

SURF Report

From ORCID Pilot to a PID-centric framework for Research Information

Clifford Tatum [preprint version, not approved by SURF]

Acknowledgments: this report benefits greatly from colleagues' reviews, comments, critiques, and suggestions:

John Dove
Magchiel Bijsterbosch
Gül Akcaova
Pim Slot
Nick Veenstra
Maurice Vanderfeesten
Henk van den Hoogen
Maurice Bouwhuis

Table of contents

1.0 Introduction	2
1.1 Executive Summary	2
2.0 Background, Author IDs in The Netherlands	2
The DAI, a National ID	3
ORCID, an International ID	5
Towards an ORCID pilot	5
3.0 ORCID-NL Pilot Initiative (2016-2017)	7
Proceedings, pilot working group	9
Insights, Lessons Learned	10
4.0 ORCID-NL Business Consortium and new explorations (2018-2019)	15
ORCID-NL working group, first year	16
Integration of ORCID consortium at SURF	18
ID Resolver: a solution to identifying coauthors	20
5.0 Identifiers for FAIR Research Information (2019-2020)	22
FAIR and Open Science: The Duality of PIDs	23
Background	23
NL PID landscape	24
PID Coordination	26
National ORCID Campaign	26
NWO PID Strategy	26
International Engagement	28
Concepts and Investigations	29
Research Activity Identifier (RAiD)	29
Infrastructure ID	29
Openness Profile	31
6.0 Towards a National PID Strategy (2021 and beyond)	32
Open, Community-Owned Scholarly Infrastructure	32
Emerging market structure for data and analytics	33
Next steps	34
Appendix 1, Open Call: ORCID pilot initiative	36

1.0 Introduction

This report brings together the outcomes of several SURF projects aimed at mobilizing use of Persistent Identifiers (PIDs) in The Netherlands. This work was motivated by the opportunity to implement the Open Researcher and Contributor Identifier (ORCID) at research universities, first as a pilot among a few adventurous universities, which was then reformulated as the ORCID-NL consortium in which all research universities are now members. Follow-up projects endeavored to build on the growing installed base of ORCID embedded in university databases for research information (CRISs). Accordingly, the report begins with a brief history of author IDs in the Netherlands (section 2), which is followed by a description of the ORCID-NL pilot initiative (section 3), transition to the ORCID-NL consortium model (section 4), and the Identifiers for FAIR research information (section 5). The final section (6) concludes with concerns about the emerging market structure for data and analytics and present work on a national PID strategy for 2021 and beyond.

1.1 Executive Summary

- TBA

2.0 Background, Author IDs in The Netherlands

[cut from the intro section, perhaps useful here?] In 2005, development of the Digital Author Identifier¹ (DAI) in the Netherlands was inspired in part to improve the validation of Dutch research impact. Seven years later, recognition of the need for a global researcher identifier led to the 2012 launch of the Open Researcher and Contributor Identifier (ORCID).

In 2005, SURF developed and deployed the Digital Author ID (DAI), a persistent identifier for researchers, for all researchers in the Netherlands. This early innovation anticipated content organisation challenges ushered in by the digitization of virtually all aspects of research. Linking a researcher's name to a persistent identifier, such as the DAI or ORCID, enables the possibility to accurately monitor and analyze research outputs associated with this researcher across different databases. Although the DAI was progressive for its time, it was a national ID and therefore its utility was limited to national research information systems. And while the International Standard Name Identifier² (ISNI), an author identifier used in library cataloging workflows, was considered as a potential DAI replacement, it was not a good fit for the purposes of tracking researcher outputs over the course of a research career. This researcher-centric context called for the possibility of adding information to the identifier, such as affiliation(s), research funding, and published outcomes. Ultimately, The Netherlands adopted a hybrid approach towards the adoption of researcher identifiers³, in

¹ https://en.wikipedia.org/wiki/Digital_Author_Identifier

² <https://isni.org>

³ A URL for the "Nota" Author ID report (2013) is presently not available

which both ISNI and ORCID would fill this need. SURF was subsequently asked to consider the feasibility of implementing ORCID in close collaboration with the university research libraries. This section of the report provides a background of author IDs in the Netherlands.

The DAI, a National ID

The DAI provided a unique identification number for Dutch academic authors, which was validated with two sources: the Netherlands Author Thesaurus (NTA) and the METIS research database used by most or all Dutch organizations conducting academic research. A DAI working group of stakeholder organisations, including KB⁴, OCLC⁵, DANS⁶, SURF⁷, institutions and the UKB, was formed. The working group's rationale for the DAI was summarized as follows:

Bibliographical works, published by authors in their professional lives, are described in bibliographic databases and/or repositories. Since the advent of the Internet, such resources are increasingly interconnected. This enables the possibility of listing the productive output of individuals. Therefore, a directory of an author's output can be used to increase the visibility of an author or his or her affiliation(s) for different evaluation purposes. When used for evaluation purposes at the level of an organisation or institution, it is important to establish and maintain accurate links between authors and their respective outputs. (DAI Working Group 2015)

The DAI working group addressed and adjudicated many issues, including replacement of the DAI with an internationally compatible solution:

A recurring point of criticism of the Dutch DAI infrastructure in recent years was that it is a national infrastructure, while scientific research is international. A second point of criticism is that the DAI infrastructure is removed from the daily practice of researchers. (DAI Working Group 2015)

Both ORCID and ISNI were considered as successors to the DAI. Support for ISNI was strong among libraries as it was and remains an important part of the library resources' work practices. However, the working group recognized ORCID as more relevant to working researchers, as an ORCID record was designed to facilitate links to additional information such as a researcher's affiliation and outputs:

ORCID focuses directly on the researcher on the basis of direct self-registration (so-called author claim). The possibility is offered to keep a profile with an overview

⁴ <https://www.kb.nl/en>

⁵ <https://www.oclc.org/en/home.html>

⁶ <https://dans.knaw.nl/en>

⁷ <https://www.surf.nl/en>

of [ones] work. The management of this profile can be delegated to third parties, for example the employer. (DAI Working Group 2015)

The DAI working group therefore recommended a hybrid solution:

1. *Choose ISNI as sustainable identifier for Author Names (libraries)*
2. *Use ORCID as an identifier within the scientific communication (researchers)*
3. *Motivate researchers to claim their ORCID*
4. *Register ORCID (and ISNI) in research information systems*
5. *Consider, with sufficient demand, to negotiate a national license for ORCID*
6. *Close examination of information to the ISNI infrastructure*
7. *Establish a Dutch Registration Agency for ISNI (KB)*
8. *Universities / UKB libraries and Academy / DANS-Narcis to be supported by OCLC; need to achieve a transition from DAI to ISNI*
9. *A number of universities is associated with SURF in 2015/2016 in a pilot project investigating the possibilities of ORCID for strengthening the Dutch research infrastructure*

As noted in the position paper produced by the DAI working group (2015), ISNI is the preferred ID for library collections and ORCID is the preferred ID for research. As such, both ISNIs and ORCIDs will be present in the national infrastructure. Because the DAI system was being discontinued, all DAIs were in the process of being converted to ISNIs. The OCLC and Royal Library had taken up this task. At present, completion of this project remains in progress. Table 1 outlines a timeline of the shift from DAIs to ORCID in the Netherlands.

Table 1: Timeline, Author IDs in The Netherlands

2005	Netherlands' Digital Author Identifier (DAI) launched
2013-15	Dutch institutions update CRISs; ORCID noted as upcoming international ID
2015/2016	DAI working group proposes hybrid solution (ORCID and ISNI) ⁸
2016	UKB endorses recommendations and the start of an ORCID pilot led by SURF
2016-17	ORCID pilot initiative working group ⁹ established
2017	Transition to ORCID-NL (business) consortium
2018	ORCID-NL contract compliance addressing new GDPR requirements

⁸ DAI Working Group. 2015. "Nota Toekomst Nederlandse Infrastructuur Voor Auteursidentifiers." SURF.

⁹ <https://www.surf.nl/orcid>

2019	ORCID-NL new contract with 10 consortium members
2020	ORCID-NL consortium with 16 members

ORCID, an International ID

At the time of the DAI working group hybrid proposal, there was insufficient information about ORCID adoption in the Netherlands, so it was not feasible to proceed with a predefined national plan. Instead, the pilot initiative was conceived in discussion with universities who already had interest in implementing ORCID, which was in part linked to the recent purchase of new CRISs among nearly all research universities. Anecdotally, the rationale for this approach was based on expected synergy from upgrading the database and researcher ID at same time. This process was not as straightforward as expected, as will be discussed in the next section. Nevertheless, the two upgrade tasks (CRIS and researcher ID) were often situated among the same Library/ICT staff members who participated in the ORCID pilot initiative.

In 2015, as SURF investigated the potential for an ORCID pilot initiative, researchers worldwide were already claiming ORCIDs and ORCID was quickly gaining international traction and support. The United Kingdom, for example, had recently published the outcomes of their national and centrally-supported pilot¹⁰. Because ORCID offered different implementation models ranging from a free service for authors to a premium intuitional license with added services (use of the ORCID API, improved support, national uptake statistics etc.), a common question was concern about the possibility of multiple ORCIDs for a single author. ORCID clarified that multiple IDs assigned to one author can be hierarchically linked in the ORCID database. The author would then select the primary ORCID and the others would be deprecated, and thus still accessible for historical purposes.

Towards an ORCID pilot

In spite of numerous other researcher ID systems in use¹¹, ORCIDs' rapid international adoption has resulted in its recognition as the de facto standard¹². However, since the DAI had already been implemented throughout the Netherlands, in many universities, research institutes, and national databases, the transition to ORCID was, and remains, a complicated endeavor.

¹⁰ <https://orcidpilot.jiscinvolve.org/wp/>

¹¹ see below in table 3 for list of researcher IDs in use

¹² Bryant, Rebecca, Annette Dortmund, and Malpas Constance. 2017. "Convenience and Compliance: Case Studies on Persistent Identifiers in European Research Information." Dublin, Ohio: OCLC Research. <https://doi.org/10.25333/C32K7M>.

Chief among the challenges for ORCID adoption is a chicken/egg dilemma. There is general acceptance of ORCID as a concept: that a not for profit, community-governed, global researcher ID system has the potential to bring some needed order (and interoperability) to the increasingly complicated domain of research information. However, most of this benefit requires a critical mass of adoption. Sitting between the ideal of ORCID and the need for mass adoption is a special feature that is both crucial to the value of ORCID and a substantial barrier to rapid adoption: researcher consent. Unlike other researcher IDs, a researcher controls access to the content of their ORCID record. Many of the other researcher IDs in use, especially proprietary IDs such as Clarivate's ResearcherID¹³ and Elsevier's ScopusID, are assigned without researcher consent.

ORCID and research IDs in general operate in the broader domain of metadata -- machine-readable information about a research object -- that is embedded in all sorts of digital objects. A researcher's ORCID, for example, is a unique string of characters that can be attached to publications, datasets, software programs, or digital representations of material resources such as infrastructure or funding sources. In this way, information about research, often referred to as 'research information,' can be collected and analyzed. Increasingly, research universities use specialized databases (e.g. CRISs) in support of aggregation and analysis of these metadata. To be clear, research information is information about research rather than the output of research itself. Although research IDs, such as DOIs and ORCIDs, are typically included in a research object's metadata, they also provide additional features, such as uniqueness, persistence, and the ability to convey relational information when linked to other IDs.

