An Examination of Faceted Searching in Discovery Systems and the Impact on Information Discovery

By Katie Lai

Abstract
This study compares the performances of the resource type facets and format facets in Primo and WorldCat Discovery respectively. Through looking at librarians’ perceived understanding of selected facets, the information retrieval mechanisms employed, and the search results yielded, the author reveals gaps between users’ perception and the information actually retrieved. The goals are to see how successful Primo and WorldCat Discovery are in making themselves a one-stop shop for music information, and to determine whether the resource type or format facets in these tools facilitate information discovery. The findings also prompt librarians to reflect on what can be done to enhance information discovery through the teaching of information literacy and through collaboration with information and systems providers.

Introduction
Discovery services have seen a high uptake in academic libraries since their introduction in the mid-2000s.\(^1\) While they are meant to provide a one-stop place for searching library catalogues and electronic resources,\(^2\) are these discovery tools performing as they were intended to? To what extent have they simplified the discovery process and enhanced users’ search experiences? In Canada, the two discovery indexes that are most used by academic libraries as of 2023 are Ex Libris Central Discovery Index (CDI) and OCLC WorldCat Discovery.\(^3\) Thus, with a focus on music resources, this paper critically examines two discovery tools, namely Ex Libris Primo VE (which uses Ex Libris

---

Katie Lai (katie.lai@mcgill.ca) is an Associate Librarian and Liaison for Music at the Marvin Duchow Music Library of McGill University. She is currently a member of McGill’s WorldShare and WorldCat Discovery Management Working Group. Prior to moving to Quebec, she has project-managed and participated in library systems migrations to Ex Libris Alma from Innovative Millennium, SirsiDynix Horizon, and SirsiDynix Workflows, as well as discovery tool trial/implementation including ProQuest Summon and Ex Libris Primo and Primo VE.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

---

CDI) and OCLC WorldCat Discovery. It aims to determine the success of these tools as one-stop shops for music information, and whether the resource type or format facets in these tools facilitate information discovery. Through a questionnaire survey, this study further reveals Canadian music library communities’ perceived understanding of selected resource type or format facets and how well they align with the original design of the facets. The findings not only unveil how the facets and retrieval mechanisms could affect the discoverability of information, but also prompt librarians to reflect on what can be done in other areas, including the teaching of information literacy, to enhance information discovery.

Background

As libraries moved away from traditional online public access catalogs (OPACs), federated search started to come into play in the late 1990s. Aiming to offer a Google-like search experience, federated search systems search multiple databases and library catalogues at once with a single query and compile results from various sources under one list. However, since they do not hold any metadata and need to crawl live sources, federated search systems are often criticized for their lack of thoroughness and slow response in retrieving results. Then, in the mid-2000s, discovery systems emerged. Using a centralized index that draws metadata from library catalogues, databases and online resources, discovery systems, which still feature Google’s single search box, no longer need to connect to multiple databases when executing a search but only their pre-indexed content. The response time is therefore much improved, and they have since become the mainstream in libraries.

Primo, in its early version launched by Ex Libris in 2006, supports information discovery through its central index which contains metadata of billions of records of different resource types provided by publishers, aggregators, and content providers. Unlike traditional OPACs, which are bibliographic

---

4 According to Ex Libris’ Knowledge Center, “Primo VE is a deployment model of Primo, which is aimed at simplifying Primo’s back-end processes and further optimizing the management of Primo with Alma by utilizing the Alma platform.” For simplicity’s sake, Primo VE will be called Primo throughout this paper. https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/010Getting_Started_with_Primo_VE/005Primo_VEOverview#:~:text=Primo%20VE%20is%20a%20deployment,by%20utilizing%20the%20Alma%20platform.


8 Breeding, Index-Based Discovery Services: Current Market Positions and Trends, 14.

records-driven and allow primarily title-level searches, Primo facilitates full-text article-level searching of scholarly and academic materials worldwide through its CDI in addition to the libraries’ local holdings and repositories. It also offers features that are typical of a discovery system, such as relevancy ranking, faceted search, query recommendations, and link resolvers, to name a few.\textsuperscript{10}

WorldCat Discovery was launched in 2014, replacing its predecessor, WorldCat Local, which was introduced by OCLC in 2007.\textsuperscript{11} WorldCat Discovery is a discovery service based on OCLC’s own \textit{WorldCat} database (for title-level bibliographic records), plus article-level metadata pulled from citation information provided by publishers, aggregators, and content providers. Because of its large network of member libraries, WorldCat Discovery also helps users easily find resources available at their own libraries and other libraries worldwide through a single search.\textsuperscript{12}

Many librarians do not have the opportunity to decide which library system or discovery layer to use or implement. This might be because the decision was made before they joined the library, or because it was a collective decision made by the library or consortium as a whole after weighing the pros and cons of different products and the service requirements of different units, regardless of individuals’ preferences. Consequently, once a system or discovery tool is selected, it may be there to stay for years. Librarians may therefore develop high competency with the one at their home libraries and not have the opportunity to compare and understand the discovery technologies used in other systems and their performances. Seeing the lack of such a comparative overview, the author, who is currently an instruction librarian using WorldCat Discovery and has participated in several Primo and Primo VE migrations in recent years, offers a closer look at both services.

