MESUR: usage-based metrics of scholarly impact

Johan Bollen
Digital Library Research & Prototyping Team
Los Alamos National Laboratory
Los Alamos, NM 87545
jbollen@lanl.gov

Marko A. Rodriguez
Digital Library Research & Prototyping Team
Los Alamos National Laboratory
Los Alamos, NM 87545
marko@lanl.gov

Herbert Van de Sompel
Digital Library Research & Prototyping Team
Los Alamos National Laboratory
Los Alamos, NM 87545
herbertv@lanl.gov

Categories and Subject Descriptors

I.2.4 [**Knowledge Representation Formalisms and Methods**]: Semantic Networks; H.2.8 [**Database Applications**]: Data mining; H.3.7 [**Digital Libraries**]: Collection

General Terms

Measurement, Performance, Human Factors

Keywords

Digital libraries, usage data, scholarly evaluation, impact factor, semantic networks

1. INTRODUCTION

The assessment of scholarly impact is now largely a matter of expert opinion or metrics derived from citation data, e.g. Thomson's Scientific ISI Impact Factor. Usage data has the potential to provide a more complete picture of scholarly impact as illustrated by recent results [1, 2]. However, usage-based metrics of scholarly impact have not yet made inroads as reliable and community-accepted means of assessing scholarly impact due to sampling and cross-validation problems associated with usage data.

The MESUR[1] project, a two-year effort funded by the Andrew W. Mellon Foundation and executed by the Digital Library Research and Prototyping Team at the Los Alamos National Laboratory (LANL) Research Library (RL), seeks to survey a range of usage-based impact metrics to determine their validity and properties. It will do so by creating a large-scale reference data set in the form of a semantic network that relates usage, citation and bibliographic data at a scale that is intended to be representative of the scholarly community.

The MESUR project will proceed according to the following project phases:

1. Defining a model of the scholarly communication process represented as an RDF/OWL ontology [3].
2. Creation of a large-scale reference data set, represented as a semantic network, according to the created ontology that combines usage, citation and bibliographic data obtained from major publishers, aggregators and institutions.
3. Characterization of the structure and properties of the created semantic network to deepen our knowledge of the significant demarcations and structural features of the scholarly communication process.
4. Surveying a wide range of usage-based metrics to determine their validity, reliability and scholarly correlates on the basis of the created semantic network.

The project has now passed its phase 1 and 2 milestones. In the near future it will perform a characterization of what is shaping up to be the largest semantic network ever created to represent the scholarly community, and survey a wide range of usage-based metrics of impact. The final project outcome will be the publication of guidelines with regards to the properties of various usage-based impact metrics, and how they can be appropriately applied. This poster provides an overview of the MESUR project's workplan and architecture, and will show preliminary results relating to the characterization of its semantic network and a range of usage-based impact metrics.

2. ACKNOWLEDGMENTS

This research is supported by a grant from the Andrew W. Mellon Foundation.

3. REFERENCES

[1] J. Bollen, H. Van de Sompel, J. Smith, and R. Luce. Toward alternative metrics of journal impact: a comparison of download and citation data. *Information Processing and Management*, 41(6):1419–1440, 2005.

[2] M. J. Kurtz, G. Eichhorn, A. Accomazzi, C. S. Grant, M. Demleitner, and S. S. Murray. The bibliometric properties of article readership information. *JASIST*, 56(2):111–128, 2004.

[3] M. A. Rodriguez, J. Bollen, and H. Van De Sompel. A Practical Ontology for the Large-Scale Modeling of Scholarly Artifacts and their Usage. In *Proceedings of the Joint Conference on Digital Libraries, Vancouver, Canada, 2007.*

[1]Pronounced "measure", an acronym for "Metrics from Scholarly Usage of Resources".

JCDL'07, June 18–23, 2007, Vancouver, British Columbia, Canada.
ACM 978-1-59593-644-8/07/0006.