Implementation of ORCID involves intervening in the domain of research information, which is both ubiquitous and intangible. In the domain of research management and administration, any digital research object, or representation of that object (e.g. person, physical resource, financial resource) can be accounted for in databases within an institution. The collection and aggregation of research information provides the basis for monitoring research inputs, throughputs, and outputs and also provides the basis for evaluation of research at many levels: from the level of researcher, research project, department, and faculty, as well as at the level of the university as a whole.

Moreover, research information is collected from many sources, such as external publication databases, internal staff databases, and manual data entry. Interoperability of these data is therefore important for facilitating automated workflows, especially the exchange of

¹³ This may have changed with the acquisition of Publons, as ResearcherID has been integrated with Publons. see <https://publons.freshdesk.com/support/solutions/articles/12000038281-what-is-my-web-of-science-researcher-id->

information between databases. Although many standards for metadata are in use, application of these standards is inconsistent and unruly. Two key features of ORCID in this regard, are (1) the consistent researcher metadata fields across the global installed base¹⁴ of ORCID and (2) that consent from researchers is required for both openly publishing the persistent identifier as an identity attribute and any further information appended to the ORCID. To be clear, an individual ORCID is both an identifier (unique string of alphanumeric characters) and an editable record of e.g. affiliations and outputs.

3.0 ORCID-NL Pilot Initiative (2016-2017)

In August 2015, SURF posted an open call on the SURF website¹⁵, inviting universities to participate in an ORCID pilot. The objectives of this initiative were to (a) coordinate shared resources and best practices among participants, (b) facilitate group-level interaction with the ORCID organisation, and (c) document the configuration and outcome of ORCID implementations. The invitation was aimed at Dutch research institutions interested in joining the initiative either as pilot participants or learning more about ORCID implementation. Interested institutions were asked to submit an ORCID implementation plan using a provided template. The SURF-funded pilot was initially budgeted for 5 institutions, which was the minimum threshold to receive a consortium discount. The financial benefit for selecting the consortium model was a reduced cost per institution from \$20,000 to \$6,000 USD/year. SURF requested an implementation plan from each institution to ensure pilot members were in a position to work on implementation of ORCID within a common timeframe.

In January 2016 the pilot initiative was launched with five institutional participants. The open call also attracted another 5 institutions who were interested in observing the pilot, but they were not yet in a position to join. By the conclusion of the pilot initiative, we had 10 institutional consortium members (see table 2). Table 2 lists the initial pilot participants and a second group of five universities that joined while the pilot was underway..

Table 2: ORCID Pilot consortium members

01 June 2016	added 01 January 2017
Leiden University	University of Amsterdam
VU Amsterdam	Tilburg University
Technical University Delft	Wageningen University & Research

¹⁴ At present, the number of 'live' ORCID stands at more than 11 million worldwide: <https://orcid.org/statistics>

¹⁵ Webpage not presently available: <https://www.surfspace.nl/artikel/1848-open-call-orcid-pilot-initiative/> (see Appendix 1 for PDF version)

Maastricht University	Radboud University
Utrecht University	Technical University Eindhoven

In addition to the pilot participants and observer institutions, there was interest among the universities of applied sciences (especially, Saxion University of Applied Sciences and Hanzehogeschool Groningen) and the HKI (HBO Knowledge Infrastructure). However, an ORCID business case was at that moment difficult to justify for individual universities of applied science with a relatively low number of researchers. Nor was there a strong business case for the HKI, as a CRIS system or similar central database would have been needed for universities of applied science to make use of the ORCID API and only a handful of organisations were making use of such systems at the time.

The pilot consortium was organised using a working group framework, with meetings held to address issues and share experiences and solutions. SURF serves as consortium lead, which also fulfills an obligation of ORCID's consortium contract--that the consortium is locally lead. The working group addressed common concerns associated with Dutch research infrastructure and relevant local and international data security.. In this structure, SURF coordinated meetings on topics such as use of SURFconext federated authentication, compliance with Dutch privacy regulations and ORCID licensing at the consortium rate¹⁶. In addition, the working group provided a forum to coordinate best practices among the individual implementation projects and to foreground the potential for local solutions that would remain compatible at national and international levels.

The working group structure enabled three ways to participate in the pilot: (1) regular working group meetings to address common ORCID-related concerns associated with Dutch research infrastructure, relevant local and international data privacy, CRIS integration challenges, and ORCID implementation, (2) EDUgroepen (digital collaboration space), with a collection of information resources related to ORCID and author IDs generally, as well as contributions from pilot participants during the course of the initiative, and (3) an email list, which was used to coordinate pilot discussions and planning and was open to others who remained interested, but had not joined the pilot. Working group meetings were open to non-members, which provided real-time information for those considering ORCID and thereby also provided an open pathway to join the pilot initiative.

¹⁶ ORCID membership categories and rates: <https://orcid.org/about/membership>

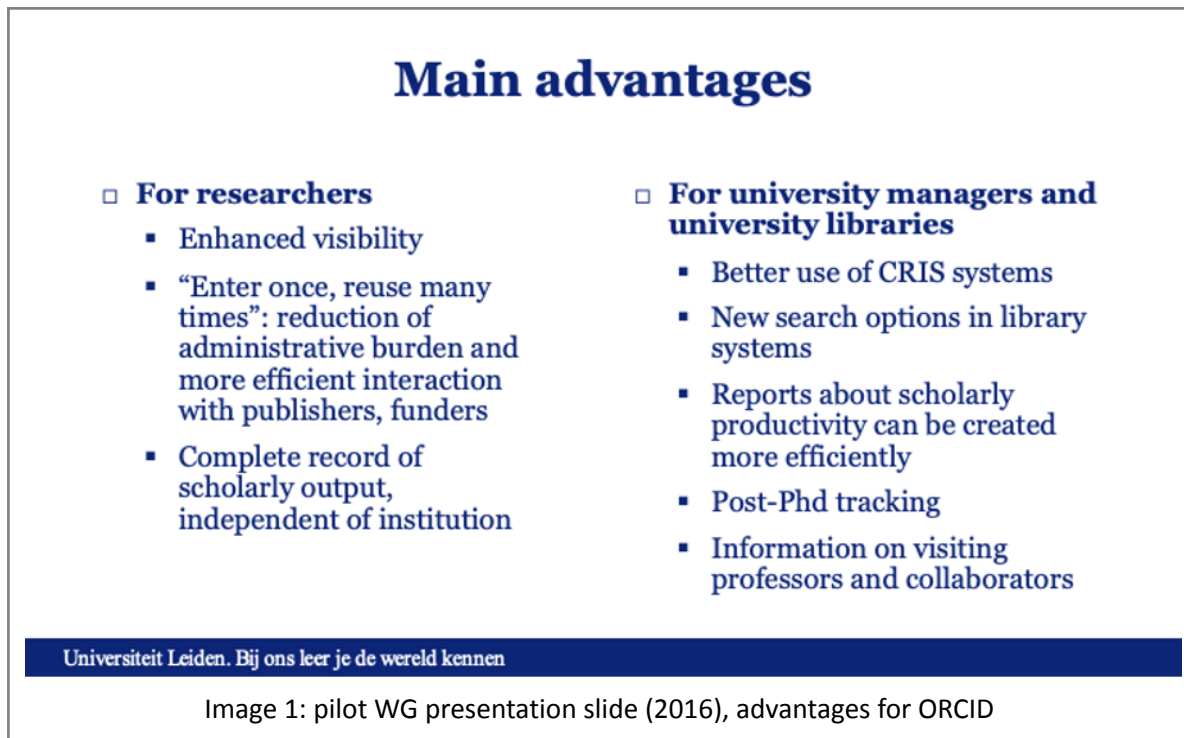


Image 1: pilot WG presentation slide (2016), advantages for ORCID

Image 1 represents the information shared in a presentation to working group members about the potential advantages of implementing ORCID. The proposed benefits of joining the working group included a group (rather than individual) interface for interacting with the ORCID, support for incorporating the use of SURFconext, support for legal issues associated with author IDs and data privacy/security, and the opportunity to draw on the collective expertise of the group to solve complex problems. The working group was also useful in facilitating alignment with other related developments in the national research information context (new CRISs, national OA policy, etc.).

Proceedings, pilot working group

As noted above, the hybrid solution for Dutch author IDs meant that ORCID implementations would be configured for active researchers and collected in institutional CRISs, whereas ISNI would remain the dominant researcher ID for libraries. Operationally, this distinction has two primary implications. First, the ISNI is most appropriate for all contributors to creative works, living or dead, and not just for academic research. The continued use of ISNI is important in developing and maintaining library collections. Second, the ORCID is better suited for active researchers as it is designed as a dynamic identifier that can be updated with new research output and changes in affiliation across the course of one’s career. Whereas an ISNI identifier is normally assigned to an author by a Registration Agency¹⁷, usually upon publication and without consent or notification of the author, an ORCID must be claimed by a researcher and the information it displays requires explicit consent.

¹⁷ <https://isni.org/page/isni-registration-agencies/>

Working group meetings were well underway by the time the pilot consortium agreement was signed and active. In total, six working group meetings were held to address issues related to ORCID features, ORCID/CRIS integration (or lack thereof), and terms of the pilot consortium agreement. Participants shared implementation strategies, challenges, and successes, and ORCID was invited to join specific meetings to help with technical or contractual issues and to share expertise about implementing/operating the ORCID API. SURF colleagues also regularly contributed to these meetings, usually in relation to data privacy guidance and potential synergies with SURFconext and related ideas.

Selected list of meeting topics:

1. SURFconext/ORCID working together agreement, which enabled federated authentication for researcher login to ORCID
2. Pilot participant ORCID plans and configurations; members were invited to present slides to provide updates and/or query others about local challenges
3. Strategies for motivating researchers to claim/register an ORCID
4. Strategies for persuading university stakeholders to (financially) support ORCID
5. ORCID/CRIS workflows and challenges of CRIS integration
6. Assessment and discussion of different ORCID membership levels and features
7. Ideas/issues about transition to a national ORCID consortium
8. ORCID staff presentation of their 'Collect & Connect' feature aimed at optimizing use of ORCID in general and the ORCID API in particular
9. Data privacy, data quality, and authoritative sources
10. Towards the end of the pilot: pros and cons of the ORCID license (compared to using the open (free to use) ORCID API


All pilot participants were also in the process of either acquiring or implementing a new CRIS. There were certainly advantages to implementing ORCID at the same time, as data model considerations for both would be mutually flexible during the implementation phase. Moreover, it was common for ORCID pilot participants to also be involved in the CRIS. However, this approach also created some challenges. CRIS implementation seemed to take more resources than expected and, among pilot ORCID participants, CRIS implementation work generally had priority over ORCID implementation. In spite of high ambitions for the ORCID pilot, attention and resources were often diverted from the pilot to address ongoing challenges with the CRIS. Progress in the pilot initiative was slow in the beginning and began picking up momentum quite late in the program.

Insights, Lessons Learned

During the course of the ORCID pilot initiative, participation among members was strong and most successfully deployed the ORCID service within respective CRISs. Participants shared plans, workflows, and technical processes, and worked together on strategies for addressing

organisational obstacles. The following insights are provided as a summary of lessons learned from the pilot initiative, which describe on-going challenges and inform steps toward realising the expected benefits from implementing ORCID.

