

Towards Trusted Identities for Swiss Researchers and their Data

Julien A. Raemy
HES-SO University of Applied Sciences

René Schneider
HES-SO University of Applied Sciences

Abstract

In this paper we report on efforts to enhance the Swiss persistent identifier (PID) ecosystem. We will firstly describe the current situation and the need for improvement in order to describe in full detail the steps undertaken to create a Swiss-wide model. A case study was undertaken by using several data sets from the domains of art and design in the context of the ICOPAD project. We will provide a set of recommendations to enable a PID service that could mint Archival Resource Key (ARK) identifiers or a flavour of Research Resource Identifiers (RRIDs) as complement to Digital Object Identifiers (DOIs). We will conclude with some remarks concerning the transferability of this approach to other areas and the requirements for a national hub for PID management in Switzerland.

Received 06 December 2018 ~ *Revision received* 18 October 2019 ~ *Accepted* 06 January 2020

Correspondence should be addressed to René Schneider, HEG-Genève, Rue de la Tambourine 17, CH-1227 Carouge, Switzerland. Email: rene.schneider@hesge.ch

An earlier version of this paper was presented at the 14th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: <http://www.ijdc.net/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

Persistent Identifiers (PIDs) are a necessary tool to assure referenceability and – as their name says – identifiability in a bi-unique manner upon which further mechanisms can be built upon, e.g. citability to mention only the probably most important one (Hakala, 2010; Van de Sompel, Sanderson, Shakar and Klein, 2014).

Their persistence is a real added value compared to the cool URIs of the Linked Data approach, which is not in opposition to a joint use of both PIDs and Linked Data. For a deeper understanding of the relationship between cool URIs and PIDs see Bazzanella, Bortoli and Bouquet (2013). In the meantime, it is often forgotten that PIDs encompass a large variety of heterogeneous and ever-growing denominators for which the metaphor of ecosystem seems to be appropriate (see Espasandin, Jaquet and Lefort (2018) for a complete overview), and that the combination of several PIDs bears a high potential for the creation of what are called Trusted Identities in research data management, where their use is primordial both for sharing and archiving data, no matter if the focus is set on open access or long term archival. As a matter of fact, PIDs seem to work as a hinge or even as a fulcrum between these two approaches in data curation.

Nevertheless, a short glance at the real existing use of persistent identifiers shows that instead of diverse ecosystems preference is given to a monoculture. In this paper we will illustrate this and provide solutions to enhance this situation with examples from the Swiss scientific landscape.

The Swiss PID Landscape

Currently the situation in Switzerland is dominated by one and only one PID, namely the prominent Digital Object Identifier (DOI), which is attributed since 2010 by the library of the Swiss Federal Institute of Technology Zürich (ETH) – a DataCite member. The PID service is used for documents and data and is only free of charge for the scientists working for the ETH; a constraint which can be seen for the inhibited use of the service.

To overcome this constraint, some institutions recommend their scientists to publish their data on Zenodo¹, which is hosted at the European Organization for Nuclear Research (CERN), being an extra-territorial organisation located on the Swiss-French border near Geneva. The specific circumstance of extra-territoriality makes other institutions forbid their scientists to publish the data there, since their policies do not allow data publication outside Swiss borders.

A third Swiss data repository is FORSbase², which is only open to data from the Social Sciences and has no strict policy concerning identifiers (with some data sets having DOIs and others not having persistent identifiers at all).

Parallel to these well-established services, the allocation of Archival Resource Key (ARK) identifiers is in preparation at the University of Basel's Digital Humanities Lab as part of the development of the Data and Service Center for the Humanities

1 Zenodo: <https://zenodo.org/>

2 FORSbase: <https://forsbase.unil.ch/>

(DaSCH)³, a long-term archive for data from (digital) humanities (Rosenthaler, Fornaro and Clivaz, 2015). These ARKs will hence only be given to data from a restricted number of scientific disciplines during and after completion of the project.

Another ongoing Swiss project named Data Life-Cycle Management (DLCM)⁴ is still under development and plans to also attribute DOIs to the data to be hosted in the near future.