\underline{Literature review}

Google Scholar and the Library, \textit{Information Science and Technology Abstracts} database show that much has been written on discovery systems and information literacy. Literature from the 2010s displays strong confidence and enthusiasm for this new generation of search tools, explaining how they would positively change librarians’ teaching of information literacy skills and students’ search experience. Fagan writes that the broad subject and format coverage and the diversity of information available in discovery systems allow users to more easily explore general information sources to increase familiarity with the topic.\textsuperscript{13} Cmor and Li share that discovery tools enable users to uncover the interdisciplinary aspects of academic topics and provide timely opportunities for librarians to teach the nature and use of different information sources.\textsuperscript{14} Rose-Wiles and Hofmann

\begin{itemize}
  \item \textsuperscript{10} Breeding, \textit{Index-Based Discovery Services: Current Market Positions and Trends}, 6.
  \item \textsuperscript{11} Breeding, 19.
  \item \textsuperscript{12} WorldCat Discovery, OCLC, accessed March 18, 2024, \url{https://www.oclc.org/en/worldcat-discovery.html}.
  \item \textsuperscript{14} Dianne Cmor and Xin Li, “Beyond Boolean, Towards Thinking: Discovery Systems and Information Literacy,” \textit{Library Management} 33, no. 8/9 (2012): 455, \url{https://doi.org/10.1108/01435121211279812}.
\end{itemize}
comment that by breaking the silos of different resources, discovery systems eliminate the need to determine whether a user is to start the search in a library catalogue or a database, or to learn which subject-specific database to use. Instead, users could now move to the “higher order questions” such as how much has been written on a given topic, who wrote it, in what ways, and at what level, and librarians no longer have to teach Boolean searching which, as Cmor and Li express, students have often found difficult to apply. Librarians’ pedagogical approach can also “shift from being explanatory to exploratory.” This shift is affirmed by Buck and Steffy, who write that with discovery tools, librarians can spend more time teaching transferable skills such as evaluating and refining search results. Seeber also concurred that “teaching a database” is no longer obligatory in information literacy sessions, and that librarians can in fact take the time to focus on the critical evaluation of information, which is particularly necessary given the large amount of information that may result from a discovery search. Such critical thinking skills, along with good understanding of the scholarly publishing process continue to be helpful to users, regardless of advancement of future search technologies. Thus, discovery tools generally received a warm welcome and were perceived as a one-stop shop that could simplify the discovery process and enhance users’ search experience.

Similarly, faceted searching, which is a feature of discovery tools and which traditional OPACs did not quite support, has also been positively regarded. Not only do facets enable users to see a clearer picture of the wide range of information types available, but they also allow librarians to more easily discuss the distinctions in scope and level of various sources, such as books versus articles, or scholarly articles versus newspaper articles. Furthermore, from a searching point of view, facets allow the user to filter and narrow down a long list of search results, offering an experience that is “more user-friendly than the traditional ‘advanced searches,’ because users can elaborate queries progressively, rather than constructing an elaborate set of limits from the start.” For music, it is particularly useful to employ facets for genre, form, medium of performance, and

---


17 Cmor and Li, 451.


21 Cmor and Li, “Beyond Boolean, Towards Thinking,” 452.

22 Cmor and Li, 452.

format to quickly find needed music materials, as Iseminger et al. discussed regarding faceted music vocabularies in resources discovery.\textsuperscript{24} Hence, interactive information retrieval such as faceted searching is seen in relevant literature as providing more efficient information-seeking support.\textsuperscript{25}

When it comes to the examination of individual discovery services, Belford offered methods and examples to evaluate discovery tools using catalogue records of music materials.\textsuperscript{26} In addition, several usability studies have been conducted to assess specific discovery tools based on previous versions of Ex Libris and OCLC products such as Primo and WorldCat Local. For example, in the study of Primo conducted by Kliwer et al., students appreciated being able to search for a large amount of information from multiple source types via a single search box and also liked the convenience of finding relevant materials.\textsuperscript{27} When assessing WorldCat Local, Bertot et al. concluded that users generally found its design and the navigation between different search features clear and intuitive, but the large amount of search results produced was overwhelming.\textsuperscript{28} While many of the general themes of these earlier studies still apply and provide useful background for this research, these studies evaluate previous generations of discovery tools which have since undergone substantial development and are therefore less relevant than research of current discovery tools. Indeed, up to this point, no research has been found that methodically examines Primo VE and WorldCat Discovery, the most current versions of discovery tools from Ex Libris and OCLC. There is also no comparative study that examines the information discoverability and the performance of faceted searches using these two systems. The lack of such a comparison prompted this study.

Methodology
A seven-question online survey (see Appendix) using LimeSurvey was conducted in the Canadian Association of Music Libraries, Archives and Documentation Centres (CAML) Annual Conference held in May 2023. At the beginning of the author’s presentation titled “How discoverable are your music resources: A critical examination of OCLC WorldCat Discovery and Ex Libris Primo and their impact on teaching information literacy,” conference attendees were invited to participate in the survey by using the QR code or URL projected on the screen. The goal was to capture participants’ most instinctive responses before the author showed the comparative findings of the two discovery tools. The link to the online survey was also subsequently emailed to the CAML Listserv inviting

\textsuperscript{24} Iseminger et al., 415–16.
subscribers who did not attend the author’s presentation to complete the questionnaire by June 30, 2023.

The first two questions asked about the job title or function of the survey participants and the discovery tools they used at their institutions. The third question asked about the types of materials participants would expect a discovery tool to be able to retrieve. Questions 4 to 7 were open-ended. Participants were shown screenshots of search results from Primo and WorldCat Discovery and were asked to share their perceived understanding of various resource type or format facets in both systems. All answers were collected anonymously.

Findings

Question 1: Participants’ role
Thirty-one responses were received, and 30 (96.8%) were valid. One (3.2%) was invalid as a substantial portion of the survey was not answered. Of the valid responses, 24 (80.0%) identified themselves as a librarian, archivist, professor, or a staff member working in a library/archive/library science faculty, four (13.3%) were full-time students, one (3.3%) was a non-library faculty member/researcher, and one (3.3%) was in the “Other” category.

