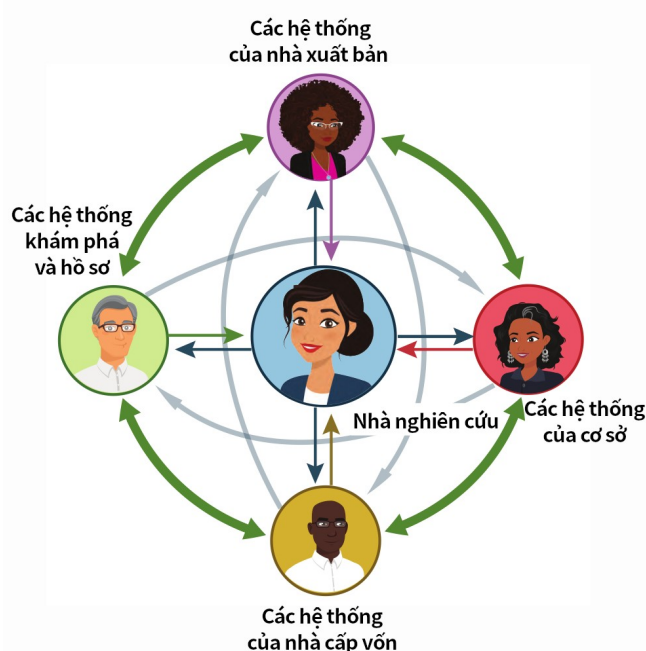


Vòng đời nghiên cứu được tối ưu hóa bằng các mã nhận diện thường trực (PID)

Mục đích của ‘Hồ sơ Tính mở’[1] là để giúp cho các nhà nghiên cứu và những người đóng góp vào nghiên cứu được đánh giá chuẩn xác nhất về kết quả làm việc của họ, qua đó có được những cơ hội về tài trợ, thăng tiến sự nghiệp, lương bổng và quan trọng là có thể tiến hành các hoạt động nghiên cứu một cách thuận tiện nhất, tiết kiệm nhiều thời gian nhất có thể. Một trong các đặc tính cốt lõi và thú vị nhất của Hồ sơ Tính mở là khả năng tích hợp và kết nối (tương tác) với các (hệ thống) mã nhận diện thường trực - PID (Persistent Identifier) khác nhau, đặc biệt là xung quanh ORCID, và vì thế mở rộng khả năng tham chiếu tới nhiều đối tượng nghiên cứu khác cũng như các thông tin/dữ liệu cùng với các công cụ ứng dụng/dịch vụ đi kèm với chúng trong môi trường (kỹ thuật) số qua Internet. Trên thực tế, một số nơi đã và đang triển khai cách tiếp cận tích hợp và tương tác này.

ORCID với hướng dẫn các quy trình để tích hợp với nó

Để có sự tích hợp và tương tác của ‘Hồ sơ Tính mở’/ORCID với các bên liên quan khác, ORCID đã đưa ra nhiều hướng dẫn cụ thể các quy trình tích hợp và tương tác[2] dựa vào các giao diện lập trình ứng dụng - API (Application Programming Interface), chẳng hạn như các quy trình cho: các quy trình làm việc; bài báo trên tạp chí; sách; máy chủ preprint; bình duyệt ngang hàng; giáo dục; việc làm; tài nguyên nghiên cứu; nhà cấp vốn và các trợ cấp; các hệ thống kho tài nguyên; cùng nhiều tiện ích/hướng dẫn khác. Tất cả các quy trình đó đều có mục đích để thể hiện các mối liên hệ lấy nhà nghiên cứu làm trọng tâm của ORCID với các bên liên quan chính khác[3], như trên **Hình 1**.