No PIDs are allocated so far to persons or organizations, although another initiative, known as SWITCH edu-ID⁵, has been started to create IDs for all people involved in higher education institutions in Switzerland, developed by SWITCH – an organization for nation-wide IT-support in higher education. In this context it should be seen critical that these identifiers are not really based on a social contract (like, for example, ORCID, where each researcher willingly allows the creation of an ID for him or herself) but automatically attributed for a lifetime as soon as a person enters an academic institution; a circumstance that might bear some juridical problems if used for other things than merely administrative matters.

This mixture of monopolized attribution, contradictory policies for non-payable deposits, work in progress, and unclear legal issues, leads to an unsatisfying and confusing situation, far beyond a scenario that is needed to create trusted identities for dynamic and ready-to-ingest research data at all granularity levels including scientists and their organizations.

For this reason, the project ICOPAD (= *Identités de confiance pour les données de l'art et du design*)⁶ started in 2017 to reflect on a suitable PID model for the Swiss scientific landscape, specifying all necessary requirements and workflows.

The ICOPAD Project

Since the situation described above can only be improved through collaboration and partnership, ICOPAD started as a joint project between four Swiss institutions (namely the HEG-Genève as instigator and managing authority; the Swiss Institute for Cultural Sciences, the Swiss Institute for Art Research, the Zurich Central Library as well as the Zurich University of the Arts). All institutions worked together successfully in a previous project on Linked Data for Art and Design in Zurich (LOD-Z)⁷ (Prongue, Ricci, Schneider and Schurte, 2017). This project ended with the development of a prototypical system and the desire to further work on PIDs to loosen the dependency from cool URIs.

The goal of this 18-month project (July 2017 to December 2018) was also to work out all details underlying this need and to give an exemplary showcase that proved the structural feasibility of this approach dedicated to the disciplines of art, design, and digital humanities in order to derive general conjectures and to prove the transfer of the model and workflow created to other disciplines that share the same needs for their scientists and the data they generate. The ultimate goal of this project was to clarify all questions concerning the creation of a Swiss PID infrastructure, namely the theoretical, organisational, technical, and financial aspects. A more concrete implementation will be done in a follow-up project on a national level.

3 DaSCH: <http://dasch.swiss/>

4 DLCM: <https://www.dlcm.ch/>

5 SWITCH edu-ID: <https://www.switch.ch/edu-id/>

6 ICOPAD project: http://campus.hesge.ch/id_bilingue/projekte/icopad/default.asp (in German)

7 LODZ project: http://campus.hesge.ch/id_bilingue/projekte/lodz/default.asp (in German)

Preliminary Studies

The project started with a first study of the situation in Switzerland, which resulted in a detailed description of the situation depicted roughly above as well as a first overview of the PID universe.

For the sake of generalization and comprehensibility to the outer community, a flow chart was developed that can be understood as a decisional path to the attribution of one or more complementary PID(s) in Switzerland. This work was strongly inspired by prior work done by the Australian National Data Service and other policy checklists (ANDS, 2017) and led to the following diagram:

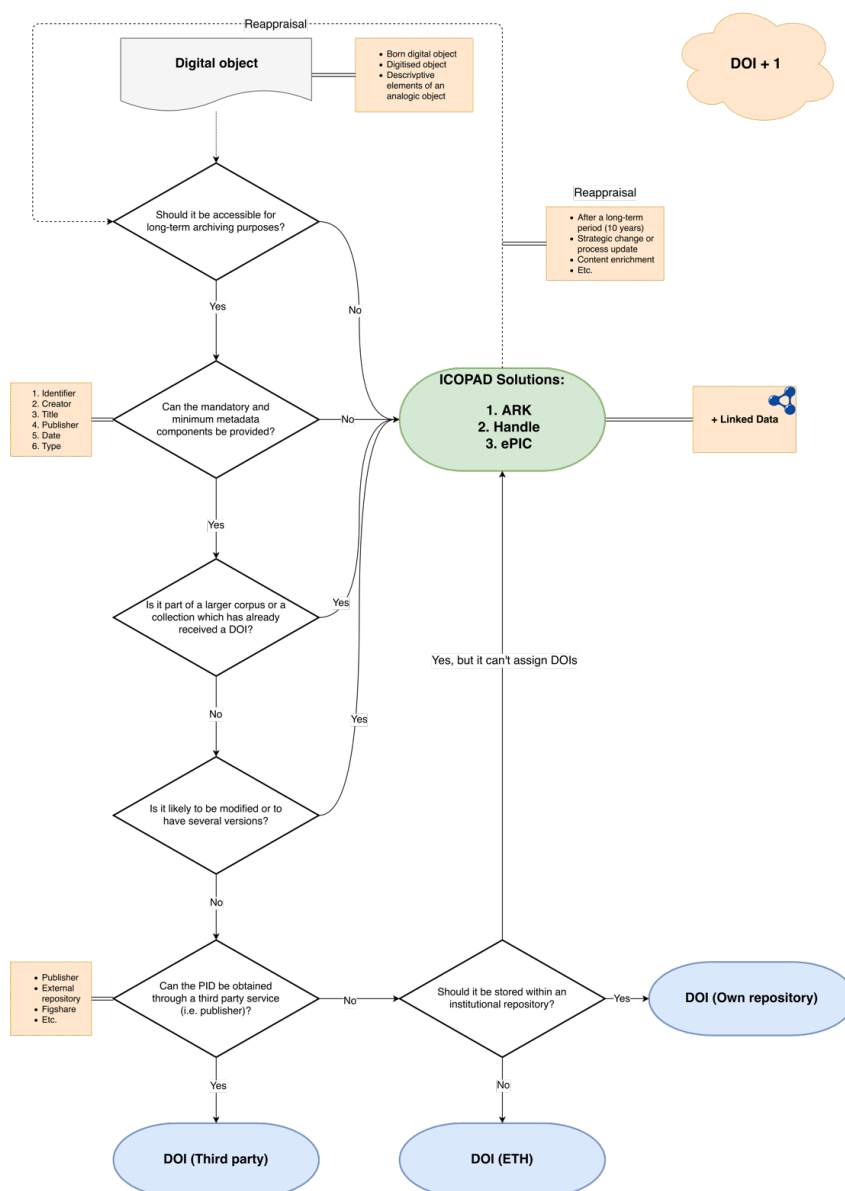


Figure 1. Generic PID Attribution Approach for the ICOPAD project.⁸

⁸ Thanks to the International Image Interoperability Framework (IIIF) Image API, this figure can be better viewed at: <https://free.iiifhosting.com/iiif/e698c0d3beb68cb6764a359849c9b161d3177f1a9c8571cb73ad525b1e943ce2/>

Based on our research, we decided that dynamic data sets (Treloar, Groenewegen and Harboe-Ree, 2007), i.e. those not yet being considered for long-term preservation, and those needing a fine level of granularity, as well as data likely to be often modified (versioning) could be better handled with PIDs other than DOIs.

Synthetic Overview and First Recommendations

To gain more insights and know more about all possibilities that the PID universe offers, a second exploratory work was initiated. This work led to a synthetic overview of almost all PIDs being practically at disposal and their graphical representation as a disc⁹.