Question 2: Discovery system at home institution
Regarding the discovery tool used at their institutions, 18 (60.0%) participants used Ex Libris Primo, six (20.0%) used OCLC WorldCat Discovery, one (3.3%) used EBSCO Discovery Service, two (6.7%) indicated uncertainty on which system they used, and three (10.0%) used other systems, namely Polaris Leap, BiblioMondo, and CAP (see Figure 1).

![Discovery System Used at Home Institution](image)

**Figure 1.** Question 3: Retrieval expectation
Among the 14 material types listed, i.e. books, book chapters, journals, journal articles, databases, conference proceedings, newspapers, reference entries, streaming videos/audios, dissertations and theses, maps, web resources, archival materials/manuscripts, and government documents, 14 (46.7%) participants expected a discovery tool to be able to retrieve them all. When looking at each material type closely, books were expected by all participants to be retrievable by a discovery system. Other material types that 90% or more participants believed to be retrievable included book chapters, journals, journal articles, conference proceedings, newspapers, and dissertations and theses. However, for reference entries, maps, and archival materials/manuscripts, only 60% to 67% participants expected them to be retrievable. Figure 2 shows a summary of participants’ expectations.

The five open-ended sub-questions grouped under Questions 4 and 5 aim to learn about survey participants’ perceived understanding of the resource type facets in Primo. Based on a keyword search for *Beethoven*, participants were shown a screenshot of the first page of the search result with facets and were asked what the “Articles,” “Book Chapters,” “Scores,” “Text Resources” and “Web Resources” facets meant to them. The first three are the resource type facets that are found in both Primo and WorldCat Discovery. The latter two are Primo-specific facets and were included because the author found them unclear and thus wanted to see what participants’ perceived understanding was.

**Figure 2. Questions 4-5: Perceived meanings of selected resource type facets in Primo**
For the “Articles” and “Scores” facets, nearly all participants described them as articles from journals/magazines/serials/newspapers and musical scores respectively. For “Book Chapters,” most participants referred this facet to chapters/sections of books. Interestingly, five (16.7%) responses specifically denoted “Book Chapters” as chapters from ebooks. This data makes one wonder whether participants perceived this facet as covering only ebook chapters, and not chapters from print books.

For the “Text Resources” and “Web Resources” facets, participants gave a wide variety of answers and showed apparent confusion. While four (13.3%) participants were unsure about the meaning of “Text resources,” others made creative guesses such as “Something written... but not a book or article???,” “any written, unpublished text perhaps? Like a white paper or report?,” “Anything published in print format,” “Textual matter, so either books or serials, but this excludes audio and audio-visual resources [...] However it certainly could include theses and dissertations, and online textual content (textual web sites, blogs etc.),” and “Something not able to be described by other facets” (see Table 1). In short, the common descriptions for this resource type facet are something textual and in print/physical format.

### Table 1. Open-ended responses to Question 5a – “What does ‘Text Resources’ mean to you (in Primo)?” (n=30)

<table>
<thead>
<tr>
<th>1</th>
<th>Resources other than published volumes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>materials from printed sources</td>
</tr>
<tr>
<td>3</td>
<td>Anything like book, article, news article</td>
</tr>
<tr>
<td>4</td>
<td>Books, articles, book chapters</td>
</tr>
<tr>
<td>5</td>
<td>Something text-based-physical?</td>
</tr>
<tr>
<td>6</td>
<td>Fixed</td>
</tr>
<tr>
<td>7</td>
<td>I have no idea! Something written... but not a book or article???</td>
</tr>
<tr>
<td>8</td>
<td>Printed resource</td>
</tr>
<tr>
<td>9</td>
<td>Anything published in print format</td>
</tr>
<tr>
<td>10</td>
<td>Physical holdings found in the library</td>
</tr>
<tr>
<td>11</td>
<td>Textual material</td>
</tr>
<tr>
<td>12</td>
<td>Physical texts – e.g. books, magazines, newspapers</td>
</tr>
<tr>
<td>13</td>
<td>Print matériels</td>
</tr>
<tr>
<td>14</td>
<td>Manuscript</td>
</tr>
<tr>
<td>15</td>
<td>literature, not visual/audio material</td>
</tr>
<tr>
<td>16</td>
<td>Not really sure</td>
</tr>
<tr>
<td>17</td>
<td>no idea... any written, unpublished text perhaps? like a white paper or report?</td>
</tr>
<tr>
<td>18</td>
<td>Unsure</td>
</tr>
<tr>
<td>19</td>
<td>Words (as opposed to music or scores)</td>
</tr>
<tr>
<td>20</td>
<td>PDF</td>
</tr>
<tr>
<td>21</td>
<td>Books and scholarly paper publications</td>
</tr>
<tr>
<td>22</td>
<td>Something not able to be described by other facets</td>
</tr>
<tr>
<td>23</td>
<td>Resources that include only, or mainly, text</td>
</tr>
</tbody>
</table>
In print
Works that consist of words but do not fit into other categories.
I would assume books and journals, but I don’t think my system uses this
Textual matter, so either books or serials, but this excludes audio and audio-visual
resources. So, no CDs, no LPs, no DVDs, no scores, parts, miniature scores, etc. However it
certainly could include theses and dissertations, and online textual content (textual web
sites, blogs etc.)

Don’t know – vague and general
it means nothing but if I were to assume, I would say something in print
I am not certain what “Text Resources” means.

