Persistent Identifiers for Digital Cultural Heritage

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

Remco van Veenendaal
Marcel Ras
National Archives
P.O. box 90520
2509 LM The Hague, Netherlands
+31629451951 / +31703140180
remco.van.veenendaal@nationaalarchief.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 and what the criteria are for trustworthy identifier systems; to present the latest version of a decision tree tool that helps determine which system is the most appropriate for a particular set of needs; and to discuss the current features of PIDs and the development of future properties.

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 make 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, a Netherlands Academy of Arts and Sciences, the National Diet Library of Japan and the European Union are moving forward with interesting projects to assign persistent identifiers to their archives.

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 three 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. Recommendations of the Research Data Alliance (RDA) for reliable & trustworthy PIs [6]
   d. Review of current identifier systems

2. 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) [7]
   b. Discuss the applicability of the decision tree outside the Netherlands

3. Persistent Identifiers: current features and future properties
   a. Resolution
   b. Future developments of PIDs

SCOPE
The intention is that this tutorial be “PI-agnostic”; that is, any persistent identifier (Handle [8], DOI [9], ARK [10], URN:NBN [11], URN [12] and PURL [13]) 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 [14] [15] [16] [17].

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 and Juha Hakala of the National Library of Finland.

6. REFERENCES
[10] ARK (Archival Resource Key) Identifiers https://confluence.ucop.edu/display/Curation/ARK
[17] Introduction to persistent identifiers, The THOR Project Knowledge Hub https://project-thor.readme.io/docs/introduction-to-persistent-identifiers