

Persistent Identifiers for Digital Cultural Heritage

Jonathan Clark
International DOI Foundation
Rembrandtlaan 12
1231 AC Loosdrecht
The Netherlands
+31654733572
jonathanmtclark@gmail.com

Maurizio Lunghi
Associazione ARCA MEMORIE
Florence, Italy, 50100
Italy
+393351396371
mauriziolunghi66@gmail.com

Remco van Veenendaal & Marcel Ras
National Archives
P.O. box 90520
2509 LM The Hague, Netherlands
+31629451951 / +31703140180
Remco.van.veenendaal@nationalearchief.nl
Marcel.ras@kb.nl

Juha Hakala
National Library of Finland
Fabianinkatu 35
00170 Helsinki
Finland
+358 50 382 7678
juha.hakala@helsinki.fi

ABSTRACT

This is an introductory level tutorial. The goals of the tutorial are: to explain what persistent identifiers are and why they are so important; to discuss what the criteria are for trustworthy identifier systems; to introduce a decision tree tool that helps determine which system is the most appropriate for a particular set of needs; and to share case studies that are relevant to the Digital Preservation community.

Keywords

Persistent Identifiers; Handle System; DOI; ARK; URN; URN:NBN; PURL.

1. INTRODUCTION

Over the past years, a growing number of collections belonging to archives, libraries, media, museums, and research institutes have been digitized and made available online. These are exciting times for ALM institutions. On the one hand, they realize that, in the information society, their collections are goldmines, as “data is the new gold” [1]

On the other hand, unfortunately most heritage institutions do not yet meet the basic preconditions for long-term availability of their collections. Apart from the problems of digital preservation the digital objects often have no long lasting fixed reference. URL's and web addresses change. For instance, some digital objects that were referenced in Europeana and other portals can no longer be found. References in scientific articles can have a very short life span, which is damaging for scholarly research, as is shown in the well-known article of Herbert van de Sompel and others [2].

Thus, in this digital world there is a need to unambiguously determine what a resource is and where it can be found in a way that is persistent over time. However, the identifiers themselves are simply strings of numbers and not inherently persistent. What makes them persistent are the policies, organization and agreements that sit behind these numbers.

Many different systems have emerged over the years, some more robust than others. What is the difference between them all and how to choose between the various options available?

What are the criteria to judge whether a PI system can be trusted or not? [3]

Finally, of course a choice must be made for a PI system that meets the long-term needs of the ALM institution. In 2015, the Dutch Digital Heritage Network (NDE) [4] started a two-year work program to co-ordinate existing initiatives in order to improve the (long-term) accessibility of the Dutch digital heritage for a wide range of users, anytime, anywhere. [5]

The NDE is a partnership established as an initiative of the Ministry of Education, Culture and Science. The members of the NDE are large, national institutions that strive to professionally preserve and manage digital data, e.g. the National Library, The Netherlands Institute for Sound and Vision, the Netherlands Cultural Heritage Agency, the Royal Netherlands Academy of Arts and Sciences, the National Archive of the Netherlands and the DEN Foundation, and a growing number of associations and individuals both within and outside the heritage sector.

Meanwhile, other institutions such as the British Film Institute, the National Diet Library of Japan and the European Union are moving forward with interesting projects to assign persistent identifiers to their archives.

2. TOPICS

The tutorial on persistent identifiers for digital cultural heritage will consist of four parts

1. Introduction: what are persistent identifiers (PI) and why are they important?
 - a. Knowing what's what and who's who
 - b. The importance of Social Infrastructure
 - c. Review of current identifier systems
2. The criteria for reliable & trustworthy PIs
 - a. Insights from the APARSEN project [6] [7]
 - b. Recommendations of the Research Data Alliance (RDA)[8]

3. The NDE decision tree for PIs
 - a. A tool in form of an online questionnaire that guides cultural heritage organizations through the process of selecting a particular type of Persistent Identifier (Handle, DOI or NBN:URN) [9]
 - b. Discuss the applicability of the decision tree outside the Netherlands
4. Case studies
 - a. NDE from the Netherlands
 - b. National Diet Library, Japan
 - c. Office of Publication of the European Union

SCOPE

The intention is that this tutorial be “PI-agnostic”; that is, any persistent identifier (Handle [10], DOI [11], ARK [12], URN:NBN [13], URN [14] and PURL [15]) will be discussed.