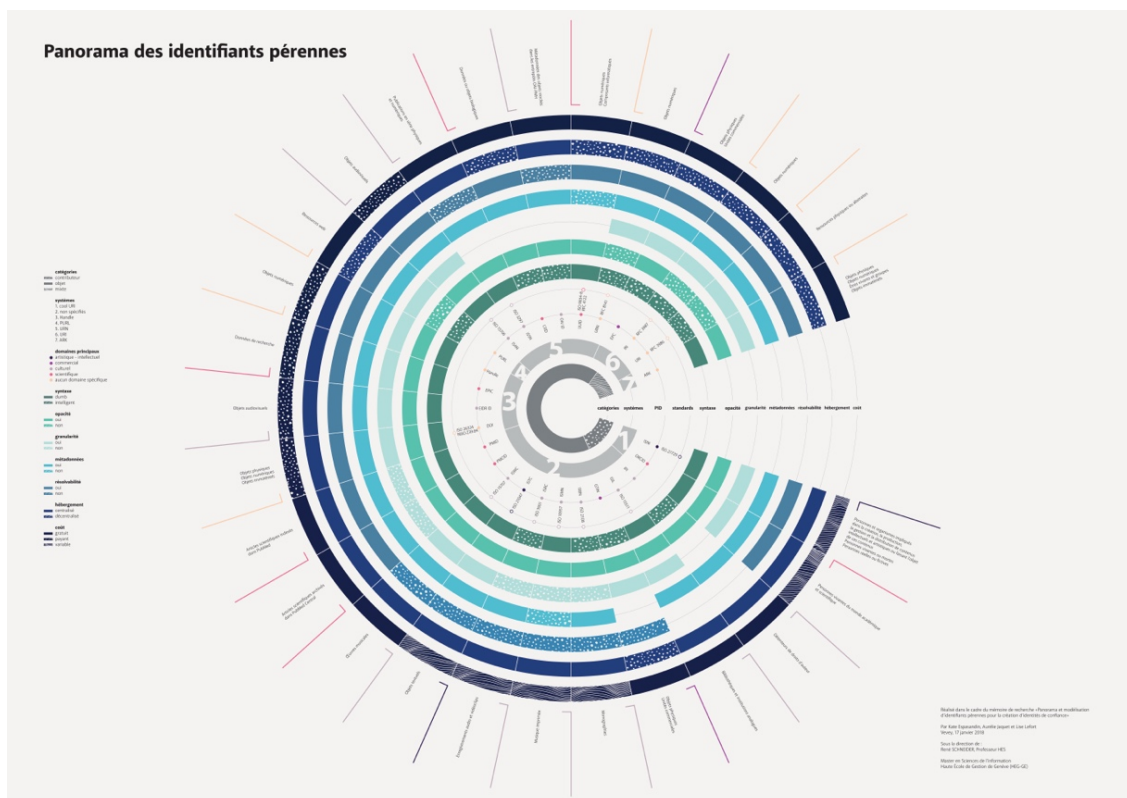


Figure 2. Panorama of PIDs (Espasandin et al., 2018).

The panorama can be viewed from the centre outwards. It is organised around ten classification criteria integrated into the circular sections (categories, systems, standards, syntax, opacity, granularity, metadata, resolvability, hosting, and cost) and two ancillary factors (main domains and the type of entities that can be described) (Espasandin et al., 2018). Based on the 27 different PIDs identified on the disc, the authors recommended an attribution system that combined ARK, DOI, and ORCID. The former could be assigned to any kind of entity and the two latter PIDs could be leveraged together.

After these two preliminary investigations the work was continued in a twofold manner:

⁹ The zoomable version can be consulted at the following URL:
<https://free.iiifhosting.com/iiif/996d2d88ad5b7901c0c770fdb6215a4eeecb360b81b45d19ed045444245910b/>

1. a detailed analysis of diverse data sets being delivered by the project partners for the exact determination of PIDs needed and;
2. the development of an overall attribution model containing one or more complements beside the DOIs attributed so far.

Data Set Analysis

The Swiss Institute for Art Research (SIK-ISEA)¹⁰, the Zurich Central Library (ZB)¹¹, and the Zurich University of the Arts (ZHdK)¹² provided nine different use cases within the ICOPAD project. The scope and the data sets of the participating institutions are summarised in the Table 1 below.

Table 1. ICOPAD participating institutions: scope and data sets.

