



Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI

Pablo de Castro, Ulrich Herb, Laura Rothfritz, Joachim Schöpfel

► To cite this version:

Pablo de Castro, Ulrich Herb, Laura Rothfritz, Joachim Schöpfel. Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI. Knowledge Exchange. 2023. hal-04245151

HAL Id: hal-04245151

<https://hal.univ-lille.fr/hal-04245151v1>

Submitted on 16 Oct 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

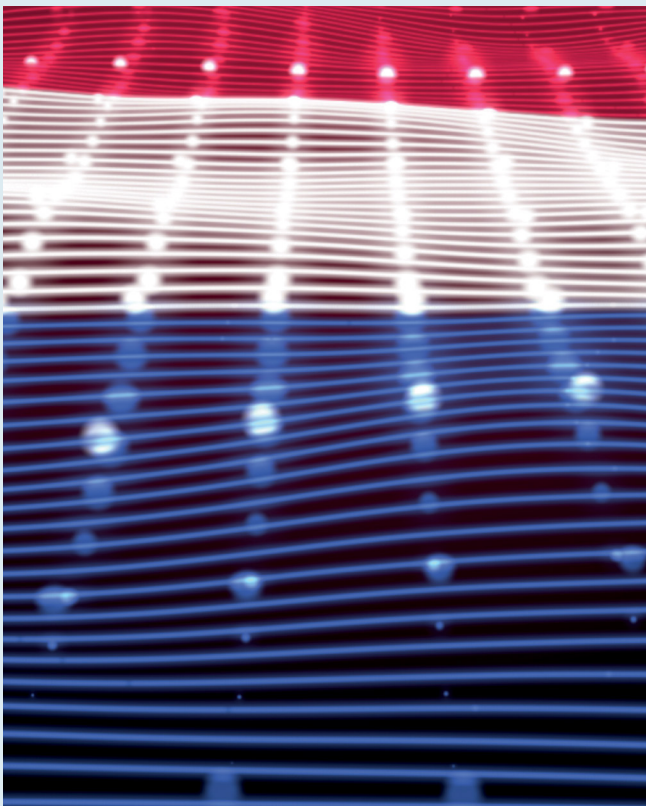
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI

February 2023



This case study is part of a series that has been produced within the study on “Risks and Trust in pursuit of a well-functioning PID infrastructure for research” commissioned by the Knowledge Exchange in July 2021. The main outcome of this study is a report examining the current PID landscape with an emphasis on its risks and trust-related issues.

This complementary series of case studies aims to provide a deeper insight into specific areas of activity, workflows and stakeholders within this wider PID landscape.

Title: Adoption of the DAI in the Netherlands and subsequent superseding by ORCID/ISNI

Authored by: Pablo de Castro, Ulrich Herb, Laura Rothfritz and Joachim Schöpfel (scidecode science consulting).

Email: office@knowledge-exchange.info

DOI: [10.5281/zenodo.7327505](https://doi.org/10.5281/zenodo.7327505)

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



Contents

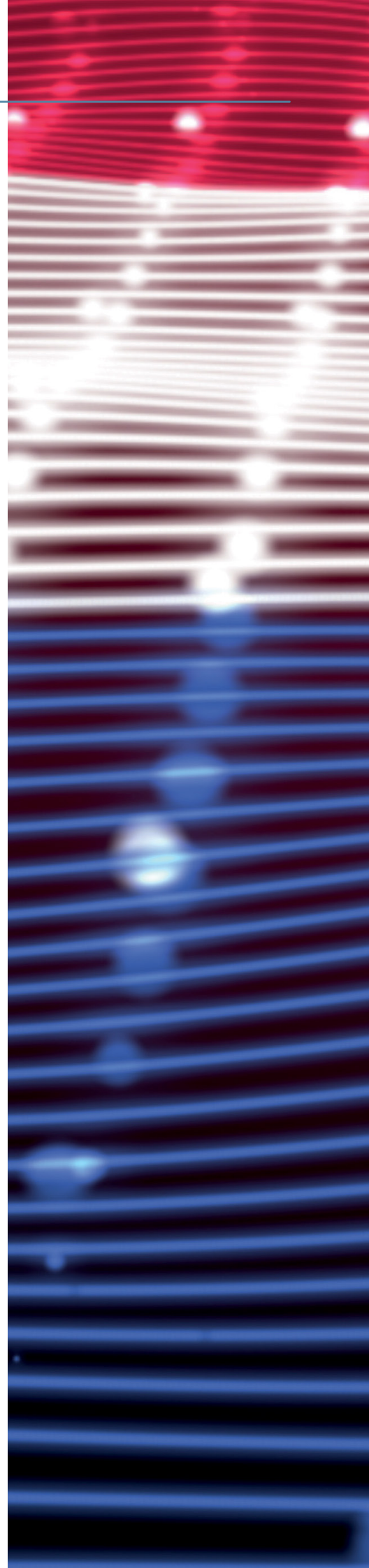
1. Rationale	4
2. The case of the DAI	6
3 Issues around risks and trust regarding the transition away from DAIs	11
3.1 The process for having DAIs replaced and/or superseded	11
3.2 Technical vs community-oriented issues regarding trust	11
3.3 Public vs 'commercial' PID infrastructures	13
4. Authorship	16
5. Literature	17