Hình 1. Các mối liên hệ lấy nhà nghiên cứu làm trọng tâm của ORCID

Thực tế sử dụng ORCID cho thấy, để tích hợp và tương tác được với sự đa dạng của các bên liên quan và các kết quả đầu ra nghiên cứu, ORCID đã tích hợp với hàng chục mã nhận diện thường trực khác nhau đủ các thể loại, như trên **Hình 2**, rõ ràng là công việc không hề dễ dàng. Điều này càng đáng chú ý khi người ta cân nhắc về chi phí/lợi ích và rủi ro/lòng tin vào các PID đó trong trung và dài hạn, đặc biệt trong bối cảnh bức tranh ứng dụng và phát triển các PID hiện đang tiến hóa và bị phân mảnh, khả năng triển khai các PID lỗi và PID không tin cậy[4] có thể gây ra các hậu quả không đáng có, thậm chí có những PID ở mức quốc gia dù trước đó là hữu ích, nhưng đã bị thay thế bởi các PID với các tiêu chuẩn toàn cầu đáp ứng các yêu cầu bảo tồn lâu dài, không bị phụ thuộc và/hoặc không bị kiểm soát bởi bất kỳ công ty hay tập đoàn nào, khi được cộng đồng sở hữu và cộng đồng điều hành, như trường hợp Mã nhận diện Tác giả Số - DAI (Digital Author Identifier) của Hà Lan[5] bị thay thế bởi ORCID/ISNI.

agr: Agrícola	authenticusid: AuthenticusID	cit: CiteSeer	emdb: Electron Microscopy Data Bank	handle: Handle
ark: Archival Resource Key Identifier	Bibcode	cstr: Science and technology resource identificat...	emplar: Electron Microscopy Public Image Archive	hir: NHS Evidence
arxiv: ARXiv	cba: Chinese Biological Abstracts	ctc: CiteExplore submission	ethos: ETHOS Persistent ID	isbn: International Standard Book Number
asin-tld: ASIN top-level domain	cgn: Culturegraph Number	dnb: German National Library identifier	grant number	ismn: International Standard Music Number
asin: Amazon Standard Identification Number	cienciauiol: Ciência-IUL Identifier	doi: Digital object identifier	hal: Hyper Articles en Ligne	issn: International Standard Serial Number. Incl...
jfm: Jahrbuch über die Fortschritte der Mathema...	lensid: Lens ID	osti: Office of Scientific and Technical Information	pmid: PubMed Unique Identifier	Scopus Identifier
jstor: JSTOR abstract	mr: Mathematical Reviews	Other identifier type	ppr: Europe PMC Preprint Identifier	ssrn: Social Science Research Network
k10plus: K10plus	Non-standard ID from work data source	Patent number	Proposal ID	uri: URI
kuid: KoreaMed Unique Identifier	oclc: Online Computer Library Center	pdb: Protein Data Bank identifier	rfc: Request for Comments	urn: URN
lccn: Library of Congress Control Number	ol: Open Library	pmc: PubMed Central article number	rrid: Research Resource Identifier	wosuid: Web of Science™ identifier
				zbl: Zentralblatt MATH

Hình 2. Vài chục mã nhận diện thường trực được ORCID hỗ trợ, bao gồm DOI

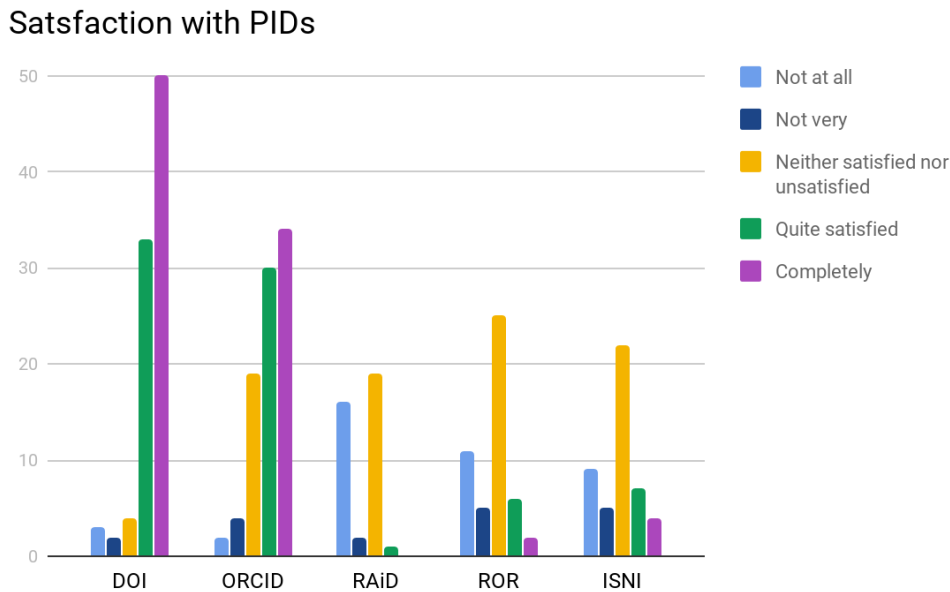
Để giải quyết sự đa dạng phức tạp nêu trên, một vài nơi đã chọn cách tiếp cận tập trung vào một vài mã nhận diện thường trực vừa có mức độ phổ biến càng rộng càng tốt, vừa đáp ứng được các tiêu chuẩn cũng như nhu cầu của cộng đồng các nhà nghiên cứu và những người đóng góp cho nghiên cứu.

