Open Discovery Initiative: Promoting Transparency in Discovery

A Recommended Practice of the
National Information Standards Organization

About NISO Recommended Practices

A NISO Recommended Practice is a recommended "best practice" or "guideline" for methods, materials, or practices in order to give guidance to the user. Such documents usually represent a leading edge, exceptional model, or proven industry practice. All elements of Recommended Practices are discretionary and may be used as stated or modified by the user to meet specific needs.

This recommended practice may be revised or withdrawn at any time. For current information on the status of this publication contact the NISO office or visit the NISO website (www.niso.org).

Contents

Foreword

About This Recommended Practice

The ODI was initiated in June 2011 by Oren Beit-Arie of Ex Libris; Jenny Walker, independent consultant working on behalf of Ex Libris; and Marshall Breeding, a library industry consultant (previously at Vanderbilt University) who, at the 2011 ALA Annual Conference in New Orleans, invited senior industry managers to meet and explore areas of mutual interest related to library discovery services. Following this initial meeting, a proposal was submitted to NISO; later in the year, the NISO Discovery to Delivery Topic Committee accepted the proposal as a new NISO work item. The original 2014 Recommended Practice was the outcome of that project.

The NISO ODI Working Group started its work in early 2012 to define best practices for the new generation of library discovery services. These services use an aggregated central index to enable searching across a wide range of library related resources—both licensed and open access—from multiple providers. They also offer more sophisticated capabilities and faster performance than those provided by systems relying on federated search technologies. Several major discovery products, based on centrally indexed search, have been released to the market since early 2009 and have been widely deployed in libraries globally.[1]

Libraries have made substantial investments in these index-based discovery services, which have become established as one of the main channels through which users discover and access content. The important role of these discovery services demands a set of industry practices in several areas to ensure their optimal performance. To facilitate improved communications and clarity, the ODI established the following goals:

- Create ways for libraries to assess the level of content providers' participation in discovery services.

- Help streamline the process by which content providers work with discovery service vendors.

- Define models for fair or unbiased linking from discovery services to publishers' content.

- Determine what usage statistics should be collected.

Libraries expect their entire collection, including licensed and purchased electronic content as well as open access content, to be made available through the discovery service of their choice. When acquiring licensed content, libraries expect a clear explanation of the degree of availability of that content in that discovery service.

Based on the input from a survey done early in the project (see Section 2.7), the ODI group agreed to develop recommended practices in the following areas:

1. Technical recommendations for data format and data transfer, including method of delivery and ongoing updates.

2. Recommendations for the communication (automated or through reporting) of libraries' rights to distribute or display specific content (e.g., restricted to subscribers versus open to all users). These recommendations are to include technical specifications on how data will be exchanged and procedural specifications regarding update frequency and other logistical details.

3. Clear descriptors regarding the level of indexing performed for each item or collection of content and the level of availability of the content.

---

[1] Marshall Breeding. "Automation Marketplace 2013: The Rush to Innovate." *The Digital Shift*, April 2, 2013. Available at: www.thedigitalshift.com/2013/04/ils/automation-marketplace-2013-the-rush-to-innovate/

4. Definition of fair linking from discovery service to the published content.

5. Determination of what usage statistics should be collected, for whom, and how this data should be disseminated.

Further, the Working Group agreed to develop mechanisms to evaluate conformance with the Recommended Practice.

To work towards the development of recommended practices to address the above issues, five subgroups were formed as follows:

1. Technical Formats

2. Communication of Library's Rights

3. Level of Indexing

4. Fair Linking

5. Usage Stats

Early in the process, Subgroups 2 and 3 merged.

A further goal of the ODI was to develop mechanisms to evaluate conformance with the Recommended Practice. To help libraries better understand the position of content providers and discovery service providers, these organizations can take measures to demonstrate the extent to which they conform with the recommended practices issued by the Open Discovery Initiative. Conformance statements will be voluntarily issued by content providers and by discovery service providers, respectively. In the absence of voluntary statements, libraries can use the presence or absence of these factors to infer conformance.

About the 2020 Revision

NISO constituted a new Open Discovery Initiative Standing Committee following the approval of the ODI Recommended Practice. This standing committee has worked to facilitate the adoption of the principals of the Recommended Practice and to promote the adoption of conformance statements from discovery service providers and content providers. The committee has extended the work of the ODI Working Group by conducting additional surveys addressed to Content Providers, Discovery Service Providers, and to libraries to gather more extensive and recent data regarding the content discovery environment and to identify interest in enhancements to the Recommended Practice. Informed by these survey responses, the ODI Standing Committee has developed a revision to the Recommended Practice.

In 2018, the ODI Standing Committee members prioritized seven high-priority areas to review as part of the ODI RP revision.

- Library Responsibilities in ODI

- Handling of Open Access Content, Including Hybrid OA Content

- More Meaningful Usage Statistics for Content Providers

- Fair Linking

- Identifying the Source of the Record in the Discovery Interface

- Content Coverage Disclosure (Reporting on Discovery Service Content at a Collection Level)

- Identification of Additional Metadata and Content Elements

The work item proposal was approved in January 2019. The Standing Committee engaged in information-gathering activities in the spring and summer of 2019, including surveying stakeholder groups

(summaries available in Appendices D–G). Input was analyzed in fall 2019, and the draft Recommended Practice was submitted for public comment in January of 2020.

Significant updates in the revised Recommended Practice include updates to the metadata elements recommended for content providers to deliver to discovery services (Section 3.2.1), fair linking (3.2.5.1 and 3.3.2.1), discovery service content listings (3.3.1), updated statistical reporting recommendations (3.3.4), treatment of open access content (3.3.5), authentication (3.3.6), alternative coverage lists (3.3.7), record display (3.3.8), ranking algorithm disclosure (3.3.9), and use of content provider metadata (3.3.10). A new section, Best Practices for Libraries, related to system maintenance, library advocacy, and training and communication has been added (3.4). The revised version also addresses areas that the original ODI Working Group placed out of its scope, such as more detailed treatment of Abstracting and Indexing (A&I) content products and Fair Linking. In addition, all sections of the original document were reviewed and updated for currency and alignment with the revised/new recommended practices.

NISO Information Discovery & Interchange Topic Committee Members

The following individuals served on the NISO Information Discovery & Interchange Topic Committee that had oversight for this work and approved this Recommended Practice:

Scott Bernier
EBSCO Information Services

Robert Boissy
Springer Nature

Mark Dehmlow
University of Notre Dame Libraries

John Dove
Paloma & Associates

Peter Murray *(Co-chair)*
Index Data

Doralyn Rossmann
Montana State University

Bob Schulz
OCLC

Christine Stohn *(Co-chair)*
Ex Libris

Jan Waterhouse
University at Albany Libraries

Julie Zhu
IEEE

ODI Working Group Members

The following individuals served on the NISO Open Discovery Initiative Working Group that developed and approved this Recommended Practice:

Marshall Breeding *(Co-Chair)*
Independent Consultant (formerly with Vanderbilt University)

Jenny Walker *(Co-Chair)*
Independent Consultant (formerly affiliated with Ex Libris)

Anya Arnold (through July 2012)
Orbis Cascade Alliance

Scott Bernier (from December 2013)
EBSCO Information Services

Jamene Brooks-Kieffer
University of Kansas Libraries (formerly with Kansas State University Libraries)

Sara Brownmiller
University of Oregon Libraries

Lettie Conrad
SAGE Publications

Michael Gorrell (through November 2013)
EBSCO Information Services

Jeff Lang
Thomson Reuters

John Law
ProQuest

David Lindahl
University of Missouri-Kansas City Libraries
(formerly with eXtensible Catalog Project)

Laura Morse
Harvard University

Michele Newberry
Unaffiliated (formerly with University of Florida)

Peter Noerr (through June 2012)
MuseGlobal

Roger Schonfeld
Ithaka S+R

Ken Varnum
University of Michigan Library

Aaron Wood
Alexander Street Press

Open Discovery Initiative Standing Committee members

The following individuals served on the NISO ODI Standing Committee that developed the revision to the ODI Recommended Practice:

Rachel Kessler *(Co-Chair)*
ProQuest

Laura Morse *(Co-Chair)*
Harvard University

Scott Bernier
EBSCO Information Services

Marshall Breeding
Independent Consultant

Noah Brubaker
Private Academic Library Network of Indiana
(PALNI)

Teresa Hazen
University of Arizona Libraries

Jay Holloway
OCLC

Geoffrey Morse
Northwestern University Libraries

Pappaparvathi (Bobbi) Patham
Springer Nature

Don Sechler
Clarivate Analytics

Maria Stanton
American Theological Library Association

Ken Varnum
University of Michigan Library

Julie Zhu
IEEE

Acknowledgments

Trademarks and Service Marks

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

<div style="background:#1F4E79;color:white;padding:8px;text-align:center;">

Section 1: Introduction

</div>

1.1 Purpose and Scope

1.1.1 Purpose

The Open Discovery Initiative (ODI) aims to facilitate progress through exploration of relevant issues and the development of recommended practices for the current generation of library discovery services based on centrally indexed search. The domain of index-based discovery services involves a complex ecosystem of interrelating issues and interests among content providers, libraries, and discovery service creators.

This model of discovery relies on an index populated with metadata, full-text, or other representations of the content items—such as journal articles, book chapters, e-books, research reports, reference sources, images, maps, datasets, audiovisual materials, and other selected material—that a library provides to its users. The content comes from a range of information providers and products, such as commercial and nonprofit publishers, universities and other research institutions, and many other types of organizations. The content of interest to ODI includes any materials that libraries would consider within their collection, regardless of the business model for acquisition or the type of license, such as commercially restricted or open access.

Several major discovery products have been released to the market since early 2009 that are based on the model of centrally indexed search—largely influenced by the Google search model and users' expectations for a single, unified discovery experience. An increasing number of libraries, especially those that serve academic or research institutions, have invested in index-based discovery services. These products serve as one of the interfaces through which the library's patrons gain access to the rapidly growing breadth of information that may be available to them. These discovery services play an increasingly strategic role in the way that libraries provide users with access to their collection, they represent a growing segment of the library technology industry, and they may become a factor in how libraries select content products. These factors draw attention to the discovery services arena for any improvements that might be gained through this Recommended Practice.

To work effectively, discovery services need to be as comprehensive as possible in their content coverage. Libraries expect their uniquely licensed and purchased electronic content to be indexed within their discovery service of choice. Further, they require comprehensive and clear representation of each category of content in the discovery service. Content items not represented in a discovery service present a challenge to libraries in how they might otherwise ensure that these materials are discovered and accessed. Libraries have an interest in knowing whether any content providers are excluded or underrepresented in any given discovery service.

The Open Discovery Initiative aims to facilitate increased transparency in the content coverage of index-based discovery services and to recommend consistent methods of content exchange or other mechanisms. Full transparency will enable libraries to objectively evaluate discovery services and to deal with daily operational issues surrounding these products.

Discovery services depend on the cooperation of content providers with discovery service creators to provide access to metadata or full text of information resources in order to create effective indexes. The inclusion of data in the indexes of the current slate of discovery services is based on private agreements and ad hoc exchange methodologies between information providers and discovery service creators. Index-based discovery can potentially benefit content providers through enhanced exposure of their materials. It also presents some concerns, such as enabling library patrons to bypass the specialized interfaces created by content providers, potentially reducing or eliminating branding and losing control in how content is presented to the end user. And, as libraries' uptake of these services increases, the usage (and perceived

value) of publisher products can be greatly influenced by how successfully discovery services drive readers to content providers' assets.

ODI investigated the need for standard protocols to make the transfer of data from content providers to discovery service creators. Consistent practices in the exchange and formats of data aim to lower the level of complexity as content flows through this ecosystem, mitigating technical issues that might hinder broader participation by content providers or potential discovery service creators.

Libraries need a clear understanding of the degree of exposure for the content that they have acquired as represented in a discovery service. This understanding is essential as libraries evaluate and select a discovery service, and on an ongoing basis once it is implemented. Libraries require specific information on exactly which articles, databases, and other sources are represented; whether they are indexed in full text, by citations only, or both; and whether the metadata derives from aggregated databases or abstract and indexing (A&I) resources.

In the operation of an index-based discovery service, many different factors contribute to how it presents and orders results and how it connects users to content resources. For any given item of content, multiple metadata elements contributed from different content providers may be indexed by the discovery service. For a journal article, for example, its full text might be contributed by the primary publisher, citation data from the provider of an aggregated database, and abstracts or controlled vocabulary terms may be provided by yet another provider. Content providers are motivated to contribute to discovery services in order to gain more access from the patrons associated with the libraries that implement the discovery services. It is therefore important to each type of content provider that its contributions are appropriately recognized. If a record contributed by an A&I service, for example, leads to the selection of a full-text resource from another provider, how does the A&I service gain benefit from the discovery transaction? A subgroup of ODI on Fair Linking was established to explore and make recommendations on these issues.

The Open Discovery Initiative recognized and aimed to address perceptions regarding bias and concerns about the possibility of bias in discovery services. Special concerns surround the possibility of bias when discovery services are owned by the same corporate parent as content products or services. Concerns also arise through exclusive arrangements or other business relationships made by a discovery service with a content provider that might introduce bias. Some of these recommended practices were developed with the intent of helping discovery service providers mitigate concerns that exist in the community about conflicts of interest and other relationships that create bias. By explaining the nature of their business connections with related content providers and third parties alike, and affirming the neutrality of their discovery offerings, these services will be positioned to reassure both libraries and content providers about the nature of their practices.

1.1.2 Scope

In broad terms, ODI focuses primarily on the issues related to the composition of the central index associated with these discovery services and not with the design of user interfaces. The initiative does not seek to intrude into areas of proprietary innovation that distinguish each of the discovery service products.

The arena of index-based discovery spans many different issues, some of which lend themselves to a more open and standard treatment, while others remain in the realm of product development. The Open Discovery Initiative recognizes that even among the issues that might potentially benefit from its attention, some rank at a higher priority and others may need to be addressed through possible follow-up activities.

In Scope

A primary area of interest for ODI involves the arena of content coverage of the discovery services represented in their central index. The content used to create the indexes of a discovery service comes from a range of information providers and content products, such as commercial and nonprofit publishers,

universities and other research institutions, and many other types of organizations. The content of interest to ODI includes any materials that libraries consider within their collection, regardless of the business model for acquisition or the type of license, such as commercially restricted or open access.

This initiative aims to address the following questions in the realm of content coverage:

- The quantity of content that a provider makes available to discovery service providers relative to its total offerings

- The form of that content, such as whether it consists of citation-level metadata or if it also includes full-text representations

- Whether the discovery services operate in a way that results presented to the user do not favor or disfavor items from any given content source or material type

- The specific metadata fields provided within metadata records

- The specific metadata fields indexed in the discovery services

- Whether any controlled vocabularies or ontologies are included

- How A&I products relate to discovery services

- How branding of content providers is presented in a discovery service

This initiative aims to address the following technical issues:

- The transfer mechanisms or protocols by which data are delivered from content providers to discovery service creators

- The format in which the records are delivered by content providers to discovery service creators

One area of focus deals with the definitions of the metadata delivered by content providers to discovery services, as well as the data made available to licensed customers. A perceived lack of transparency across these data flows prompted the need to develop definitions of the data points and to clarify what metadata or data elements are made available to which parties and under which conditions. For example, a content provider might allow certain metadata elements to be included in the search index for retrieval purposes, but not allow those elements to be displayed in the final user interface. Conversely, libraries might not understand what elements from which full-text or A&I services are available and in which circumstances. Some elements might be displayed to authenticated users, and some not, but definitions of these distinctions are sometimes vague, if they are described at all.