1. Rationale

The Dutch Digital Author Identifier remains the best example to date for a successful superseding/replacement of an existing PID layer by a new, more comprehensive solution.

Initiatives like ORCID and ISNI are fairly consolidated by now and no-one would consider a national-level author ID project to be worth the effort, but both these international initiatives are relatively recent and before they arrived it made sense to try and implement a national-level author ID keeping in mind the fact that it might eventually need to be superseded.

There are several reasons why a DAI case study makes sense in the context of this work around risks and trust issues for a well-functioning PID landscape. Not only this initiative has been little documented in the literature thus far (if at all), but the successful process for its superseding provides a valuable blueprint for a possible way forward in other PID areas where the landscape is rather fragmented. This is particularly relevant for emerging PIDs for organisations (OrgIDs), instruments and facilities (PIDINSTs) and even for grant IDs.



The lesson that this case study provides is that a fragmented PID landscape with potentially diverging technical solutions may not be a critical issue as long as there is a contingency plan for an eventual replacement or superseding of a given solution in a way that ensures the interoperability of the end result.

2. The case of the DAI

The Digital Author Identifier (DAI) in the Netherlands was the first initiative in the world to implement a national-level persistent identifier for researchers.

The DAI was developed in 2005 by a working group involving the Royal Library (KB), OCLC, DANS, SURF, the Dutch HE institutions and the UKB under the SURF DARE programme¹. This means that this national-level author identifier scheme not only predated the emergence of ORCID as an international standard for author identification, but it also took place at a time when the all-European OpenAIRE infrastructure for Open Science implementation was still at a very early stage of development – namely while it was still the DRIVER project². The DAI was just one of the Open Science domains where SURF and the Netherlands played a forerunner role in those early days, others being the development of the NARCIS national portal as a metadata harvester of institutional repository content and the implementation of the METIS-based institutional research information management system network for all Dutch universities.

Other countries in Europe such as France³ and Greece (see screenshot on the right-hand side of the page) also eventually developed their own national author ID scheme, but the Dutch one had already been in operation for quite a few years before these other national-level author ID schemes were conceived.

Greek Researcher Directory - *researchers.gr*

- ▶ Database of Greek Researchers
 - › Researchers in Greek organisations and Greek researchers abroad
- ▶ Access only to researchers - registration only by invitation
- ▶ Extremely easy interface for researchers to update their profile (one-click functions)
- ▶ Pre-loaded information for permanent personnel of Greek public research performing organisations
- ▶ Automatic loading of information from sources:
 - › CrossRef, PubMed, Microsoft Academic Search, DBLP, Greek Books in Print
 - › Link publications to researchers - automatic mechanisms
- ▶ Automatic recognition of references from researchers CVs
- ▶ Export profile in CERIF XML

N Houssos (2014). "Using CERIF-based CRIS to support the academic and research community: emerging services in Greece". euroCRIS Strategic Membership Meeting Autumn 2014 (KNAW, Amsterdam, Nov 11-12, 2014), <http://hdl.handle.net/11366/346>. The Greek Researcher Directory has now been deprecated.