Diversity of relevant stakeholders related to researcher IDs – Participation in the pilot initiative entailed discussion of issues that typically crossed multiple organisational boundaries. For example, ORCID use cases can involve local group reporting workflows, data privacy regulations (ie. GDPR), staff employment contracts, technical implementation in university CRISs, and financial planning. Changing local practices in an effort to incorporate ORCID is also complicated by different levels within a university. Using ORCID as the primary identifier for research staff members has administrative implications for researchers, research managers, department heads, faculty deans, and university rectors.



To conclude

- ORCID has potential, but more stakeholders have to embrace it;
- Doubts about some added values;
- No (top down) support at THIS moment;
- PURE implementation and registration of research output

⇒ Increase awareness/added value;
⇒ Work on system integration with PURE / HR
⇒ Final plan after summer 2016;
⇒ Actions on voluntary base (bottom up)

University Library

Image 2: pilot WG presentation slide (2016), need for stakeholder support

As the pilot initiative progressed, we were often confronted with the need to understand university procedures not specifically represented among our participants. This included issues related to data privacy, infrastructure budgeting, research evaluation practices, administration of externally-funded research, adherence to open access policies, and internal communication policy. Mapping the diverse stakeholders within member institutions, who were affected by adoption of ORCID, was necessary to understand which functions might be impacted by implementation of ORCID. It was not necessary to include all

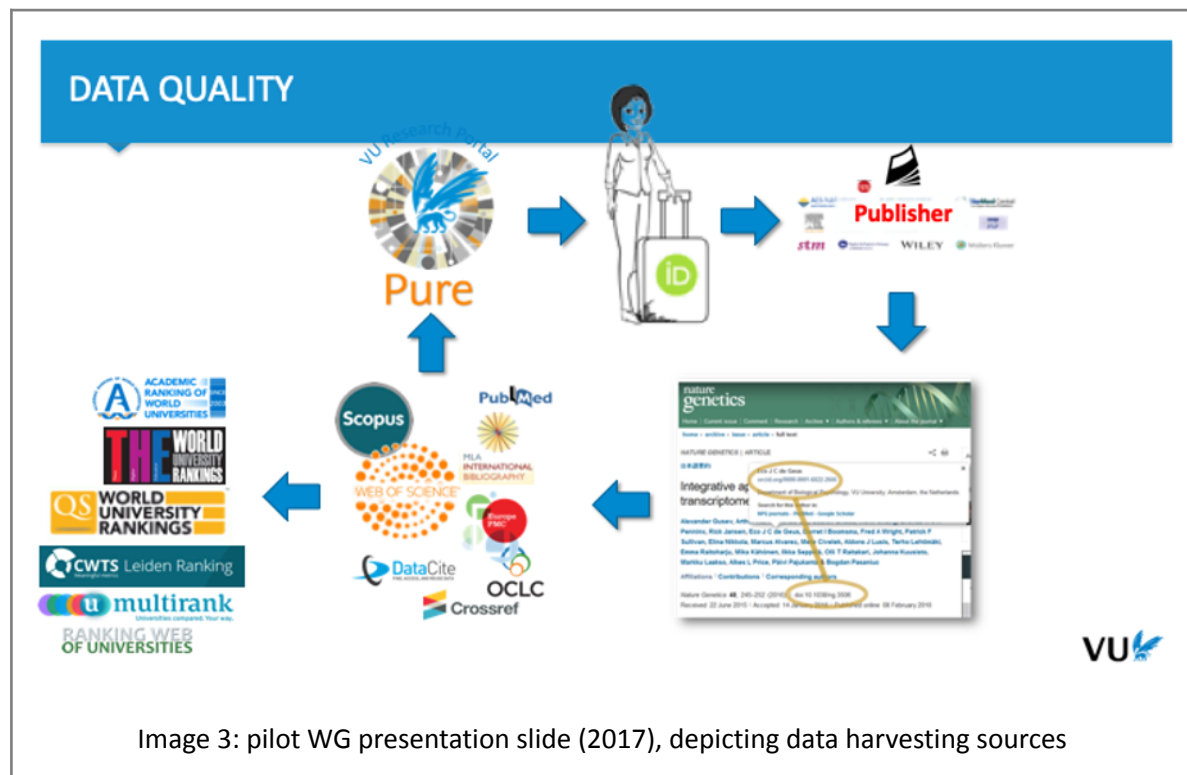
stakeholders in the pilot, but common sense suggests it would be prudent to inform and seek reaction from relevant stakeholders. Image 2, a slide from a 2016 presentation to the working group, makes clear this need for support from all stakeholders at participating institutions.

Authority for author IDs (and Research Information generally) is ambiguous and varies across institutions – A key obstacle emerged in the identification of university staff who had the authority to approve changes to administrative processes or to make financial commitments. At that time, no one had organisational ownership of ORCID. The library might have had responsibility for author IDs (such as the DAI and ISNI), but the ORCID was new and differed in ways that made it difficult to place organisationally: data privacy (it required researcher consent), the financial and contractual implications of licensing the service, and its important role in CRIS workflows and operations. These factors posed two challenges for the ORCID implementation. First, it was difficult to find a particular staff member who had the authority to make organisational or financial decisions about ORCID, such as whether to pay for an ORCID license in the future. And second, it was quite difficult to find the appropriate level of management (i.e. faculty Dean or research director) who was aware of the value proposition for implementing ORCID and could authorise some level of commitment.

Friction with CRIS suppliers regarding ORCID integration – Nearly all research universities in the Netherlands had recently licensed new CRISs as a response to increased demand for all kinds of analytics (e.g. research evaluation events, assessment related to hiring and promotion, open access compliance, and strategic management of university resources). Most universities selected Elsevier's Pure system. Leiden selected Clarivate's Converis system and Radboud retained METIS. The CRIS suppliers were slow and sometimes reluctant to make changes to better accommodate ORCID. This necessitated work-around solutions that slowed progress and were often suboptimal. In this situation, Radboud had an advantage as METIS is also housed at Radboud (at the 'Universitair Centrum Informatievoorziening' department).

ORCID success depends greatly on local system integration – To reach the full potential of ORCID within institutions, base administration systems responsible for most of the CRIS system's master data (e.g. staff, organization structure, funding, projects) will need to be adapted to handle ORCIDs as personal identifiers. In particular the HR and Finance systems will need to adopt ORCID: HR as the source for personal identifiers and subsequent administration (e.g. onboarding), and Finance for the mapping of researchers from funding and projects to relevant output in the CRIS system. Each institution's ecosystem of base administration systems linked to the CRIS varies greatly, and even if there is a solid base where all master data is coming from dedicated source systems, it takes a lot of effort to convince the entire research information chain within an institution to adopt ORCID. Without

this support, the full potential of ORCID or any other identifier cannot be achieved. Table 3, a presentation slide depicting data harvesting sources, demonstrates just how complex these systems are and how much infrastructure institutions are required to adapt.



The messy state of Researcher IDs – Even with an optimal implementation of ORCID, the evolving domain of researcher IDs is itself a data problem that the CRIS is not equipped to address. There are three aspects to the domain of researcher IDs. First, there are multiple researcher ID systems that are operated independently of each other. Populating a CRIS with bibliographic metadata typically involves harvesting external databases, many of which use local and sometimes proprietary researcher ID systems. The following table provides a list of common researcher/author IDs in circulation.

Second, is the so-called ‘external coauthor problem’. Imported publication data generally includes coauthors. This is useful information, especially as it shows collaboration between universities in the Netherlands. However, it is difficult to identify external coauthors unambiguously. Pilot participants indicated that identification of external coauthors requires substantial manual labor. The external coauthor problem compounded differences in the variety of (meta)data sources used by individual universities to enrich CRISs. Many sources exist for bibliographic metadata--open sources such as OpenAIRE and proprietary sources such as Web of Science. The differences between these data sources tend to further complicate interoperability among CRISs in the Netherlands.

Table 3: Sample of Author IDs presently in use

ORCID	Researcher/Creator ID, community governed ¹⁸
Researcher ID	Clarivate, proprietary ¹⁹
Scopus ID	Elsevier, proprietary ²⁰
ISNI	International Standard Name Identifier ²¹
ARXIV ID	Recently adopted DOI as their document ID ²²
DAI	NL national author ID developed by SURF ²³

Data privacy issues are significant – ORCID makes data privacy central to their researcher ID concept. When registering for an ORCID, a researcher must explicitly choose which content is exposed to the public. ORCID’s data privacy policy²⁴ is stringent and ORCID is diligent about remaining compliant with the European Commission’s (EC) evolving General Data Privacy Regulations (GDPR). However, the GDPR applies additional scrutiny to data stored in the United States—which is the case for ORCID. With increased attention to data privacy related to researcher profiles, we observed a general need for increased data-privacy expertise related to ORCID adoption. Data privacy expertise exists at universities, but typically in relation to research involving human subjects. In this sense, the ORCID pilot revealed a need for increased regulatory attention to data privacy in relation to research staff. This finding intersected with the universities themselves ramping up data privacy capacities with regard to GDPR.

Business case for ORCID is largely based on future benefits, while stakeholders need near-term solutions – A question often addressed in consortium meetings was something to the effect: *‘what are the use-cases we can use to help explain why ORCID is important?’* There are two common variations to the question: (a) *why is it important for researchers to register an ORCID id?* And (b) *why is it important for a university to adopt ORCID?* Both variations of the question are challenging to answer because of a chicken and egg paradox; most of the benefits come when a critical mass of researchers and universities have/use ORCIDs. As such, it is difficult for institutions to see a compelling argument for adopting ORCID before there is a critical mass. There are a couple of useful approaches to responding

¹⁸ <https://orcid.org>

¹⁹ <https://researcherid.com/#rid-for-researchers> - accessed 04 February 2021 (implications are not clear regarding the link between researcher ID and [Publons](#))

²⁰ https://service.elsevier.com/app/answers/detail/a_id/11212/supporthub/scopus/ - accessed 04 February 2021

²¹ <https://isni.org/>

²² <https://zenodo.org/record/4470084#.YBwDey1h2wc>

²³ DAI webpage (not presently available)

²⁴ <https://info.orcid.org/privacy-policy/>

to this chicken/egg paradox. First, to explain the relevance of ORCID it is helpful to include it in the broader category of persistent identifiers for research. And second, it helps to extend this discussion into the international context. However, this explanation still relies on future benefits and we hear regularly that the demand for analytics is based on near-term needs.

The basic consensus from the pilot participants was that ORCID would not cease to exist and the perceived (future) benefits would justify setting up a national consortium: ORCID-NL.

Taking these lessons learned into account has resulted in a proposal to extend ORCID activities. The main goal would be to transition from the pilot phase into a ORCID-NL consortium that could be managed through the regular procurement teams within SURF.

The lessons learned led four goals:

- Supporting the regular ORCID working group meetings that focussed on implementation, whilst also extending to more future oriented, innovative applications of ORCID (possibly through another working group);
- Connecting to the relevant expertise in terms of privacy to optimise the accession agreement (through the local privacy officer networks);
- Connecting to stakeholders such as funders and the VSNU to broaden the awareness of the value proposition of ORCID;
- Identifying additional ways to incentivise Dutch researchers to claim and enrich their ORCID IDs/profiles.

The next section provides an account of the transition of ORCID-NL from a pilot (an innovation project) to part of SURF's normal business practices associated with standard service offerings.

