc gia của các trường đại học. Ngăn ngừa việc bị khóa trói vào nhà cung cấp về khía cạnh các hệ thống tri thức.

Sự chú ý đặc biệt cho sử dụng các ứng dụng được các nhà cung cấp thương mại và các tập đoàn công nghệ lớn cung cấp. Một tập hợp các nguyên tắc được xác định sẽ là một phần của các thương thảo với các bên liên quan đóng vai trò trong các hệ thống kiến thức. Ở mức địa phương, các công nghệ nguồn mở là một phần của vài chương trình Khoa học Mở của các trường đại học.

7 Tóm tắt các trường hợp PID của KE

Được đề xuất rằng một chuyên gia tư vấn hợp tác chặt chẽ với đại diện các quốc gia thành viên của KE và Nhóm Nhiệm vụ & Hoàn thành PID của KE phác thảo một tóm tắt các trường hợp PID của KE.

7.1 Rủi ro

- Các rủi ro được thừa nhận ở mức tích lũy tổng thể là gì?
- Các rủi ro trong hạ tầng PID hiện hành có thể được hiểu như thế nào, dưới ánh sáng của các chiều phân tích khác nhau (Khung Uyên thâm Mở)?

- Rủi ro có thể được giảm nhẹ như thế nào (ví dụ, bằng việc xây dựng một quỹ ổn định được xây dựng dựa vào lòng tin, hoặc một tổ chức quốc tế, các thỏa thuận, các tiêu chuẩn; hợp đồng .v.v.)?

7.2 Lòng tin

- Khái niệm về lòng tin có liên quan như thế nào, với việc hoạt động tốt của một hạ tầng PID tương hợp được và được quốc tế thừa nhận trong tương lai?
- Lòng tin kéo theo sự chấp nhận ‘một vài’ rủi ro

8 Kết luận các vấn đề về rủi ro và lòng tin của PID

Được đề xuất rằng một chuyên gia tư vấn hợp tác chặt chẽ với các kết luận phác thảo của Nhóm Nhiệm vụ & Hoàn thành PID về phân tích được tiến hành này.

9 Các khuyến nghị

Được đề xuất rằng một chuyên gia tư vấn hợp tác chặt chẽ với các khuyến nghị phác thảo của Nhóm Nhiệm vụ & Hoàn thành PID trong ngắn hạn và dài hạn đối với hành động tiếp theo sau của KE nhằm vào các tổ chức đối tác của KE, cũng như các tổ chức nghiên cứu của các quốc gia thành viên của KE.

~#~

10 Các tài liệu hỗ trợ

10.1 Các dự án và báo cáo liên quan

Các dự án của châu Âu

- CLARIN/DARIAH : several documents available re PIDs (<https://www.clarin.eu/sites/default/files/pid-CLARIN-ShortGuide.pdf>, <https://zenodo.org/record/3744091#.YAqsHOhKg2w>)
 - notably P. Wittenburg. “Persistent end Unique Identifiers”. CLARIN-2008-2.
- FREYA project (https://www.project-freya.eu/en/deliverables/freya_d3-1.pdf)

- FREYA PID Federation Scoping Study: Final report (<https://zenodo.org/record/4059557>)
- THOR project <https://project-thor.readme.io/docs/project-deliverables>
- ODIN project <https://odin-project.eu/project-outputs/deliverables/>
- DEFF Opera project: <https://deffopera.dk/>
- PID Graphs: <https://doi.org/10.1016/j.patter.2020.100180>

EOSC

- Hellström, Maggie, André Heughebaert, et al. ‘Second Draft Persistent Identifier (PID) Policy for the European Open Science Cloud (EOSC)’. 2020, May 1. <https://doi.org/10.5281/zenodo.3780423>. Or <https://www.eoscsecretariat.eu/eosc-liaison-platform/post/launch-initial-persistent-identifier-policy-eosc>
- A Persistent Identifier (PID) policy for the European Open Science Cloud (EOSC): <https://op.europa.eu/da/publication-detail/-/publication/35c5ca10-1417-11eb-b57e-01aa75ed71a1/language-en>
- EOSC PID Architecture <https://docs.google.com/document/d/1T-bpNsmuxQewsLq48XTyUJoe0lsV7poaXohpgDo9W34/edit>