Another topic concerns factors related to whether or not a discovery service functions with a bias towards certain databases or content products based on business relationships rather than user needs or library preferences. It was deemed important to propose practices that facilitate the presentation of unbiased links to a user following the execution of a search through a discovery service and in support of this objective to ensure transparency about discovery service practices.

Out of Scope

This initiative does not address issues related to performance or features of the discovery services, as these are inherently business and design decisions guided by competitive market forces.

Aspects of index-based discovery not deemed within the scope of ODI include:

- **User interface issues** – The user interface components of a discovery service may depend on the content indexes, but are out of scope for this ODI initiative.

- **Relevancy ranking** – The specific methods that a discovery service employs to order items in a search result set may fall within the realm of proprietary technologies used competitively to differentiate commercial offerings.

  Further, the demands of relevancy ranking—necessary for satisfying user expectations in keeping pace with open web search applications—require continuous enhancement of supporting technologies and algorithms. As such it would be both impractical and an impediment to require that service providers continuously publish the highly dynamic and substantially detailed workings of their search relevancy algorithms. Therefore, the ODI Working Group concludes that, while disclosure of the broad aspects of a relevance ranking algorithm is encouraged, there should be no expectation that a discovery service provider would explain in any level of detail the ranking algorithms it applies. (As noted above, however, whether rankings result in the bias of results presented to the end user is within scope.)

- **APIs exposed by discovery services** – Several discovery service providers offer access to their discovery index through an application program interface (API). This is a set of protocols that allows a computer program to query the index and receive search results. Some libraries build their own interfaces to the search index; others use the API to embed data retrieved from the search index into other applications, like pull lists of citations into course sites or other web pages.

  The ODI Working Group determined that APIs were out of scope for this initial foundational stage of the initiative. This decision was made somewhat reluctantly based on the growing need for libraries to have access to data via API. Because the provision of an API to customers is largely a business decision by the vendor, it was felt the existence of an API could not be suggested as a best practice. Additionally, the desire to recommend standardization of API results would have added significant complexity to the workgroup's recommendations. Thus, the current committee concluded that best practices concerning discovery service APIs should be deferred until a later round of work. See Section 4.2 for recommended next steps.

- **Protocols for data exchange** – In the arena of the technical mechanisms involved in the transfer of data between content providers and discovery services, the ODI concludes that the existing protocols and methodologies previously defined provide adequate options and that it is not necessary to create a new protocol specifically for use within the discovery services ecosystem. There has been much work and standards development in the area of file formats, schemas, naming practices, transport mechanisms, etc.

1.2 Terms and Definitions

In order to ensure consistent use of terminology, one of the early tasks of ODI was to create a glossary of definitions relevant to the field of investigation. The terms as used in this recommended practice have the meanings indicated below.

Term	Definition
Search Models and Related Concepts	
central index	The result of storing and indexing content in a central location. Disparate content sources are aggregated with consistent formatting, indexing, and ranking algorithms. Also referred to as *centrally indexed search engine*.

Term	Definition
federated search	Method for searching multiple disparate content sources with one query. Federation of searching involves restructuring queries to appropriate forms and sending them to multiple remote search engines. Results are coordinated and displayed to the user.
metasearch	Method for searching multiple disparate content sources with one query. Metasearch is sometimes synonymous with federated search; however, metasearch can also describe searching across disparate content sources stored at a single location.
full-text search	Method for searching the text of the work rather than metadata about the work.

Data Definitions

full text	The complete text of the work. The text may be in display format or be a simple character representation of the complete text of the work.
index only	Limitation of rights of the search service or aggregator to only allow indexing of portions of the content with limited or no rights to display content to users. Most often applied to full text requiring the user to be directed to another source for display capabilities.
snippet	A small portion of the text of the work, often including text immediately around terms matching the query. Used for display of search results to give the end user context of the retrieved result.
core metadata	Descriptive metadata and textual descriptions extracted directly from the work including abstracts, tables of contents, and full text.

Methods of Data Exchange

harvest	Method of extracting indexing and/or full text from remote web-accessible sites for the purpose of providing search and/or display from a different location.
syndication	Method of pushing content to remote indexing, abstracting, or display services.
RSS (really simple syndication)	Standard method for advertising the availability of frequently published content that includes metadata, publication date, and authorship information.
AtomPub (Atom publishing protocol)	Standard method for publishing syndication feeds similar to RSS. AtomPub provides a more robust method than RSS for publishing beyond blog and webpage entries.
screen scraping	Simplest method of harvesting content that places no technical burden on the publishing site. The human-readable text is extracted and indexed, formatted, and searched by various engines.
linked data	Method for publishing structured data intended for further processing by machines rather than directly by human readers.
API access (application programming interface access)	Generic description of a method for exposing services and data through an advertised programming interface. Specification of the programming interface through either a proprietary or published standard is required.

Term	Definition
Actors	
end user	The final consumer in an information retrieval session.
licensee	The institution or individual who has acquired rights to access others' content or services.
licensor	The rights holder granting search and/or access rights to others. Also referred to as *rights owner*.
content provider	The organization providing dissemination of the content (literature or information). May be a publisher, aggregator, open access or institutional repository, or A&I service provider. The same content may be available from multiple content providers.
publisher	A person or organization whose predominant activity is to commission, create, collect, validate, and edit information in printed and/or in electronic form. A publisher may also act as a content provider for its material.
aggregator	The organization that collects information from varied sources and provides consistent search, presentation, and/or access.
open access repository/ institutional repository	The entity that collects and disseminates content created during the research process at academic institutions.
A&I service provider	The organization that provides abstracting and indexing services, including controlled vocabularies, for published content.
License Terms	
authenticated access	Access to search, metadata display, or content display via a known method for verifying end user identity or institution affiliation.
subscription	Provision of access to search, metadata, or content through specified license terms for a period of time, usually annually.
mutual subscriber	A collaboration between two service organizations agreeing to provide equivalent access terms to a single individual or institution.
market product	A specific, defined collection of resources made available for license or use by a content provider to a library. This could be a citation database, a set of e-journals or e-books, or other constellation of titles or objects commonly licensed as a single entity.

1.3 Stakeholders

The Open Discovery Initiative worked to include participation and input from each of the categories of organizations involved in discovery services. These organizations were considered in three stakeholder groups:

- **Content Providers (CPs)** – These organizations offer content products or services primarily intended for access by library patrons or the general public. The content provided by these organizations is used to generate the central indexes associated with the discovery services. Content providers include commercial and nonprofit organizations. Many different access and

license models may apply, including those restricted to individuals affiliated with subscribing organizations and those based on open access licenses.

- **Discovery Service Providers (DSPs)** – These organizations create index-based discovery services intended to enable end users to search the broad universe of content made available through their library.

- **Libraries** – These organizations—which may be affiliated with universities, research institutes, or commercial firms—acquire content from a variety of content providers and may also implement an index-based discovery service. Libraries represent particular user communities including staff, students, researchers, etc.

Section 2: The Evolution of Discovery and Delivery

Libraries have relied on different tools and technologies in recent decades to provide their users the ability to search for materials within their collection and gain access to items of interest. The successful discovery and delivery of resources has been a key library mission from the time when collections were primarily composed of print materials until the current time, when library collections include massive amounts of electronic content. A succession of products has addressed these tasks in response to evolving demands for discovery and retrieval of library materials.

2.1 Catalogs and Indexes

During the time when library collections were composed primarily of print materials, the online catalog module of the integrated library system was the primary tool for patrons to search for books and journals. A variety of indexes were published in print to find journal articles. The online catalog provided search and request features for the content managed by the integrated library system (ILS), which was broadly the content the library owns and holds locally, but did not include materials managed through other platforms. It also does not provide the most granular view of resources. Journal titles, and the volumes and issues received by the library, were included, but not the articles contained within each issue. Libraries acquired A&I services, individual e-journal subscriptions, and aggregated databases of articles, each with their own search interface, to provide access at the article level. Working through each of the possible electronic resources that might have material for a given research topic was often unwieldy for library patrons, which led to the creation of utilities able to search across these resources simultaneously.

2.2 Abstracting and Indexing Resources

A&I services form an important component of the development of resource discovery tools and the current index-based discovery ecosystem. In their original form, A&I services produced printed indexes that allowed researchers to find articles according to topics of interest. These A&I products predate online technologies. H.W. Wilson's *Reader's Guide to Periodical Literature*, for example, began publication in the early 1900s.

The production of A&I services depends on manual efforts, usually performed by individuals with specific knowledge in a given domain, to compose summaries or abstracts of each article and to assign appropriate subject terms. These cover general categories of content, but many address narrow scholarly or professional disciplines, with specialized thesauri, indexing techniques, or other value-added methods. The production of these resources comes at significant business cost, but is a service that researchers value.

A&I products today are delivered primarily though web-based platforms, some offered directly by the organization that produces the resource and others by organizations that license and aggregate multiple resources into products that span broader areas of subject coverage. Aggregators may also license from the primary publishers the full text of the articles covered by the A&I products, allowing them to provide a comprehensive service that includes both search and full-text delivery within their interface. These products may also use proprietary or OpenURL linking mechanisms to link from citations to the full text of articles not embedded within their products. An A&I product functions much like a discovery service, though operating within a narrow scope of content.

Libraries function as the primary distributors of these A&I products, subscribing to the resources that best match the requirements of their researchers. Libraries also subscribe to the e-journals referenced by the A&I products, either directly from the primary publishers or as part of aggregated products. To enable coverage of all of the disciplines represented in their collections of articles, libraries subscribe to multiple A&I products. Library users, in turn, must know which of these products they need to use to complete

their literature search. To a very large extent, index-based discovery services emerged to simplify the research process relative to the complex process of identifying and using many different individual content products, each with its own interface and search procedure.

A&I resources have historically had limited participation with discovery services. As they both serve similar functions, enabling researchers to find the information they need, A&I services have viewed discovery services with some trepidation. Furthermore, A&I services may view their highly curated records as their primary asset, which they have been hesitant to provide without additional input about how their content is used. However, libraries value discovery services in part because they include these high-value records and want their discovery services to be reflective of all their subscribed materials, including A&I resources. While challenges remain, discovery services have universally taken strides to accommodate A&I concerns and to encourage A&I participation, and many A&I databases participate in one or more discovery indexes. (See the ODI report, *Bridging the Gap Between Abstracting & Indexing Provider Needs and Discovery Service Approaches*, for more information.) This recommended practice also outlines steps that discovery service providers should undertake (if not already in place) in order to ensure increased A&I participation with all discovery platforms.

2.3 Metasearch Utilities

One of the techniques that emerged in the early 2000s to provide simultaneous search across a library's electronic resources consisted of a utility that would prompt the researcher for a query and then transmit that query to multiple targets. The utility would accept the results returned by each of the resource targets and present them to the researcher. How they might be presented would vary according to the tool's capabilities and configuration and might include grouping by source or interfiling and ranking results. This approach, called metasearch or sometimes federated search, comes with some inherent limitations, such as the number of live simultaneous sessions that could be maintained with resource targets. One technique implemented to address this limitation involved the creation of disciplinary sets that would be selected by the user in order to direct the query to the best set of targets for the topic at hand.

On a technical level, metasearch initially worked through communication sessions with each target that essentially emulated the search session presented by a user. This approach required parsing the HTML delivered by each target into a structured form that could then be handled by the metasearch utility for sorting, presentation, and other tasks. This method of HTML parsing proved to be fragile, with any change made by a resource provider requiring updates to the parsing algorithms. Specialized connectors were developed, with some organizations specializing in creating and maintaining connectors on behalf of metasearch providers. Some content targets offered Z39.50 or SRU/SRW servers, enabling metasearch products to make use of these well-established library standard protocols. Other techniques such as XML gateways were also developed for some resource targets designed to respond to metasearch requests with structured data and with server processes insulated from production instances of their content products. Metasearch, though a pragmatic approach, had inherent performance limitations due to the multiple communications streams and processing performed in real time in response to user queries. The slow performance, limited number of content targets that could be included in a query, and the limited number of results returned were factors that impacted the success and satisfaction with metasearch.

The NISO Metasearch Initiative was launched in May 2003 to define standards and recommended practices in support of this search model.

2.4 Enhanced Library Catalogs

The online catalogs provided with ILSes were often not easily understood by library users, both in terms of operating the interface and in the scope of content addressed. In contrast, search engines such as Google were used effortlessly and helped to establish much higher expectations for the simple and

comprehensive nature of any search facility offered by libraries. Enhanced library catalogs, or discovery interfaces, that included characteristics such as improved user interfaces; simplified and advanced query options; results ordered according to relevancy rankings; alternative query recommendations; faceted navigation; and enriched content such as cover art, summaries, or tables of contents emerged in the mid-2000s. Most of these discovery interfaces were designed to work with any of the major ILS products and were based on local indexes built from the contents of the ILS as well as other local content repositories. The scope of these products could be expanded to include article-level content through the use of optional metasearch components.

2.5 Index-Based Discovery Services

Index-based discovery services, initially introduced in 2009, aim to provide an interface that provides a very broad representation of a library's collection through a single search box accessing an index populated with the article-level metadata and full-text content offered by the library—whether in print or via other resources managed in the local automation environment. The creators of these index-based discovery services make arrangements with library-oriented content providers to receive citation and/or full-text metadata of their resources. The indexes are constructed to include as much of the potential content as possible to which libraries might subscribe or otherwise add to their online collections. The discovery service also includes a mechanism to profile the content according to the subscriptions associated with any given library implementing that discovery service. Metadata from the library's local ILS or other local repositories may also be included in the index and periodically synchronized. Library patrons who use the discovery service are presented with lists of results that match their query. When the user selects an entry and requests to view the associated content item, the discovery service links to the content provider's platform/server, potentially via the library's OpenURL link resolver. Items selected from the library's print materials, managed through the ILS, are either presented as links to the corresponding entry in the local online catalog or to the appropriate service options integrated into the discovery service interface.

This model of index-based discovery generally bypasses the native interface of the content providers' services, but brings library users to those same content items for viewing, downloading, printing, or other available access options. The purpose of these discovery services is not to republish materials represented in the index, but to provide an additional channel for connecting library users with that content.

This model of index-based search offers advantages over previous approaches implemented in libraries. By indexing content in advance, discovery services have the ability to deliver more sophisticated services with instant performance, compared to the metasearch techniques used previously.

2.6 Related Initiatives

A number of initiatives produced recommendations related to the use of discovery services. The most relevant of these to the ODI work are described below. Links to the projects and reports described are available in the Bibliography.

COUNTER (https://www.projectcounter.org/)

> COUNTER (Counting Online Usage of Networked Electronic Resources) is an international initiative serving librarians, publishers, and intermediaries by setting standards that facilitate the recording and reporting of online usage statistics in a consistent, credible, and compatible way.

> COUNTER included recommendations for discovery service reports in *Code of Practice*—Release 4—in June 2016 (https://www.projectcounter.org/wp-content/uploads/2016/02/APPM.pdf). These recommendations continue in the most recent COUNTER release from January 2019 (Code of Practice—Release 5 (https://www.projectcounter.org/code-of-practice-five-sections/abstract/).