4.0 ORCID-NL Business Consortium and new explorations (2018-2019)

Transition from the pilot initiative, an innovation project, to offering ORCID to universities as a standard service, was operationally straight forward. Some pilot members did not immediately join the new consortium; conversely, over time universities that did not participate in the pilot began joining. Table 4 outlines these ORCID consortium members as of January 2021. In addition to normal consortium business, the transition involved three interrelated activities. First was formalizing the ORCID-NL consortium agreement and embedding it in SURF's normal procurement services portfolio. Second was facilitating a proof-of-concept for the ID Resolver concept with participation of a subset of pilot consortium participants. Although proof of concept²⁵ was achieved and the idea had merit, the ID Resolver concept was not further developed. And third was engaging the international community associated with the domain of persistent IDs, a dynamic context in

²⁵ Project website on SURFnet wiki: <https://wiki.surfnet.nl/display/P3GFel2015/ID-resolver> (apparently migrated to new location as part of 'one SURF' consolidation)

which the ORCID organisation both served as an influential participant and was influenced by changes enacted by others.

Table 4: ORCID consortium members as of January 2021

Leiden University	University of Amsterdam
VU Amsterdam	Tilburg University
Technical University Delft	Wageningen University & Research
Maastricht University	Radboud University
Utrecht University	Technical University Eindhoven
Open University	University of Groningen
Erasmus University Rotterdam	University of Twente
Géant	KNAW

The ORCID-NL pilot consortium was active through mid 2017. The pilot was organized as an ORCID consortium with 5 institutional members. During the course of the pilot, membership grew to 10 at its peak. Upon conclusion of the pilot, beginning June 2017, a new formalized business consortium began, also with 5 partners. With the next consortium contract starting January 2019, membership had again grown to 10 members. Since 2020, the ORCID-NL consortium membership is 16, which includes all research universities plus the KNAW and Géant. An expansion of the number of members is foreseen in the Spring of 2021 with the addition of Hanze University of Applied Sciences Groningen.

ORCID-NL working group, first year

The working group model was retained as the operational framework for the transition from the ORCID pilot to the ORCID-NL consortium. As consortium lead, SURF organised quarterly consortium meetings to provide a venue for sharing best practices and for collectively addressing common challenges. Consortium members provided the following priority topics for discussion for the first four meetings of the new ORCID-NL consortium.

1. Substantial expansion of registered ORCID IDs
2. Monitoring the use of the ORCID
3. Approaches to lobby large funders and publishers to adopt ORCID
4. Functional wishes: e.g. authentication, CRIS-content in ORCID, linking pubs in ORCID

Each priority topic was addressed in turn as the main agenda point for consecutive consortium meetings. At the 08 February 2018 ORCID-NL consortium meeting, members reported their respective ORCID adoption approaches and statistics. Members were invited to share their respective approaches to address the goal of ‘substantial expansion of registered ORCID’s’ (i.e. recruiting researchers to register an ORCID),. SURF presented ORCID’s published resources available for engaging researchers and encouraging adoption of ORCID. Among the consortium members, Leiden University’s (LU) approach, which adapted ORCID’s resources, was the most developed at that time.

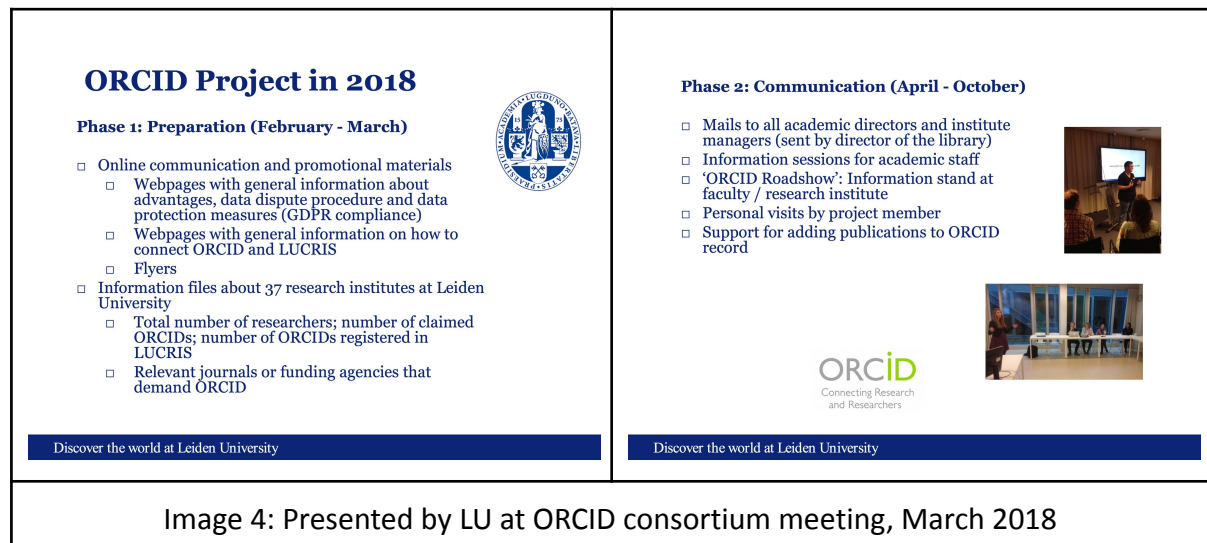


Image 4: Presented by LU at ORCID consortium meeting, March 2018

Both the Leiden University recruiting approach (see Image 4) and VU’s monitoring approach (see Image 5) were subsequently adapted for a national ORCID campaign²⁶ sponsored by the UKB Coordination Point Research Impact²⁷ and carried out by the University Libraries in collaboration with SURF.

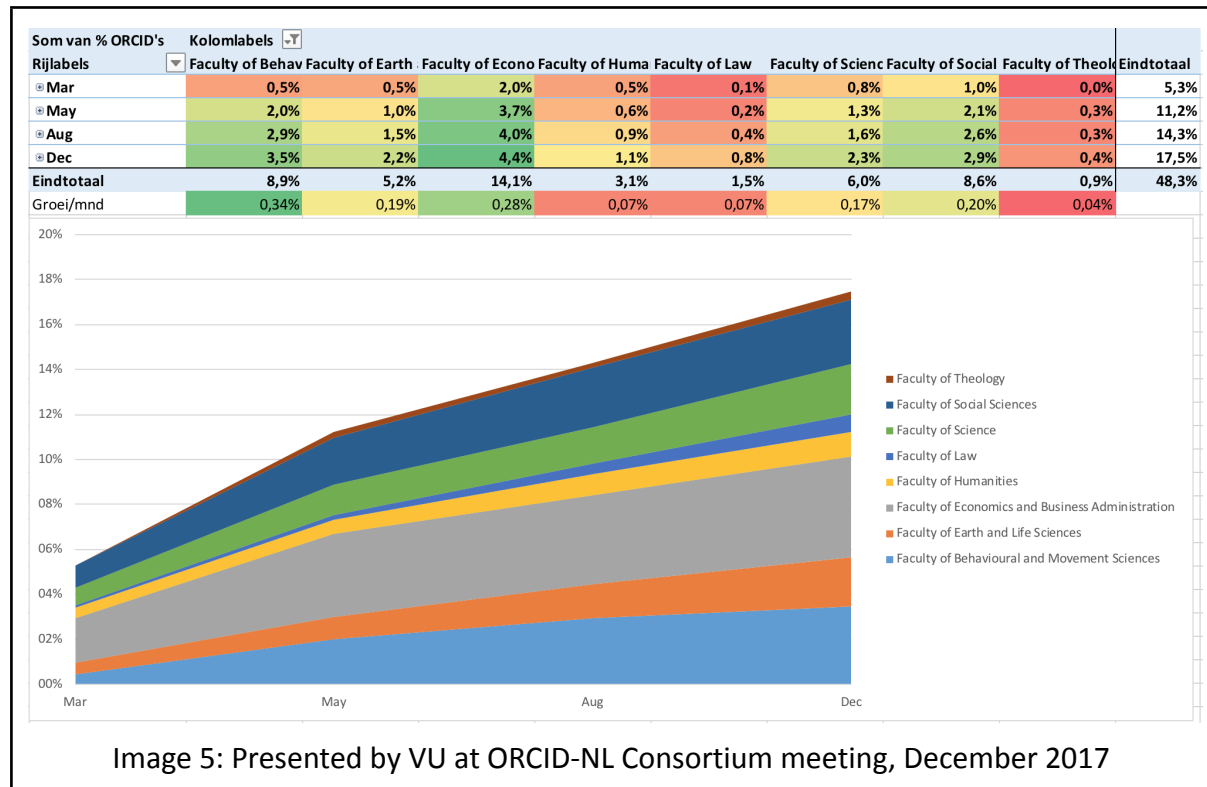
The third and fourth topics were less operational and thus organised as sub-projects within the working group. The issues were first clarified and focused via email discussion in the time leading up to the relevant consortium meeting, and then developed further at the meeting itself. The third topic, approaches to lobby large funders and publishers to adopt ORCID, was reformulated to focus on strategic alignment (in relation to ORCID workflows) among key stakeholders. Through discussion, two key stakeholders were identified, The Dutch Research Council (NWO) and Dutch Language publishers.

The possibility of NWO implementing ORCID's in their funding workflows was a topic of interest among consortium members. There was general agreement on Acomment from a consortium member demonstrates the general consensus.

²⁶ <https://www.ukb.nl/nieuws/nederlandse-universiteiten-starten-orcid-campagne>

²⁷ <https://www.ukb.nl/werkgroepen-commissies/research-impact>

[If] NWO supports or requires ORCID, this is an important new argument that we can use in our promotional campaigns. But it would be nice if an ORCID could actually help in making reports on research projects. Perhaps a scenario is conceivable in the future in which NWO assigns a "grant id" to a new research project, and in which all publications that are written within that project can also (automatically) be linked to that project. (ORCID-NL meeting minutes, 22 May 2018).

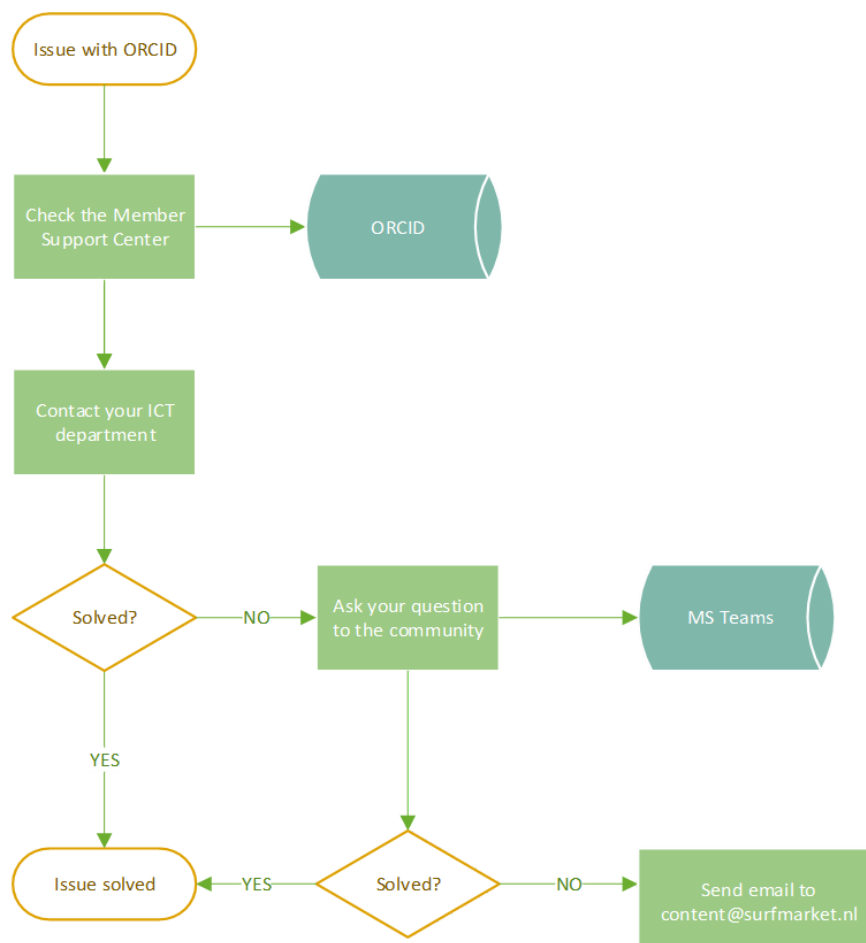


Integration of ORCID consortium at SURF

A concept for integrating the ORCID consortium at surf was developed in early discussions with the SURFmarket contract administration group. However, because of the pending reorganization of SURF, a decision to hold off on this integration was made. Later, transition from the ORCID pilot involved a series of additional discussions with the SURF(market) contract administration group. The central challenge was that the ORCID service did not fit easily into the two services portfolios at SURF. At that time, in the area of procurement services, software licences (like Microsoft, Oracle and Adobe) were managed through a software procurement unit. This unit delivers their services mostly to the ICT department of the SURF members. Content licences (e.g. journals or software applications), on the other hand were managed by a content management unit, which delivers their services mostly to the libraries of SURF's members. ORCID, as a non-content service that is usually managed through libraries, straddles these two units. While the ORCID system is free to use, membership offers advanced functionality via the ORCID Application Interface (API).