Dự án khảo sát các PID được ưu tiên của Vương quốc Anh

Dự án này do Ủy ban các Hệ thống Thông tin Chung - JISC (Joint Information Systems Committee) của Vương quốc Anh triển khai suốt năm 2020 với mục đích để phát triển và tinh chỉnh chiến lược và lộ trình của Vương quốc Anh đối với các PID nhằm cải thiện quyền truy cập tới nghiên cứu của mình. Như một phần của dự án này, nhóm đã phát hành một khảo sát[6] trên trực tuyến để điều tra các mức hiện hành về nhận thức, áp dụng và tích hợp PID vào các hệ thống chủ chốt dành cho nghiên cứu, dựa vào 5 PID được ưu tiên: (1) cho các trợ cấp (các mã nhận diện Đối tượng Số - DOI, do Crossref cung cấp); (2) cho các tổ chức (mã nhận diện Đăng ký Tổ chức Nghiên cứu - ROR); (3) cho các kết quả đầu ra (các mã nhận diện Đối tượng Số - DOI, do Crossref và DataCite cung cấp); (4) cho con người (Mã nhận diện Người đóng góp và Nhà nghiên cứu Mở -

ORCID); (5) và cho các dự án (Mã nhận diện Hoạt động Nghiên cứu - RAiD, do Research Data Commons của Úc cung cấp).

Kết quả khảo sát cho thấy mức độ thỏa mãn của người sử dụng đối với các PID được ưu tiên là như trên **Hình 3**, và nhấn mạnh rằng Đăng ký Tổ chức Nghiên cứu – ROR (Research Organisation Registry) đã có sự tiến bộ nổi bật về các khía cạnh cả áp dụng và nhận thức trong khoảng thời gian khá ngắn (từ 2019), tiến gần hoặc vượt qua các mã nhận diện tổ chức hiện có.

