LA-UR- 08-4725

Title: A ranking and exploration service based on large-scale usage data

Author(s): Johan Bollen
Herbert Van de Sompel
Lyudmilla Balakireva
Ryan Chute

Intended for: JCDL '08: Proceedings of the 8th ACM/IEEE-CS joint conference on Digital libraries

Los Alamos
NATIONAL LABORATORY
———— EST.1943 ————

A ranking and exploration service based on large-scale usage data.

Johan Bollen, Herbert Van de Sompel, Luda Balakireva and Ryan Chute
Digital Library Research & Prototyping Team,
Los Alamos National Laboratory
Los Alamos, NM 87545
jbollen@lanl.gov, herbertv@lanl.gov, lbalakireva@lanl.gov & rchute@lanl.gov

ABSTRACT

This poster presents the architecture and user interface of a prototype service that was designed to allow end-users to explore the structure of science and perform assessments of scholarly impact on the basis of large-scale usage data. The underlying usage data set was constructed by the MESUR project which collected 1 billion usage events from a wide range of publishers, aggregators and institutional consortia.

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

Digital libraries, usage data, impact, scholarly evaluation, architecture, standards

Poster content

Numerous interactive services now offer the ability to assess the impact of journals on the basis of citation statistics. Thomson Scientific's Web of Knowledge allows users to retrieve the citation counts, impact factor values (average citation rate per article), citation decay and various other key characteristics of nearly 8,000 selected journals. Several new services[1] have recently been introduced that offer other impact indicators, e.g. citation PageRank [1] and allow users to explore various citation-based journal relationships.

The MESUR[2] project is examining how large-scale usage data can augment citation data, and yield reliable and valid maps of science as well as novel indicators of impact. MESUR has implemented an interactive service that combines two

---

[1]http://www.eigenfactor.org/, http://www.scimagojr.com/

[2]http://www.mesur.org/

usage-based functionalities; it visualizes the major structural components of science in terms of an interactive usage network graph [2] and allows users to compare the usage-derived rankings of selected journals to those in their network neighborhood. Selecting any journal in the usage graph results in an interactive scatterplot that displays the impact of a selected journal in relation to those that are most similar. The axii of the scatterplot can be any combination of 2 metrics out of a set of more than 50 MESUR usage- and citation-based metrics. Any journal in the scatterplot links back to the usage network graph. The result is a compelling interactive environment that enables users to explore the relationships and rankings of journals that have been determined on the basis of their usage.
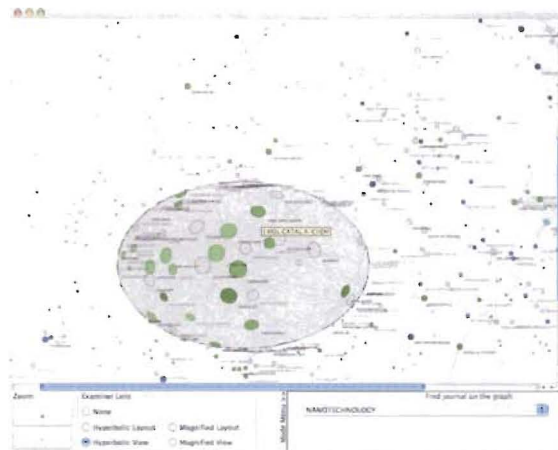

Figure 1: Prototype of MESUR's usage data explorer.

Acknowledgments

1. REFERENCES

[1] J. Bollen, M. A. Rodriguez, and H. Van de Sompel. Journal status. *Scientometrics*, 69:669–687, 2006.

[2] 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.