Moreover, ORCID's consortium membership model provides discounted licensing fees in exchange for in-kind contributions in the form of local support and community management. According to the ORCID contract, the consortium must identify a consortium lead who is the primary interface with ORCID and who is expected to coordinate among consortium members²⁸.

As consortium lead, SURF organised four consortium meetings in 2020. In early 2020 Microsoft Teams was introduced to the consortium as a tool for a shared information and communication system that would replace the mailing list. This shift also meant that SURF was prepared to host the meetings virtually, since meetings on-site were no longer possible due to COVID-19. SURF also created a page on the website providing information about ORCID and the ORCID-NL consortium. At this point, ORCID was no longer a project but a service offered by SURF, with a consortium of 16 members with a three-year contract. In order to make the transition from a SURF innovation project to a SURF service offering, the working group and support functions were passed on to the SURF contract manager.



²⁸ <https://info.orcid.org/wp-content/uploads/2020/03/Roles-and-Responsibilities-of-ORCID-Consortia.pdf>

Image 6: Workflow solution, enacted in 2020: SURF(market) license desk provides first-line support for the consortium members. The community manager was appointed by SURF.

SURF colleagues from Team Content-Inkoop were appointed to offer first-line support to the members, and participated in the web-based support provided by ORCID in July 2020. Image 6 outlines the workflow for consortium members seeking such support. Since september 2020 there has been a monthly call between the content team and project team to discuss progress on the transition. By the end of 2020, after the SURF reorganisation, Team Content-Inkoop was ready to take over the activities from the project team. Since 2021 Team Content-Inkoop has been responsible for the ORCID-NL consortium and is coordinating its activities. Nevertheless, the monthly calls will continue through 2021.

ID Resolver: a solution to identifying coauthors

The ID Resolver concept emerged from the ORCID-NL pilot working group. The specific idea was proposed by Nick Veenstra (Eindhoven University of Technology), who was also chair of the NL Pure user group at the time. The ID Resolver concept was aimed at improved identification of researchers (especially coauthors) in the context of institutional Research Information systems (CRISs). The scope of the proposed solution is limited to coauthors within the Netherlands, but affiliated with different universities. Image 7 maps out the Resolver Concept process.

Situation (excerpted from group working document²⁹)

Metadata on research publications is already used for different accountability purposes at multiple aggregation levels, which requires a large degree of interoperability of the information systems involved. Generally, publishers are the main source of metadata. Publication metadata is used by universities for local reporting on e.g. research output and compliance with open access policies. The university data must also be transferred to other aggregation databases such as NARCIS, research funders, OpenAire and the KB repository. This results in metadata transfers across several databases, and at each step matching issues arise and some quality loss is incurred. Some degree of metadata maintenance is required to retain a certain amount of quality in the CRIS systems.

Universities often require researchers to manage their own transfer of metadata from publisher to university, which involves some manual labor as a result of person and organisation matching issues. These issues exist because of limited identifier management in the CRIS system and insufficient international standards. Some degree of human interaction is needed by database managers to ensure a correct overview of research data which needs to be downloaded from several sources (i.e. publishers) with different metadata and ID standards. Even though the data is correct in itself at the publisher, researchers often need to verify the information is correct once it is downloaded to the university CRIS.

Several organisations aim to solve these identifier issues by proposing a “universal” identifier (e.g. ORCID and OCLC’s ISNI) to correctly identify publications and staff across providers. In the current

²⁹ https://docs.google.com/document/d/1GgmDvNgevsZnhMHzsnsz1BMwedB_ryrUuAQM3gldHOH4/edit#

landscape a universal ID is not a feasible solution to the overall problem; it is, however, a key component.

Proposed solution

To resolve this issue, multiple institutions export CRIS author data to a central database (ID Resolver). In this approach, each institutional CRIS is the authority for identification of their respective authors. Information is often incomplete for coauthors from other universities. All available author IDs for individual authors with a Dutch affiliation are linked in the ID Resolver. In this way, unknown Netherlands-based coauthors are identified for ID Resolver participants. This is a form of author disambiguation, which is intended to be GDPR compliant.

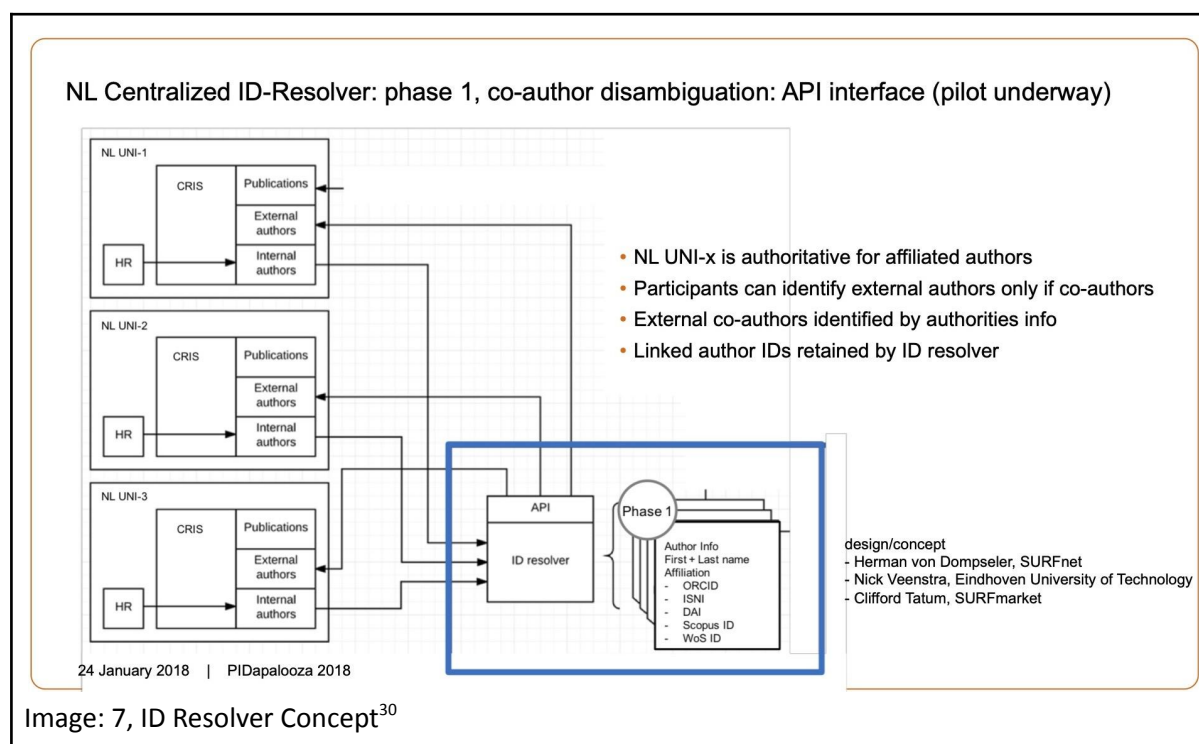


Image: 7, ID Resolver Concept³⁰

The primary use case is identification of external NL (co)authors. The expected benefits would include reduced manual tasks for CRIS managers (metadata management), reduction in manual tasks for duplicate external coauthors (with each new import), improved quality of national reporting/aggregation (VSNU also has a coauthor problem), and vendor-independent software.

While the proof of concept was successful, we did not proceed to the pilot phase. Contributing factors included three key limitations to the concept. First, this solution requires all Dutch universities to participate in order to be effective. Second, the approach was limited to Netherlands-based coauthors, whereas international research collaboration (and coauthorship) is on the rise. Third, GDPR greatly limited identification of persons through aggregation of available data (i.e. building person profiles). The ID Resolver was seen as an interim solution until ORCID's were more widely adopted. In the longer-term, this problem would be solved by a 'critical mass' of researchers with registered ORCID's. And finally, Elsevier had most likely become interested in developing a similar

³⁰ <https://doi.org/10.6084/m9.figshare.5817309.v1>

solution. During the project, Elsevier began recruiting for a Senior Data Engineer to “contribute to building Person Hub applications that will help to identify and match researchers' platforms and services within the Elsevier ecosystem and then make this data accessible” (see Appendix 2).

As adoption of ORCID among individual researchers continues, there is increased adoption and use of identifiers for publications, datasets, and other research objects, as well as a new identifiers for research organisations³¹ and for research grants³². The capacity to link research IDs within and across databases would enable embedding of relational information in ad hoc (or durable) PID assemblages, which is quite useful for monitoring and analysis of research as well as for ensuring ORCID's enduring value proposition.

5.0 Identifiers for FAIR Research Information (2019-2020)

The third phase of identifier activities, Identifiers for FAIR Research Information, aimed to build a broader understanding of the international PID landscape and create broader awareness of the current and future possibilities of PID applications and ensure that information about research will be more Findable, Accessible, Interoperable and Reusable. Due to SURF itself being in transition in the previous phase, this plan also entailed finalising the transition of the ORCID-NL consortium to the (new) services department. This innovation project explored and developed research identifier concepts for increased efficiency and precision of research information within the Dutch academic landscape. The ORCID identifier for researchers, for example, began as a pilot, transitioned to a national consortium, and recently graduated to a SURF service.

Implementation of ORCID in university Research Information Systems (CRISs) throughout the Netherlands addresses an enduring research information problem (author disambiguation) while also establishing increased capacity for, and accuracy of, research information more broadly. In addition, the growing installed base of researcher-consented ORCIDs across universities in the Netherlands provides a crucial information foundation to build upon. When planning this project, we posed an open question to ORCID-NL consortium members [*Imagine two years from now, what would you like to have in place?*]. Following are the responses provided (ref ORCID-NL meeting minutes, 01 October 2018):

- Grant ID for finding/linking outputs related to funded research
- Plan S³³, anticipating substantial monitoring during implementation
- ORCID in workflow for reports (grey literature in general), theses, other
- Perhaps we can make use of the CRediT taxonomy?
- Software sustainability (PIDs seen as part of the solution)
- Project iD would also be helpful, as projects are difficult to define/use in PURE

³¹ <https://ror.org>

³² Crossref Grant ID <https://www.crossref.org/documentation/content-registration/content-types-intro/grants/>

³³ <https://www.coalition-s.org>

- Incorporate ORCID in NARCIS, cooperation/coordination for annual reporting.
- ID for Data management Plans; managing DMPs becoming a serious undertaking at institutions.