This case study looks into how the arrival of the international ORCID identifier for persistent author identification impacted the DAI initiative in the Netherlands and how the DAI was eventually superseded by a combination of ORCIDs and ISNIs. From a perspective of risks and trust issues around persistent identifiers, the DAI provides an interesting case study for several reasons:

- ▶ It is an example for a successful superseding/replacement of an early PID implemented at a national level by an international initiative that would quickly consolidate as a well-established solution worldwide. This is particularly relevant for some of the ongoing early initiatives to implement emerging PIDs that could eventually face a similar replacement or superseding.
- ▶ The process to add an ORCID layer for persistent author identifiers on top of the DAI is relevant not just from a technical perspective, but from a community-oriented one too. Questions like how Dutch researchers took such an early overhauling of the national DAI or how the replacement/superseding process was communicated to them are very pertinent from a trust perspective.
- ▶ This case study also raises the issue of community-owned vs 'commercial' PID infrastructure: while ORCID has been implemented as a persistent author identifier on the basis of a non-profit membership organisation, the interviews with PID experts conducted as a background to this study have repeatedly included discussions on sustainability and whether it might be possible for such PID infrastructures to remain public with the appropriate funding mechanisms in place.

The following text providing some background on the implementation process for the Dutch DAI is taken from KNAW's Elly Dijk's "*NARCIS: linking CRISs and OARs in the Netherlands. A matter of standards and identifiers*" presentation at the I Workshop on CRIS, CERIF and Institutional Repositories (Rome, 2010), <http://hdl.handle.net/11366/37>:

“

To connect the information in NARCIS coming from all these different databases and repositories the first step has been to assign a Digital Author Identifier (DAI) to researchers in the Netherlands. This identifier acts like a digital glue in realising the integration of information regardless of how the researcher's name, initials or surname are spelled in the different systems. Within the OCLC library system a thesaurus of author names with corresponding DAIs was created. All the involved institutions have matched the DAIs and author names in their own CRIS (Metis). This work was finished at the end of 2008. The next step is the implementation of the DAI in the repositories of the participants, in the NOD, and in EASY. We expect this to be ready at the end of 2010 (...)

”

It makes sense for a persistent author identifier to be – at least initially – defined at a national level, since this is the level at which most research policies operate in a given country. This includes a research information management policy, a 'repository policy' to make research outputs openly available and/or a research funder policy to clearly and uniquely identify researchers applying for project funding, reviewing project proposals or becoming involved in the projects they fund. Some means of internal author identification tends moreover to be available at a national level, which can in turn be crosslinked with the (usually also available nationally) directory of research-performing organisations. The effort to make available and consolidate international standards for the persistent, unique identification of researchers and research organisations is a highly ambitious endeavour and requires a solid technical foundation, a clear set of use cases for their implementation and a strong community involvement to succeed. ORCID would probably not have succeeded – or at least not so quickly – had it not been able to rely on the [already developed ResearcherID platform](#) that ThomsonReuters made available for the initiative and on the grant funding provided by the US Alfred P Sloan Foundation or the National Science Foundation.

While DAIs are no longer in use for persistent author identification purposes in the Netherlands – they have been replaced by ISNIs and ORCIDs – the identifiers are still kept in the databases and it's still possible to retrieve them by using the *info:eu-repo* application profile that was defined when they were first created. DAIs are expressed as 9-digit strings and are usually stored in the national-level NARCIS portal alongside the authors' ORCIDs.

The best way to find the researcher with DAI 110254589 – as an example – would be to search for the string *info:eu-repo/dai/nl/110254589* on a browser. Because this is the namespace and format for having DAIs stored in NARCIS⁴, the browser will be able to locate the NARCIS record for the researcher, see below, even if the DAI itself is not being used anymore.

<https://www.narcis.nl> › PRS1243208 › Language › nl ▼

Prof.dr. W.J.M. (Wim) Voermans - NARCIS.nl

Legislative studies, administrative and constitutional law. Digital Author ID, *info:eu-repo/dai/nl/110254589*. ISNI ...

