

IOTA Working Group Summary of Activities and Outcomes

*A Technical Report of the
National Information Standards Organization*

Approved: April 26, 2013

About NISO Technical Reports

NISO Technical Reports describe the work and conclusions of a working group chartered by the organization to study a particular problem or issue that is related to NISO's standardization interests. However, unlike NISO standards or recommended practices, they do not have requirements for compliance or recommend "best practices"—although their conclusions may contain some recommendations. Additionally, no consensus process of the NISO voting membership was used in their creation or approval. They are strictly informative in nature.

Published by

National Information Standards Organization (NISO)
3600 Clipper Mill Road, Suite 302
Baltimore, MD 21211
www.niso.org
301-652-2512

Copyright © 2013 by the National Information Standards Organization

All rights reserved under International and Pan-American Copyright Conventions. For noncommercial purposes only, this publication may be reproduced or transmitted in any form or by any means without prior permission in writing from the publisher, provided it is reproduced accurately, the source of the material is identified, and the NISO copyright status is acknowledged. All inquiries regarding translations into other languages or commercial reproduction or distribution should be addressed to:

NISO, 3600 Clipper Mill Road, Suite 302, Baltimore, MD 21211.

ISBN: 978-1-937522-17-9

Contents

1	Introduction	1
1.1	Purpose and Scope	1
1.2	Terms and Definitions	1
2	IOTA Use Cases	4
2.1	Librarian Use Case	4
2.2	Content Provider Use Case.....	5
2.3	Link Resolver Use Case.....	5
3	IOTA Project Approach	5
4	The IOTA reporting system	6
4.1	Components	6
4.2	Architecture	7
4.2.1	Ingest.....	7
4.2.2	Database Structure	7
4.2.3	Reporting	8
5	Completeness Index	10
5.1	Description	10
5.2	Method 0: Descriptive Reports (Source, Metrics).....	10
5.3	Method 1: Qualitative Judgment.....	10
5.4	Method 2: Correlation Method.....	11
5.5	Method 3: Stepwise Regression.....	12
6	Recommendations and Next Steps.....	13
	Appendix A : OpenURL Elements in IOTA Repository	15
	Bibliography	21

Foreword

About this Technical Report

In November 2009, Adam Chandler submitted a proposal ([Chandler, 2009a](#)) to NISO's Business Information Topic Committee to investigate the feasibility of creating industry-wide, transparent and scalable metrics for evaluating and comparing the quality of OpenURL implementations across content providers. The proposed project was to build on work developed under a 2008/2009 Mellon Foundation grant, for which Chandler was the lead investigator. An existing reporting system and database were to be ported to a NISO-supported server and the new working group would collect further data, analyze the data, and develop metrics to evaluate OpenURL quality.

The proposal was approved by the Business Information Topic Committee in December 2009 and the Working Group that was formed was presented with the following problem statement:

The OpenURL standard is a widely deployed technology to facilitate linking to resources across the library supply chain. The OpenURL-formatted URL carries the data about an item to the link resolver of the library. The resolver compares the metadata embedded within the OpenURL with what is held in the library's collection and presents the available options in a results page. For a book, there is usually a link to the library's catalog card; for an article, ideally this is a link directly to the full-text of the article. At a typical academic library thousands of OpenURL requests are initiated by patrons each week. The problem is too often these links do not work as expected, leaving patrons frustrated by a lower than desired quality of service. Periodically mention is made in the library literature to problems with OpenURL linking, but since the OpenURL standard was introduced a decade ago, no systematic method has been designed and carried out to benchmark it. This work is intended to fill the gap.

The study was to be conducted over a two-year period, concluding with a report reviewing the results and making recommendations about how or whether to continue the IOTA research. This technical report is the project's promised deliverable.

NISO Discovery to Delivery Topic Committee Members

The Discovery to Delivery Topic Committee had the following members at the time it approved this Technical Report.

Tim Babbitt

Cambridge Information Group

Pascal Calarco, Co-chair

University of Waterloo Library

Peter Murray

Lyrasis

Jeff Penka

OCLC Online Computer Library Center

Tim Shearer

University of North Carolina Chapel Hill Libraries

Lucy Harrison, Co-chair

Florida Virtual Campus

Chris Shillum

Reed Elsevier

NISO IOTA Working Group Members

This technical report was developed by the following members of the Improving OpenURL Through Analytics (IOTA) Working Group:

Adam Chandler (Chair)
Cornell University Library

Clara Ruttenberg
University of Maryland

Rafal Kasprowski
Rice University

Elizabeth Winter
Georgia Tech Library

Susan Marcin
Columbia University

Jim Wismer
Thomson Reuters

Oliver Pesch
EBSCO Information Services

Aron Wolf
Serials Solutions

Acknowledgements

The IOTA Working Group would like to acknowledge the important contributions made by Ellen Rotenberg during the first year of the initiative. We would also like to acknowledge the contribution of Phil Davis, who provided crucial advice during our experiments with statistical approaches to measuring OpenURL metadata quality.

Trademarks, Services Marks

Wherever used in this technical report, all terms that are trademarks or service marks are and remain the property of their respective owners.