Liên minh Dữ liệu Nghiên cứu

- RDA PID Kernel WG (<https://www.rd-alliance.org/groups/pid-kernel-information-profile-management-wg>)
- Research Data Alliance/FORCE11 Software Source Code Identification WG; Allen, Alice; Bandrowski, Anita; Chan, Peter; di Cosmo, Roberto; Fenner, Martin; Garcia, Leyla; Gruenpeter, Morane; Jones, Catherine M.; Katz, Daniel S.; Kunze, John; Schubotz, Moritz; Todorov, Ilian T. “Software Source Code Identification Use cases and identifier schemes for persistent software source code identification”. 2020. <https://zenodo.org/record/4312464#.X9BuzGgza70>
- Wittenburg, Peter, Hellström, Margareta, Zwölf, Carlo-Maria, Abroshan, Hossein, Asmi, Ari, Di Bernardo, Giuseppe, Couvreur, Danielle, Gaizer, Tamas, Holub, Petr,

Hoof, Rob, Häggström, Ingemar, Kohler, Manfred, Koureas, Dimitris, Kuchinke, Wolfgang, Milanesi, Luciano, Padfield, Joseph, Rosato, Antonio, Staiger, Christine, van Uytvanck, Dieter, and Weigel, Tobias (2017) “Persistent identifiers: Consolidated assertions. Status of November 2017.”
<https://doi.org/10.5281/zenodo.1116189>

FAIRsFAIR WP2+WP3

- FAIRsFAIR 2nd Report on FAIR requirements for persistence and interoperability
<https://doi.org/10.5281/zenodo.4001631>

JISC

- Developing a persistent identifier roadmap for open access to UK research (https://repository.jisc.ac.uk/7840/2/PID_roadmap_for_open_access_to_UK_research.pdf)
- UK PID Consortia: UK Open Access to Research publications - Adam Tickell (2018) (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/774956/Open-access-to-research-publications-2018.pdf)

KE’s report on the Openness Profile:
<https://www.knowledge-exchange.info/event/openness-profile>

10.2 Các tài nguyên khác

- Pidforum.org <https://www.pidforum.org/>
- Pid service registry <https://pidservices.org/>
- Fair Principles: <https://www.go-fair.org/fair-principles/>
- Systematize information on journal policies and practices - A call to action <https://leidenmadtrics.nl/articles/systematize-information-on-journal-policies-and-practices-a-call-to-action>
- PID Commons (Datacite) <https://blog.datacite.org/power-of-pids/>

- Choosing and implementing persistent identifiers: Guide for research organisations <https://doi.org/10.5281/zenodo.4395767>
- The use of Persistent Identifiers for Research Datasets: Recommendation by the Finnish Scientific Community for Open Research <https://doi.org/10.5281/zenodo.3560738>
- GAIA-X: Technical Architecture: Release June 2020 <https://www.data-infrastructure.eu/GAIA/Redaktion/EN/Publications/gaia-x-technical-architecture.pdf>

10.3 Tham khảo

- C. Allen. Self-sovereign identity principles. 2016. <https://github.com/ChristopherA/self-sovereign-identity/blob/master/self-sovereign-identity-principles.md>
- I. Barclay, A. Preece, I. Taylor et al. Certifying provenance of scientific datasets with Self-sovereign identity and verifiable credentials. (2020). <https://arxiv.org/pdf/2004.02796>
- A. Dappert et al. “Connecting the persistent identifier ecosystem: Building the technical and human infrastructure for open research”. 2017. <https://datascience.codata.org/articles/10.5334/dsj-2017-028/>
- P. Golodoniuc et al. “PID Service. An advanced persistent identifier management service for the semantic web”. 2015. https://www.researchgate.net/profile/Pavel_Golodoniuc/publication/284087065_PID_Service_-_an_advanced_persistent_identifier_management_service_for_the_Semantic_Web/links/564bf3a908ae3374e5ddec58/PID-Service-an-advanced-persistent-identifier-management-service-for-the-Semantic-Web.pdf
- Martin Klein, Lyudmila Balakireva. “On the Persistence of Persistent Identifiers of the Scholarly Web”. 2020. <https://arxiv.org/pdf/2004.03011.pdf>

- Jens Klump, Robert Huber. “20 years of persistent identifiers. Which systems are here to stay?” 2017. <https://datascience.codata.org/articles/10.5334/dsj-2017-009/>
- Alexander Mühle et al. “A Survey on Essential Components of a Self-Sovereign Identity”, 2018. <https://arxiv.org/abs/1807.06346>
- N. Naik & P. Jenkins. Self-sovereign identity specifications: Govern your identity through your digital wallet using Blockchain technology. 2020. https://publications.aston.ac.uk/id/eprint/41998/1/SSI_Specifications_uPort_SovrinDrNitinNaik.pdf
- Aurelia Vasile et al. “Le DOI, une impérieuse nécessité? L'exemple de l'attribution de DOI à la collection Pangloss, archive ouverte de langues en danger”. 2020. https://halshs.archives-ouvertes.fr/halshs-02870206/file/DOI_VersionAuteur_HAL.pdf