By clicking on the browser result in NARCIS we can then retrieve the individual record for this researcher. The section devoted to identifiers within that person record shows the DAI together with other connected author IDs such as the ISNI and the ORCID. Both these 'live' author IDs (ISNI and ORCID) are hyperlinked to their metadata pages for this author, namely <https://isni.org/isni/0000000109163015> and <https://orcid.org/0000-0002-3045-5570> while the DAI is not hyperlinked anymore.

ADMINISTRATIVE CONSTITUTIONAL LEGISLATIVE STUDIES

Expertise	Legislative studies, administrative and constitutional law.
Expertise (NL)	Wetgevingsleer; bestuursrecht en staatsrecht
Digital Author ID	info:eu-repo/dai/nl/110254589
ISNI	>  ISNI 0000 0001 0916 3015
ORCID	>  https://orcid.org/0000-0002-3045-5570
Addition	Portefeuillehouder onderzoek, directeur E.M. Meijers Instituut

The ISNI webpage for this author still includes a reference to the original DAI in the section devoted to "sources", even if again it's not hyperlinked in the same way other sources are such as the ARK record <http://ark.bnf.fr/ark:/12148/cb155030473> kept at the Bibliothèque nationale de France, or the nid in the Catalogue of the German National Library (DNB), <https://d-nb.info/gnd/173885896>.

isni.oclc.org/xslt/DB=1.2//CMD?ACT=SRCH&IKT=8006&TRM=ISN%3A0000000109163015

Sources: [VIAF](#) [VIAF](#) [DNB](#) [LNB](#) [NUKAT](#) [SUDOC](#) [WKD](#)
[BNF](#)
[DAI](#)
[LCNACO](#)
[NLL2U](#)
[NLN](#)
[NTA](#)
[OCLCT](#)



The ‘note on the future of the Dutch infrastructure for author identifiers’⁵ issued in 2015 by the DAI Working Group (composed of representatives from SURF, KNAW/DANS, Royal University Groningen, Tilburg University, TU Delft, the Dutch Royal Library, OCLC and Erasmus University Rotterdam) describes the implementation process for the DAI and explores the possible ways forward in the light of the arrival of new international alternatives for author identification such as VIAF, ISNI and ORCID and of the key changes in the research information management landscape in the Netherlands. A presentation⁶ delivered later in the year by Niels van Dijk (SURFnet) and John Doove (SURFmarket) summarised the findings of the report above.

The emergence of the DAI as a first solution for unique and persistent author identification addressed the two main issues that ORCID would later tackle as well: the parallel problems of (i) duplicated names (i.e. different researchers with the same name) and the difficulties in attributing the publications to the right author and of (ii) duplicated author entries whereby the same person would be recorded with different names due to variations in pen name on the publications, due to name changes upon marriage or other factors.

The tightly integrated nature of the Dutch research information management landscape at the time the DAI was conceived – where all Dutch universities and a number of research centres were using the same CRIS solution, namely METIS – together with the existence of

the OCLC/PICA Dutch Thesaurus of Author Names (NTA) containing entries for all published authors in the country made it possible to devise a workflow to register unique and persistent author identifiers – coupled to NTA entries – whose metadata would be provided by institutions via their METIS CRIS systems.

This workflow was disrupted by the decision to move away from METIS – a home-grown CRIS originally developed at the Radboud University Nijmegen – as a common CRIS solution for all Dutch universities and by the realisation that the new international author identifiers made more sense in an area like research where national borders are largely meaningless. This note on the future of the Dutch infrastructure for author identifiers explores the possible options to keep minting and maintaining DAIs on the basis of ‘new’ CRIS solutions such as Pure (Elsevier) and Converis (ThomsonReuters at the time) and at the same time to merge the DAI layer into ISNI or ORCID.

On this latter aspect, both the note and the presentation by Van Dijk and Doove highlight several key aspects that were kept in mind when deciding which way to go in terms of superseding the DAIs:

- ▶ The International Standard Name Identifier (ISNI, ISO 27729) was launched in 2011 and contains 8,6 million person records as of Feb 2015, 2,25 million of which belong to researchers

- ▶ The ORCID database was launched on 16 October 2012 and contained 1.591.795 records worldwide at the time the presentation was delivered in 2015.