Digital Library Federation ILS Discovery Interface Task Group (ILS-DI)
(https://old.diglib.org/architectures/ilsdi/)

> "In 2007-2008, the DLF convened a Task Group to recommend standard interfaces for integrating the data and services of the Integrated Library System (ILS) with new applications supporting user discovery." The group's technical recommendations define an API for "effective interoperation between integrated library systems and external discovery applications."

International Coalition of Library Consortia (ICOLC) (https://icolc.net/)

> From its *Revised Statement on the Global Economic Crisis and Its Impact on Consortial Licenses*: "Principle 3 (added June 2010). We encourage publishers to allow their content to be made available through numerous vendors appropriate for their subject matter. We also encourage online providers and aggregators to allow their metadata to be included in emerging discovery layer services on a non-exclusive basis. Multiple access platforms will permit libraries and consortia to select content and discovery tools that are suitable and affordable for their constituents. We encourage vendors to provide options that match the range of needs that libraries have for any particular content as to degree of importance, currency, interfaces, access, archiving, preservation, and metadata. It is in the common interest of publishers, database vendors, consortia, libraries, and information consumers to work collectively to provide affordable access to licensed content, while preserving the businesses integral to our collective success."

JISC Discovery Programme (http://discovery.ac.uk/)

> "The JISC-funded Discovery programme was launched in May 2011 to create 'a metadata ecology' to support better access to vital collections data in libraries, archives and museums and facilitate new services for UK education and research." The project work included the publication of *Discovery Open Metadata Principles* and invited stakeholders to sign up in support of the principles and to enact them.

KBART–Knowledge Bases And Related Tools (https://www.niso.org/standards-committees/kbart)

> The NISO and UKSG joint Knowledge Bases and Related Tools (KBART) project was initiated to "develop and publish guidelines for best practice to effect smoother interaction between members of the knowledge base supply chain." The initial Recommended Practice (NISO RP-9-2010) provided best practices for formatting and distributing title lists by content providers to knowledge base developers to improve the accessibility of resources obtained through the use of OpenURL link resolvers. A revised Recommended Practice is now available, which expands the original recommendations and focuses on the more granular, complex issues that cause problems in metadata supply. Since the metadata supplied for knowledge bases is often the same as that supplied to indexed-discovery service providers, implementation of the KBART recommendations should also improve the data used in discovery services.

Music Discovery Requirements (http://committees.musiclibraryassoc.org/ETSC/MDR)

> The Music Discovery Requirements "explore discovery needs specific to and especially important for music materials, particularly scores and recordings. Music materials pose unique demands that must be considered for successful discovery." First released 2012, with a revised version released 2017, with official support and endorsement from the Music Library Association. The document discusses the issues and when possible gives concrete recommendations for display and indexing. Two appendices compile technical details of the specific indexing recommendations in spreadsheets.

NFAIS Recommended Practices: Discovery Services
https://wayback.archive-it.org/12684/20190905184506/https://nfais.memberclicks.net/assets/docs/BestPractices/recommended_practices_final_aug_2013.pdf)

The National Federation of Advanced Information Services (NFAIS) "members agreed that it would serve the community to develop a code of practice that would outline the rights and obligations of all participants in a discovery service relationship in order to move forward with an understanding of mutual expectations." The Recommended Practice is intended to assist their Content Owner members who contribute their content to discovery services by providing guidelines that will help maintain an equitable balance of the interests of all participants in the information distribution process. The NFAIS Recommended Practice was developed to be "the ideal towards which the Information Community should strive as technology advances, even if today's technology presents barriers to full implementation."

2.7 Information Gathering and Results

An important component of the ODI work involved gathering information from the key stakeholders: libraries, content providers, and library discovery service providers. Although the NISO ODI membership represented all three of these key stakeholder groups, the Working Group conducted a wider survey of all three constituencies; survey results helped each ODI subgroup prioritize its work and informed these recommendations.

Each of the ODI subgroups recommended suitable questions for a survey, following a series of interviews used to assess the most pressing needs of the stakeholders. The full survey was announced on September 11, 2012 and closed on October 4, 2012. The survey questions and results can be found in Appendix D.

The merged Communication of Library's Rights and Level of Indexing subgroups compiled a set of data elements useful for all parties in understanding what data are ingested into discovery indexes, what data are used to generate search results, what data are displayed on the screen, and what uses discovery service consumers might make of these elements. These elements set the stage for determining the merged subgroup's survey questions for its stakeholders (libraries, content providers, and discovery service providers).

During the process of reviewing the survey results, it was clear that some of the questions posed by the merged Communication of Library's Rights and Level of Indexing subgroup caused confusion for many respondents. In particular, responses from content providers were considered insufficient for informing best practices and led to a follow-up survey for content providers, both to validate the initial survey responses and to explore further the potential barriers to content provider participation in discovery services. The major barriers related largely to business modeling concerns (how to continue to prove their value as a service within the larger discovery environment), acknowledgement for their content/services within the discovery interface via appropriate branding, lack of metrics related to use of provided content, "unfair" linking, and concerns about lack of advance options for their content in a discovery system. Many comments throughout the research revealed a need for more transparency and educational dialog between discovery service providers, content providers, and libraries.

The 2020 Recommended Practice revision was informed by information gathering, through surveys and interviews with stakeholders. In June 2017, a subgroup of the Open Discovery Initiative Standing Committee conducted a survey of A&I providers to understand their concerns. This was followed in February 2018 with a related survey of discovery service providers. Results were summarized in the 2019 report, *Bridging the Gap Between Abstracting & Indexing Provider Needs and Discovery Service Approaches* (https://groups.niso.org/apps/group_public/download.php/21877/). Additional surveys for stakeholders were completed in 2019, available in Appendices E–G.

Section 3: Recommendations

3.1 General Recommendations

1. Create an oversight group (Standing Committee or Working Group) to promote educational opportunities about adoption of these recommended practices, provide support for content providers and discovery service providers during adoption, provide a forum for ongoing discussion related to all aspects of discovery platforms for all stakeholders (content providers, discovery service providers, libraries), and determine timing for next steps for ongoing work.

2. Content providers, discovery service providers, and libraries can take specific measures to assert their conformance with the recommended practices of the ODI. These measures will be voluntarily made by the organizations. Self-check conformance lists are provided in Appendices A–C. Conformance will not be assessed by NISO or any other formal organization, but through general community review.

3.2 Best Practices for Content Providers

3.2.1 Metadata Elements Provided to Discovery Services

The trend towards index-based discovery requires cooperation between content providers, discovery service providers, and libraries to ensure that the broadest spectrum of materials can be fully exposed through discovery platforms.

The elements listed in 3.2.1.1 and 3.2.1.2 are based on the KBART (NISO RP-9-2014) metadata encoding schema—which is intended primarily for serial and monograph titles to improve A–Z lists, link resolution, and OpenURL—but has been extended here to capture elements for subject, abstract/description, content type, and content format. This extension is intended to serve as a minimum basis from which all types of content may be described.

3.2.1.1 General Requirements

1. Content providers should make available to discovery service providers core metadata and underlying full-text/original content for complete offerings, for the purposes of indexing to meet licensed customers' and authenticated end users' needs.

2. To this aim, all content providers should make available to discovery service providers, at a minimum, the core set of metadata elements and content item (full text, transcript, etc.) (see Section 3.2.1.2) for each item they submit for indexing.

3.2.1.2 Core Metadata

The fields shown in Table 1 collectively constitute the minimum that must be provided by Content Providers (CPs) to Discovery Service Providers (DSPs) for each item provided to the DSP for indexing. Fields must be provided where they are applicable to the content type. Examples are provided for each metadata element in Table 2. Table 2 shows possible data entries with labels for purposes of illustration but is not meant to be representative of the actual data format as it is transferred. Actual data transmitted to the discovery service provider should follow established data transfer protocols as described in Section 3.3.3.

Note: It is recognized that many content providers merge Content Type and Content Format in their systems. Providing separate fields for this data is preferred, but the current practice of a single field may continue if separating the data is too burdensome.

Providing the full range of available metadata to the discovery service improves the discovery experience for users, particularly for librarians and advanced researchers who find particular value in controlled vocabularies and the added-value content created by A&I providers.

Table 1: Core metadata elements to be provided by content providers

Field Name	Definitions
Title	The main title of the item.
Authors	The author(s) of the item. Individual authors should be listed in lastname, firstname order.
Author Identifier	One or more standard identifiers for the authors of the item (e.g., ORCID/ISNI VIAF, etc.). The identifier should be preceded by a label indicating the type of identifier.
Publisher Name	The name of the publisher of the item.
Volume	Volume number of the resource, where applicable.
Issue	Issue number of the resource, where applicable.
Page(s)	Page number(s) of the resource, where applicable.
Date/Date Range	The date of publication. For a serial run, coverage dates included for the serial.
Item Identifier	One or more standard identifiers for the print or online version of the item (e.g., ISSN, OCLC number, ISBN, DOI, etc.). The identifier should be preceded by a label indicating the type of identifier.
Component of Title	Describes the publication or serial of which the individual item is a part (e.g., for journal articles, the serial title; for tracks on a CD, the album title; etc.).
Component of Title Identifier	Provides a standard identifier for the component title defined above (e.g., ISSN, OCLC number, ISBN, DOI, etc.). The identifier should be preceded by a label indicating the type of identifier.
Item URL	Either an OpenURL or a direct link for the specific item's full text.
Open Access Designation	To comply with the NISO Open Access Metadata and Indicators group's recommendations, if an item is open access, this status should be indicated with "free_to_read" and otherwise left blank. See https://www.niso.org/standards-committees/access-and-license-indicators.
Full Text Flag	A yes/no statement describing whether the content provider makes this item available in full text (or for nonprint media, a full-length or high-resolution version) to the DSP for the purpose of indexing. It is expected that this will be disclosed by DSPs to libraries in future when describing indexing coverage for a title or collection.

Field Name	Definitions
Content Type*	Intended to be used to identify whether the content being described is textual, a visual recording, a sound recording, etc. The textual descriptors from the controlled list established in the MARC 21 Type of Record position (06) of the leader field is recommended to be used for this field's content.
Content Format*	Intended to be used to indicate whether the nature of the content being described is monographic, serial, a component part, collection, etc. The textual descriptors from the controlled list established in the MARC 21 Bibliographic Level position (07) of the Leader field is recommended to be used for this field's content.
Language	Language of the content item. The list of MARC language codes at https://www.loc.gov/marc/languages/language_code.html is recommended to be used for this fields content.
Indexing Data	One or more keywords (from controlled or uncontrolled vocabularies) to describe the content of the item.
Abstract/ Description	Either a text summary on the content or (for non-text materials) a description of the item.
* It is recognized that many content providers merge Content Type and Content Format in their systems. Providing separate fields for this data is preferred, but the current practice of a single field may continue if separating the data is too burdensome.	

Table 2: Examples of core metadata elements provided by content providers

Field Name	Monograph & Serials	Articles/ Documents	Manuscripts/ Unpublished Content	Media Content Image	Media Content Recording	Media Content Motion Image
Title	Nature	Derivation of pluripotent stem cells horn cultured human primordial germ cells	The evolution of cooperative signal comprehension in the domestic dog (Canis familiaris)	Portrait of Françoise en costume tailleur	Panama	The Scent of Green Papaya
Authors		Shamblott, MJ; Axelman, J; Wang, SP; Bugg, EM; Littlefield, J; Donovan, PJ; Blumenthal, PD; Huggins, GR; Gearhart, JD	Wobber, Victoria Elizabeth	Picasso, Pablo	Armstrong, Louis	Tran, Anh Dung; Lu, Man San; Truong, Thi Loc; Yen-Khe, Tran Nu; Delhomme, Benoit; Abril, Jean-Philippe; Dedieu, Nicole; Roques, Jean-Pierre; Rossignon, Christophe; Negre, Alain; Ton-That, Tiet
Author Identifier		ORCID 0000-0003-0843-4164 ORCID 0000-0002-2861-7438 ORCID 0000-0002-1965-7614		VIAF ID 15873 ISNI 0000 0001 2275 5381	VIAF ID 195226 ISNI 0000 0001 1019 8071	ISNI 0000 0000 3572 4552 ISNI 0000 0001 1946 9376 ISNI 0000 0000 0129 8647 ISNI 0000 0001 2096 3574 *[Ed note: Full content truncated in example table]*

Field Name	Monograph & Serials	Articles/ Documents	Manuscripts/ Unpublished Content	Media Content Image	Media Content Recording	Media Content Motion Image
Publisher Name	Macmillan Journals Ltd.	New York Academy of Sciences			Future Noise Music Ltd	Kino International
Volume		95				
Issue		23				
Page(s)		13726				
Date/Date Range (first issue for serials)	19950105-	19981110	20060929	1946	2008	1993
Item Identifier	ISSN 0028-0836 eISSN 1476-4687	DOI 10.1073/pnas.95.23.13726	OCLC 157003171		ISRC GBSUW0716131	LCCN 97505399
Component of Title	N/A	Annals of the New York Academy of Sciences.	N/A	N/A	Classic Song Book, Vol. 4	N/A
Component of Title Identifier	N/A	0077-8923; 749-6632	N/A	N/A	UPC 884385621520	N/A
Item URL	http://www.nature.com/nature/index.html	http://<linkresolver>/?pid=username:password&id=doi:10.1073/pnas.95.23.13726&noredirect=true		http://www.bridgemaneducation.com/ImageView.aspx?balid=166841	http://jazz.alexanderstreet.com/View/1030219	http://feat.alexanderstreet.com. View/1825104
Open Access Designation			free_to_read			
Full Text Flag	N	N	N	N	N	N

Field Name	Monograph & Serials	Articles/ Documents	Manuscripts/ Unpublished Content	Media Content Image	Media Content Recording	Media Content Motion Image
Content Type	Book, Language Material	Article, Journal Article	Manuscript	Image, DVD, Movie	Musical sound recording	Projected medium
Content Format	Serial	Item	Monograph	Item	Monographic component part	Monograph
Language	eng	eng	eng	zxx	zxx	vie
Indexing Data	Science	Stem Cells; Regenerative Medicine; Mesenchymal Stem Cells; Differentiation; Therapy	Human-animal communication. Domestication	Female; Picasso's Mistress, Gilot, Drawing, Lover	Music & Performing Arts—Jazz	Music & Performing Arts—Film Romantic Relationships
Full Text/Transcript	[full text]	[full text]	[full text]		[transcript]	[transcript]

Field Name	Monograph & Serials	Articles/ Documents	Manuscripts/ Unpublished Content	Media Content Image	Media Content Recording	Media Content Motion Image
Abstract/ Description		Aging is associated with a progressive failing of tissues…				An Academy Award nominee for Best Foreign Language Film, Tran Anh Hung's "luxuriant, visually seductive debut" (New York Times) recreates antebellum Vietnam through both the wide eyes of childhood and the deep blush of first love. In 1951 Saigon, 10-year-old Mui (Lu Man San) enters household service for an affluent but troubled Vietnamese family. Despite her servile role, Mui discovers beauty and epiphany in the lush physical details that envelope her, while earning the fragile affection of the household's grieving matriarch. As she comes of age, the now grown Mui (Tran Nu Yen-Khe) finds her relationship with a handsome pianist she has admired since childhood growing in depth and complexity. The Scent of Green Papaya is "a film to cherish." (Roger Ebert).

3.2.2 Metadata Elements Provided by Content Providers to Libraries

The ODI recommends that content providers disclose on a publicly available website their level of participation in discovery services to library subscribers. For each market product (journal collection, A&I database, etc.), content providers should disclose the coverage depth and content depth provided, as described in Table 3.