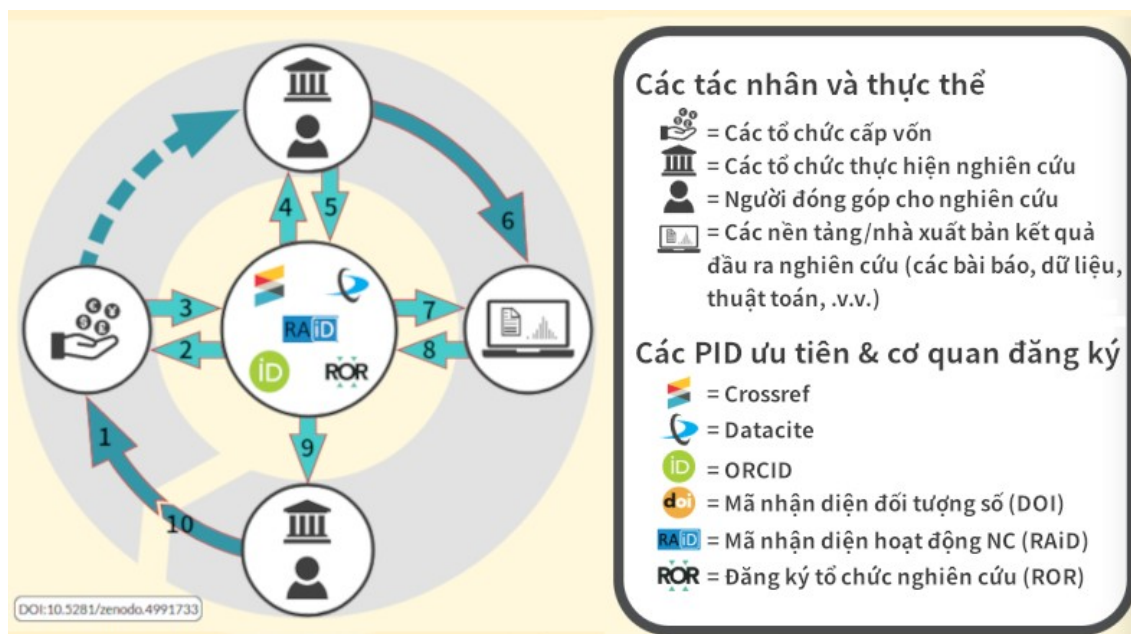


Hình 3. Mức độ thỏa mãn với các PID của người sử dụng

Vòng đời nghiên cứu được tối ưu hóa bằng các mã nhận diện thường trực

Các PID được ưu tiên trong cách tiếp cận của Vương quốc Anh ở trên gần trùng khớp với các PID trong giải pháp của MoreBrains về ‘Vòng đời nghiên cứu được tối ưu hóa bằng các mã nhận diện thường trực’[7], như minh họa bằng lược đồ trên **Hình 4**.

Lược đồ này trình bày vòng đời nghiên cứu được tối ưu hóa bằng các PID, nơi các PID - cho tất cả con người, địa điểm, dự án, và các đối tượng có liên quan tới nghiên cứu - được đăng ký, được sử dụng, và được chia sẻ ở từng giai đoạn được đánh số thứ tự từ 1 đến 10, có tính tới lợi ích cho tất cả các bên tham gia trong từng giai đoạn, như trong **Bảng 1**. Các PID được ưu tiên, dựa vào nhu cầu đầu vào từ cộng đồng. Cùng với các siêu dữ liệu của chúng, các PID được các nhà cấp vốn, các cơ sở nghiên cứu, và các nhà xuất bản thu thập và sử dụng tại thời điểm sớm nhất có thể trong quá trình đó.



Hình 4. Lược đồ Vòng đời nghiên cứu (NC) được tối ưu hóa bằng các PID

Các cơ quan đăng ký PID nằm ở trung tâm cùng hành động như một kho lưu trữ siêu dữ liệu trung tâm có liên quan tới tất cả các đối tượng và các bên liên quan nghiên cứu.

Vòng tròn ngoài gồm các biểu tượng đại diện cho các bên liên quan và những người tham gia chính trong vòng đời nghiên cứu: (1) các tổ chức cấp vốn; (2) các tổ chức thực hiện nghiên cứu; (3) những người đóng góp cho nghiên cứu/nhà nghiên cứu; (4) các nền tảng/nhà xuất bản đầu ra nghiên cứu (các bài báo, dữ liệu, thuật toán, .v.v.).