Institution	Object of investigation and epistemological interest	Data set types/entities
SIK-ISEA	The SIK-ISEA use case is an artist's dictionary entry containing biographical information, images of his artworks as well as hyperlinks to Wikipedia and the GND (<i>Gemeinsame Normdatei</i>), which is the German integrated authority file. The interest of SIK-ISEA is – according to their website – to provide a daily updated online-encyclopaedia of biographic information concerning artists and their works in Switzerland.	Artists Artworks Dictionary entries
ZB	The provided data sets by the ZB had already DOIs but it wasn't sufficient enough to their needs as they wanted to have their manuscripts and ancient books to be citable with a finer level of granularity that could point to an area of interest for the end user (page, folio, segment, paragraph, etc.).	Digital surrogates
ZHdK	The ZHdK have several platforms (<i>eMuseum</i>) ¹³ , subdomains (<i>Medienarchiv</i>) ¹⁴ , as well as dedicated webpages on which are hosted their datasets. They all have their own (ephemeral) identifiers which are derived from their different databases. Therefore, the same instance can have several identifiers depending on which platforms it is on. They would like to unify this by attributing PID to all kinds of entities (a person, an object, an event, etc.) that they are hosting. The ZHdK would like as well to implement a PID solution with a Linked Data approach to migrate existing applications such as the eMuseum or the Medienarchiv to a both stable but nevertheless agile environment and to continue the development of their own products based on the principles of PID and LOD.	Artists Artworks Events Films Glossary entries Projects Research Data

¹⁰ The Swiss Institute for Art Research: <https://www.sik-isea.ch/>

¹¹ Zurich Central Library: <https://www.zb.uzh.ch/>

¹² Zurich University of the Arts: <https://www.zhdk.ch/>

¹³ eMuseum: <https://www.emuseum.ch/>

¹⁴ Medienarchiv der Künste: <https://medienarchiv.zhdk.ch/>

Modelling

After the period of data exploration, a modelling process started with the axiom-like assertion that the DOI-only approach followed so far in Switzerland is not sufficient, which lead to the conclusion that a complement must be found. A circumstance that was named “DOI+” internally and formally might be described as:

$$(0) \ C(doi) = x$$

This assertion leaves two options open, with $x = 1$ single identifier as complement or many, i.e. n identifiers as a complement, internally designated as “DOI + 1” or formally noted for our purpose as:

$$(1) \ C(doi) = a$$

and

$$(2) \ C(doi) = (x_1, x_2, \dots, x_n)$$

with a being a distinctly chosen PID and n the number of possible identifiers combined in a vector or n tuple.

Option (2) may further be differentiated into a sequential or parallel approach, whereas sequential means that an order of priority is assigned to the different identifiers existing for several entities and parallel means that no further distinction is given and that as soon as an identifier for an entity may be found, it is assigned.

Finally, it was decided to opt for a solution that combines both approaches and creates a link both to the cool URIs of the Linked (Open) Data world as well as to all other PIDs. Internally this model was noted (without claiming it to be a correct mathematical formula) as

$$(3) \ C(doi) = a \rightarrow owl:sameAs(x_1, x_2, \dots, x_n)$$

which means that firstly one specific chosen PID a was assigned to all entities before transforming them into a Linked Data representation using the OWL predicate `sameAs` to assign further PIDs to an entity. At the end of the modelling process, the only remaining question was to choose a proper PID for a .

Choosing the Appropriate PID for All Entities

Based on the preliminary study made by Espasandin, Jacquet and Lefort (2018), preference was given to the ARK identifier and the corresponding Name-to-Thing (N2T) resolver. The ARK anatomy is given in Figure 3.

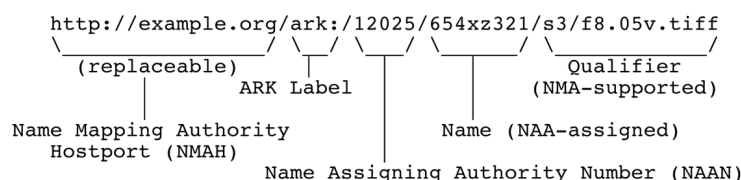


Figure 3. ARK Anatomy.¹⁵

It can be seen that ARKs can provide a fine level of granularity through the use of parallel subdomains or Name Mapping Authority Hostports (NMAH) as well as qualifiers. Each organisation is able to implement their own ARK policy and naming practices. For instance, Kunze and Rodgers (2008) suggest that ARK identifiers should not be reassigned once it has been made public, should be opaque (i.e. no widely recognisable semantic information) and should contain a check character to guard against common transcription errors.