Table 3: Summary of metadata elements to be provided by content providers to libraries

For Each Market Product Included	Disclose the Following			
	Coverage	Content Provided to Discovery Service Providers for Indexing	Content Provider Type	Content Provided To
For example: • Journal Collection • E-book Collection • Image Collection • Audio Collection • Video Collection • Database • Data sets	For journals: • One or more date range(s) For other products: • All OR • Part (describe scope of part)	List all that apply to content provided: • Metadata • Indexing data • Abstract/ Description • Full text	The type of agent the content provider acts as for the Market Product. Options are: • Publisher • Aggregator • A&I service • Open Access Repository/ Institutional Repository For full-text items, this is generally the publisher, aggregator, or open access repository/ institutional repository. For A&I-provided items, this is generally the A&I service.	Discovery services the content is made available to. Options are: • Available to all DSP agents – when content is made available for harvest by any interested party. • Available to DSPs by agreement [provide list of DSPs under agreement].

3.2.3 Disclosure

We recognize that some discovery service providers and/or content providers may wish to incorporate non-disclosure agreements (NDAs) and similar provisions into the contracts and other agreements that they enter into with one another. But NDAs should not be used to avoid the disclosures and transparency recommended in Section 3.3.2 and elsewhere in this document.

3.2.4 Technical Formats

The ecosystem of index-based library discovery includes the transmission of data from content providers to discovery service providers for the purpose of building a central index. Although the number of current providers of index-based discovery services remains fairly limited, the number of potential content providers is immense. Given the vast amounts of data being exchanged, it is in the interest of all stakeholders to employ the most efficient transfer mechanisms available.

Section 3.3.3 describes the major metadata encoding schemas, file format conventions, and methods of transfer currently in use within the discovery services ecosystem.

The ODI recommends that the transfer of data from content providers to discovery service providers should make use of existing standards where applicable. Some of the standards and protocols most directly applicable include the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH) and KBART. ResourceSync also has strong potential as a mechanism for data transfer for discovery services once it is published by NISO and software tools become available for its implementation.

3.2.5 Link Resolution

3.2.5.1 Fair Linking

ODI recommends that content providers implement the following protocol to ensure fair linking:

Content providers should continue to provide at a minimum data to support OpenURL Resolution; see ANSI/NISO Z39.88-2004 (R2010) *The OpenURL Framework for Context-Sensitive Services* (https://www.niso.org/publications/z3988-2004-r2010). The resulting OpenURL should link to the item itself, or as close to it as possible in the content provider's site structure. Additional data points needed to support direct linking mechanisms should not replace the metadata needed to support OpenURL resolution, as this would remove library choice in using OpenURL to the platform of their choice for resolution.

3.2.6 Support Tools

Content Providers should create clear channels for reporting discovery related support issues. It is preferable that these tools be available online.

3.3 Best Practices for Discovery Service Providers

3.3.1 Discovery Service Content Listings

Discovery service providers should make available to prospective and current customers sufficient information about the content of their central index to ensure an adequate evaluation of that content against the customers' needs.

3.3.1.1 Metadata Elements – Collection Level

At the market product level, metadata should be made available to both libraries and content providers. This view provides both content providers and libraries with a broad understanding of the current content represented in the central index. The following metadata should be made available:

- Provider

- Market product

- Titles in knowledge base (KB)

- For the following, the count or the proportion of overall coverage and coverage from records provided directly by the provider should be indicated:

    o   Titles in central index

    o   Number of unique records in central index

    o   Percent of records full-text searchable in central index

    o   Percent of records abstract searchable in central index

- o Percent of records subject searchable in central index (author provided, publisher provided, controlled vocabulary, etc.)

- o Percent of articles that are free to read

- Date of last market product update

- Date of report update

Additionally, it is strongly recommended that a summary page be attached to each monthly report, which shows the changes in content provider directly provided record counts for each market product over time (see Table 5).

Table 4: Definitions of collection-level data elements to be provided by discovery service providers

Field Name	Definitions
Provider	Entities providing metadata to discovery services (e.g., MLA, APA, etc.).
Market Product	A specific, defined collection of resources made available for license or use by a content provider to a library. This could be a citation database, a set of e-journals or e-books, or other constellation of titles or objects commonly licensed as a single entity.
Titles in Central Index	The number of titles in the market product available within the central Index.
Number of Unique Records in Central Index	The total number of unique records in the market product excluding duplicate records for the same source.
Number of KB Titles	The number of titles in the market product listed within the central index's associated link resolver knowledge base (KB).
Percent of Records Full-Text Searchable in Central Index	Percentage of records in the market product allowing searching of the full text.
Percent of Records Abstract Searchable in Central Index	Percentage of records in the market product allowing searching of the abstract.
Percent of Records from Provider that are Subject Searchable in Central Index	Percentage of records that contain subject headings (sometimes referred to as descriptor terms). These subject headings can be provided by authors or publishers and can be but are not necessarily reflective of controlled vocabulary.
Percent of Articles that are Free to Read	Percentage of records that will allow the user to access full-text content available at no charge to the user or library.
Date of Last Market Product Update	Date of most recent update of the market product within the central index.
Date of Last Report Update	Date this report was last updated.

Table 5: Example of summary table to be provided by discovery services

Provider	Market Product	Date 1	Date 2	Date 3	…
Publishing, Ltd.	Full-text Journals	12,000	11,000	10,000	

Table 6: Example of collection-level data elements to be provided by discovery service providers

Provider	Market Product	Titles in KB	Titles in Central Index		Number of Unique Records in Central Index		Percent of Records Full-Text Searchable in Central Index		Percent of Records Abstract Searchable in Central Index		Percent of Records from Provider that are Subject Searchable in Central Index		Percent of Articles that are Free to Read		Date of Last Market Product Update (content provider only)	Date of Last Report Update
			Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider		
Publishing, Ltd.	Full-text Journals	100	98	90	15,000	12,000	90%	95%	90%	95%	90%	95%	75%	50%	1-Jan-19	18-Jun-19

3.3.1.2 Metadata Elements – Title Level

From the market product level, customers should be able to drill down to a detailed title list within a collection. The following metadata (where relevant for each content type) should be made available:

- Title

- Standard identifier(s), e.g., ISSN, ISBN

- Content format

- Content type

- For the following, the count or the proportion of overall coverage and coverage from records provided directly by the provider should be indicated:
  - Dates of coverage (journals) or date of publication
  - Number of records in central index
  - Percent of records full-text searchable in central index
  - Percent of records abstract searchable in central index
  - Percent of records from provider that are subject searchable in central index (author provided, publisher provided, controlled vocabulary, etc.)
  - Percent of articles that are free to read
- Date of last report update

If content provided by the content provider does not contain a standard identifier, it may be difficult to provide counts and proportions across content providers.

Table 7: Definitions of collection-level data elements to be provided by discovery service providers

Field Name	Definitions
Title	The main title of the item.
Standard Identifier(s), e.g., ISSN, ISBN	One or more standard identifiers for the print or online version of the item (e.g., ISSN, OCLC number, ISBN, DOI, etc.). The identifier should be preceded by a label indicating the type of identifier.
Content Format	Intended to be used to indicate whether the nature of the content being described is monographic, serial, a component part, collection, etc. The textual descriptors from the controlled list established in the MARC 21 Bibliographic Level position (07) of the Leader field is recommended to be used for this field's content.*
Content Type	Intended to be used to identify whether the content being described is textual, a visual recording, a sound recording, etc. The textual descriptors from the controlled list established in the MARC 21 Type of Record position (06) of the Leader field is recommended to be used for this field's content.*
Dates of Coverage (journals) or Date of Publication	The date of publication. For a serial run, coverage dates included for the serial.
Number of Records in Central Index	The number of individual bibliographic records in the central index.
Percent of Records Full-Text Searchable in Central Index	Percentage of records allowing searching of the full text, not just the bibliographic metadata and abstract. For text items, the entirety of the document. For audio or video content, a full transcript of the spoken content of the material. May not be relevant for all indexed content.
Percent of Records Abstract Searchable in Central Index	Percentage of records allowing searching of the abstract in addition to bibliographic metadata. Abstract is either a text summary of the content or (for non-text materials) a description of the item.
Percent of Records from Provider that are Subject Searchable in Central Index	Percentage of records that contain subject headings (sometimes referred to as descriptor terms). These subject headings can be provided by authors or publishers and can be, but are not necessarily, reflective of controlled vocabulary.
Percent of Articles that are Free to Read	Percentage of records that will allow the user to link to content available without a paid subscription.
Date of Last Report Update	Date this report was last updated.

* It is recognized that many content providers merge Content Type and Content Format in their systems. Providing separate fields for this data is preferred, but the current practice of a single field may continue if separating the data is too burdensome.

Table 8: Examples of title-level data elements to be provided by discovery service providers

Title	Standard identifier(s) e.g., ISSN, ISBN	Content format	Content type	Dates of coverage (journals) or date of publication		Number of records in Central Index		Percent of Records Full-Text Searchable in Central Index		Percent of Records Abstract Searchable in Central Index		Percent of Records from Provider that are Subject Searchable in Central Index		Percent of Articles that are Free to Read		Date of Last Report Update
				Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	Overall	Content Provider	
Full-text book	978-1234567890	Monograph	Language material	2012	2012	12	12	0	0	100	100	100	100	0	0	18-Jun-19
Full-text journal	1234-5678	Serial	Language material	1980-2019	1980-2015	2,000	1,500	90	95	100	95	100	95	50	40	18-Jun-19

3.3.1.3   Distribution of Metadata Elements

The metadata in Section 3.3.1.1 should be provided in a downloadable form. It is recommended that the KBART practices be followed where relevant. However, the quantity and breadth of data handled in the discovery service supply chain necessitate some variance from the practices recommended by KBART.

1. **Format** – ODI, like KBART, recommends the use of tab-delimited text files for the transmission of metadata.

2. **File Naming** – For files provided by discovery services, files should be named as follows, encoded in UTF-8: ProviderName_YYYY-MM-DD.csv with the date information indicating the date of file generation, e.g.: ExLibrisPrimo_2013-01-28.csv.

3. **Frequency of Updates to Metadata** – Current customers should receive regular updates of the metadata specified in Sections 3.3.1.1 and 3.2.1.2 from their discovery service provider on a monthly basis. Prospective customers should be provided equivalent data on demand.

4. **Access to the Reports** – Access to reports by library staff may be secured using Internet Protocol address controls or password protection if CP, DSP, or library requests this. Discovery Service Provider is responsible for securing access to the reports.

3.3.2   Link Resolution

3.3.2.1   Fair Linking

ODI recommends that discovery services implement the following protocols to ensure fair linking:

1. Discovery services should not discriminate based on business relationships among content providers or products (especially their own) in the methods that are used to generate results, relevance rankings, or link order.

2. In cases where the same content is available through multiple content platforms (such as a primary publisher and a secondary database, or multiple secondary databases), discovery service providers should provide mechanisms to enable libraries implementing the service to establish preferences regarding which platforms to present to users as link targets, and in what order or priority. Discovery service providers should continue to support the OpenURL standard in their outbound links.

3. Discovery service providers should issue a statement annually to all current and potential customers (or generically on their website in an area available to all customers) explaining their business connections with content providers, including those with direct or indirect ownership relationships and those with which they have negotiated paid or other agreements for metadata deposit, direct linking, or other special arrangements.

4. Discovery service providers should offer an affirmative statement of the neutrality of their algorithms for generating result sets, relevance rankings, and link order with respect to ideological/political viewpoint, content provider source (especially but not limited to any content that it or its parent organization may provide), and any other relevant factors.

5. Discovery service providers should make other aspects of link presentation associated with a given result configurable options by libraries (including the number of links presented, the order in which links are presented within a given result, and how libraries' "get the full text" link is labeled/branded).

6. Discovery service providers should supply content providers and libraries with information when material changes are made to the discovery service that could impact the result set or relevance rankings or link order of results.

3.3.3 File Formats and Methods of Transfer

ODI recommends that discovery services implement the following practices for the transfer of data from one party to another:

1. Discovery service providers should clearly describe their capabilities, limitations, and preferences regarding how content providers should transfer data to them for the most effective indexing by the discovery service.

2. Discovery service providers should communicate to content providers when format, schema, or transport mechanisms will have an impact on features or performance of the discovery service.

3. Where possible, it is recommended that robust metadata encoding schemas be used to describe a content provider's metadata. Appropriate, supported schemas should be agreed to by the content provider and discovery service provider, and preferably disclosed to customers. Table 9 lists the primary metadata encoding schemas in wide use in the library and publishing communities. ODI will maintain an updated list of such schemas on its workroom website (https://www.niso.org/standards-committees/odi).

Table 9: Primary metadata encoding schemas

Schema	Schema Link
MARC	http://www.loc.gov/marc/
MODS	http://www.loc.gov/standards/mods/mods-schemas.html
METS	http://www.loc.gov/standards/mets/
VRA	http://www.loc.gov/standards/vracore/
DC	http://dublincore.org/documents/dcmi-terms/
KBART	http://www.niso.org/workrooms/kbart/
EAD	http://www.loc.gov/ead/
ONIX	http://www.editeur.org/8/ONIX/

4. The method and schedule of delivery of metadata to the discovery service should also be agreed to by both parties. The file formats in Table 10 and methods of transfer in

5. Table 11 are recommended.

Table 10: File formats

File Format	Example
text	csv, mrc, txt, etc.
xml	MARCXML, MODS XML, DC XML, etc.

Table 11: Methods of transfer

Transfer Method	Description
oai-pmh	Protocol for Metadata Harvesting from the Open Archive Initiative
resourcesync	Defined in the forthcoming NISO standard *ResourceSync Framework Specification* (Z39.99)
api	Standardized or proprietary application programming interface
ftp	Standard IETF File Transfer Protocol

3.3.4 Usage Statistics

3.3.4.1 Recommendations to Support Content Providers

3.3.4.1.1 Recommended Metrics Provided to Content Providers

Discovery service providers should offer usage metrics that comply with COUNTER: Provider Discovery Reports (https://www.projectcounter.org/wp-content/uploads/2016/02/APPM.pdf) as a standard operating practice. The COUNTER-defined reports provide consistent metrics that enable content providers to understand how their content is being used on various platforms. These reports add value to the relationship between the discovery service provider and the content provider.

The COUNTER reports are similar to the reports discovery service providers currently provide to libraries. The Provider Discovery Reports detail the usage of a specific content provider's metadata and content by all mutual customers.

The COUNTER statistics help content providers answer many important questions:

- Is the provider's content being found?

- Is the metadata effective (or could it be more effective)?

- Who is using the provider's content? For example, is restricted A&I metadata only being used by licensed institutions?

The process for setting up delivery of the usage reports should be transparent. Discovery service providers should designate individuals responsible for setting up and delivering monthly usage reports. This information should be shared with all content providers that partner with the discovery service provider. The timing for the monthly delivery of the reports should be predictable.

3.3.4.1.2 Link Source Identification

In order for content providers to identify, and therefore quantify, the type of user traffic originating natively from a discovery service to their own content platforms, it is recommended that discovery service provider links include an identifier for their discovery service in links to content providers in at least one of these ways: For OpenURL links generated by the discovery service provider, use the referrer SID as described in section 18.4 of the OpenURL specification For pre-calculated or direct links, append an argument, sid=<referrer>, to the URL In either case, the sid should be consistent for all links provided by the discovery platform. (The expectation is that intermediary link resolvers would pass along this referrer identification to the content provider with the OpenURL request.)