Dutch authors unknown

- ▶ “ORCID sits closer to researchers while ISNI sits closer to libraries” – meaning that the use cases for ORCID usage such as identifying themselves in manuscripts or project proposals are specifically targeting authors while ISNI offers better, more controlled metadata quality even if the associated workflows for issuing persistent unique author identifiers are actually removed from researchers themselves
- ▶ Other relevant factors raised in the report are how (and which) author identifiers are promoted and used by publishers and whether persistent unique author identifiers may also be issued for deceased researchers (which is not the case for ORCID)
- ▶ New workflows to develop an automatic coupling to emerging institutional CRIS solutions like Pure or Converis would be required for the author identifier minting mechanism to continue operating as originally designed
- ▶ The report stresses that there is not just a single solution available (i.e. replacing DAIs with ISNIs or with ORCIDs) but there is indeed room for both systems

The report finally provides a set of recommendations on how to move ahead, among them:

- ▶ To choose ISNIs as a sustainable identifier for authors’ names
- ▶ To use ORCID as an identifier within the domain of scholarly communications
- ▶ To motivate researchers to claim their ORCIDs
- ▶ To register ORCIDs (and ISNIs) in the [institutional] research information management systems

These recommendations would eventually be followed, as shown in the current NARCIS record for an example researcher above that includes his DAI, ISNI and ORCID identifiers.

The “NWO Persistent Identifier Strategy”⁷ released in 2021 by the Dutch Research Council (NWO, the main public research funder in the Netherlands) and SURF provides a roadmap for implementing a range of persistent identifiers – including Grant IDs and organisational IDs – to serve the needs of the research landscape from a research funder perspective. Not surprisingly, this roadmap does no longer include any reference to the DAIs, which have effectively been left behind by now.

Aanbevelingen

1. **Kies ISNI als duurzame identifier voor Auteursnamen.**
2. **Gebruik ORCID als identifier binnen de wetenschappelijke communicatie.**
3. **Motiveer onderzoekers hun ORCID te claimen.**
4. **Registreer ORCID (en ISNI) in het onderzoeksinformatiesysteem.**



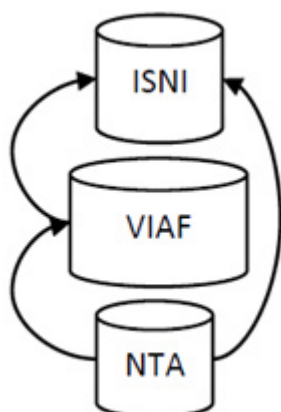
3 Issues around risks and trust regarding the transition away from DAIs

As mentioned above there are several aspects in which the replacement/superseding of Digital Author Identifiers in the Netherlands constitutes an interesting case study for PID management regarding the risks and trust issues that are the focus of this study.

3.1 The process for having DAIs replaced and/or superseded

The recommendations made in the note on the future of the Dutch infrastructure for author identifiers were largely followed. New coupling mechanisms were eventually devised for ‘new’ institutional CRIS solutions like Pure or Converis and the suggested workflow was also implemented to create (or update) ISNIs for existing DAIs via the integration of the Dutch Thesaurus of Author Names (NTA) into the Virtual International Authority File (VIAF, <https://www.oclc.org/en/viaf.html>) then transferring these into the ISNI database to create ISNI IDs that could later be completed with additional metadata like institutional affiliation by directly exporting DAIs into the ISNI database.

Inlezen van NTA in ISNI via VIAF.