Bảng 1. Vòng đời nghiên cứu được tối ưu hóa bằng PID qua từng giai đoạn

Giai đoạn	Hành động	Sử dụng PID	Lợi ích
1. Đơn xin trợ cấp	Các nhà nghiên cứu và các cơ sở chuyển các PID về các trợ cấp, kết quả đầu ra, các tổ chức, con người, và các dự án trước đây tới các hệ thống xin trợ cấp.	DOI, ORCID, ROR, RAiD	<ul style="list-style-type: none"> - Ít phải gõ dữ liệu bằng tay - Dữ liệu chính xác hơn - Nhiều thời gian hơn để nhập các thông tin theo ngữ cảnh
2. Xin trợ cấp và bình duyệt	Các nhà cấp vốn nhập dữ liệu về các trợ cấp, kết quả đầu ra, các tổ chức, và các dự án từ các sổ đăng ký PID	DOI, ORCID, ROR, RAiD	<ul style="list-style-type: none"> - Ít phải gõ dữ liệu bằng tay - Dữ liệu chính xác hơn - Nhiều thời gian hơn để nhập thông tin theo ngữ cảnh - Dễ hơn để bình duyệt lựa chọn và công nhận
3. Trao trợ cấp (1)	Các nhà cấp vốn đăng ký các DOI cho các trợ cấp mới và các siêu dữ liệu có liên quan	DOI	<ul style="list-style-type: none"> - Các ID trợ cấp duy nhất toàn cầu và siêu dữ liệu để giúp cho việc báo cáo - Siêu dữ liệu có thẩm quyền và các kết nối tới các PID khác để giúp quản lý dự án - Truyền thông về và tuân thủ với các chính sách được đơn giản hóa (ví dụ, Truy cập Mở, Kế hoạch Quản lý Dữ liệu, .v.v.)

Giai đoạn	Hành động	Sử dụng PID	Lợi ích
4. Trao trợ cấp (2)	Các cơ sở nhập dữ liệu về các trợ cấp và những người và tổ chức có liên quan	DOI, ORCID, ROR	<ul style="list-style-type: none"> - Dữ liệu chính xác hơn - Tiết kiệm thời gian/công sức - Việc báo cáo được cải thiện
5. Đăng ký dự án	Các cơ sở đăng ký các RAiD cho các dự án và bổ sung/cập nhật các đường liên kết cho các trợ cấp, trang thiết bị, con người, và tổ chức có liên quan	DOI, ORCID, ROR, RAiD	<ul style="list-style-type: none"> - Dữ liệu chính xác hơn - Tiết kiệm thời gian/công sức - Tính trực quan của cộng tác xuyên khắp các cơ sở và quốc gia được cải thiện
6. Đệ trình kết quả	Các nhà nghiên cứu chia sẻ ORCID iD của họ khi đệ trình các kết quả đầu ra mới và kết nối các ROR ID với các cơ sở liên kết và cung cấp các DOI về vốn cấp.	DOI, ORCID, ROR	<ul style="list-style-type: none"> - Dữ liệu về các đồng tác giả, cơ sở liên kết, và vốn cấp chính xác hơn - Tiết kiệm thời gian/công sức trong bình duyệt và tạo lập siêu dữ liệu - Đăng nhập dễ dàng hơn - Quản lý các khoản chi phí xử lý bài báo (APC) dễ dàng hơn - Lựa chọn và công nhận người bình duyệt dễ dàng hơn
7. Xuất bản kết quả	Nhà xuất bản/kho tài nguyên nhập dữ liệu về các trợ cấp, con người, dự án, và tổ chức có kết nối với các kết quả đầu ra	DOI, ORCID, ROR, RAiD	<ul style="list-style-type: none"> - Giảm chi phí hành chính cho các nhà xuất bản - Kinh nghiệm của tác giả được cải thiện - Siêu dữ liệu phong phú hơn trợ giúp cho khám phá, phân tích, và báo cáo
8. Đăng ký kết quả	Nhà xuất bản/kho tài nguyên đăng ký các DOI cho các kết quả đầu ra mới và nhập siêu dữ liệu	DOI	<ul style="list-style-type: none"> - Các ID kết quả đầu ra duy nhất toàn cầu và siêu dữ liệu để trợ giúp cho việc phân tích, trích dẫn, khám phá, và báo cáo - Các PID cho các thực thể có liên quan (các tác giả, vốn cấp, .v.v.) được nhúng sẽ hợp lý hóa việc báo cáo và phân tích - Khả năng lớn hơn để theo dõi nghiên cứu và tác động của nội dung
9. Báo cáo (1)	Các cơ quan đăng ký PID gửi bản cập nhật tự động của các ấn phẩm mới, v.v. tới các cơ sở nghiên cứu	DOI, ORCID	<ul style="list-style-type: none"> - Giảm chi phí hành chính cho cơ sở - Dữ liệu kịp thời hơn cho báo cáo và phân tích - Nhập dữ liệu ít thủ công hơn cải thiện tính chính xác - Tiết kiệm thời gian/công sức cho các nhà nghiên cứu
10. Báo cáo (2)	Các nhà nghiên cứu và các cơ sở chuyển các PID về các kết quả đầu ra, các tổ chức, con người, và dự án tới các hệ thống báo cáo của các nhà cấp vốn	DOI, ORCID, ROR, RAiD	<ul style="list-style-type: none"> - Nhập dữ liệu đầu vào ít thủ công hơn - Dữ liệu chính xác hơn - Nhiều thời gian hơn cho việc bổ sung thông tin theo ngữ cảnh