ARKs were chosen for the following reasons, to name only a few:

- ARK identifiers are free once you have found an institution that is willing to connect to the California Digital Library (CDL)'s Name Assigning Authority Number (NAAN) registry¹⁶, thus they seem to be the appropriate alternative to the pay-per-ID approach practiced so far in Switzerland;
- ARKs are built using a completely different theoretical model, consisting of a decentral and domain (i.e. DNS) agnostic approach, which allows a considerable freedom for the internal management of the data sets;
- ARKs allow easy use with LOD – a circumstance that has proven its feasibility in a number of projects, e.g. the data.bnf.fr¹⁷ project for creating a Linked Data application for several data bases hosted by the French National Library (BnF);
- ARKs can be combined effortlessly with other specifications, such as the International Image Interoperability Framework (IIIF) canonical URI syntax¹⁸.

Recommendations

In the context of the ICOPAD project, several alternatives or approaches have been selected to mint ARKs or to design their own identifiers.

For the latter, the PID that has been considered to be emulated is a mimic of the Research Resource Identifier (RRID)¹⁹ used for giving consistent links to biological data in the following format: `RRID:Identifier`. It is leveraged by, among others, the University of California²⁰ and these identifiers can be found on Google Scholar or PubMed. For the purpose of the ICOPAD project, the authors and the participating

¹⁵ The ARK syntax can be synthesized as follows: `[http://NMAH/]ark:/NAAN/ARK Name [Qualifier]`

¹⁶ NAAN registry: https://www.cdlib.org/services/uc3/naan_registry.txt

¹⁷ The data.bnf.fr project: <http://data.bnf.fr/>

¹⁸ The Canonical URI Syntax in the IIIF Image API 2.1: <https://iiif.io/api/image/2.1/#canonical-uri-syntax>

¹⁹ Resource Identification Initiative: <https://www.force11.org/group/resource-identification-initiative>

²⁰ A presentation entitled “*Adventures in rolling your own persistent identifier system*” given at PIDapalooza 2018 by Tom Gillespie, UC San Diego, inspired the authors of this article.

institutions have decided to label this RRID-like identifier as such: The Art and Design Identifier (ADID).

1. **ARK via its own means:** Each organisation could ask a NAAN to the CDL and mint their own ARKs. They could decide whether they want to deploy one or several NMAH. This solution could also be conducted for some time before switching to the ones proposed by the implementation of a national Hub minting either ARKs (3) or RRID-like identifiers (5).
2. **ARK via DaSCH:** As stated in the Swiss PID landscape section, the DaSCH project attributes ARK identifiers for the needs of its research data archival system. The internal structure of the system could theoretically and practically allow the minting of ARKs to third members. The NMAH (ark.dasch.swiss) and NAAN (72163) would be then the ones that DaSCH has deployed or received. Unfortunately, DaSCH does not see itself as the future national hub at the moment.
3. **ARK via a national hub:** The creation of a Swiss Hub that is able to mint ARKs and offer services to organisations that have participated or are interested in the ICOPAD project effort. This national hub could deploy one or several NMAH.
4. **RRID-like (ADID) via its own means:** Each organisation could mint their own RRID-like identifiers (ADIDs). This solution could also be conducted for some time before switching to the ones proposed by the implementation of a national Hub minting either ARKs (3) or ADIDs (5).
5. **RRID-like (ADID) via a national hub:** The creation of a Swiss Hub that is able to mint ADIDs rather than ARKs. They could also deploy one or several hostname services (the equivalent to NMAH in the ARK anatomy) to resolve PIDs.

Figure 4 shows a matrix that combines two types of variables: hostname (or the sub-domain that can resolve PIDs) and the PID authority (in other words the organisation that is able to mint PIDs). In the matrix, the five approaches are exemplified with three projected and hypothetical PID URLs from the ICOPAD contributing institutions. The “core numbers” of the digital objects are the following:

- SIK-ISEA: *lex4000336z* – a lexicon entry in Sikart²¹, the online information system on historical and contemporary art in Switzerland
- ZB: *ema45784v* – a digitised manuscript
- ZHdK: *bfic02001t* – an art exhibition stored in their Medienarchiv²² database

For the first and third approaches, *11111*, *22222*, *33333*, and *99999* have been taken as NAANs and the different sub-domains or NMAs used in the matrix are speculations based either on existing platforms or supposed ones.