For “Web Resources,” all participants referred this facet to websites, internet resources or things
that are available online or electronically. One participant mentioned “libguides” in addition to
websites. Others wrote “databases and ebooks,” “Less-scholarly sources. Anything that contains the
searched info,” and “anything that involves a link to a web site, whatever the content…” (see Table
2). Hence, it seems that where the information is present was the main focus, regardless of what
content the web resources contained.

**Table 2. Responses to Question 5b – “What does ‘Web Resources’ mean to you (IN PRIMO)?” (n=30)**

<table>
<thead>
<tr>
<th></th>
<th>Libguides, websites pointed to by the library chosen by librarians</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>anything web-based, including online encyclopedias, etc.</td>
</tr>
<tr>
<td>3</td>
<td>Online materials</td>
</tr>
<tr>
<td>4</td>
<td>Electronic resources</td>
</tr>
<tr>
<td>5</td>
<td>Something available through a website –ie through a browser, not necessarily free</td>
</tr>
<tr>
<td>6</td>
<td>Updating</td>
</tr>
<tr>
<td>7</td>
<td>Internet resources</td>
</tr>
<tr>
<td>8</td>
<td>Internet resources [sic]</td>
</tr>
<tr>
<td>9</td>
<td>Anything published electronically</td>
</tr>
<tr>
<td>10</td>
<td>Online digital resources, either open access or subscription based</td>
</tr>
<tr>
<td>11</td>
<td>Websites</td>
</tr>
<tr>
<td>12</td>
<td>Online journals, articles, streaming music</td>
</tr>
<tr>
<td>13</td>
<td>Databases and ebooks</td>
</tr>
<tr>
<td>14</td>
<td>Websites</td>
</tr>
<tr>
<td>15</td>
<td>available online</td>
</tr>
<tr>
<td>16</td>
<td>Resources that can be found from the web that are pertinent to the search</td>
</tr>
<tr>
<td>17</td>
<td>A public, free website</td>
</tr>
<tr>
<td>18</td>
<td>Websites that discuss Beethoven or mention his name</td>
</tr>
<tr>
<td>19</td>
<td>Web sites –only available on the web</td>
</tr>
<tr>
<td>20</td>
<td>Web sites or other web material (HTML)</td>
</tr>
<tr>
<td>21</td>
<td>Less-scholarly resources. Anything that contains the searched info.</td>
</tr>
<tr>
<td>22</td>
<td>Generally accessible content</td>
</tr>
<tr>
<td>23</td>
<td>Online resources, openly available</td>
</tr>
<tr>
<td>24</td>
<td>Online</td>
</tr>
<tr>
<td>25</td>
<td>Materials available on a website that do not fit into the existing categories.</td>
</tr>
</tbody>
</table>
eresources, but I hate that term for them
I guess, anything that involves a link to a website, whatever the content – visual, audio, text etc. If this catalogue is the McGill catalogue, for instance, I would imagine that it could include both web resources hosted at McGill (either made by McGill or not), and external resources - in which case sometimes the links might no longer work, but that's life.

Similar to a Google search
something off the internet
Web resources are likely websites that are “important enough” have been catalogued.

Questions 6-7: Perceived meanings of selected format facets in WorldCat Discovery

Questions 6 and 7 examine the format facets in WorldCat Discovery. Participants were asked to state what the facets “Articles,” “Chapters” and “Computer file” meant to them. They were also asked to indicate their perceived differences between “Musical Score” and “Downloadable Musical Score” and between “Internet Resource” and “Website.” The first two facets serve as a comparison against those in Primo. The last facet and the two comparison questions were included as the author found them ambiguous and thus wanted to identify participants’ understanding.

For “Articles” and “Chapters,” all participants shared a similar understanding of the facets in WorldCat Discovery as they had in Primo, i.e. “Articles” are articles from journals/magazines/serials/newspapers and “Chapters” are chapters from books/ebooks. Regarding the difference between “Musical Score” and “Downloadable Musical Score,” 26 (86.7%) participants thought that the former refers to physical scores while the latter means online scores. One (3.3%) participant wrote that “Musical Score” refers to physical or downloadable scores whereas “Downloadable Musical Score” refers to scores that are not physical. One (3.3%) believed that “Musical Score” can be “any type of score – either printed or paper, or available in microform, or in electronic format” and “Downloadable Musical Score” is a score that is “definitely downloadable.” One (3.3%) perceived that the two facets refer to “Older scanned music vs. purely online.” One (3.3%) did not know how the facets classify the difference (see Figure 3).
As for the “Computer file” facet, a variety of responses were received. In general, 12 out of 28 (42.8%) participants responded that “Computer file” is a digital file that can be downloaded or are stored in a computer/server. Eight (28.6%) participants perceived it as a disk/CD-ROM/file in a physical medium. Four (14.3%) expressed that they did not know what the facet meant. Two (7.1%) perceived it as an archival file. One (3.6%) wrote that it is unpublished online resource, and one (3.6%) guessed that it refers to a non-textual file. In sum, the majority saw “Computer file” as referring to a digital file that exists in a computer/server or in a physical format.

When being asked to state the difference between “Internet Resource” and Website,” obvious confusion was again observed. Nine of 28 (32.1%) participants expressed that there was no difference. Five (17.9%) wrote that they were not sure. Five (17.9%) indicated that “internet resource” is broader and “website” is a specific type of internet resource. Three (10.7%) participants wrote that these terms refer to subscription and free web content respectively. Six (21.4%) gave varying responses such as “Scholarly and credible vs less-so” and “Internet resource is deeper in content” (see Table 3).