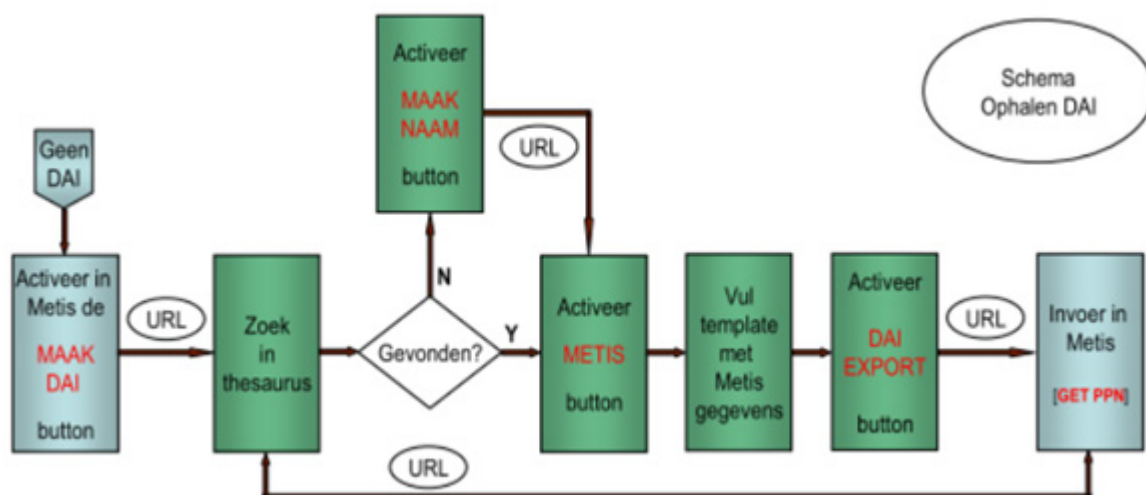
At the same time the ORCID solution was promoted across researchers and institutions, the former being prompted to create, self-claim and maintain their ORCID profiles at no cost. This would eventually become a mainstream effort by Dutch institutions when SURF – as it was also recommended in this report – became a member of ORCID⁸ and started running pilots with pioneering institutions to make ORCIDs available for all their staff. This is the same way the ORCID adoption process worked in other countries.

3.2 Technical vs community-oriented issues regarding trust

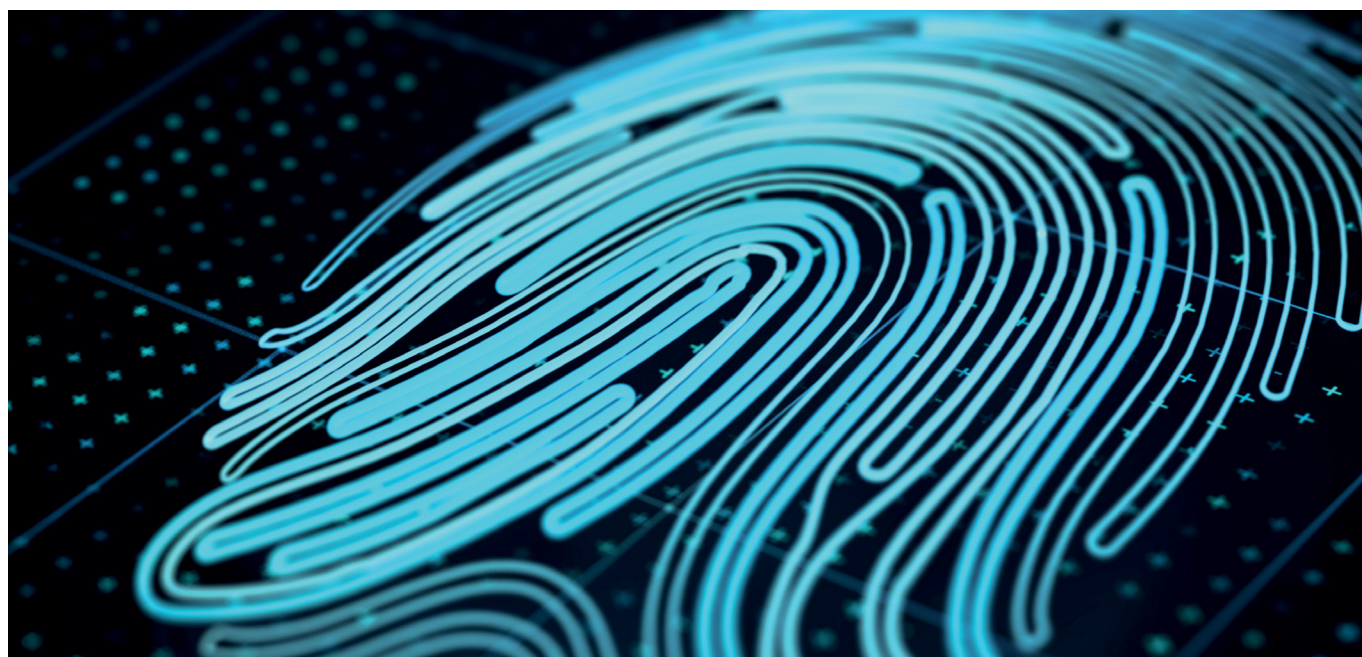
While the technical superseding of DAIs was reasonably straightforward, there was also a need to explain the update to the end-users of the persistent identifier scheme, i.e. the researchers themselves, who had previously been told that DAIs would become the default approach to author identification in the Netherlands.