Taking the consortium members' input as both general interest in optimizing the present state of information with further use of PIDs and specific points of interest, in this project we investigated promising identifier systems and use cases. Three key objectives aimed at leveraging the growing installed-base of ORCIDs in the Netherlands: (a) increased automation through machine readability, (b) improved capacity to monitor/evaluate research activities and outcomes, and (c) potential for future demand or requirements for sustained (meta)data and analytics infrastructure. In addition to the above motivations, we operate from the principle that research identifiers are fundamental to the possibility of FAIR.

This section of the report begins by situating the role of PIDs in facilitating Open Science, which also provides additional background on the relationship between the FAIR and openness. This background is followed by an outline of the main project activities divided in two areas: coordination of PID developments (both in the Netherlands and internationally) and specific PID concepts.

FAIR and Open Science: The Duality of PIDs

The emergence and ascendance of ORCID as the de facto standard for author IDs³⁴ has shed new light on the potential of persistent identifiers in bringing a degree of order to a rapidly evolving field of research information. This is in part due to ORCID's efforts in establishing the ORCID record as a user-controlled hub of human and machine-readable information about a research contributor. The combination of openness and granularity of user control over what information is linked to an ORCID, and who has access to it, illustrates the potential of persistent identifiers more broadly.

Background

Persistent Identifiers (PIDs) used in academic research live in two worlds: the world of open science and the world of strategic decision making³⁵. Whereas the objectives of open science aim to share the resources and outcomes of publicly funded research as much as possible (discoverability), strategic management of research operates in the competitive space of

³⁴ Bryant, Rebecca, Annette Dortmund, and Malpas Constance. 2017. "Convenience and Compliance: Case Studies on Persistent Identifiers in European Research Information." Dublin, Ohio: OCLC Research. <https://doi.org/10.25333/C32K7M>

³⁵ The ideas here draw on this presentation at the euroCRIS 2018 conference: Tatum, Clifford, and Josh Brown. 2018. "Principles and Pragmatics of 'as Open as Possible': Persistent Identifiers as the Interface between Research Information Commons and Closed Systems," <http://dSPACECRIS.eurocris.org/handle/11366/658>

attracting talent and winning research grants (analytics). This tension is often overlooked in the context of national sharing of research information³⁶.

While the publishing industry provides many crucial services to the academic enterprise, the move to increased openness complicates this centuries-old relationship. Open, community-governed persistent identifiers provide a fundamental building block for retaining enduring access (and preferably also ownership) to information about research funded and/or produced in the Netherlands, while also setting the foundation for realising a more open scientific enterprise.

An open, persistent identifier is FAIR by definition, but the object it points to can be made findable, irrespective of accessibility or openness. Once the existence and whereabouts of research information is signposted by a persistent identifier, it can be made actionable by the consistent provision of a core set of metadata, including at least:

- Source data - the location and nature of the system/profile that holds the info
- Rights information - access restrictions defined
- Negotiation - ability to request or determine access electronically or manually

As such, a minimum required degree of openness is tied to the principles of transparency that underpin open science, or findability in terms of FAIRness. The minimum set of open information to enable others in the community to interact with a research object is *knowledge of its existence*. In this way, persistent identifiers, which are actionable by design, serve as an interface between public and private aspects of research information.

Whether or not a research object is openly accessible, persistent identification facilitates visibility, traceability and citability. Knowledge of an object's existence (findability) is generally sufficient for research analytics purposes. Analysis employing the relationships between researcher, institution, funding and outputs--for example--are important inputs for strategic planning and research evaluation.

While privacy and content ownership will likely continue to limit certain forms of access, this need not render research contributions or resources invisible. In addition to serving the interest of NL open science ambitions, this duality of PIDs suggests that broader implementation of persistent identifiers can help address the enduring challenges of creating and maintaining national research information resources.

NL PID landscape

In the Netherlands, there are several PID services embedded in specific use cases. Each PID system (listed in Table 4) operates independently from the others with a low level of coordination between them. SURF is presently developing a national framework to increase

³⁶ Proposal for focus on national level PID solutions: Tatum, Clifford, and John Doove. 2018. "Domination and Submission: The Struggle to Retain Ownership/Control of National Research Information." presented at the PIDapalooza, Girona, Spain, January 24. <https://doi.org/10.6084/m9.figshare.5817309.v1>

coordination of PID workflows and explore ways to increase interaction between PID systems.

Table 4: PID Landscape at a glance

ORCID	Researcher, contributor ID (ORCID-NL consortium)
DAI	NL Digital Author ID (retired, but still in use)
ISNI	Author ID, University Libraries (registrant)
DOI: HSS data	EASY for HSS datasets (DANS)
DOI: HSS data	DataverseNL, during research (DANS)
DOI: data	Datacite (41 repository accounts via 4TU)
ePIC: data	Handle for datasets during research (SURF)
various: objects/data	Digital Cultural Heritage (pid guide)
PURL: objects/collections	Biodiversity (e.g. Naturalis)

Whereas community-governed PID organisations support, and in some cases embrace, development of open research infrastructure, contributions from for-profit publishing/data companies complicate this effort. To be clear, this is not a tension between good and bad actors. Rather, it is a tension between competing principles. This ongoing tension between community interests and shareholder interests produces uncertainty about the long-term sustainability of PID workflows that depend on enduring access to associated metadata.

A related concern is the commercial bundling of metadata (including PIDs) within a tightly integrated set of services. For example, the combination of CRIS, data subscription, and analytics software provides an integrated bundle of services for which there are few, if any, competitors. This combination of integrated services and market dominance creates a form of vendor lock-in.

To be sure, there are two factors in this form of vendor lock-in. First, the purchasing decision. Functionally, there's the convenience of compatible, well-integrated products. Financially, this can result in cost advantages compared to using an assemblage of components from multiple suppliers. Technically, integrated products can hamper interoperability with other systems and prevent integrations and migrations. And second, there are larger systemic aspects. Patronizing dominant players in an oligopoly market further weakens potential alternatives through both interoperability as well as the pricing strategies as laid out above.

In general the Dutch Research Institutions want to engage with players in a healthy market, fostering joint innovation, while protecting the sovereignty of scholarly capital and preserving academic autonomy through a community-owned governance system. The following general principles for the management of research information

are designed to guide the future collaboration with commercial third party providers.
--[Dutch Taskforce](#) on Responsible Management of Research Information and Data.³⁷

In this context, coordinated utilization of PIDs can help to overcome the technical and market effects, while also providing a means to bridge between disparate databases and analytics tools. In this way, PIDs facilitate a layer of metadata independence (or sovereignty).

PID Coordination

In this section we first provide an outline of the PID landscape in the Netherlands. We then outline outcomes of PID coordination activities on the basis of national and international priorities.

National ORCID Campaign

Now that most universities had made the decision to join the consortium (most often via the research libraries), a national approach to increase ORCID adoption was feasible. SURF and the Coordination Point Research Impact proposed to devise a campaign for ORCID adoption. The UKB reserved budget for the campaign and in doing so initiated this one-year endeavour that was coordinated by the Coordination point Research Impact³⁸ in collaboration with SURF.

The campaign was carried out in two stages. First, promotion materials were collected and created to promote national awareness of ORCID. Second, the university libraries organized local outreach events for their respective research communities. To entice competition, an ORCID implementation contest was launched. This resulted in a significant growth in ORCID uptake and a growing awareness amongst researchers and research supporters alike. The campaign concluded with a festive ORCID event on 4th of November in 2020, at which the results of the competition were presented and prizes were awarded.³⁹

NWO PID Strategy

Central to the PID-coordination activity was outreach to the Dutch Research Council (NWO). As open identifiers can play an important part in exchanging information on research information -- information about researchers, research grants and research outcomes -- NWO was interested in exploring how ORCID and other PIDs could help with exchanging and enriching information about NWO-funded research. SURF served as liaison between the interests of NWO and both national and international developments around Identifiers.

³⁷ Guiding Principles on Management of Research Information and Data, Final version 11 May 2020
https://www.vsnul.nl/files/documenten/Nieuwsberichten/Guiding%20Principles%20on%20Management%20of%20Research%20Information%20and%20Data_11May.pdf

³⁸ <https://www.ukb.nl/werkgroepen-commissies/research-impact>

³⁹ <https://www.surf.nl/en/news/5-universities-win-an-award-at-the-end-of-orcid-campaign>

This NWO and SURF collaboration resulted in an NWO PID strategy⁴⁰ focused on the benefits of Open Persistent Identifiers implemented in funder workflows. . The strategy, approved by the NWO board⁴¹, was informed by lessons learned in the Netherlands as well as international experiences with identifiers. Although funder oriented, the strategy has potential to inform policy makers on the implementation of identifiers. The following provides a summary overview of the strategy and recommendations.

Summary of Recommendations⁴²

NWO works with three fundamental kinds of information that form the basis for most workflows related to funded projects: information about researchers, about organisations, and about grants. Thus we recommend the implementation of three corresponding identifiers into NWO's information architecture. Implementing these individual PIDs, and making explicit links between them, enables analysis of funded research at many levels of aggregation.

Recommendation 1	Implement ORCID ID for researchers into grant application, peer review, and project reporting workflows.
Recommendation 2	Implement Crossref Grant ID in grant application and project reporting workflows.
Recommendation 3	Implement research organisation IDs in grant application and project reporting workflows.

No stakeholder – be it funders, publishers, research performing organisations, or infrastructure providers – is able to cover the entire information spectrum on their own. Given its connecting ('nexus') role and ambition, NWO can play a crucial role in promoting the use of PIDs in the wider national and international research landscape by engaging with key stakeholders. We propose NWO participate both nationally and internationally to help shape the PID ecosystem, within which funders are both beneficiaries and enablers of change.

Recommendation 4	Contribute to shaping the national PID landscape by participating in the ORCID-NL consortium and in a future PID Advisory Board.
Recommendation 5	Collaborate with other funders in the international PID landscape, for instance within the context of Science Europe.

With the recommendations adopted by the NWO board, NWO will be entering the PID domain with a cohesive strategy, whereas many funders are implementing PIDs piecemeal. Such a cohesive strategy will help maximise the benefits of implementing PIDs, not just for NWO, but also for other key partners in the national and international landscape. In entering the PID domain, NWO will be joining

⁴⁰ NWO Persistent Identifier Strategy, <https://doi.org/10.5281/zenodo.4674513>

⁴¹ <https://www.nwo.nl/en/news/nwo-persistent-identifier-strategy-will-lead-increased-efficiency-and-insight>

⁴² Excerpted from the unpublished NWO PID Strategy, draft version 0.5.

some funders that have led the way ([Wellcome](#), [FWE](#), [RCUK](#)) and has the opportunity to consider PIDs in a more holistic way.

International Engagement

Research and research infrastructures, including the systems which underpin the provision and management of PIDs, operate on a global scale. These systems operate across communities and rely on common standards (e.g. for metadata) developed through international cooperation between stakeholders in the PID domain. Although persistent identifiers have been a staple for many years--ORCID recently turned 10 years old and Crossref has been registering DOIs for 20 years--the research identifiers domain is experiencing a spike of innovation and interest. As such, it is important to engage in PID initiatives both nationally and internationally. Doing so helps to ensure local plans and commitments are compatible with international developments while also weighing in on the trajectory of international developments.