3.3.4.2 Recommended Metrics Provided to Libraries

Discovery service providers should offer a set of structured metrics to illustrate to libraries how the discovery service is utilized. More complex metrics should be considered for a future phase of this initiative. The current recommended metrics are as follows:

1. **Total number of searches per month** – Report the total number of times a user performed a search of the discovery service, providing for tracking of month-over-month usage trends. Additionally, when possible, the discovery service should provide search counts by collection for A&I databases and similar types of collections since these resources are often restricted to authenticated users; this will help to assess the use of these resources in the discovery process.

2. **Total number of unique visitors per month** – Report the total number of unique visitors by month to provide library clients with month-to-month trending of usage among their constituency.

3. **Total number of click-throughs per month** – Report the total number of times a user has clicked on a link to request full-text content. This metric provides an indication of how frequently the discovery service yielded a result perceived by the user to be valuable to their research. When possible, the discovery service should provide click-through counts by collection for items in A&I databases and similar types of collections since these resources aid discovery of items that may not be the ultimate target destination for the user; this will help to assess the contribution of these resources in the discovery process.

4. **Top 500 search queries for the last period** – Report on the 500 most frequently submitted search queries for the specified period. Search query data enables a library to consider the types of searches—and topics—being researched when using the discovery service.

5. **Top 100 referring URLs to the discovery service for the last period** – Report on the 100 most frequently used URLs to bring the user to the discovery service. This metric provides the library with an indication of how users are getting to the discovery service.

To allow libraries to analyze the usage of their discovery service, they will need to have the ability to access the aforementioned metrics and integrate them into a usage analysis or library management system. To satisfy this need, it is recommended that the reporting of the data points be provided in CSV formatted reports. Note that implementation of a discovery service might result in declining search counts from native content platforms. This is not unexpected since the library is promoting access to its content via search of the discovery service's search index. However, full-text retrievals from native product platforms typically increase upon introduction of a discovery service. As full text is the artifact patrons seek, libraries should find full-text document retrievals a better metric for assessing the value of a content product rather than search counts on native product platforms.

It is recommended that the metrics above be incorporated by COUNTER in a future version of the *Code of Practice* that specifies details of format and distribution.

3.3.5 Open Access

In order for users of discovery services to understand which materials are available open access, in the public domain, or otherwise freely available, item-level records presented to end users should indicate the access and license terms as defined by NISO RP-22-2015 *Access and License Indicators* (https://www.niso.org/standards-committees/access-and-license-indicators) and as described in Section 3.2.1.2.

3.3.6 Authentication

1. Discovery services should offer authentication mechanisms, when necessary, to ensure content from databases included in search results are only available for authenticated users of institutions that are mutual subscribers of both the database and the discovery service in which it is included.

2. Discovery services should take measures to ensure that only current subscribers activate content provider databases in discovery systems.

3.3.7 Alternative Coverage Lists

When discovery service providers make alternative coverage lists for content that is not directly indexed from specialized metadata providers, the reports should clearly indicate that the metadata fields being indexed are alternatives to specialized metadata. Reports should note potential differences in quality, depth, and currency of the alternative metadata.

3.3.8 Record Display

1. Discovery services should have the ability to exclude records from merged records that combine metadata from multiple providers and groups of records that reflect the same item such that records can stand alone if requested by the content provider.

2. Discovery services should be able to display the source provider of the record and the provider logo.

3. Records in discovery services should include a link back to the source provider's platform if supplied by the provider. It is preferable that these links be proxied so that users are authenticated upon reaching the provider's platform.

3.3.9 Ranking Algorithm

Discovery Service Providers should explain the fundamentals of how metadata is generally utilized within the relevance algorithm (mapping metadata to indexes, weighting of indexes, etc.) and how it enhances discoverability.

3.3.10 Use of Content Provider Metadata

Discovery service providers should utilize the core metadata and underlying full-text/original content for complete offerings provided by content providers, for the purposes of indexing to meet licensed customers' and authenticated end users' needs.

3.3.11 Support tools

Discovery Service Providers should create clear channels for reporting support issues. It is preferable that these tools be available online.

3.4 Best Practices for Libraries

3.4.1 System Maintenance

3.4.1.1 General System Configuration

1. Libraries should designate individuals to be responsible for configuring the Discovery Service, link resolver, proxy server, and other discovery-related tools for subscribed content from Content Providers. Designated point persons should coordinate with each other.

2. Libraries should enable linking, ranking, and authentication options available in the Discovery Service Tool for subscribed content from Content Providers, and seek clarification as needed.

3. Libraries should document configuration decisions and the reasons behind these decisions so future staff will understand why the discovery system is configured in a particular way.

4. Libraries should regularly evaluate and update Discovery Service configurations.

3.4.1.2 Discovery System Upgrades

1. Libraries should plan for system upgrades and seek assistance from vendors if needed.

2. Libraries should plan for and document discovery tool changes and seek assistance from vendors as needed.

3.4.1.3 Content Activations

Libraries should select the correct database collections available in the Discovery Service Tool for subscribed content from Content Providers and seek clarification from the Discovery Service Provider and Content Provider when needed.

3.4.1.4 Relevancy Configuration

Libraries should work with the Discovery Service Provider to learn exactly what, if any, adjustments can be made to result ranking, boosting, and weighting, and seek clarification as needed.

3.4.1.5 Linking Configuration

1. Libraries should select the correct holdings packages available in link resolver tools for the subscribed content from Content Providers and seek clarification from the link resolver providers and Content Providers when needed.

2. Libraries should regularly evaluate and update their configurations of link resolver tools for subscribed content.

3.4.1.6 Authentication/Authorization Configuration

Libraries should configure the proxy server for discovery service, link resolver, and ODI links, and seek clarification when needed.

3.4.2 Library Advocacy

3.4.2.1 Support Services

Libraries should work with vendors to create detailed agreements on what support the provider can offer during and after implementation.

3.4.2.2 ODI Conformance

1. Libraries should regularly check ODI Conformance Statements from Content Providers and Discovery Service Providers and publisher-provided configuration guides for libraries and evaluate the conformance statements made by vendors.

2. Libraries and user communities should advocate with vendors to increase their conformance with ODI Conformance Statements for Content Providers and Discovery Service Providers, encouraging providers to release conformance statements and engaging in conversation about conformance statements that do not align with library experiences.

3.4.2.3 Content Improvements

Libraries should use the vendor-provided tools to report missing, incomplete, or incorrect database collections, first to the Discovery Service Provider, and next to the Content Provider.

3.4.3   Training and Communication

3.4.3.1   Training

1. Libraries should establish staff training programs for one or more staff members on managing the Discovery Service, link resolver, proxy, and other discovery-related tools.

2. Libraries should provide regular training sessions to patrons and staff on how to use Discovery Service Platforms.

3. Library staff should read the official documentation, announcements, and updates, and make relevant adjustments in configurations.

4. Library staff should join the online communities (listservs and forums) for the library service providers, engage in dialogue, offer ideas and suggestions, and share innovative approaches created or implemented.

3.4.3.2   Communication

Library staff for licensing, electronic resources, systems and front-end services should regularly meet to discuss configuring Discovery Service Platforms, link resolvers, proxy servers, and other discovery-related tools for subscribed content.

Section 4: Recommended Next Steps

The ODI Standing Committee has identified a number of items that need to be addressed once the revised recommendations, outlined above, are in place.

Some items identified for further work in the original 2014 Recommended Practice did not rise in significance in reviewing discovery needs and have been removed from this updated version. Other items continue to be relevant and will be the focus of future work. The descriptions have been updated in this revision.

4.1   Collaborative Discussion

The ODI Standing Committee should continue in key activities:

- Monitor the discovery landscape and identify emerging needs among all ODI stakeholders.

- Provide educational opportunities to promote fundamental understanding of discovery platforms and adoption of the ODI recommended practices.

- Provide support for content providers and discovery service providers during adoption.

- Provide open dialog mechanisms to discuss ongoing issues related to all aspects of discovery platforms.

- Determine timing for next steps for ongoing work.

4.2   Application Programming Interfaces

Since the release of the original 2014 Recommended Practice, there has been a significant increase in the number of libraries using application programming interfaces (APIs) to access data managed in discovery services. Libraries need to have a clear understanding of the usage rights that come with discovery service records and, indeed, with individual data elements contained in a record. Some libraries (e.g., the University of Michigan, Villanova University, and North Carolina State University) are providing access to discovery services exclusively or primarily through APIs. Other libraries are considering ways to use citation information pulled dynamically from discovery services as feeds into other applications. There is no current standard around access rights to data other than a distinction between "authenticated user" and "unauthenticated user" attached to the entire record. When is it appropriate for a library to use selected fields retrieved from a discovery service outside the context of a discovery interface? And, where appropriate, what fields may be displayed to authenticated and unauthenticated users? These are some of the issues that should be discussed and potentially standardized.

4.3   Identifying the Source of a Record within a Discovery Service

The content provider and librarian surveys raised questions about how discovery services can provide attribution within individual discovery service records for the source(s) that created the discoverable record. ODI recommends follow-on exploration to understand the needs of content providers and libraries and the capabilities of discovery service providers, as a path to a future Recommended Practice.

Appendix A:
Content Provider Conformance Checklist

When requested by libraries, Content Providers can use the table below to indicate their ODI compliance. A "Y" (for Yes) in column 1 indicates compliance with the indicated paragraph of this Recommended Practice. A "P" response indicates partial compliance for which explanatory comments should be entered in the last column. An "N" (No) response indicates that the content provider does not comply with the recommendation. Explanatory comments may be added for any response.

Y/N/P	Recommendation	Paragraph	Comment
	Content Provider makes core metadata and underlying full-text/original content for complete offerings available to Discovery Service Providers.	3.2.1.1 (1)	
	Content Provider makes the core set of metadata elements and content item (full text, transcript, etc.) for each item submitted for indexing available to Discovery Service Providers.	3.2.1.1 (2)	
	Title	3.2.1.1 (2)	
	Authors	3.2.1.1 (2)	
	Author Identifier	3.2.1.1 (2)	
	Publisher Name	3.2.1.1 (2)	
	Volume	3.2.1.1 (2)	
	Issue	3.2.1.1 (2)	
	Page(s)	3.2.1.1 (2)	
	Date/Date Range	3.2.1.1 (2)	
	Item Identifier	3.2.1.1 (2)	
	Component of Title	3.2.1.1 (2)	
	Component of Title Identifier	3.2.1.1 (2)	
	Item URL	3.2.1.1 (2)	
	Open Access Designation	3.2.1.1 (2)	
	Content Type	3.2.1.1 (2)	
	Content Format	3.2.1.1 (2)	
	Language	3.2.1.1 (2)	
	Indexing Data	3.2.1.1 (2)	
	Full Text/Transcript	3.2.1.1 (2)	
	Abstract/Description	3.2.1.1 (2)	
	Content Provider provides libraries, on request, with a statement of participation in the discovery services, including disclosure of coverage depth and content depth.	3.2.2	

	Content Provider's agreements with Discovery Service Providers do not include NDAs.	3.2.3	
	The transfer of Content Provider's data to Discovery Service Providers makes use of existing standards where applicable and uses one of the metadata encoding schemes listed in 3.3.3.	3.2.4	
	The access and license indicator (NISO RP-22-2015 *Access and License Indicators*) is applied to all materials with an appropriate designation.	3.2.1.2	
	Content Provider provides data to support OpenURL resolution (ANSI/NISO Z39.88-2004 (R2010) *The OpenURL Framework for Context-Sensitive Services*). The resulting OpenURL should link to the item itself, or as close to it as possible in the content provider's site structure.	3.2.5	
	Content Provider provides clear channels for reporting discovery related support issues.	3.2.6	

Appendix B:
Discovery Service Provider Conformance Checklist

When requested by libraries, Discovery Service Providers can use the table below to indicate their ODI compliance. A "Y" (for Yes) in column 1 indicates compliance with the indicated paragraph of this Recommended Practice. A "P" response indicates partial compliance for which explanatory comments should be entered in the last column. An "N" (No) response indicates that the content provider does not comply with the recommendation. Explanatory comments may be added for any response.

Y/N/P	Recommendation	Reference	Comment
	Discovery Service provides collection-level content listing for library customers.	3.3.1.1	
	Provider	3.3.1.1	
	Market product	3.3.1.1	
	Titles in KB	3.3.1.1	
	Titles in Central Index	3.3.1.1	
	Number of unique records in Central Index	3.3.1.1	
	Percent of records full-text searchable in Central Index	3.3.1.1	
	Percent of records abstract searchable in Central Index	3.3.1.1	
	Percent of records subject searchable in Central Index	3.3.1.1	
	Percent of articles that are free to read	3.3.1.1	
	Date of last market product update	3.3.1.1	
	Date of last report update	3.3.1.1	
	Discovery Service provides title-level content listing for library customers.	3.3.1.2	
	Title	3.3.1.2	
	Standard identifier	3.3.1.2	
	Content format	3.3.1.2	
	Content type	3.3.1.2	
	Dates of coverage or date of publication	3.3.1.2	
	Number of records in Central Index	3.3.1.2	
	Percent of records full-text searchable in Central Index	3.3.1.2	
	Percent of records abstract searchable in Central Index	3.3.1.2	
	Percent of records subject searchable in Central Index	3.3.1.2	
	Percent of articles that are free to read	3.3.1.2	

Y/N/P	Recommendation	Reference	Comment
	Date of last report update	3.3.1.2	
	Distribution of metadata elements	3.3.1.3	
	Provide data in downloadable form	3.3.1.3 (1)	
	Provide data in tab-delimited text files	3.3.1.3 (2)	
	Use ProviderName_YYYY-MM-DD.csv naming convention	3.3.1.3 (3)	
	Provide monthly reports	3.3.1.3 (4)	
	Secure access to the reports	3.3.1.3 (5)	
	Linking	3.3.2	
	Discovery Service does not discriminate among Content Providers contributing to the service.	3.3.2 (1)	
	Mechanisms are offered to enable libraries to establish preferences regarding which platforms to present to users as link targets, and in what order or priority.	3.3.2 (2)	
	Discovery Service confirms non-bias with regard to content indexed and results presented to the user. A statement in this regard is published annually.	3.3.2 (3)	
	Discovery Service uses an algorithm that is non-preferential with regard to the Content Provider for generating result sets, relevance rankings, and link order.	3.3.2 (4)	
	Link presentation associated with a given result is configurable by libraries.	3.3.2 (5)	
	A seamless link (not necessarily a full-text link) is provided from the discovery screen(s) to the A&I interface(s) utilized in the Discovery Service index.	3.3.2 (6)	
	Information about link behavior changes is sent to content providers and libraries when materials changes that could impact the result set, relevance rankings, or link order of results.	3.3.2 (7)	
	File formats and methods of transfer	3.3.3	
	Provides description of capabilities, limitations, and preferences regarding how content providers should transfer data to them for the most effective indexing by the discovery service.	3.3.3 (1)	
	Communicates to content providers when format, schema, or transport mechanisms will have an impact on features or performance of the discovery service.	3.3.3 (2)	
	Uses mutually agreed-upon robust metadata encoding schemas.	3.3.3 (3)	

Y/N/P	Recommendation	Reference	Comment
	Uses mutually agreed-upon method and schedule for delivery of metadata.	3.3.3 (4)	
	Discovery Service provides COUNTER-compliant usage reports on schedule to all content providers.	3.3.4.1.1	
	Discovery Service includes a referrer URL identifying the discovery service when linking to content platforms or link resolvers.	3.3.4.1.2	
	Discovery Service provides usage metrics to libraries.	3.3.4.2.	
	Report on total number of searches per month.	3.3.4.2 (1)	
	Report on total number of unique visitors per month.	3.3.4.2 (2)	
	Report on total number of click-throughs per month.	3.3.4.2 (3)	
	Report top 500 search queries for the last period.	3.3.4.2 (4)	
	Report top 100 referring URLs to the Discovery Service for the last period.	3.3.4.2 (5)	
	Discovery Service displays the access and license indicator for all content to which it has been applied.	3.3.5	
	Discovery Service provides mechanism to only allow certain databases to be searchable by authenticated users.	3.3.6 (1)	
	Discovery Service takes measures to ensure that certain databases are only activated by mutual subscribers if desired by the content provider.	3.3.6 (2)	
	Discovery Service alternative coverage lists clearly indicate the metadata fields indexed are alternatives to specialized data, and note potential differences in quality, depth, and currency.	3.3.7	
	Discovery Service should be able to exclude records from merged or grouped records if desired by the content provider.	3.3.8 (1)	
	Discovery Service should be able to display the source provider of the record and their logo.	3.3.8 (2)	
	Records in Discovery Service should include a link back to the source provider's platform if supplied by the provider.	3.3.8 (3)	
	Discovery Service makes available documentation that describes how metadata is generally utilized within the relevance algorithm and how it enhances discoverability.	3.3.9	
	Discovery Service Provider provides clear channels for reporting support issues.	3.3.11	

Appendix C:
Library Conformance Checklist

Libraries can use the table below to indicate their ODI compliance. A "Y" (for Yes) in column 1 indicates compliance with the indicated paragraph of this Recommended Practice. A "P" response indicates partial compliance for which explanatory comments should be entered in the last column. An "N" (No) response indicates that the content provider does not comply with the recommendation. Explanatory comments may be added for any response.