**Table 3. Responses to Question 7b – “What are the differences between ‘Internet Resource’ and ‘Website’ (in WorldCat Discovery)?” (n=28)**

<table>
<thead>
<tr>
<th></th>
<th>1 No difference here</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>website is a specific type of internet resource</td>
</tr>
<tr>
<td>3</td>
<td>Not sure...</td>
</tr>
<tr>
<td>4</td>
<td>Internet resources seem to point to electronic scores. I’d expect a website to be an open access site. I admit that it’s highly confusing however!</td>
</tr>
<tr>
<td>5</td>
<td>First is an electronic object accessible via a subscription? Second available via a website/through a browser?</td>
</tr>
<tr>
<td>6</td>
<td>Internet resource is deeper in content.</td>
</tr>
<tr>
<td></td>
<td>I don’t know!</td>
</tr>
<tr>
<td>---</td>
<td>----------------</td>
</tr>
<tr>
<td>8</td>
<td>Website is specific. Internet resource is broad</td>
</tr>
<tr>
<td>9</td>
<td>Subscription or open access resource vs free web content</td>
</tr>
<tr>
<td>10</td>
<td>I don’t know</td>
</tr>
<tr>
<td>11</td>
<td>Nothing – both could mean a paid subscription or a free online resource</td>
</tr>
<tr>
<td>12</td>
<td>Not clear</td>
</tr>
<tr>
<td>13</td>
<td>I’m not sure</td>
</tr>
<tr>
<td>14</td>
<td>Internet resource includes websites but could also include files and random things in different formats on the internet</td>
</tr>
<tr>
<td>15</td>
<td>I would expect these are the same thing</td>
</tr>
<tr>
<td>16</td>
<td>An internet resource is a general category, while a website denotes a specific location for retrieving information.</td>
</tr>
<tr>
<td>17</td>
<td>Not sure what is meant</td>
</tr>
<tr>
<td>18</td>
<td>Naxos for example</td>
</tr>
<tr>
<td>19</td>
<td>Scholarly and credible vs less-so</td>
</tr>
<tr>
<td>20</td>
<td>Subscription vs generally accessible content</td>
</tr>
<tr>
<td>21</td>
<td>Not clear</td>
</tr>
<tr>
<td>22</td>
<td>More or less the same</td>
</tr>
<tr>
<td>23</td>
<td>Internet resource is a broad category and website is a subset of possible types of internet resources</td>
</tr>
<tr>
<td>24</td>
<td>Internet resource suggests that a vendor has put together a product (database, collection, exhibit, etc) while a website is just a website</td>
</tr>
<tr>
<td>25</td>
<td>Hm. I guess that “website” would simply point to a given website, probably to the intro page or whatever. “Internet resource” might be any given resource available through a link - perhaps a PDF, or a video, whatever, but not a complete web site?</td>
</tr>
<tr>
<td>26</td>
<td>unknown</td>
</tr>
<tr>
<td>27</td>
<td>I am not sure</td>
</tr>
<tr>
<td>28</td>
<td>I do not feel there should be a difference between these two categories.</td>
</tr>
</tbody>
</table>

**Discussion**

From the survey results, it is apparent that both Primo and WorldCat Discovery exhibit clarity problems in how certain resource type and format facets are named, and librarians displayed difficulty in understanding what they meant. What follows is a discussion of the two discovery systems’ treatment of the terms that cause confusion, and the impact of these issues on discovery.

**Resource type facets in Primo**

In Primo, the items shown in the result list under each of the “Articles,” “Book Chapters,” and “Scores” facets largely reflect the same resource types perceived by participants. One possible inaccurate understanding of the “Book Chapters” facet, however, is the specification of ebooks rather than all books in both print and electronic formats. According to Ex Libris’ Knowledge Center, things will fall under the “Book Chapters” facet if the MARC Leader/06-07 is aa (i.e. language material and monographic component part), or, Leader/06-07 is ab (i.e. language material and
serial component part) and 008/21 is w (i.e. monographic series),\textsuperscript{29} or if it is provided in the central index. Hence, the facet “Book Chapters” is not restricted to digital or electronic format, but would cover all books in both print and electronic versions based on the MARC mapping.

For the facets of “Text Resources” and “Web Resources,” there was evident confusion among participants. According to Ex Libris’ Knowledge Center, “Text Resources” is mapped to materials that have MARC Leader/06 as m (i.e. computer file/digital material) and 008/26 as d or e (i.e. document or bibliographic data), or as provided in the central index. A look at the documentation for the previous version of Primo hints that this resource type facet includes “text resources that cannot be identified as a book, journal, or article.”\textsuperscript{30} Yet, nearly all participants failed to determine how “Text Resources” differed from other types of textual materials such as books, journals, and dissertations. Only two (6.7%) participants more closely associated “Text Resources” with “something not able to be described by other facets” and “works that consist of words but do not fit into other categories.”

Regarding “Web Resources,” Ex Libris’ Knowledge Center indicates that materials with MARC Leader/06 as m (i.e. computer file/digital material) and tag 008/26 as j (i.e. online system or service), or, MARC Leader/06 as a (i.e. language material), Leader/07 as b or i or s (i.e. serial component part, integrating resource, or serial) and tag 008/21 as w (i.e. updating web site),\textsuperscript{31} or as provided in the central index, would go to this “Web Resource” facet. Thus, most participants correctly connected this facet with online resources and websites.