This potential trust-related issue was successfully addressed via the direct involvement of institutions – mainly via their CRIS administrators – in the process for issuing author identifiers of one sort or another. The original workflow for minting DAIs involved the coupling between the institutional METIS CRIS and the Dutch Thesaurus NTA as shown in the figure on the next page. This was a manual process carried out by each METIS CRIS administrator. Multiple or successive researcher’s affiliations with various Dutch institutions could thus overlap into a single DAI that would receive input from different CRIS systems.

All researchers affiliated with Dutch institutions were assigned a DAI but because of the early nature of the development there were no use cases yet for actually using them for identification purposes with publishers or funders. This meant that even if researchers may have been aware of the new author ID infrastructure, there were actually very few or no opportunities for effectively using them.



This changed with the arrival of ORCIDs, which is an author identification initiative heavily supported by publishers and CrossRef. ORCID had thus a very clear use case right from the outset in making sure a means became available for publishers to be able to tell a given Jan Janssen in a specific publication from his namesake researcher working in a different research field. Likewise, the ORCID registration workflow made it necessary for the researcher to confirm his/her identity via an email address, meaning that – as stated in the report above – ORCIDs were actually sitting much closer to researchers than DAIs or ISNIs had ever been and would ever be.



3.3 Public vs ‘commercial’ PID infrastructures

“

When you mentioned all these different types of identifiers for all these different types of things, and then thinking about the infrastructure and the manpower, and all it would require to put that in place, promoted, maintained over the long term? I'm getting the shivers. And I've shared this many times with people. I say: "Yeah, absolutely. We need persistent identifiers for all of that. But are we going to create an organisation for each of those that we then need to become a member of?" I think that's really just not sustainable.

”

As mentioned above, this is an issue that has been frequently raised in the interviews with PID experts from various organisations that provide the backbone to this study. This issue is also clearly present in the recommendation included in the report on the future of the Dutch author identifier infrastructure to have DAIs “transferred” into the ISNI database. Although the report also mentions contacts with ORCID back in 2012 to ensure that a DAI holder would be able to couple the new ORCID to the pre-existing DAI upon ORCID registration, these conversations saw no follow-up and the 2015 report subsequently endorsed the ISNI/VIAF option instead.

“

Again, for me, [the issue] is way more [about] the long-term sustainability. And basically, the budget that comes with all that, and especially [with] this notion of really long term, I mean, we're talking not five years, 10 years, we're talking really long term here. And when I think about [the] really long term, I keep coming back to the cultural heritage organisations, the national libraries, and so on, that have been with us for 100, 200 years as kind of good places for that kind of infrastructure. Then on the other hand, just for the sustainability of it, I cannot imagine those organisations to be equally innovative, and hip, and all, you know, in the way that these PID organisations are. So from a marketing perspective, I think it would not be a great idea. From a long term sustainability perspective, I think it's way more logical to have this kind of very significant long term responsibility in the hands of multiple cultural, long-term cultural heritage organisations.

”

The interviews with PID experts on this topic mainly addressed the issue of sustainability. Despite the enormous progress made by ORCID in securing integrations with many different stakeholders in the scholarly communications domain, there was a widespread perception among interviewees that if for instance National Libraries were able to maintain this sort of infrastructure (which would of course need the appropriate funding allocation), it might be possible to better address certain issues such as how to maintain DOIs from publishers that cease operations. There is also an occasional – if not widespread – perception among the user community that ORCIDs are tools specifically designed to serve the needs of publishers, a perception that carries some weight even if it's not entirely true – research funders are increasingly benefitting from ORCID use as well as institutions themselves. Finally the perceived lack of mechanisms devoted to ensure data quality is a topic that gets regularly raised by the community.^{9, 10}

“

This is why the economic model of ORCID is weak. ORCID cannot sell services of high quality because of the original model of ORCID: it's free, it's for every researcher, there is no quality control, there is no hassle to people. So it's a limitation of the development of ORCID for the future. I think there is no solution. Because you don't ask people to prove who they are. And if you don't manage the unicity of ORCID, then one person can have 10, 20 ORCIDs and we can do nothing against that. For ORCID, the problem is not a technical problem, it's a conceptual problem. But this is the price to pay for a free system that people are keen to adopt.