For this activity, we generally operationalize international engagement in two modes: participating in working groups and presenting the SURF coordinated concepts in international conferences. Following are examples of these international engagement activities:

Working groups

- A. PID workshops (Singapore and London): this activity, and the interest it created, was eventually overtaken by a broader effort to organize a PID federation⁴³
- B. ORCID person citation working group (Appendix 3)⁴⁴
- C. Knowledge Exchange Open Scholarship Expert Group⁴⁵
- D. Knowledge Exchange Task and Finish group: Openness Profile⁴⁶
- E. Knowledge Exchange Task and Finish group: Trust and Sustainability of PIDs⁴⁷
- F. FREYA ambassador⁴⁸

Conference/Workshop venues

- PIDapalooza⁴⁹
- euroCRIS⁵⁰
- FORCE11⁵¹

⁴³ Freya report, <https://doi.org/10.5281/zenodo.4059557>

⁴⁴ Unpublished report, available upon request.

⁴⁵ <https://www.knowledge-exchange.info/projects/project/open-scholarship>

⁴⁶ <https://www.knowledge-exchange.info/event/openness-profile>

⁴⁷ Sub-group of the Knowledge Exchange Open Scholarship Expert Group (in progress)

⁴⁸ <https://www.project-freya.eu/en/ambassadors/our-ambassadors>

⁴⁹ <https://zenodo.org/record/4439098#.YDzogy1h3gg>

⁵⁰ <https://dspacecris.eurocris.org/handle/11366/658>

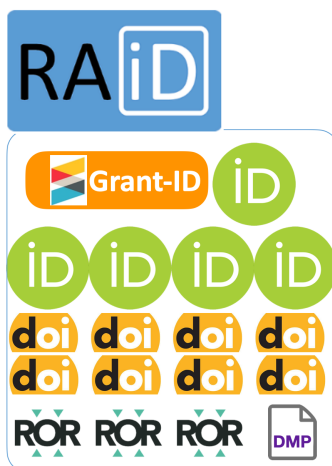
⁵¹ <https://zenodo.org/record/3501793#.YDzosC1h3gg>

Concepts and Investigations

The concepts presented here are focused on increased efficiency in collecting research information and increased precision in collecting research information within the Dutch academic landscape. In this section, the Infrastructure ID and Openness Profile concepts are introduced. Both employ the Research Activity Identifier (RAiD), developed by ARDC⁵².

Research Activity Identifier (RAiD)

The RAiD project identifier provides a means of collecting all project information in one place and sets up the possibility of providing explicit links between project-related resources. Developed by Australian Research Data Commons (ARDC), the RAiD is a 'new generation' identifier that provides a means to collect information about project inputs, outputs and associated research activity over the life of a project. As such, the RAiD thereby also serves as a project information archive after the project is completed. RAiD offers flexible administration of openness, whereby sensitive data can remain closed. Public access to project information can also remain embargoed until an appropriate time (usually after completion of the project).



Like the ORCID, the RAiD includes a modifiable record and the capacity to add or change content without versioning the ID itself. This is in contrast with object identifiers, such as DOI, that use versioning to clearly track changes. The RAiD was designed to capture diverse information about research projects. As new information is added to the project, such as a new collaborator's ORCID or the DOI from a publication, RAiD time-stamps each new entry. Additionally, the source of new information is captured when available, which is how RAiD establishes provenance.

Although the RAiD is presently operating only in Australia, ARDC is in the process of launching international RAiD services. Expanding the RAiD internationally⁵³ involves completing the ISO standards process (expected completion by May 2021), which is nearing completion, and a commensurate increase of staff (which is underway), further technical development for increased scale, and working with the global community to establish an appropriate governance model.

⁵² <https://ardc.edu.au>

⁵³ ARDC, Research Projects: the fulcrum of the research world <https://zenodo.org/record/4477755#.YGHTE50Rrgg>

Infrastructure ID

In 2019 the Infrastructure ID concept was developed to address current limitations in assessing the outcomes from national supercomputer resources. The SURF infrastructure ID concept uses three linked identifiers: ORCID, Crossref's Grant ID, and ARDC's Research Activity ID (RAiD). In this approach, we aim to leverage the unique features of each of these three identifier systems, to provide a flexible, scalable solution that can be implemented in existing workflows and established reporting procedures.

The specific use case is allocation of time on SURF-managed supercomputers for the purpose of research, which is administered as a form of research funding. The Dutch Research Council (NWO) delegates some of the administration of this activity to SURF. For a funded project, the Infrastructure ID use-case covers the full life cycle of the research with clear start and end points. This project structure provides a clear and reliable information collection pattern that is suitable for demonstrating the utility of creating relational links between researchers, resources, and outcomes. This sort of linking of project-related entities also underpins the proposed NL Open Knowledge Base⁵⁴ (OKB) concept.

Development of the Infrastructure ID concept involved a series of meetings with SURF staff members involved in this process. Individual meetings focused on: (a) grant application workflows, (b) the present CUA identity management processes, (c) supercomputer log information flows, and (d) technical interface with the supercomputer application platform.

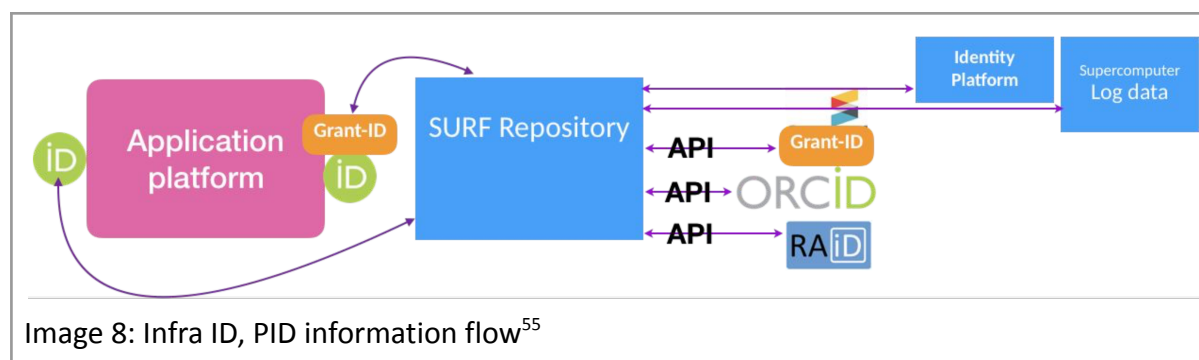


Image 8: Infra ID, PID information flow⁵⁵

The main objective of the Infrastructure ID is Improved capacity to monitor/evaluate infrastructure utilization and associated research outcomes. The infrastructure ID concept consists of three linked identifiers: ORCID, Grant ID, and Research Activity ID (RAiD). In this concept, we leverage ORCID's delegated assertion technology, whereby information is exchanged among the three Identifiers. As noted in the diagram above (Image 8), this process begins with the applicant's ORCID, which can be authorized (by the researcher) to allow automated assertions to their ORCID record. Once the link is made between the ORCID and the Grant ID, we would then register a RAiD to collect ORCID's from project

⁵⁴ <https://openworking.wordpress.com/2020/05/29/what-is-an-open-knowledge-base-anyway/>

⁵⁵ Source: <https://zenodo.org/record/3632958#.YGHSgy0Rrgg>

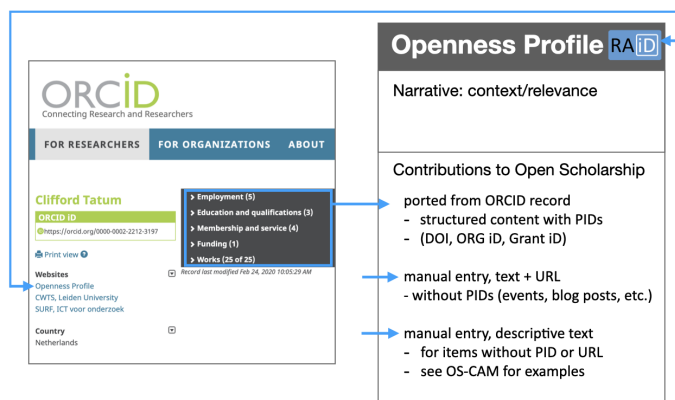
collaborators, DOIs from datasets and publications, as well as resource utilization log data related to the particular grant (i.e. from supercomputer log data).

To illustrate this, if the funder asserts the Grant ID to a researcher's ORCID, the funder is identified as the source of that association in the ORCID record. When the RAiD collects information from the researcher's ORCID ID, say a publication related to the funded project, RAiD collects both the publication DOI and its association with the particular grant. In this example, the relationship between ORCID ID, publication DOI, and Grant ID is also captured.

But not all PIDs will include this embedded relationship information. Linking PIDs in metadata and asserting related PIDs to an ORCID record are still new and evolving practices. In such cases, when PIDs with incomplete information are collected in the RAiD, they are related by association to the RAiD. Documenting this information in the RAiD record creates implicit relationships among entities, which can be exploited for analytical purposes.

Openness Profile

The Knowledge Exchange Openness Profile is conceived as a format for documenting contributions to open scholarship, procedures for self-publishing these contributions as a digital object with a persistent identifier, and strategic use of contemporary research information infrastructure to establish prominent placement of the published contributions.



The Openness Profile (OP) is a portfolio approach for organising and presenting contributions to open scholarship, which is linked to, or embedded in, one's ORCID record. The RAiD is adapted here to serve the portfolio function of the Openness Profile. The RAiD's editable record feature and that it is itself an identifier makes it particularly suitable for this purpose.

Creating an Openness Profile in relation to an ORCID ID is motivated by two key affordances. First, it establishes a human readable collection of contributions that draws on the content associated with ORCID ID. In addition to being the de facto standard for researcher and contributor IDs, ORCID has also established a high degree of trust associated with information facilitated through the platform. Second, linking to the ORCID system enables machine readability of one's contribution to open scholarship. Machine readable information increases automation, thereby reducing administrative burden, which is increasingly important in many reporting and evaluation workflows, such as CRISs, Grant applications, performance reviews, and job applications.

6.0 Towards a National PID Strategy (2021 and beyond)

In this final section, we reflect on the emerging market structure for data and analytics, which provides important context for present work on a National PID Strategy. Interest in and the relevance of Persistent Identifiers as essential building blocks for open and interoperable information about research has grown over the past five years. Peter Wittenburg adequately captured this relevance of PIDs in a recent publication.

Using PIDs offers thus a number of great advantages such as clear and stable identities allowing humans and machines to exactly refer to the right data even after many years, to have easy ways to prove identity, integrity, and authenticity, to provide stable references also as basis for citations, to easily find descriptive metadata, and information needed for authorization, for reuse tracing information, on versioning, etc. We realise, however, that we are increasingly dependent on a stable PID system... (Wittenburg, 2019, p. 13⁵⁶)

Throughout the move toward increased openness of research practices and outcomes, there has been a debate of sorts about the implications for commercial services. In the early days of the open access movement, for example, Steve Harnad's subversive proposal⁵⁷ called for sharing academic texts on computer servers instead of publishing in established journals. The aim was to dramatically reduce both the costs of publishing and the time it takes to share new knowledge. This debate, a form of which continues today within the broader concept of open science, centers on increased efficiency and lower cost of dissemination via the internet and internet related technologies. Peter Wittenburg's account above is the PID version of this argument.