In summary, survey participants were confused with some of the facets presented, and there are gaps between their understanding and what the facets would retrieve. So, are the five Primo facets examined here helping users quickly find relevant information, as intended by faceted searching? There are undoubtedly designated retrieval criteria in the documentation for “Text Resources” and “Web Resources,” for example. However, how helpful are these criteria, when both “Text Resources” and “Web Resources” retrieve webpages, blog posts, online articles on magazine websites, etc.? This is puzzling and creates confusion for users.

Format facets in WorldCat Discovery

While naming and definitions of resource type facets in Primo is perplexing, the same can be said for WorldCat Discovery. While most survey participants expected the “Chapter” facet to retrieve book chapters, a close examination of the search result was shocking. Based on a keyword search

\textsuperscript{29} “Mapping to the Display, Facets, and Search Sections in the Primo VE Record,” Ex Libris Knowledge Center, Ex Libris, accessed March 18, 2024, https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/120Other_Configurations/Mapping_to_the_Display%2C_Facets%2C_and_Search_Sections_in_the_Primo_VE_Record.

\textsuperscript{30} “The Display Section,” Ex Libris Knowledge Center, Ex Libris, accessed March 18, 2024, https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/Primo/Technical_Guide/010The_PNX_Record/d/040The_Display_Section.

\textsuperscript{31} “Mapping to the Display, Facets, and Search Sections in the Primo VE Record.”
for *Beethoven* in the author’s home library, which uses WorldCat Discovery, “Chapter” does not point users to book chapters but to papers from conference proceedings! All 41 items retrieved were conference proceedings on science/engineering topics from an IEEE subscription (see Figure 4). There were no resulting book chapters or books, nothing about Beethoven, the composer, and nothing related to arts or humanities. What does OCLC say about this facet? According to the OCLC Support webpage, when a record is not in WorldCat and MARC tag 949 subfield x is o, the item will be mapped to the “Chapter” facet. However, tag 949 is a local field defined by OCLC internally, and there is no documentation on the open web explaining what “o” refers to. It thus remains a mystery why conference proceedings would be treated as book chapters in WorldCat Discovery. This facet simply yielded irrelevant results, and in this instance, the performance of this discovery tool proved unacceptable.

**FIGURE 4. EXCERPT OF A KEYWORD SEARCH RESULT FOR BEETHOVEN USING THE “CHAPTER” FACET IN WORLDCAT DISCOVERY.**

The facets “Musical Score” and “Downloadable Musical Score” provide similarly disconcerting outcomes. Looking at the search results for *Beethoven*, the “Musical Score” facet covers both

---

physical and online scores, not just physical scores. This is confirmed by WorldCat Discovery’s documentation that scores are grouped under this facet when MARC Leader/06 is c or d (i.e. notated music or manuscript notated music) and there is no mention of format as a facet criterion. For a general facet such as “Musical Score,” it may make sense to cover scores of all formats. However, there is misalignment with participants’ perceived understanding of the more peculiar facet “Downloadable Musical Score.” Incompatible with the literal meaning of the word “downloadable,” the “Downloadable Musical Score” facet retrieves not only downloadable scores but also online scores that are view-only and cannot be downloaded, e.g. those from the database Music Online: Classical Scores Library by Alexander Street Press. This is because scores that have MARC Leader/06 as c or d (i.e. notated music or manuscript notated music), tag 008/23 as s (i.e. electronic) and the presence of a link in tag 856 are all placed under the “Downloadable Musical Score” facet. There is no additional element in place to distinguish whether an online score is downloadable or view-only. Search results will therefore automatically return both downloadable and non-downloadable digital scores. “Downloadable Musical Score,” in this case, is misnomer that retrieves results that will surely baffle users.

The “Computer File” facet, although a variety of responses were received, does not pose as much of a discrepancy. According to OCLC Support, if MARC Leader/06 is m (i.e. Digital material consisting of computer software, numeric data, computer-oriented multimedia, or online services or systems) and the document type is not game, interactive multimedia, or serial, then the item goes to the “Computer File” facet. Participants’ responses were therefore generally correct in determining that this facet refers to digital files stored in a medium/computer/server.

Lastly, the facets “Internet resource” and “Website” are noticed to create the most ambiguity among all format facets studied. In the OCLC documentation, “Internet Resource” refers to items that are not in WorldCat. On the other hand, “Website” is used for records that have MARC Leader/06 as a or t (i.e. language material or manuscript language material), Leader/07 as i (i.e. integrating resource), 008/23 is s (i.e. electronic) and a web link is present in tag 856. “Website,” then, should direct users to online resources that primarily have textual content and are accessible via web links. However, when examining search results, there is no apparent difference between the two facets. A keyword search for Beethoven in the WorldCat Discovery from the author’s home library yielded only one result under the “Internet Resource” / “Website” facets. When doing a broader keyword search for Music, 131 catalogue records were retrieved. However, these records were again covered by both facets, and nothing was unique to the “Website” facet. This raises the question: what is the nature of an “Internet Resource” as compared to a “Website”? The purpose of distinct format facets is to help users specify the format of resource they need, but WorldCat Discovery has failed to provide the clarity necessary to make this facet useful.