3. INTENDED AUDIENCE

Anyone interested in learning more about persistent identifiers for digital information. This tutorial has a strong practitioner focus and will be especially interesting for those working with Digital Archives and Digital Collections. This tutorial also acts as an introduction and level set for the Workshop on Smart Persistent Identifiers. Participants may be interested to read some background articles [16] [17] [18] [19]

4. EXPECTED LEARNING OUTCOMES

Participants will leave this tutorial with a clear understanding of what persistent identifiers are and why they are important. They will have an overview of the different identifier systems in use today and have ample opportunity to answer their questions.

5. SPEAKERS

The tutorial speakers will be Jonathan Clark, Managing Agent for the DOI Foundation, Remco van Veenendaal of the Dutch National Archives, Marcel Ras of the National Coalition for Digital Preservation and Maurizio Lunghi of the Associazione ARCA MEMORIE and Juha Hakala of the National Library of Finland.

6. REFERENCES

- [1] “Data is the new gold”. Is the title of a speech from Neelie Kroes as Vice-President of the European Commission responsible for the Digital Agenda. Opening Remarks on a Conference on Open Data Strategy in Brussels, 12th December 2011. http://europa.eu/rapid/press-release_SPEECH-11-872_en.htm?locale=en
- [2] Herbert van de Sompel et al, Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot, Plos One, December 26, 2014. <http://dx.doi.org/10.1371/journal.pone.0115253>
- [3] iPRES 2012, ‘Interoperability Framework for Persistent Identifier systems’, iPRES 2012, Toronto, urn:nbn:it:frd-9491 <https://ipres.ischool.utoronto.ca/sites/ipres.ischool.utoronto.ca/files/iPres%202012%20Conference%20Proceedings%20Final.pdf>; LIBER 43 Annual conference,

- “Implementing an Interoperability Framework for Persistent Identifier systems” in Riga, Latvia, on 2-5 July 2014, ISBN 978-9984-850-23-8
- [4] The Digital Heritage Network in the Netherlands <http://www.ncdd.nl/en/ncdd-projects/digital-heritage-network/>
 - [5] Persistent Identifier project information / National Coalition for Digital Preservation <http://www.ncdd.nl/en/ncdd-projects/digital-heritage-network/project-persistent-identifiers/>
 - [6] Alliance for Permanent Access to the Records of Science Network <http://www.alliancepermanentaccess.org>
 - [7] APARSEN project, Deliverable 22.1 ‘Persistent Identifiers Interoperability Framework’ <http://www.alliancepermanentaccess.org/index.php/about-aparsen/aparsen-deliverables/>
 - [8] Research Data Alliance Recommendations in the field of Persistent Identifiers for digital objects <https://rd-alliance.org/group/data-fabric-ig/wiki/persistent-identifier-bundle.html>
 - [9] Persistent Identifier decision tree: <http://www.ncdd.nl/en/pid>
 - [10] Handle.Net <https://www.handle.net>
 - [11] The DOI® System <http://www.doi.org>
 - [12] ARK (Archival Resource Key) Identifiers <https://confluence.ucop.edu/display/Curation/ARK>
 - [13] Using National Bibliography Numbers as Uniform Resource Names, RFC 3188, J. Hakala <https://tools.ietf.org/html/rfc3188>
 - [14] Uniform Resource Names, <https://datatracker.ietf.org/wg/urnbis/charter/>
 - [15] Persistent URLs <http://www.purlz.org> and <https://purl.oclc.org/docs/index.html>
 - [16] Implementing Persistent Identifiers: Overview of concepts, guidelines and recommendations, Hans-Werner Hilse and Jochen Kothe, 2006 <http://nbn-resolving.de/urn:nbn:de:gbv:7-isbn-90-6984-508-3-8>
 - [17] Persistent identifiers in the Digital Preservation Handbook, Digital Preservation Coalition, <http://handbook.dpconline.org/technical-solutions-and-tools/persistent-identifiers>
 - [18] Persistent Identifiers (PIDs): recommendations for institutions, ATHENA WP3 Working Group, April 2011 <http://www.athenaeurope.org/getFile.php?id=779>
 - [19] Introduction to persistent identifiers, The THOR Project Knowledge Hub <https://project-thor.readme.io/docs/introduction-to-persistent-identifiers>