²¹ Sikart: <http://www.sikart.ch/>

²² ZHdK – Medienarchiv der Künste: <https://medienarchiv.zhdk.ch/>

		PID Authority =Name Assigning Authority Number (NAAN)	
		One PID Authority for all organisations	Each institution is its own PID Authority
Hostname - Identifier Service Provider =Name Mapping Authority Hostport (NMAH)	One hostname for all types of data	2) ARK via DaSCH http://ark.dasch.swiss/ark:/72163/lex4000336z http://ark.dasch.swiss/ark:/72163/ema45784v http://ark.dasch.swiss/ark:/72163/bfic02001t	1) ARK via its own means http://n2t.sikart.ch/ark:/11111/lex4000336z http://n2t.zb.ch/ark:/22222/ema45784v http://n2t.zhdk.ch/ark:/33333/bfic02001t 4) RRID-like (ADID) via its own means http://n2t.sikart.ch/adid:lex4000336z http://n2t.zb.ch/adid:ema45784v http://n2t.zhdk.ch/adid:bfic02001t
	One hostname per institution	3) ARK via a national Hub http://sikart.icopad.ch/ark:/99999/lex4000336z http://zb.icopad.ch/ark:/99999/ema45784v http://zhdk.icopad.ch/ark:/99999/bfic02001t 5) RRID-like (ADID) via a national Hub http://sikart.icopad.ch/adid:lex4000336z http://zb.icopad.ch/adid:ema45784v http://zhdk.icopad.ch/adid:bfic02001t	N/A
	One hostname per domain/activity	3) ARK via a national Hub http://artist.icopad.ch/ark:/99999/lex4000336z http://manuscripts.icopad.ch/ark:/99999/ema45784v http://medienarchiv.icopad.ch/ark:/99999/bfic02001t 5) RRID-like (ADID) via a national Hub http://artist.icopad.ch/adid:4000336z http://manuscripts.icopad.ch/adid:ema45784v http://medienarchiv.icopad.ch/adid:bfic02001t	1) ARK via its own means http://artist.sikart.ch/ark:/11111/lex4000336z http://manuscripts.zb.ch/ark:/22222/ema45784v http://medienarchiv.zhdk.ch/ark:/33333/bfic02001t 4) RRID-like (ADID) via its own means http://artist.sikart.ch/adid:lex4000336z http://manuscripts.zb.ch/adid:ema45784v http://medienarchiv.zhdk.ch/adid:bfic02001t

Figure 4. Hostname/PID Authority Matrix.

Conclusions and Future Work

We have seen in this paper that the importance of persistent identifiers for research data is still underestimated. Briefly said, they can be seen as the *conditio sine qua non* for adhering to the FAIR principles (Wilkinson et al., 2016). Despite the diversity of PIDs, the Swiss landscape heavily depends on DOIs, (with some alternatives and/or complements starting to be at disposal). The major problem may consist in the fact that whenever people think of PIDs they chiefly think of DOI and it remains difficult to convince them that alternatives or PIDs for other entities exist. Besides that, the DOIs are mainly assigned to publications and data sets with a rough level of granularity. They are also less suitable for linked data enrichment and do not permit decentralised management.

Secondly, people often see linked data or cool URIs as an alternative for PIDs, without thinking of them as possibly being two sides of the same coin in order to bring research data to their full potential. The state-of-the-art work done for PIDs in general, their use in Switzerland, as well as the still unfulfilled scientific needs (illustrated using the example of several data sets from the faculty of Art & Design) has shown that there is still lots of work to do.

The peculiarity of the Swiss scientific landscape, as being quite developed and advanced but not fully connected to the rest of Europe requires some common efforts

with respect to PIDs. So far, Swiss research data are far from being FAIR. This paper's aim was to show that minor efforts, such as the choice of a second and complementary PID, namely ARK, may lead to path-breaking solutions for infrastructures and services. With respect to this, it should not be forgotten, that despite all technical issues

'Persistence is not dependant on the identifier itself, but on legal, organisational and technical infrastructure'

and that the

'(...) best that an identifier can do is to lead users to the services that support robust reference.'