Open, Community-Owned Scholarly Infrastructure

Persistent Identifier systems, particularly those organized as *open and community governed*, are presently undergoing a period of dynamic growth and innovation. Many of these PID organizations see their technologies and services as contributing to open research infrastructure. Crossref, for example, recently reformulated their mission from PID-centric to the broader idea of open infrastructure.

Crossref is committed to the collaborative development of open scholarly infrastructure for the benefit of our members and the wider research community [promoting] the development and cooperative use of new and innovative technologies to speed and facilitate scientific and other scholarly research.

⁵⁶ Wittenburg, Peter. 2019. "From Persistent Identifiers to Digital Objects to Make Data Science More Efficient." *Data Intelligence* 1 (1): 6–21. https://doi.org/10.1162/dint_a_00004.

⁵⁷ Okerson, Ann, and James O'Donnell, eds. 1995. *Scholarly Journals at the Crossroads: A Subversive Proposal for Electronic Publishing*. Washington, DC: Association of Research Libraries. <https://archipel.uqam.ca/170/1/subversive.pdf>.

--Ed Pentz, July 2020⁵⁸

The Crossref board subsequently adopted the Principles of Open Scholarly Infrastructure⁵⁹ (POSI)⁶⁰, which provided an explicit framework for . POSI is a list of sixteen commitments that will now guide the board, staff, and Crossref's development as an organisation into the future

Steve Harnads subversive proposal was but one of many steps toward the present day formulation of open access, which has dramatically expanded but not yet fully realized. Although not all commercial publishers have fully embraced open access, they remain important creators of PIDs for published research and important disseminators publication metadata that includes related PIDs (such as ORCIDs and Grant IDs). However, there remains a tension between the community principles of open scholarly infrastructure and the profit imperatives of commercial publishers.

In the Netherlands it has become increasingly important for the academic community to ensure all information about research is openly available and can be exchanged. In order to maintain the needed control and ownership of that information, community owned open persistent identifiers provide an essential component. In this context, persistent identifiers serve as an open interface to research objects regardless of their present state of openness. Open PIDs point to an object that may or may not be open. In this way, the coverage of national research information is made more complete. Even in circumstances that prevent datasets from being made openly accessible, a full and accurate account of national contributions to science serves many research assessment activities.

Emerging market structure for data and analytics

Control and ownership of 'the' authoritative dataset about NWO-funded projects for example, creates a new mode of accountability, enables a core dataset for understanding the effects of funded research, and provides a measure of self assurance in the face of emerging oligopoly market characteristics among commercial data and analytics service providers. This latter point is particularly relevant in relation to academic publishers acquiring new data-related assets (often by purchasing start-up companies). The fact that only a few publishers hold the vast majority of published content and related bibliographic data results in commercial control rather than academic community control.

⁵⁸ <https://www.crossref.org/blog/meet-the-new-crossref-executive-director/>

⁵⁹

<https://www.crossref.org/blog/crossrefs-board-votes-to-adopt-the-principles-of-open-scholarly-infrastructure/>

⁶⁰ Bilder G, Lin J, Neylon C (2020), The Principles of Open Scholarly Infrastructure, retrieved [07 June 2021], <https://doi.org/10.24343/C34W2H>

In this light it is crucial for the academic community to be attentive to what control they want to exert over the information that is essential for strategic planning on the basis of past research outcomes. Inaction toward sovereignty of national research information assets risks leaving the door open for commercial actors to control and profit from them by offering a paid subscription to information (bundled with analytical services) about national scholarly capital. Aspesi and Brand have argued to ensure the research information on which these services are based is kept openly available.

“The healthy functioning of the academic community, including fair terms and conditions from commercial partners, requires that the global marketplace for data analytics and knowledge infrastructure be kept open to real competition.”

--Aspesi and Brand 2020⁶¹

Steps to be taken to come to a workable solution in which the control and ownership of the information about publicly funded research remains with the academic community. The Dutch Taskforce on Responsible Management of Research Information and Data⁶² was set up that formulated guiding principles⁶³ to do just that in early 2020. A recent study assessing the need and feasibility⁶⁴ of a so-called Open Knowledge Base (OKB), calls for an incremental approach to establishing an NL OKB based on the guiding principles, in which all data from research information systems are openly available. If such an OKB becomes a reality, open, community owned Persistent Identifiers will be a sure part of this database. A national roadmap as described in the previous paragraph would support such an endeavour since it would be able to identify the needed identifiers as well as validate their compliance to the guiding principles that would govern such a database.

Next steps

Building on the outcomes of projects outlined in this report, we begin 2021 working towards a National PID roadmap. In 2020 we focused on the added value of Identifiers for the national research funder NWO and efforts to further the dissemination of ORCID in the Netherlands (see section 5 above). The PID strategy co-developed with and for NWO provides a foundation for a national PID strategy. Scaling this strategy to the national level benefits not only from NWO’s commitment to implement PIDs in their information workflows, and the resulting capacity for NWO to publish project related PIDs, but also from NWO’s authoritative position within the Dutch research community.

⁶¹ Aspesi, Claudio, and Amy Brand. 2020. “In Pursuit of Open Science, Open Access Is Not Enough.” *Science* 368 (6491): 574–77. <https://doi.org/10.1126/science.aba3763>.

⁶² https://www.vsnu.nl/en_GB/os_onderzoekinformatiesystemen-open-knowlegde-base.html

⁶³ https://www.vsnu.nl/files/documenten/Nieuwsberichten/Guiding%20Principles%20on%20Management%20of%20Research%20Information%20and%20Data_11May.pdf

⁶⁴ https://www.vsnu.nl/files/documenten/Domeinen/Open_Access/Dialogic%20Feasibility%20study%20Open%20Knowledge%20Base.pdf

The aim is to bring together the views from the different stakeholders in research information on the use of Identifiers and propose a roadmap to work on implementation of the most relevant Identifiers. Specifically, we are looking at Identifiers for Grants, Data management plans, research institutions and research projects. To ensure a nationally supported roadmap, we are in the process enacting a NL PID advisory group consisting of senior managers from the various stakeholder groups, along with a working group of local experts to produce the roadmap.

Appendix 1, Open Call: ORCID pilot initiative

Open Call: ORCID pilot initiative - Artikel - SURFspace
27/11/2018, 13:59

SURFspace maakt gebruik van cookies. [Lees meer.](#) OK

Inloggen | Profiel aanmaken

HOME SIG'S FORUM LEDEN
Zoeken

ARTIKEL

Artikel toevoegen | Delen »

Tatum | 13-08-2015 | 18:04
Laatst bewerkt door John Doove op 09-09-2015 18:04

👍 2
💬 0
📄 1

Open Call: ORCID pilot initiative

SURFmarket is planning a pilot initiative for implementation of Open Researcher and Contributor IDs (ORCID) in existing research information contexts. The objectives of this initiative are to (a) coordinate shared resources and best practices among participants, (b) facilitate group-level interaction with the ORCID organization, and (c) document the configuration and outcome of each installation. At present, we are in discussion with a few institutions already planning implementation of ORCID. This open call is an invitation to all Dutch research institutions to join the initiative either as a pilot participant or to learn more about ORCID implementation.

Working Group
We will employ a working group framework to coordinate among the ORCID implementation projects. In this way, we can address common concerns associated with Dutch research infrastructure and relevant local and international data security. In this structure, SURFmarket will coordinate issues such as the use of SURFconext federated authentication[1], compliance with Dutch privacy regulations (legal framework for Higher Education in the Netherlands)[2], and the possibility of obtaining ORCID licensing at the consortium rate[3]. In addition, the working group will provide a forum to coordinate best practices among the individual projects and to ensure that local solutions are compatible at national and international levels. For example, the ORCID pilot projects will build on and/or maintain compatibility with the ISNI ID system presently being hosted by the Royal Library (KB).

Hybrid Approach
As noted in the position paper produced by the DAI working group[4], ISNI is the preferred ID for library collections and ORCID is the preferred ID for research. This means we will eventually see both ISNIs and ORCIDs in the national infrastructure. As the DAI system will be discontinued (at the appropriate time) all DAIs are being converted to ISNIs. The OCLC and Royal Library have already taken this up.

ORCID Implementation
Researchers worldwide are already claiming ORCID and ORCID is quickly gaining international traction and support[5,6,7]. ORCID has different possible models ranging from a free service for authors to claim their own ORCID to a premium national license with added services (use of the ORCID API, improved support, national uptake statistics etc. [3]). There is a possible shortcut for implementation of ORCID since ISNIs and ORCIDs share a base ID-number. Once the DAIs have been replaced by ISNIs, the ISNIs could then be translated to (or displayed as) ORCIDs. However, the translation from ISNI to ORCID must include author consent because ORCIDs are open to the public. Regarding the possibility of multiple ORCIDs for a single author: according to Josh Brown (ORCID), multiple IDs assigned to one author can be hierarchically linked in the ORCID database. The author would then select the primary ORCID and the others would be deprecated, thus still accessible for historical purposes.

Discussion About National Adoption of ORCID
National adoption has been discussed in the DAI working group and at present there is insufficient data/knowledge to proceed with a predefined national plan. (The position paper by the DAI working group provides additional background on international trends and the use of author IDs in the Netherlands[4].) In addition to supporting local implementation of ORCID, the ORCID pilot initiative provides an opportunity to explore whether there is potential for a national solution.

Expression of Interest
On the basis of feedback from this open call, we plan to organize a kick-off meeting for the ORCID working group. For institutions already planning to use ORCID and interested in joining the working group, please provide the following information via email to Clifford Tatum (Clifford.Tatum@surfmarket.nl):

- Level of interest in ORCID implementation
- Number of academic staff at your institution
- Primary focus or objective of ORCID
- System(s) used/interfaced at your institution
- Is there institutional support for ORCID

GERELATEERDE ARTIKELEN

Korting op inschrijving OE Global conferentie
Gouwehand

Year of Open 2017
Gouwehand

*Werkconferentie rond open acces tot praktijkgericht on
Eva Woertman

*Input wanted! Onderzoek SURFmarket over Open acc
Carljin Huijsmans

*Quality Open Access Market heeft wetenschappelijke l
Leo Waaijers

POWERED BY SURF ©2015 SURF

[Over SURFspace](#) | [Community Guidelines](#) | [Gebruiksvoorwaarden](#) | [Contact](#)

<https://www.surfspace.nl/artikel/1848-open-call-orcid-pilot-initiative/>
Page 1 of 2

- Preferred ORCID membership
- Lead project participant/contact person

The kick-off meeting will be open to all interested parties. We will establish a date, preferably in September, and organize the meeting content on the basis of the information provided in your responses to the above questions. Please email responses by end of August.

Referenced Resources

- [1] SURFconext
- [2] Legal framework for Higher Education in the Netherlands
- [3] ORCID Membership and Subscription
- [4] ↪ Nota toekomst Nederlandse infrastructuur voor auteursidentifiers
- [5] Italy's national adoption of ORCID
- [6] UK's ORCID Pilot Study
- [7] UK's ORCID Pilot Study (final report): Institutional ORCID Implementation and Cost-Benefit Analysis Report

Contacts:

Clifford Tatum, Clifford.Tatum@surfmarket.nl
John Doove, John.Doove@surfmarket.nl

Tags: open access, CRIS, ORCID Tags volgen

SIG's en subthema's: Research Support, Research Information

REACTIES Reactie toevoegen

Geen reacties gevonden.

PLAATS REACTIE

[Plaats reactie](#)