33 “Format Display in Search Results.”
34 “Format Display in Search Results.”
Information discoverability

In addition to studying the resource type and format facets identified in Questions 4 to 7, it is also worth examining how the two discovery tools handle known item searches. In the case of a book chapter, users will have two rather different experiences. In a chapter search for “Beethoven Overture King Stephen” for instance, Primo shows the book chapter record at the chapter level and the link provided takes users straight to the specific chapter needed (see Figure 5a). This is straightforward and requires no guessing on the part of the user. However, things are not the same in WorldCat Discovery. Users will not see the chapter title “Beethoven Overture King Stephen” as the prominent information in the entry. WorldCat Discovery will instead show the book title “Orchestral Masterpieces under the Microscope” in the result list, and the chapter “Beethoven Overture King Stephen” will only appear as part of the content notes (see Figure 5b). Hence, users may easily miss it and think that the library does not have the book chapter queried if they do not recognize the need to also look at the content notes displayed further down in the record. This book title-level display of book chapter results is possibly due to WorldCat Discovery’s catalogue record-driven information retrieval approach which relies on the book-title level MARC records in the WorldCat database. As a result, WorldCat Discovery retrieves results based on the data in a specific MARC tag rather than on a central index or knowledge base. Thus, if the MARC record of a book does not have data entered in the formatted content notes in tag 505, none of the chapters may be discovered. Book chapters are therefore not as easily discoverable in WorldCat Discovery as they are in Primo.

**Figure 5a. A known item search for book chapter “Beethoven Overture King Stephen” in Primo. The emphasized title is the chapter title searched.**
Another comparison worth looking at is a known database search, since providing easy access to subject databases is an essential first step in a scholarly search, and users may likely start with the single search box in the discovery layer. In Primo, this search involves the “Databases” facet. By contrast, in WorldCat Discovery, one must make use of the “Website” facet. Using the major music database RILM as an illustration, a keyword search for RILM in the discovery layer of University of Western Ontario, University of Toronto, and Queen’s University, all of which use Primo, returns RILM as the eighth, twelfth, and twenty-seventh hit in the result list respectively as of March 18, 2024. This is not ideal as users would usually look only at the first results page or the first few results on the list, and any results beyond the first page run the chance of being overlooked. Of course, using the “Databases” facet could immediately separate RILM from the rest, but this would require an extra step by the users. One convenient solution is to use the Resource Recommender function in Alma, which is not only a library services platform itself, but also manages functions in the Primo discovery layer. By assigning a searchable tag to RILM in Alma, Primo would display RILM as a suggested database at the top of the brief result list upon a simple keyword search, and this

---


could put a spotlight on RILM and direct users to the database effectively (see Figure 6).37 Regarding WorldCat Discovery, a keyword search for RILM will result in it being listed as the sixth item in the search result at the author’s library. Since there is no facet for databases in WorldCat Discovery, users have to creatively select the “Internet Resource” facet or the “Website” facet in order to arrive at RILM as their first result. This additional step does not help offer a smooth information discovery experience. Furthermore, there is no resource recommendation customization similar to Primo to enhance ease of access. Of course, libraries often provide a browsing option via the Database A-Z list, but this requires users to navigate beyond the single-search box. Thus, from the search functionality perspective, Primo is demonstrated to provide better database discovery than WorldCat Discovery.

**Figure 6. A Known Item Search for Database RILM in Primo in Which RILM Has Been Configured as a “Best Bet” to Facilitate Easy Access to the Database**

![Image of Primo search for RILM](image)

While looking at RILM, it is also useful to see how discoverable RILM contents are in both discovery tools. Since RILM is exclusively available on EBSCO and its metadata is not shared with other vendors and system providers, RILM content is deemed not retrievable *prima facie*, except in

37 “Resource Recommender for Primo VE,” Ex Libris Knowledge Center, Ex Libris, accessed March 18, 2024, [https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/120Other_Configurations/010Resource_Recommender_for_Primo_VE](https://knowledge.exlibrisgroup.com/Primo/Product_Documentation/020Primo_VE/Primo_VE_(English)/120Other_Configurations/010Resource_Recommender_for_Primo_VE).
EBSCO’s own discovery system. This is true for Primo. For example, in a known journal article search in Primo for “Recession, reflation: Skempton, Finnissy and musical modernism’s classical roots,” which is available in the database RILM with Full Text, the record retrieved does not point users to RILM due to the unavailability of RILM metadata in Primo. Instead, the record includes links to other databases that also have this article indexed, such as Music Periodicals Database. However, when doing the same search in WorldCat Discovery, a RILM link will in fact turn up, despite the lack of RILM metadata in its central index. A detailed examination revealed that the retrieval was a result of WorldCat Discovery populating information from its search box and executing an indexed field search on the EBSCO platform, i.e. (AU whittall) AND (IS "00274666") AND DT 2018. WorldCat Discovery did not rely on a central index or knowledge base; the results were simply the work of a Boolean search. However, a few more test searches showed that such “copy-searching” was not consistently performed by WorldCat Discovery and the RILM links did not always surface in the search results. The reason for this inconsistency cannot be ascertained at this time. However, this creative workaround is a bonus, enabling the discovery of exclusive content that would otherwise remain undiscovered. This analysis also demonstrates that because of the exclusiveness of metadata and the protective policies of information providers, it is unfortunately still necessary to introduce subject databases to users and to teach database searching, contrary to the ideal world envisioned by earlier researchers. Instead of teaching all resources, it now becomes a question of what “extra” databases librarians have to teach on top of teaching the discovery tools, taking into account what is not discoverable there. In fact, bypassing the discovery tool altogether and going directly to the databases may still be common practice by some librarians.