Without a *camino real* being at hand actually due to the dynamic development of the realm, this study should be seen as a first pavement of a road with a view to a unified framework. The latter will enable us to support research in Switzerland and to allow a reasonable management of research data within Switzerland, as well as the interconnectivity of the data to create links to the outer world.

In order to leverage upon the work of the ICOPAD project, an exploration project which will model a hub prototype, in particular by producing a blueprint that takes into account the financial and technical aspects, will be carried out by Master students in Information Science in Geneva between February 2019 and January 2020. In parallel, a request for funding from the swissuniversities-p5 fund to launch a Hub in Switzerland in the year 2020 will also be carried out.

The models of the ICOPAD project, which had many heterogeneous use cases in the field of art and design, can be perfectly applied and transferred to all scientific fields because the need for persistent identification is universal and also because ARK does not consider the type of data that are being identified.

The latter project will develop a complementary infrastructure for the low-cost attribution of PIDs with high granularity based on the ARK identifier. This incremental approach will allow data curation for all scientific domains comprising short as well as long tail data sets at an early stage, i.e. during active research and thus long before the transfer to a long-term archive and even before transmission into an open or closed repository.

Acknowledgements

The authors would like to thank all involved partner institutions and their collaborators, i.e. Paul Brunner (SIK-ISEA), Matthias Oberli (SIK-ISEA), René Schurte (ZB), Susanne Schumacher (ZHdK) and Birk Weiberg (ZHdK) as well as Nicolas Prongué, Kate Espasandin, Lise Lefort, and Aurélie Jacquet for their valuable contribution to this project.

References

ANDS. (2017). *Persistent identifiers: Expert level*. Australian National Data Service. Retrieved from <https://www.ands.org.au/guides/persistent-identifiers-expert>

- Bazzanella, B., Bortoli, S., & Bouquet, P. (2013). Can persistent identifiers be cool? *International Journal of Digital Curation*, 8(1), 14–28. doi:10.2218/ijdc.v8i1.246
- Espasandin, K., Jaquet, A., & Lefort, L. (2018). *Panorama et modélisation d'identifiants pérennes pour la création d'identités de confiance* (Research paper). Haute école de gestion de Genève, Geneva, Switzerland. Retrieved from <https://doc.rero.ch/record/309479>
- Hakala, J. (2005, June). *Persistent identifiers: The 7 levels of identification*. Presented at the ELAG (European Library Automation Group) 2005, CERN, Geneva. Retrieved from https://indico.cern.ch/event/408136/sessions/163227/attachments/815321/1117183/Persistent_identifiers_elag2005.ppt
- Hakala, J. (2010, October 13). Persistent identifiers: An overview. Retrieved from <http://www.metadaten-twr.org/2010/10/13/persistent-identifiers-an-overview/>
- Kunze, J., & Rodgers, R. (2008). The ARK Identifier Scheme. Retrieved from <https://escholarship.org/uc/item/9p9863nc>
- Prongué, N., Ricci, F., Schneider, R., & Schurte, R. (2017). Art and design as linked data: the LODZ project (Linked Open Data Zurich). *Libellarium: Journal for the Research of Writing, Books, and Cultural Heritage Institutions*, 9(2). doi:10.15291/libellarium.v9i2.256
- Rosenthaler, L., Fornaro, P., & Clivaz, C. (2015). DASCH: Data and Service Center for the Humanities. *Digital Scholarship in the Humanities*, 30(suppl_1), i43–i49. doi:10.1093/llc/fqv051
- Treloar, A., Groenewegen, D., & Harboe-Ree, C. (2007). The data curation continuum: Managing data objects in institutional repositories. *D-Lib Magazine*, 13(9/10). doi:10.1045/september2007-treloar
- Van de Sompel, H., Sanderson, R., Shankar, H., & Klein, M. (2014). Persistent identifiers for scholarly assets and the web: The need for an unambiguous mapping. *International Journal of Digital Curation*, 9(1). doi:10.2218/ijdc.v9i1.320
- Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016, March 15). The FAIR Guiding Principles for scientific data management and stewardship [Comments and Opinion]. doi:10.1038/sdata.2016.18