1 Introduction

1.1 Purpose and Scope

The purpose of the IOTA (Improving OpenURL Through Analytics) project was to analyze the quality of the metadata that is passed to the link resolver from the OpenURL source and develop a standardized metric for such quality evaluation. The project focused most of its attention on a specific genre of OpenURLs—those intended to provide access to journal articles.

Essentially there are three parts to OpenURL linking: the metadata in the incoming OpenURL, matching metadata in the OpenURL to the library's knowledge base, and constructing accurate links into content provider websites.

The quality of the data in the link resolver knowledge base itself was outside the scope of IOTA. Also outside of the scope was the quality of the linkages between link resolvers and full-text content providers.

A related NISO initiative, [Knowledge Base And Related Tools](#) (KBART), is concerned with accurate knowledge base content in the link resolver. Standardizing the link syntax between the link resolver and content provider remains unaddressed by any industry effort. Chandler's 2011 *Against the Grain* article describes the relationship between these initiatives in more detail.

1.2 Terms and Definitions

The following terms, as used in this technical report, have the meanings indicated. Terms in boldface in a definition indicate the term is also defined in this section.

<u>Term</u>	<u>Definition</u>
A&I database Abstract and Index database	A content discovery product that provides descriptive information about content items but not the full text of the content. A&I databases are frequently a source in OpenURL linking.
completeness	The number of metadata elements provided in the OpenURL out of a desired or core number.
Completeness Index	A number that is attributed to a content provider (OpenURL referrer or source) to measure the completeness of the provider's OpenURLs in aggregate. It is essentially an average of the Completeness Scores of OpenURLs coming from that content provider.
Completeness Score	The measure of the completeness of a single OpenURL .

Term

Definition

DOI[®]

Digital Object Identifier

A unique and persistent identifier for a digital content object, such as an online article, journal, chapter, book, or image, using a syntax defined in ANSI/NISO Z39.84 and ISO 26324. The DOI provides current information about the object, including where to find it on the internet. Typically registered with an authorized registration agency, e.g., CrossRef (www.crossref.org), whereby it is possible to use the DOI to look up descriptive metadata about a content item using a provided free service. Sometimes called a DOI name.

discovery platform

A website or product through which users can discover content items.

element

Descriptive metadata such as ISSN, Volume, Issue, Start Page, etc. used in an **OpenURL**.

Element Weight

A value assigned to an **element** and used in calculation of a **Completeness Score**. The value of the Element Weight represents the relative importance of the **element** to the success of the **OpenURL** in providing the user access to full text; the higher the value the more important the **element**.

Enhanced OpenURL

The set of **OpenURL elements** available after the **link resolver enhancers** have been run on the incoming **OpenURL**. Typically an Enhanced OpenURL will have more data **elements** than the original **OpenURL**.

enhancer

An automated process that enhances an **OpenURL** by using internal and external sources to supplement the **OpenURL** data **elements**. An example would be an enhancer that looks up article-level metadata from CrossRef using the **DOI** presented on an **OpenURL**.

fail

Describes the state of an **OpenURL** that does not generate any item-level links to full-text items. *See also success.*

NOTE: Within the context of this technical report the following are not considered links to full-text items: links to a journal homepage, the table of contents for an issue, or a search page at the vendor site.

full-text target link

A link to the complete text—including all references, figures, and tables—of an article on the **target** site. If the referenced item is a journal article, the full-text article link would direct the user directly to that article at the content provider's site without requiring further navigation or searching.

Term**Definition**

knowledge base

A database used with **OpenURLs** that contains information about what **targets** are available to users of the **link resolver**. Data within the knowledge base includes but may not be limited to link syntaxes and holdings with coverage details. The **link resolver** environment for a given institution can be customized to reflect that institution's collection and to only provide links to **targets** to which the library subscribes.

link resolver

Technology that controls the linking between **sources** and **targets**. The link resolver accepts and deconstructs the **OpenURL** describing a content item from a **source** and uses its **knowledge base** and associated programs to determine full text and other **targets** appropriate for the user and create predictable links to these. The role of the link resolver is context sensitive linking to the appropriate copy of a content item.

OpenURL

As defined in ANSI/NISO Z39.88-2004, *The OpenURL Framework for Context-Sensitive Services*, a URL designed to transport metadata and thus enable linking from information resources such as **abstracting and indexing databases (sources)** to library services (**targets**), such as academic journals, whether online or in printed or other formats. The linking is mediated by **link resolvers**, or link-servers, which parse the elements of an **OpenURL** and provide links to appropriate **targets** available through a library by the use of a **knowledge base**. *Source: Wikipedia [boldface notation added]*

PubMed ID

A unique identifier assigned to a record in the PubMed databases (www.pubmed.org). A PubMed record typically describes an article or other content item. Using a PubMed ID it is possible to look up descriptive metadata of the content item through a free online service provided by the National Library of Medicine.

referrer

The identity of a website or **discovery platform** that is the **source** of the **OpenURL**. In an **OpenURL** this is represented as the SID (Source ID) or the Referrer ID. In some contexts, **referrer** is used interchangeably with **source**.

required data elements

The data **elements** that must be available to create a successful link to an **OpenURL target**. For example, if a publisher's link to a full-text article is constructed from ISSN, Volume, Issue, and Start Page values, then these fields are considered required data elements since the link will fail to access the full text if any one of them is missing.