Việc tích hợp các PID và tương tác với các cơ quan đăng ký PID thể hiện cách tiếp cận tương tự như của ‘Hồ sơ Tính mở’. Sự khác biệt trong giải pháp này của MoreBrains nằm ở chỗ nó có thể khắc phục được sự phức tạp mà ORCID gặp phải là thay vì phải làm việc với vài chục PID, giải pháp của MoreBrains chỉ tập trung vào vài PID có tính phổ cập cao theo các tiêu chuẩn toàn cầu của các cộng đồng nghiên cứu, cũng như phải thỏa mãn các yêu cầu được đặt ra cho các hạ tầng nghiên cứu khoa học mở, biết rằng tất cả các tổ chức đăng ký PID như Crossref (DOI), Datacite (DOI), ORCID, ROR, và RAiD đều là các tổ chức phi lợi nhuận, do cộng đồng điều hành.

Ngoài ra, giải pháp của MoreBrains còn mang lại nhiều lợi ích cho tất cả các bên liên quan và những người tham gia chính trong từng giai đoạn của vòng đời nghiên cứu, và vì thế, rất có khả năng, giải pháp này sẽ còn được nhắc đến nhiều trong tương lai.

Các chú giải

- [1] Lê Trung Nghĩa (2023): *Hồ sơ Tính mở - Nhân tố không thể thiếu của Khoa học Mở*: <https://tiasang.com.vn/doi-moi-sang-tao/ho-so-tinh-mo-nhan-to-khong-the-thieu-cua-khoa-hoc-mo/>
- [2] Trang Giáo dục Mở - Tài nguyên Giáo dục Mở: *ORCID - Quy trình làm việc*: <https://giaoducmo.avnuc.vn/khoa-hoc-mo/orcid-quy-trinh-lam-viec-1007.html>
- [3] ORCID: *From Vision to Value: ORCID's 2022–2025 Strategic Plan*: <https://doi.org/10.23640/07243.16687207>. Bản dịch sang tiếng Việt: <https://doi.org/10.5281/zenodo.8213227>
- [4] Pablo de Castro, Ulrich Herb, Laura Rothfritz and Joachim Schöpfel (2023): *Failed PIDs and unreliable PID implementations*: <https://zenodo.org/record/7330527>. Bản dịch sang tiếng Việt: <https://zenodo.org/record/8015956>
- [5] Pablo de Castro, Ulrich Herb, Laura Rothfritz và Joachim Schöpfel (2023): *Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI*: <https://doi.org/10.5281/zenodo.7327505>. Bản dịch sang tiếng Việt: <https://zenodo.org/record/8057937>
- [6] Alice Meadows and Josh Brown (2020): *Persistent identifiers adoption and awareness survey report*: <https://repository.jisc.ac.uk/8107/1/PIDs%20for%20OA%20project%20community%20survey%20report.pdf>. Bản dịch sang tiếng Việt: <https://zenodo.org/record/8105182>
- [7] MoreBrains: *The PID-optimised research cycle*: <https://resources.morebrains.coop/pidcycle/>



Giấy phép nội dung: [CC BY 4.0 Quốc tế](https://creativecommons.org/licenses/by/4.0/)



Lê Trung Nghĩa, ORCID iD: <https://orcid.org/0009-0007-7683-7703>

Bài được biên tập và đăng trên tạp chí Tia Sáng số 18 năm 2023, xuất bản ngày 20/09/2023, các trang 33-35 với tiêu đề: '*Hồ sơ tính mở: Cần các mã nhận diện?*' Phiên bản điện tử của Tia Sáng có [ở đây](#).