Y/P/N	Recommendation	Reference	Comment
	Library designates individuals to be responsible for configuring Discovery Service, link resolver, proxy, and other discovery-related tools for subscribed content from Content Providers. Designated point persons should work or coordinate with each other.	3.4.1.1 (1)	
	Library enables linking, ranking, and authentication options available in the Discovery Service Tool for subscribed content from Content Providers, and seeks clarification if needed.	3.4.1.1 (2)	
	Library documents configuration decisions and the reasons behind these decisions.	3.4.1.1 (3)	
	Library regularly evaluates and updates its configurations of Discovery Service configurations.	3.4.1.1 (4)	
	Library plans for system upgrades and seeks assistance from vendors if needed.	3.4.1.2 (1)	
	Library plans and documents for discovery tool changes and seeks assistance from vendors if needed.	3.4.1.2 (2)	
	Library selects the correct database collections available in the Discovery Service tool for subscribed content from Content Provider and seeks clarification from the Discovery Service Provider and Content Provider when needed.	3.4.1.3	
	Library works with the Discovery Service Provider to learn exactly what, if any, adjustments can be made to result ranking, boosting, and weighting, and seeks clarification if needed.	3.4.1.4	
	Library selects the correct holdings packages available in link resolver for the subscribed content from Content Providers and seeks clarification from the link resolver provider and Content Provider when needed.	3.4.1.5 (1)	
	Library regularly evaluates and updates its configurations of library link resolver configurations.	3.4.1.5 (2)	
	Library configures the proxy server for discovery service, link resolver and ODI links, and seeks clarification if needed.	3.4.1.6	
	Library works with vendors to create detailed agreements on what support the provider offers during and after implementation.	3.4.2.1	
	Library regularly checks and evaluates ODI Conformance Statements from Content Providers and Discovery Service Providers and publisher-provided configuration guides for libraries.	3.4.2.2 (1)	

41

	Library advocates increasing ODI conformance for Content Providers and Discovery Service Providers.	3.4.2.2 (2)	
	Library uses the vendor-provided tools to report missing, incomplete, or incorrect database collections, first to the Discovery Service Provider, and next to the Content Provider.	3.4.2.3	
	Library developed and executes a staff training program for one or more staff members on managing the Discovery Service, link resolver, proxy, and other discovery-related tools.	3.4.3.1 (1)	
	Library provides regular training sessions to patrons and staff on how to use Discovery Service tool.	3.4.3.1 (2)	
	Library staff read official documentation, announcements, and updates, and make relevant adjustments in configurations.	3.4.3.1 (3)	
	Library staff join the online communities (listservs and forums) for the library service providers, engage in dialogue, offer ideas and suggestions, and share innovative approaches created or implemented.	3.4.3.1 (4)	
	Library staff on licensing, electronic resources, systems and front service regularly meet to discuss configuring Discovery Service, link resolver, proxy server, and other discovery-related tools for subscribed content.	3.4.3.2	

Appendix D:
2012 General Landscape Survey Summary

During September and October 2012, the ODI Working Group surveyed librarians, content providers, and discovery service providers to learn more about the current state of satisfaction with the new research tools and to measure the value of various requirements in cross-sector practice. The survey addressed current levels of scholarly metadata delivery/indexing, technical successes/opportunities in these data exchanges, and potential benefits of greater development/distribution of discovery tool usage data.

More than 2,000 individuals were contacted by ODI Working Group members and NISO staff in order to recruit participants for this survey (drawing on NISO's *Newsline* email list and registrants of 17 industry mailing lists). Cross-sector contributors were targeted for participation, drawing primarily on those based in the United States and the United Kingdom.

A total of 871 completed survey responses were logged: 782 identified as librarians, 74 identified as publishers, and 15 identified as discovery service providers. The ODI believes these numbers reflect an active engagement by libraries, which demonstrates the value of these new tools to library patrons and to the future of library services as a whole. However, the total number of responses relative to the global library population was not evaluated.

Survey findings and analysis are provided in greater detail in the previously published report, available at: https://groups.niso.org/apps/group_public/download.php/9977/NISO ODI Survey Report Final.pdf.

Appendix E:
2019 Content Provider Survey Summary

Open Access

Fourteen content providers responded to the survey. Not all respondents answered every question.

Do you provide open access content to discovery service providers?
Of the 14 content providers who answered this question, 12 said "Yes" and two said "No."

How do you indicate content is open access?
In the eight responses to this question, a variety of methods were provided. Three used KBART, while one said simply "in meta data." One used the "free to read" tag. Three others conveyed open-access status in different ways at the item level.

Have you enabled the open access "free to read" indicator recommended by NISO RP-22-2015, Access and License Indicators?
Of the eight content providers who answered this question, three indicated "yes" and five "no."

Do you have any observations about or suggestions for improving the open access indicator, as described in NISO RP-22-2015, Access and License Indicators?
Of the eight content providers who answered this question, five said they had no suggestions. One of the remaining three noted, "It is a big deal to change our XML style for the open access tag, from F to *[sic]*. Is there any evidence or study that shows complying with ALI will improve the discovery and linking of open access publications and articles?"

Are there differences in the way you tag "open access" vs. "public domain" content?
Of the eight content providers who answered this question, seven responded "No." Of these seven, three did not publish public domain content and felt that the question was not applicable. The one provider who did distinguish between open access and public domain content noted, "Yes, in that 'public domain' might apply to US Government, Crown Copyright, etc. licenses. We tag CC-x and these other license types specifically in our data."

Discovery System Use of Metadata

What metadata elements are not consistently consumed by discovery service providers that you would like to see made available in discovery systems?
The response rate for content providers was low, with only 10 providers answering questions. When asked which metadata elements are not consistently consumed by discovery service providers, we were not able to determine a pattern of issues. Two respondents noted that full text is not always consumed. Others noted that identifiers for records (OCLC numbers) or authors (ORCID/VIAF URLs) were not ingested. It was also noted that there are sometimes difficulties in providing nontraditional material types (e.g., archival material) for use by discovery systems. Transparency of what is consumed and how links to content are created were reported as issues as well.

Options for Linking to Full Text

The response rate for content providers was low, with only 10 providers answering questions.

Please rank your preferred method for inbound linking to your content, and note why for each choice. Note, this is what you would like to see, not what currently may be in place in your systems.
The responses for ranking of linking methods did not show a distinct preference for any of the linking methods.

	RANKING			
	1	**2**	**3**	**4**
OpenURL	3	4	3	0
Precalculated link	4	1	3	0
DOI/Crossref	3	5	1	0
Other (please describe below)	0	0	0	2

Other method for linking, if any:
No other method for linking was noted.

What tools/data do you find useful for troubleshooting when linking fails?
Content provider respondents reported using metadata points to analyze customer reports of linking failure, replicating the user experience to pinpoint the failure, either ad hoc or through a formal checklist process. One noted use of proactive tools prior to providing data to reduce occurrence of data issues.

What linking data element(s) do you provide to discovery service providers/link resolver vendors?
Responses for data elements provided largely noted KBART is provided, with a few noting direct links or data supporting direct links is provided. A few providers noted preference for direct links over OpenURLs due to success rate with OpenURLs.

Do you have other thoughts or comments about full-text linking?
There were no overarching themes in the responses, but truly unique IDs for e-content, vendor transparency, more use of direct links to native hosting platforms (rather than to alternative sources of full text, abstract, or indexing services), and less reliance on potentially faulty OpenURLs or DOIs were noted.

Usage Statistics

Does your organization currently receive usage statistics reports that specify usage for your content access via a vendor's discovery platform?
Nine of the respondents answered this question: three answered "Yes," three answered "Not Sure," and three answered "No." Most of the respondents who answered "No" or "Not Sure" did not answer any of the other questions.

What information can you share regarding why usage reports are not available for your content via the discovery service? (Only respondents who answered "No" to the first question)

Two of the respondents answered the question.

One representative who responded "No" to the previous question noted they have agreements with all the main DSPs and that they have asked on many occasions and have been told that the other DSPs do not provide usage metrics.

Another representative who also responded "No" later noted they felt that internally they had a

knowledge gap related to the larger DSPs.

From which discovery service vendors do you receive usage statistics? (Only respondents who answered "Yes" to the first question)
Two responded "EBSCO Discovery Service (EDS)," two responded "Ex Libris Summon," two responded "Ex Libris Primo," and one responded "WorldCat Discovery."

What format(s) are provided (e.g., csv, pdf, Excel, txt)? (Only respondents who answered "Yes" to the first question)
One respondent whose content is available through Ex Libris Primo noted that they receive PDFs. The other two respondents indicated that they receive Excel files. (Between them, the respondents who receive Excel files have content available on all major platforms: Ex Libris/Primo, Summon, EBSCO Discovery Service, and WorldCat Discovery).

What format would you prefer to receive? (Only respondents who answered "Yes" to the first question)
All of the respondents indicated that they want to receive Excel, CSV, or text. Clearly, they want a file they can manipulate.

What data do you receive (e.g., number of searches, result clicks, full-text usage, click-throughs and referrals, article-level usage)? (Only respondents who answered "Yes" to the first question)
Two responded "result clicks," two responded "full-text usage," one responded "article-level usage," one responded "number of searches," and one responded "click-throughs and referrals."

One respondent shared the concern that "each vendor provides different metrics and dimensions in the usage reports which they provide." Instead of indicating what they receive, the respondent used this opportunity to highlight that vendors might use the same terminology, but the basis for the count is different. Later, the respondent further noted that this limits the organization's ability to use the information.

What additional data do you want to receive? (Only respondents who answered "Yes" to the first question)
The following metrics and metadata elements were mentioned:
- Content counts – Knowing the number of articles/chapters indexed is helpful in maintaining currency and assessing the need for re-indexing.

- Standard Unique Identifier for Libraries – This respondent wants a unique identifier that can be used to efficiently link the usage data to customer information within the organization's business systems.

Do you have specific feedback regarding the usage reports you currently receive? (Only respondents who answered "Yes" to the first question)
The respondent who receives PDF-formatted reports expressed concerns that they cannot use the information given in the format they receive. The other two respondents essentially shared the same concern that the lack of standardization makes it difficult to combine the information from different vendors to compare and analyze the information effectively.

Do the usage statistics you receive add value to your relationship with the discovery service provider? (Only respondents who answered "Yes" to the first question)
The two respondents who receive Excel files answered "Yes." The respondent who receives PDFs answered "No."

Do you recommend any enhancements to the usage statistics you receive? (Only respondents who

answered "Yes" to the previous question)
Two responded "to provide standard format (structure/columns)," one responded "to provide consistent metrics/dimensions across discovery vendors," one responded "to include a primary key (customer/library) that the content provider can use to connect data," and one responded "to remove trial/demo data from the reports that are provided."

How could the reports be improved to add value? (Only respondents who answered "No" to the previous question)
The respondent who receives PDFs answered that a format for reuse and analysis would add value.

Source of Discovery Content

Nine content providers responded to these questions; none of the respondents provide A&I collections.

In discovery systems that merge records, please indicate the importance of understanding where the elements came from.
Four responded "very important," four responded "important," and one responded "not important."

In discovery systems that display single records, please indicate the importance of understanding why the single record displayed was chosen.
Five responded "very important" (including the provider who responded "not important" to the previous question), and four responded "important."

Please indicate the importance of identifying the source of a record/metadata element on a result list in discovery.
Four responded "very important," four responded "important," and one responded "not important;" the "not important" response was from a consultant (not a provider of content).

In discovery systems that group/cluster records, please indicate the importance of understanding why the main record displayed was chosen.
Five responded "very important," and four responded "important."

In discovery systems that group/cluster records, please indicate the importance of understanding how the records are ordered within a group/cluster of records.
Five responded "very important," three responded "important," and one responded "not important."

Please indicate the importance of understanding how usage statistics are attributed to the sources of metadata (e.g., if one click attributes a "usage" for all of the associated suppliers of the record rather than more discrete credit being given by match of actual search terms to metadata elements provided by each supplier).
Four responded "very important," and five responded "important."

If you have any comments regarding the previous set of questions and the importance of any area, please share them.
Three content providers responded to this question, all pointing to knowledge gaps. Two mentioned not knowing which discovery services merge or cluster, and one would like more transparency around if and how their records are enriched.

If the record displayed as the "main" record is from your product (and the record ranking resulted from use of your company's metadata), is it acceptable to you that records from other products

may be grouped with it?
Eight responded "yes," and one responded "no."

Role of Libraries in Configuration

Valid responses from 11 content providers were received. Some respondents did not fill out answers for all the questions.

What types of discovery and linking related support questions do you receive from library customers and how often?
Six respondents mentioned linking issues (broken DOIs, OpenURL problems, missing links, incorrect links); five mentioned KB issues (missing items, collections not up to date); four mentioned metadata issues (missing or incorrect); two mentioned missing content or confusing collection names; and one mentioned declines in usage. The frequency ranged from daily to weekly.

What are the most common support issues that libraries could fix themselves?
The most common issues were regarding activating the correct content, with three content providers mentioning selecting the correct targets/holdings in the KB, and two mentioning the same for the discovery index. Two respondents mentioned issues with proxying links, one mentioned enabling multiple linking options, and one mentioned issues with reporting issues.

Are there ways in which libraries could be a more proactive partner or communicate better with discovery services?
Three mentioned providing specific, verifiable, and technical details when reporting problems. Two mentioned contacting the discovery service provider about discovery/linking problems and then letting the discovery service contact the content provider (one also suggested contacting both the discovery service and the content provider). One suggested that librarians check discovery service and content provider sites for best practice guides on how to configure and one suggested that libraries provide more clarity for who is responsible for the different parts of the process.