”

So, while DAIs have in fact been superseded by a combination of ISNIs and ORCIDs, the SURF report raises some concerns on the way forward and firmly chooses to rely on an international standard (ISNI) “that foresees a central quality control guaranteed by the Bibliothèque National de France and the British Library”.

“

However, in terms of quality control, ORCID has gotten better at that, but especially in those years, they weren't as proficient at that as they are now I would say. So they didn't do a lot of checking for double email, for double ORCIDs that were made by researchers, or checking whether the organisational email was indeed someone that worked there. So the validation steps were not as thorough as they are now. So you couldn't fully rely on ORCID to take all the functionality, including the quality standards that we wanted to meet from the DAI part. This metadata quality is something that for libraries is their bread and living, that's there. So, the libraries specifically asked us to not just say, okay, ORCID is going to be the de facto standard in research information, so just do that. They wanted both.

”



4. Authorship

This case study has mainly been written by Pablo de Castro (University of Strathclyde and euroCRIS, ORCID <https://orcid.org/0000-0001-6300-1033>) within a team of consultants including Ulrich Herb (Saarland University, ORCID <https://orcid.org/0000-0002-3500-3119>), Laura Rothfritz (Humboldt University Berlin, ORCID <https://orcid.org/0000-0001-7525-0635>) and Joachim Schöpfel (University of Lille and euroCRIS, ORCID <https://orcid.org/0000-0002-4000-807X>) under the umbrella of scidecode science consulting (ROR <https://ror.org/02c0bjd31>). The work has been overseen by the Knowledge Exchange Task & Finish Group whose composition is listed at <https://www.knowledge-exchange.info/event/pids-risk-and-trust>.

5. Literature

1. DAI wikipedia page, https://nl.wikipedia.org/wiki/Digital_Author_Identifier
2. Project record for DRIVER-II: "Networking European Scientific Repositories" (2007-2009) at DTU Orbit portal, <https://orbit.dtu.dk/en/projects/driver-ii-networking-european-scientific-repositories>
3. IdRef: Identifiants et Référentiels pour l'enseignement supérieur et la recherche, <https://www.idref.fr/autorites.jsp>. See an author IdRef example at <https://www.idref.fr/130971367>
4. DAI spec in the info:eu-repo application profile, <https://wiki.surfnet.nl/display/standards/info-eu-repo#infoeu-repo-DigitalAuthorIdentifier>
5. "Nota toekomst Nederlandse infrastructuur voor auteurs identifiers", <https://docplayer.nl/8892670-Nota-toekomst-nederlandse-infrastructuur-voor-auteursidentifiers.html>
6. "ORCID: een unieke, persistente identifier voor onderzoekers" (Niels van Dijk, SURFnet and John Doove, SURFmarket), <https://docplayer.nl/14113393-Orcid-een-unieke-persistente-identifier-voor-onderzoekers-niels-van-dijk-surfnet-john-doove-surfmarket.html>
7. Cruz M, Tatum C (2021) "NWO Persistent Identifier Strategy", <https://zenodo.org/record/4695367>
8. SURF BV as ORCID-NL Consortium lead, <https://orcid.org/members/001G000001rkowJIAQ-surf-bv>
9. Baglioni, M., Manghi, P., Mannocci, A. and Bardi, A., 2021. We Can Make a Better Use of ORCID: Five Observed Misapplications. Data Science Journal, 20(1), p.38. DOI: <http://doi.org/10.5334/dsj-2021-038>
10. Ritter F (2021) "Problems with ORCID, or 10 reasons why I don't have an ORCID number", <http://www.frankritter.com/problems-with-orcid.html>

Knowledge Exchange
C/O Jisc
4 Portwall Lane,
Bristol, BS1 6NB
United Kingdom

T +44 203 697 5804
E office@knowledge-exchange.info