One more comparison that shows the difference in information discoverability between discovery systems is the retrieval of reference entries. In the survey, reference entries received the lowest retrieval expectation, with only 18 (60%) participants stating positive expectations. So, can reference entries be retrieved via discovery tools? When using the same keyword search for Beethoven in Primo and filtering the results using the “Reference entries” facet, Primo retrieves encyclopedia articles about Beethoven in reference resources from various disciplines such as the Oxford Dictionary of Music, New Oxford Rhyming Dictionary, Oxford Essential Quotations, etc. If the keyword is changed to the name of a living composer such as Kaija Saariaho, who has been written about by music scholars but less so by researchers in other disciplines, Primo would suitably point users to major music reference sources such as the Oxford Dictionary of Music, the SAGE International Encyclopedia of Music, Contemporary Musicians, and New Grove Dictionary of Music and Musicians, as shown in Primo at the University of Toronto Library. However, in WorldCat Discovery, there is simply no related facet that could enable such filtering or metadata in its central index or knowledge base to facilitate discovery of entries from reference sources. This is an area of WorldCat Discovery that needs improvement.
Limitations
This paper focuses on specific resource type facets in Primo and format facets in WorldCat Discovery. The selections were made based on the author’s frustration when using these facets. Hence, this study does not attempt to examine all resource type or format facets in a comprehensive manner; instead, it offers a snapshot view of some of the problems found as a way to ignite critical thinking about and evaluation of the two discovery systems Canadian libraries use the most. Future research could certainly look at all facets in detail and provide a comprehensive comparison. Another limitation of this study is the relatively small sample size and the focus on music librarians’ perspectives rather than that of general users. Thus, research capturing a broader audience’s views could offer a complementary view on these facets.

Another consideration is that technologies used in discovery tools change rapidly, and the contents available in the CDIs of both Primo and WorldCat Discovery may change or expand on a daily or weekly basis. Issues reported in this study are limited to the state of things in moment of active research. Nonetheless, the findings can serve as a good comparative overview of the current state of Primo and WorldCat Discovery.

Conclusion
In this study, it was found that the meaning of several of the facets under examination is rather unclear and ambiguous and does not align with users’ perceived understanding. This ambiguity has led to confusion among survey participants who expressed uncertainty regarding the types of information that would be retrieved in various situations and hindered a good search experience. In addition, the MARC mappings and retrieval mechanisms currently in place do not always retrieve the types of information suggested by the facets. As McGrath writes, the performance of facets intertwines with facet vocabularies, metadata, and interface design, and a carefully designed faceted search increases both precision and recall.38

The analyses and observations described in this study show that there is opportunity for significant improvement in faceted searching for both discovery tools. They also reveal the inadequacy of Primo and WorldCat Discovery to serve as a true one-stop shop for music information due to the indiscernability of certain types of resources, such as reference entries and contents from databases like RILM that work only with selected vendors. While exclusiveness of metadata and the noncooperation between system and information providers may be attributed as one of the causes, reliance on cataloguing records alone without the addition of the central index or knowledge base can also have a negative impact on information discovery.

So, what does this mean for librarians? No system is perfect, and areas for improvement are to be expected. For librarians who are tasked to teach searching skills, it is important to be aware of the strengths and weaknesses of each discovery system so that workarounds and alternative solutions can be planned and introduced to users. If librarians themselves are not familiar with the limitations of their discovery tools and are not aware of what can and cannot be retrieved, it would not be realistic to expect users – especially students – to navigate their own research journeys with few hurdles and little stress, or not to abandon the library catalogue and resort to Google. Awareness of the capabilities of various discovery tools also allows librarians to make enhancement requests to system vendors for the betterment of the whole community. Another way to expand information discoverability is to work with database vendors and compel them to demonopolize metadata. This can be done regardless of what system a library uses. Library personnel can also advocate within user communities such as Music OCLC Users Group (MOUG), Ex Libris Users of North America (ELUNA), and the Program for Cooperative Cataloging Consultation Group for Library Systems and Vendors. With concerted effort, librarians can change the (information) world!

Author Declaration
I confirm that no funding was received for this work. There are no known conflicts of interest associated with this publication and there has been no financial support for this work that could have influenced its outcome. The questionnaire survey mentioned in this paper has been reviewed and approved by the Research Ethics Board of McGill University.
Appendix

Questionnaire Survey

Q1. I am:
   a librarian, archivist, professor or a staff member working in a library/archive/library science faculty
   a non-library faculty member/researcher
   a full-time student
   other (please specify) ______________________

Q2. Which discovery system is currently used by your institution?
   EBSCO Discovery Service
   Ex Libris Primo
   Ex Libris Summon
   OCLC WorldCat Discovery
   SirsiDynix Enterprise
   I am not sure
   Other (please specify) ______________________

Q3. Which of the following do you expect a discovery system to be able to retrieve? (multiple selections allowed)
   Books
   Book chapters
   Journals
   Journal articles
   Databases
   Conference proceedings
   Newspapers
   Reference entries
   Streaming videos/audios
   Dissertations and theses
   Maps
   Web resources
   Archival materials/manuscripts
   Government documents
   All of the above
Q4. Here are the search results based on a keyword search for Beethoven in Primo. Looking at the filters on the left-hand menu of the screenshot below:

a. What does “Articles” mean to you?
b. What does “Book Chapters” mean to you?
c. What does “Scores” mean to you?
Q5. Here are the search results based on a keyword search for Beethoven in Primo. Looking at the filters on the left-hand menu of the screenshot below:
   a. What does “Text Resources” mean to you?
   b. What does “Web Resources” mean to you?

Q6. Here are the search results based on a keyword search for Beethoven in WorldCat Discovery. Looking at the filters on the left-hand menu of the screenshot below:
   a. What does “Article” mean to you?
   b. What does “Chapter” mean to you?
   c. What are the differences between “Musical Score” and “Downloadable Musical Score”?
Q7. Here are the search results based on a keyword search for *Beethoven* in *WorldCat Discovery*. Looking at the filters on the left-hand menu of the screenshot below:

a. What does “Computer File” mean to you?

b. What are the differences between “Internet Resource” and “Website”?
Bibliography