What are the most common support issues that are caused by misunderstandings about configuration options or how published content is described?
The main issue was the naming of collections. Two cited this issue for the discovery indexes, and two mentioned it for the KB.

What do you feel libraries' responsibilities vis-à-vis discovery services and content providers should be?
Several themes emerged from the responses to this question. The most common area of focus was configuration of linking and the central index. Three respondents mentioned selecting the correct collections in the discovery index; three mentioned the same issue in the link resolver; two mentioned enabling linking, ranking, and authentication options; two mentioned regularly evaluating and updating configurations in the discovery index and link resolver; and one mentioned checking NISO's ODI, KBART, ALI websites, vendor sites, and publisher sites for configuration best practices.

Staff best practices were also frequently mentioned. Three respondents suggested that libraries designate individual(s) to be responsible for configuring Discovery Service and link resolver tools; two suggested providing regular training sessions to patrons and staff on how to use the discovery service; and one suggested that licensing, electronic resources, systems, and service desk staff regularly meet to discuss configuring the discovery service and link resolver.

Are there best practices libraries could be doing that would create a better user experience? If so, what?
The suggestions included:

- Staff training, which would go to training in the discovery system, workshops, coding/programming, and user testing.

- Improving discovery interface based on patron feedback.

- Reporting issues to the discovery service before reporting to the content provider and then approaching the content provider with a ticket number.

- Performing audits to make sure content is represented in their discovery service.

Appendix F:
2019 Discovery Service Provider Survey Summary

Note: We received responses from two organizations responsible for three of the discovery service platforms currently marketed to libraries.

Discovery System Use of Metadata

What metadata elements are not consistently provided by content providers that you would like to consume in your discovery system?
Both discovery service providers who responded to this question cited consistency in how metadata is tagged across content providers as an area of difficulty. Both also mentioned they would like to see open access indicators and ORCIDs being used more commonly.

One provider cited that book data is particularly problematic. Even consistency across whether they are called "books" or "e-books" would be helpful. They would also like clear and consistent metadata with mapping between chapter and title-level records.

In general, more clear relationships between records would be helpful. One discovery service provider mentioned that more references and connections between articles and research data would be useful.

One discovery service provider also mentioned a lack of inclusion of industry-accepted identifiers to be problematic, i.e., ISXN, DOI, etc.

Options for Linking to Full Text

How do you link to content (e.g., OpenURL, precalculated link, DOI, etc.) from full-text and A&I providers? Please note the record characteristics that determine which type of link is used.
We received responses from two discovery service providers responsible for three of the discovery service platforms currently marketed to libraries. They both noted using a variety of linking mechanisms, including OpenURL, precalculated links, and DOIs, preferring the most stable linking service for each content provider. One also noted offering proprietary linking methods.

When records are grouped or clustered, is a link to full text available in every record display? What kinds of links are used in these circumstances?
One discovery service provider respondent does not group or cluster records, so the question was not applicable. The other noted that libraries determine the best available linking method for records in a cluster (each record could have its own link) through settings in the link resolver. For merged records, a variety of links are displayed.

Usage Statistics

Does your organization currently provide usage statistics reports to content providers with information/data specific to usage generated via your discovery layer interface?
We received responses from two organizations responsible for three of the discovery service platforms currently marketed to libraries. The answer was "yes" for all three platforms.

If Yes:

What data elements are included in the reports?
The survey instrument did not provide a pick list, which makes comparing the responses challenging. In some cases, respondents did not include data we can reasonably assume that they all provide (e.g., the customer or accessing institution name). Also, the respondents used their internal language to describe the data elements. This table attempts to normalize the responses, with notes whenever possible to highlight distinctions in the responses.

Data Element	# of platforms that report on this metric
Sessions	1 of 3
Searches	3 of 3
Turnaways	1 of 3
Record views (includes: PDF full-text requests, HTML full-text requests, page view count, clicks on the vendor's record, and abstract views)	3 of 3
Various "linking" requests (smart linking to/from, custom links, links to the provider's platform)	2 of 3
Times cited (for Web of Science & Scopus data)	1 of 3
Recommended resources count	1 of 3
Recommended topics count	1 of 3
Access institution market	1 of 3

Notes:

Record views: One provider indicated that they provide as many as three distinct data points related to record views (PDF full text, HTML full text, and abstract views). It was not clear from the other responses if those providers give this level of detail.

Various "linking" requests: One provider breaks down linking requests into the various types available on their platform; i.e., it is more than one data point.

What format(s) are provided (e.g., csv, pdf, Excel, txt)?
One provider offers HTML, csv, and tsv; the other provides PDF and Excel.

Do the reports conform with ODI's Recommended Metrics Provided to Content Providers (RP 3.3.4.1.1)?
One provider indicated that they conform with the current Recommended Practice. The other vendor indicated that they partially conform for both supported platforms.

If you answered Yes: Please provide any additional feedback regarding the reports. For example: Do the reports provide more data than the ODI Recommended Practice? If so, why? If you answered No: Why not?
There was no additional feedback for this question.

If you answered No to the previous question: If your organization is either in the process of complying or considering complying with the Recommended Practice, how long would it take to complete a project to make these reports available?
The vendor that indicated partial compliance is working on full compliance for one of their platforms.

If No:
If not, why?

What would influence your organization to provide discovery layer usage reports for content providers? How long would that project take?
There was no additional feedback for these questions since all respondents answered that they already provide reports.

Source of Discovery Content

How does your discovery service handle duplicate records that come from multiple sources?
All three discovery services handle duplicate records differently. One discovery service merges metadata into one record. One groups/clusters the records, and one displays a single record.

If Groups/Clusters:

What is the deciding factor in determining the source that is displayed or featured in cases where the same record is available from multiple sources?
The discovery service that groups/clusters indicated that the record that is displayed is selected randomly. The discovery service that displays a single record also responded, mentioning that the record that ranked the highest against the relevance algorithm is the one that is displayed.

Is the source of a record displayed on the result list?
This question is not applicable to the discovery service that merges records. The other two do display the source of the record on the result list.

In the event that your discovery service has the same record covered from multiple sources, how is one source selected to be displayed?
This question is not applicable to the discovery service that merges records. The discovery service that groups/clusters indicated that the record that is displayed is selected randomly. The discovery service that displays a single record responded that the record that ranked the highest against the relevance algorithm is the one that is displayed.

How do you determine which record is featured and which are shown as secondary options?
This question is not applicable to the discovery service that merges records. The discovery service that groups/clusters indicated that the record that is displayed is selected randomly. The discovery service that displays a single record responded that the record that ranked the highest against the relevance algorithm is the one that is displayed.

In the event of a group record where multiple sources are associated to the featured record, how is

the order of the display of associated records determined?
This question is not applicable to the discovery service that merges records. The discovery service that groups/clusters indicated that the record that is displayed is selected randomly. The discovery service that displays a single record responded that the record that ranked the highest against the relevance algorithm is the one that is displayed.

How is usage attributed (statistics) to providers in a group record scenario when multiple sources are visible?
This question is not applicable to the discovery service that merges records nor to the discovery service that only displays a single record. The third discovery service responded that "usage is counted based on clicks on records, not on the search result."

If merged record:

Is the source of a record displayed on the result list?
Only the discovery service that provides merged records responded, stating that if the content provider supplies metadata for the source field, the source will be displayed.

Are all the sources of the metadata in the record displayed in the record display? If so, how? If not, how do you choose which to display?
The discovery service that provides merged records responded that not all sources are displayed, despite using multiple sources in merged record.

How are usage statistics attributed as it relates to a merged record (when the user clicks on record, or launches to the full text via that record)?
All discovery services responded that this question is not applicable.

If Displays a single record:

What is the deciding factor in determining the source that is displayed or featured in cases where the same record is available from multiple sources?
The discovery service that displays a single record responded that they select the record that ranked highest against the relevance ranking algorithm.

In the event that your discovery service has the same record covered from multiple sources, how is one source selected to be displayed?
The discovery service that displays a single record responded that they select the record that ranked highest against the relevance ranking algorithm.

Role of Libraries in Configuration

What types of discovery and linking related support questions do you receive from library customers and how often?
Content-related questions include general content inquiries, title coverage, and admin support (how to add content to the account so that it appears in the search), as well as issues regarding metadata, flagging, rights, and availability.

Linking-related questions also involve setup and maintenance issues; specifically, broken links, customizations, and preferences.

Frequency: One indicated that they receive support questions daily; the other respondent did not mention a frequency.

What are the most common support issues that libraries could fix themselves?
Although described differently by both providers, thematically, mismatches in holdings data seemed to be a core area where libraries could actively work to resolve support issues for themselves. This mismatch results in inaccurate linking. The Discovery Service Provider's knowledge base needs to be correct and current, and the library needs to activate the correct collection or title coverage information. The title coverage information in a library's account needs to match their subscription/purchase. Both providers support approaches to automated holdings updates to help mitigate this issue. One mentioned a solution targeted to updating holdings for customers; the other, a solution to update the publisher information.

Are there ways in which libraries could be a more proactive partner or communicate better with discovery services?
Both respondents advocated for libraries to become actively engaged in their respective online communities (e.g., working groups, listservs, and forums) to discuss best practices and troubleshoot common issues. Other libraries can provide meaningful support related to issues they have already addressed or innovative approaches they have implemented.

Both indicated that these communities provide a valuable feedback loop to the Discovery Service Provider.

One advocated for institutions that self-publish to register the content for DOIs to support discoverability across platforms.

What are the most common support issues that are caused by misunderstandings about configuration options or how published content is described?
Both Discovery Service Providers offer customers various types of help documentation related to the customization options available within the admin console. They also both acknowledged that the configuration options can be complicated. Libraries sometimes contact support because it is easier to have someone work directly with them, as opposed to watching training videos. Libraries also call because they were unsuccessful because the configuration process involved multiple steps and they missed some of the steps.

One vendor indicated that they appreciated the complexity of the tools and that they are embarking on an initiative to better understand the user experience, which will result in the redesign of the admin console to create a more intuitive user experience.

What do you feel libraries' responsibilities vis-à-vis discovery services should be?
Three themes emerged from the responses to this question:

- Libraries should actively engage with their user community. Both vendors have active user communities that help promote open dialog, resolve issues, and promote new features. Both vendors advocated for libraries to be engaged in their respective user communities.

- Libraries should maintain the basics of their discovery layer, which includes reading documentation, providing clear examples when raising a support issue, and engaging respectfully with support staff.

- Discovery service providers should provide online and in-person support.

Are there best practices libraries could be doing that would create a better user experience? If so, what?

Both providers stressed the importance of understanding the user's behavior. They recommended regular activities aimed at studying and understanding users' activities, search behaviors, expectations, and specific experiences to determine if they are having a successful experience. The goal should be to optimize the experience for the user; the respondents mentioned pain points such as authentication, number of clicks, and unsuccessful links as reasons why libraries can lose users. Both put studying and learning from users as the most important practice to help create a better experience.

Appendix G:
2019 Library Survey Summary

Discovery Metadata

Section 3.2.1 of the Recommended Practice includes a list of core metadata and enriched content that should be provided to discovery services. Core metadata includes: Title, Authors, Publisher Name, Volume, Issues, Page, Date/Date Range, Item Identifier, Component of Title, Component of Title Identifier, Item URL, Open Access Designation, Full Text Flag, Content Type, Content Format. Enriched Content includes: Indexing Data, Full Text/Transcript, and Abstract/Description.

What additional metadata elements do you feel are not fully represented in discovery indexes?
Several metadata elements were overwhelmingly noted as not represented in discovery services in the 61 library responses received.

1. Subject (20%)

2. DOI (18%)

3. Author Identifier (11%)

4. Language (10%)

5. Edition (10%)

Most of the elements were already included in the Recommended Practice, making their absence an issue of compliance with Section 3.2.1. Two elements were not in the current Recommended Practice: Author Identifier and Language.

Open Access

Are you satisfied with the amount of expected open access content your users can find in your discovery tool?
The majority of libraries (74 percent) stated that they were either not satisfied with the amount of open access content they could find in their discovery layer or that they were not sure. We conclude from this that more open access content needs to be available, and it needs to be more obvious to libraries that it is available. It also seems likely that at least some librarians are not sufficiently aware of what open access content is, in fact, available in their discovery systems.

Discovery System Use of Metadata

How important is it to identify the source of a record/metadata element on a result list in discovery?
Thirty-eight responded "very important" (3), 40 responded "important" (2), and 10 responded "not important" (1). The average response to this question was about 2.3, which is somewhere between "important" (2) and "very important" (3).

How does your discovery service handle duplicate records that come from multiple sources?
The majority of respondents use Ex Libris Primo, which groups or clusters multiple records behind a single record that contains a link to view the other records. The preponderance of Primo users is reflected in the relatively large number of respondents (53) who report that their discovery system groups or

clusters records. It is worth pointing out that some respondents with EBSCO EDS reported that their records are merged when in fact they are not. Similarly, some Ex Libris Primo users reported the merging of records that are in fact clustered. This indicates a lack of understanding regarding how discovery services handle duplicate records.

How important is it to understand where the metadata elements came from?
Twenty-two librarians answered this question, with nine stating it was "very important" (3), 11 stating it was "important" (2), and two stating that it was "not important" (1). This averages out to 2.3, which is closer to "important" than "very important."

How important is it to understand why the single record displayed was chosen?
Only 14 respondents replied to this question, and with one exception, all answered "important" (2) or "very important" (3). This averages out to a score of 2.15, again closer to "important" (2) than "very important" (3).

How important is it to understand why the main record displayed was chosen?
Fifty-two responded to this question, three of whom said it was "not important" (1). Twenty-two felt this was "important" (2), while 27 felt this was "very important" (3). The average score for this question was 2.46, which is between "important" (2) and "very important" (3).

How important is it to understand how the records are ordered within a group/cluster of records?
There were 52 responses to this question, with an average score of 2.2 ("important"). Eight respondents felt that this was not an important consideration (1), 25 felt this was "important" (2), and 19 felt that this was "very important" (3).

How important is it to understand how usage statistics are attributed to the sources of metadata? (E.g., if one click attributes a "usage" for all of the associated suppliers of the record, rather than more discrete credit given by match of actual search terms to metadata elements provided by each supplier.)
Nine respondents felt that the question of usage statistics was "not important" (1), 25 found this "important" (2), and 49 stated that this was "very important" (3). The average score for this question was about 2.5.

Options for Linking to Full Text

Please rank your preferred mechanism for linking from your discovery system, and note why for each choice. Note, this is what you would like to see, not what currently may be in place in your systems.

Library respondents prioritized preferred mechanisms for linking. The top two choices were OpenURL and DOI/Crossref. This makes sense when viewed in the larger discovery environment; many users are following a path to content outside the discovery systems, with OpenURL resolvers being the primary tool for linking across all institutions, regardless of discovery platform.

Kind of Link	Ranking				
	1	2	3	4	No Response
OpenURL	32	30	15	1	30
Pre-calculated Link	17	20	39	1	31
DOI/Crossref	28	28	20	3	29
Other	1	1	3	52	51

Other method for linking, if any:
Eight respondents noted a preference for direct links to resources to avoid link breakdown. Three noted that systems should link to records in the ILS. Three noted that links should deliver content (PDF, etc.) in as few clicks as possible.

