



biblio.ugent.be

The UGent Institutional Repository is the electronic archiving and dissemination platform for all UGent research publications. Ghent University has implemented a mandate stipulating that all academic publications of UGent researchers should be deposited and archived in this repository. Except for items where current copyright restrictions apply, these papers are available in Open Access.

This item is the archived peer-reviewed author-version of:

Title: Digital Long-Term Preservation, Provenance and Linked Open Data.

Authors: Sam Coppens, Erik Mannens, Raf Vandesande, Rik Van de Walle.

In: Proceedings of the 35th European Librarian Automation Group conference (ELAG 2011), May 2011, Prague, Czech Republic

Optional link to the article: <http://elag2011.techlib.cz/en/808-digital-long-term-preservation-provenance-and-linked-open-data/>

To refer to or to cite this work, please use the citation to the published version:

Title: Digital Long-Term Preservation, Provenance and Linked Open Data

Name of Author(s): Sam Coppens, Erik Mannens, Raf Vandesande, Rik Van de Walle

About the presenter:

Sam Coppens is a researcher and Ph.D. candidate in computer science and engineering at the Multimedia Lab of Ghent. He received his M.Sc. degree in Engineering (specialisation on micro-electronics) from K.U. Leuven, Belgium, in 2005. As a researcher he is involved in several projects with the emphasis on semantic web technologies and cultural heritage. His research interests and major expertise is centered around metadata/ontology modeling, semantic web technologies (a.o. linked open data), archiving and aggregation/harvesting (a.o. in the context of Europeana). He is member of the W3C's provenance incubator group, IASA metadata and documenting committee, SWPM programme committee and involved in several standardization activities, e.g., PREMIS.

Erik Mannens received his M.Sc. degree in Electro-Mechanical Engineering (1992) at KAHO Ghent and his M.Sc. degree in Computer Science (1995) at K.U.Leuven University. Before joining IBBT-MMLab in 2005 as project manager, he was a software engineering consultant and Java architect for over a decade. His major expertise is centered around metadata modeling, semantic Web technologies, broadcasting workflows, iDTV and Web development in general. He is involved in several projects as senior researcher and is finishing up his PhD on Semantic News Production. He is co-chair of W3C's Media Fragments Working Group and actively participating in other W3C's semantic Web standardization activities. On all of these subjects he has published several papers and book chapters. He is also an active member of the technical committees of ACM MultiMedia, IEEE CCNC, SAMT and MARESO. His full bio can be obtained from <http://www.mmlab.be/emannens>.

Intended audience: Libraries, Digital long-term archives, Linked Open Data

Abstract:

Nowadays, the need for digital long-term preservation is growing. A lot of the digital information produced merely a decade ago is in danger of getting lost as technologies are changing. This also threatens a lot of information from heritage institutions in Belgium. Analogue carriers are degrading and digitisation is a must, but digital copies are also subjected to many risks, e.g., data carrier corruption, or file formats becoming obsolete. For this, a distributed long-term preservation platform is being developed overcoming those risks, assuring the accessibility of the information. This distributed archive harvests the data records from the institutions, disseminates the data as Linked Open Data, and archives the data for the long-term by providing and tracking its provenance information. This provenance information is described using PREMIS OWL, our developed and (LOC) approved semantic implementation of the PREMIS 2.0 data dictionary. The developed platform also enables the provenance information on the Web. For this, the long-term preservation platform relies on datetime content negotiation, an extension to the HTTP protocol, proposed by Memento, and a special HTTP link header for discovery of the provenance information. The datetime content negotiation will deliver the right version of the data to the end-user, the HTTP provenance link header points to the provenance information of that version. This provenance information links the current version of the data to its other versions.