Does link success (people getting to the correct content) outweigh resolution choice (preferred content platform)?
Ninety-three percent of library respondents did note, however, that link success (getting users to the correct content) is currently more important than guiding users to content on a preferred platform.

What tools/data do you find useful for troubleshooting when linking fails?
When investigating linking failures, 56 percent of respondents report that analysis of the underlying metadata and user path are the key tools for resolving linking issues. 27 percent of the respondents noted that searching for a different link to the content in other systems was useful.

What tools (browser extensions, bookmarklets, etc.) do you provide for linking your users to full-text content outside your main discovery interface?
The largest number of respondents noted continued use of URL resolvers to link users to full-text content. Some reported use of browser extensions and bookmarklets.

Response	EDS	Primo	Summon	WorldCat	Other	Total
URL resolvers	9	31	9	7	3	**59**
Browser extensions	5	13	0	0	1	**19**
Bookmarklets	1	11	0	2	0	**14**
Other	4	8	1	0	0	**13**

A number of commercial services were noted as being useful for end users: BrowZine, LibKey, Google Scholar, Kopernio, PubMed Central, EZproxy, Scopus, and Web of Science.

Discovery System Optimization & Configuration

Please note the tasks that were necessary for initial implementation of the system related specifically to configurations of the discovery index such as collection activations, links from discovery index records, relevance ranking, and link resolver behavior. What configuration tasks did your library perform?
Overall, 21 percent of respondents reported that they engaged in a "wide range of configuration activities," 23 percent reported that their configuration efforts focused on the discovery index and knowledge base, five percent reported that the vendor did the configuration work, five percent reported that their configuration work focused on migration of data, two percent focused on interface configurations, and two percent reported a focus on the loading of records from various platforms. Importantly, 11 percent were unable to answer this question because staff who worked on the initial configuration had departed. Although it was not stated explicitly, it may be inferred that information these staff members had is gone. Regarding libraries reporting that their configuration work encompassed a wide range of activities, it should be noted that only those at the smallest institutions (under 1,000 FTE) did not indicate they had undertaken extensive configuration work. This might indicate that extensive configuration is the norm and that it is important to be prepared for this.

Related to the previous question, what assistance was provided from the discovery service provider?
Overall, eight percent of the respondents found the service they received to be "superior," 38 percent acknowledged that assistance was provided, 10 percent felt that the support they received was not sufficient, one percent stated they received no support whatsoever, and seven percent did not know because the staff present during implementation no longer worked there. It is clear from the responses that support is needed and appreciated. While only about 10 percent of the respondents indicated they did not receive enough support, it is not clear if those reporting they received support would have liked more. Additionally, the positive responses from those feeling they received excellent support is indicative of how much this support is appreciated.

Please note the tasks that are necessary to maintain the functioning of the library discovery system.
Overall, 19 percent of respondents reported that keeping up with system upgrades was key to maintaining the discovery system, 18 percent reported that knowledge base and link resolver configurations were key, 14 percent reported that work with the central index was important, 11 percent reported that key tasks revolved around harvesting and indexing, and one percent of respondents reported that this was "not their role."

Related to the previous question, what configurations will make the links function correctly?
It seems that the most important areas amongst the respondents are OpenURL configuration and EZproxy configuration. One thing that was mentioned repeatedly was the importance of good metadata from discovery service providers in their central indexes. The answers to this question were as follows: four percent reported knowledge base configuration, 23 percent reported OpenURL link resolver configuration, 16 percent reported proxy configuration, one percent reported that this was not part of their job, and somewhat concerningly, 12 percent answered but did not know what configurations would make links function correctly. This may indicate the need for libraries to become more familiar with how these systems operate.

What configurations are required to adjust ranking of results?
There appears to be some confusion about what is possible. While it appears that no DSP allows for much configuration of relevance ranking, there are a number of partial adjustments that can be made. That said, a number of institutions report no adjustments are possible at all, while others indicate that they can make some. Most of these adjustments related to weighting certain types of results and blending local results with the central discovery index. Perhaps this indicates a need for both library training and some clarity in documentation about exactly what adjustments can be made to search results.

When new packages are added to the central index, what procedures are in place to make sure that collections are indexed?
Overall, 32 percent of the respondents did report having a specific procedure to assure indexing of collections. About nine percent of respondents reported sending this task to another library department, where presumably some procedures were in place. However, 19 percent of respondents reported no procedures and another five percent either did not know or answered "not applicable." The remaining 34 percent did not answer this question. In some of these cases the respondents indicated that they relied on the discovery service provider to take care of this.

What other tasks are needed?
Respondents reported that the majority of additional tasks are related to testing new collections, verifying that they work, and reporting and correcting errors. The relatively large number of "no answer" responses to this question may indicate that most did not consider other tasks were needed. The results break down as follows: eight percent other tech services work, 17 percent verify new collection packages work, two percent account for collections not included in metadata, five percent do not know, three percent other

tasks, and 64 percent no answer.

How do you keep abreast of discovery system updates and confirm that new features are implemented correctly?
Overall, 49 percent of the respondents reported relying on vendor release notes and announcements. Disregarding the 33 percent who did not answer this question, the next largest percentage of respondents were those that rely on discussion with other library staff members, and these were only eight percent of the respondents. Based on the above, it seems that the release notes, listserv bulletins, documentation, and other official information from the vendors is far and away the most common means of keeping up with new features. Some of the respondents also mentioned getting feedback about problems, but the vendor communications seem most important.

Were there instances where the discovery system was not initially configured correctly, and you had to go back and make corrections?
Overall, 52 percent of respondents reported "yes," that there were instances where the initial configuration was not correct, while only 22 percent answered "no" to this question (26 percent did not respond at all). It is interesting to note that the largest institutions were most likely to answer "yes" to this question.

If you answered "yes" to the previous question, please describe any lessons learned from this.
The answers to this question were varied and fell into a number of categories as follows: Specific localized needs: three percent; Ranking: one percent; Documentation: three percent; Theoretical understanding: six percent; Central Index/KB: nine percent; Vendor interactions: 10 percent; Testing: eight percent; New features: three percent; No lessons reported: seven percent; No answer: 49 percent. The responses to this question were particularly difficult to categorize. The category "Theoretical understanding" contains answers that seemed to be getting at gaining a better holistic understanding of how the discovery system worked. The responses were mostly grouped in the following four categories: "Theoretical Understanding," "Central index/KB," "Vendor interactions," and "Testing." There seemed to be a fair amount of confusion in communications with vendors, and additionally, there were a number of tips regarding testing.

What are the various system configuration roles librarians are required to perform regarding the central index?
Of the 60 respondents answering that they did need "an electronic resources librarian [who] makes sure the knowledge base is updated to reflect new content," 23 also answered that they used "a public services librarian [who] makes sure content from the central index is appearing in [the] front end."

Is there training for library staff in configuration roles?
Overall, 40 percent answered that their institution did have training of some sort while 31 percent did not (28 percent did not answer the question).

Please describe any areas that are particularly challenging in relation to discovery system configuration.

This question resulted in a very large variety of answers, as many responses were often unique to a specific situation and subsequently hard to categorize. The most common areas reported as being challenging were issues related to the central index, issues related to the complexity of the back office/back end, and lack of sufficiently skilled staff to handle configurations. Here is a list of the different categories of answers with the percentage of respondents assigned to each:

- Normalization: 4%

- Relevancy/boosting/blending: 1%

- Back office: 12%

- Vendor training documentation: 2%

- Central index issues: 17%

- Black box/Not clear what/why is happening: 2%

- Authentication: 1%

- Lack of sufficient staff technical skills: 6%

- Community zone issues: 1%

- Configuration options not yet available: 3%

- Linking issues: 4%

- No Issues/Works well: 1%

- Open access: 1%

- Harvesting: 1%

- Deduplication: 1%

- Outside security requirements: 1%

- No answer: 41%

What do you think the librarian's roles and responsibilities should be regarding discovery system configuration?

There is no overall consensus regarding librarian roles and responsibilities. It is worth noting that there is some support for relying heavily on vendors in discovery system management. The importance of local customizations was also noted by several libraries, as was the importance of multiple library units working together to manage the discovery system. A number of libraries emphasized the importance of understanding all aspects of the discovery system.

Please indicate which "point person" assignments your library has created regarding discovery system configuration:

Overall, 60 percent of respondents reported that activation of electronic resources was a task assigned to the discovery system point person. Fifty-three percent reported that configuration of links in records was one of these tasks, 28 percent reported that work with relevance ranking was one of these tasks, and 12 percent listed "other" tasks.

Please describe how communication with discovery services vendor support is handled at your institution (e.g., opening tickets, emailing a staff member). Does this communication go through a liaison or do staff members in various roles submit directly to the vendor?
Overall, 15 percent of the respondents reported that their institution had one assigned person to handle all communication with the vendor. Thirty-two percent reported having two or more specific people that handled this communication, while 21 percent reported that non-specific staff members communicated with the vendor. For the latter, this seemed to mean that anyone could open communication with the vendor. One percent reported that consortium members handled this on a library-by-library basis, while 31 percent did not answer this question.

Bibliography

Asher, Andrew D., Lynda M. Duke, and Suzanne Wilson. "Paths of Discovery: Comparing the Search Effectiveness of EBSCO Discovery Service, Summon, Google Scholar, and Conventional Library Resources." *College & Research Libraries,* September 2013, 74 (5): 464-488. Available at: http://crl.acrl.org/content/74/5/464.full.pdf+html

Baldwin, Dee, Michael Kucsak, and Alice Eng. "Don't Touch that String! There Went the Databases." *Information Outlook*, September/October 2012, 16 (5): 24-32. Originally presented at the 2012 Special Libraries Association Annual Conference, Chicago, IL, July 16, 2012. Available at: http://digitalcommons.unf.edu/library_facpub/7/

COUNTER Code of Practice for e-Resources. Release 4. COUNTER, April 2012. Available at: http://www.projectcounter.org/code_practice.html

COUNTER Code of Practice for Usage Factors. Draft Release 1. COUNTER, March 2012. http://www.projectcounter.org/usage_factor.html

DCMI Usage Board. *DCMI Metadata Terms*. DCMI Recommendation. Dublin Core Metadata Initiative, June 14, 2012. Available at: http://dublincore.org/documents/dcmi-terms/ [Also published as ANSI/NISO Z39.85-2012, available at http://www.niso.org/standards/z39-85-2012.]

Digital Library Federation ILS Discovery Interface Task Group (ILS-DI) [webpage]. http://old.diglib.org/architectures/ilsdi/

Discovery Open Metadata Principles. London: JISC Discovery Programme, May 26, 2011. Available at: http://discovery.ac.uk/businesscase/principles/

Ekins, Andy, and Lukas Koster. Unified Resource Discovery Comparison [website]. http://sites.google.com/site/urd2comparison/

Encoded Archival Description (EAD) [website]. http://www.loc.gov/ead/

ICOLC. *Revised Statement on the Global Economic Crisis and Its Impact on Consortial Licenses*. International Coalition of Library Consortia originally issued January 19, 2009; reissued June 14, 2010. Available at: http://icolc.net/statement/revised-statement-global-economic-crisis-and-its-impact-consortial-licenses

Information Retrieval: Application Service Definition & Protocol Specification. ANSI/NISO Z39.50-2003 (R2009). Baltimore, MD: National Information Standards Organization, approved 2003; reaffirmed 2009. Available at: http://www.niso.org/standards/z39-50-2003/

Kelley, Michael, "Stakeholders Strive to Define Standards for Web-Scale Discovery Systems." *The Digital Shift*, October11, 2012. Available at: http://www.thedigitalshift.com/2012/10/discovery/coming-into-focus-web-scale-discovery-services-face-growing-need-for-best-practices/

KBART: Knowledge Bases And Related Tools. NISO RP-9-2010. Baltimore, MD: National Information Standards Organization, February 8, 2010. Available at: http://www.niso.org/publications/rp/rp-9-2010/

Knowledge Bases and Related Tools (KBART) Recommended Practice. Phase II revision. NISO RP-9-2014. Baltimore, MD: National Information Standards Organization, March 27, 2014. Available at: http://www.niso.org/publications/rp/rp-9-2014/

Knowledge Base And Related Tools (KBART) [webpage]. http://www.niso.org/workrooms/kbart

Luther, Judy, and Maureen Kelly. "The Next Generation of Discovery." *LibraryJournal.com*, March 15, 2011. Available at: http://www.libraryjournal.com/lj/home/889250-264/the_next_generation_of_discovery.html.csp

MARC Standards [website]. http://www.loc.gov/marc/

Metadata Encoding & Transmission Standards (METS) [website]. http://www.loc.gov/standards/mets/

Metadata Object Description Schema (MODS) [website]. http://www.loc.gov/standards/mods/

Music Library Association's Emerging Technologies and Services Committee. *Music Discovery Requirements*. Middleton, WI: Music Library Association, April 23, 2012. Available at: http://committees.musiclibraryassoc.org/uploads/ETSC/MDRdocument.pdf

NISO Metasearch Initiative [webpage]. http://www.niso.org/workrooms/mi

NISO Metasearch XML Gateway Implementers Guide. NISO RP-2006-02. Bethesda, MD: National Information Standards Organization, August 7, 2006. Available at: http://www.niso.org/publications/rp/RP-2006-02.pdf

ONIX family of standards [website]. http://www.editeur.org/8/ONIX/

The Open Archives Initiative Protocol for Metadata Harvesting. Version 2.0. Open Archives Initiative, June 14, 2002. Available at: http://www.openarchives.org/OAI/openarchivesprotocol.html

Postel, J., and J. Reynolds. *File Transfer Protocol*. RFC 959 aka STD9. Internet Engineering Task Force, October 1985. Available at: http://www.rfc-editor.org/rfc/rfc959.txt

Quint, Barbara. "The Undiscovered Discovery." *Information Today*, July/August 2010, 27 (7): 7-11. Available at: http://www.infotoday.com/IT/jul10/Quint.shtml

Recommended Practices: Discovery Services. Philadelphia, PA: National Federation of Advanced Information Services, August 30, 2013. Available at: http://info.nfais.org/info/Recommended_Practices_Final_Aug_2013.pdf

ResourceSync Framework Specification. NISO Z39.99-2014. National Information Standards Organization and Open Archives Initiative, April 21, 2014. Available at: http://www.niso.org/standards/z39-99-2014/

Rowe, Ronda. "Web-Scale Discovery: A Review of Summon, EBSCO Discovery Service, and World Cat Local." *The Charleston Advisor*, July 2010, 12 (1): 5-10. Available at: http://dx.doi.org/10.5260/chara.12.1.5

searchRetrieve. Version 1.0. OASIS Standard, January 30, 2013. Available at: http://www.loc.gov/standards/sru/oasis/

Somerville, Mary M., Barbara J. Schader, and John Sack. *Improving the Discoverability of Scholarly Content in the 21st Century: Collaboration Opportunities for Librarians, Publishers, and Vendors*. Thousand Oaks, CA: SAGE Publications, n.d. (est. 2011). Available at: http://www.sagepub.com/repository/binaries/librarian/DiscoverabilityWhitePaper/

Vaughan, Jason. "Web Scale Discovery Services." *Library Technology Reports*, January 2011, 47 (1). Available at: http://dx.doi.org/10.5860/ltr.47n1

VRA Core [website]. http://www.loc.gov/standards/vracore/

Way, Doug, "The Impact of Web-Scale Discovery on the Use of a Library Collection." *Serials Review*, 2010, 36 (4): 214-20. Available at: http://dx.doi.org/10.1016/j.serrev.2010.07.002