

■ BERICHT DER AG PRIMO-IMPLEMENTIERUNG AN DIE VOLLVERSAMMLUNG DES ÖSTERREICHISCHEN BIBLIOTHEKENVERBUNDES ZUM MÖGLICHEN EINSATZ DES DIENSTES PRIMO CENTRAL

von *Eveline Pipp, Markus Knitel, Wolfgang Mayer, Michaela Putz*

Inhalt

1. Untersuchungsgegenstand
2. Produktbeschreibung
3. Vorgangsweise bei der Evaluierung von Primo Central
4. Testergebnisse
5. Primo Central und andere Dienste
6. Schlussfolgerungen
7. Fazit

Zusammenfassung: Die Arbeitsgruppe PrimoImplementierung des Österreichischen Bibliothekenverbundes evaluierte im Winter 2010/11 Primo Central, den Discovery und Delivery Service von Ex Libris, im Zuge eines dreimonatigen Testzuganges und mehrerer Frage-und-Antwort-Runden mit Produktentwicklern von Ex Libris. Primo Central integriert hunderte Millionen anbieterseitig gelieferter Metadaten zu vornehmlich unselbständiger Literatur mit lokalen Beständen. In Bezug auf viele Forschungsgebiete wurde Primo Central als inhaltsreich empfunden, hauptsächlich hinsichtlich wissenschaftlicher Zeitschriftenaufsätze. Es deckte die Bestände einer großen und einer mittleren Universaluniversitätsbibliothek gut ab. Das Fehlen nationaler (deutschsprachiger) und regionaler Titel schränkt die Anwendbarkeit für Suchen nach österreichischem und deutschem Recht wesentlich, für Suchen in den Bereichen Wirtschaft, Psychologie und Bauingenieurwesen teilweise ein. Primo Central wird regelmäßig aktualisiert, ist jedoch notwendigerweise weniger aktuell als Verlagsplattformen oder PubMed, demzufolge auch weniger aktuell als Federated Search-Systeme wie MetaLib. Die dauerhafte Verfügbarkeit von Datenquellen in Primo Central kann von Ex Libris nicht garantiert werden. Die Metadaten der in Primo Central indexierten Ressourcen kommen derzeit nur zu einem geringen Teil vom Verlag selbst (der Großteil über CrossRef und IngentaConnect), und enthalten daher oft keine inhaltliche Erschließung. Die Suchmöglichkeiten und die Reihung der Treffer nach Relevanz sind vergleichbar mit Google Scholar. Zusätzliche Funktionalitäten von Primo Central betreffen die Möglichkeit der Facettierung der Suchergebnisse und die Einschränkung der Treffer auf an der jeweiligen Bibliothek verfügbare Volltexte. Die Implementierung von Primo Central verlief schnell und problemlos. Voraussetzungen für den Ein-

satz von Primo Central sind die Verwendung von Primo als Bibliothekssuchmaschine oder von MetaLib als Federated Search System und der Einsatz eines Link Resolvers (nicht zwingend SFX).

Schlagwörter: *indexbasierte Suchmaschine, Discovery and Delivery Service, Evaluierung, Test, Kataloganreicherung, Bestandsnachweis, Nachweis unselbständiger Literatur, Facettierung, Relevanzranking, Deduplizierung*

APPLICABILITY OF THE DISCOVERY SERVICE PRIMO CENTRAL FOR AUSTRIAN ACADEMIC LIBRARIES. A REPORT OF THE TASK FORCE PRIMOIMPLEMENTATION FOR THE AUSTRIAN LIBRARY NETWORK

Abstract: *In winter 2010/11 the task force PrimoImplementation of the Austrian Library Network evaluated Primo Central, the Discovery and Delivery Service from Ex Libris, by means of a three month test account and several Q&A sessions with product developers from Ex Libris. Primo Central integrates hundreds of millions of metadata from various data vendors with local holdings. The product was found to be comprehensive in many areas of research (mainly scholarly journal articles), well representing the holdings of a large as well as a medium multi-faculty university library. The lack of national (German) and regional titles severely affects searches for Austrian and German law, it partly affects searches in the fields of economics, psychology and civil engineering, too. Primo Central is updated regularly, but has to be less up-to-date than publisher platforms or PubMed or real time federated search engines like MetaLib. Perpetual availability of records from publishers or database vendors cannot be guaranteed by Ex Libris. Currently, metadata of resources indexed in Primo Central mainly do not come from the publishers themselves, but from CrossRef or IngentaConnect. Therefore subject indexing is often missing for these records. Search functionalities and display of relevance ranked search results are comparable to Google Scholar, additional features of Primo Central being faceted search and the opportunity to show only those results available at the particular library. Implementation of Primo Central has been smooth and quick. Preliminaries for the use of Primo Central are the use of Primo as library search engine or MetaLib as federated search engine and a link resolver (not necessarily SFX).*

Keywords: *index-based search engine, Discovery and Delivery Service, evaluation, trial, catalogue enrichment, evidence of holdings, evidence of component literature, faceted search, relevance ranking, deduplication*

1. Untersuchungsgegenstand

Gegenwärtige Bibliothekskataloge weisen vorwiegend selbständige Literatur (Bücher, Monographien, Zeitschriften) nach. Bestrebungen zur Erweiterung von Katalogen um unselbständige Literatur (Zeitschriftenaufsätze, Buchbeiträge, Kongressbeiträge, u.a.) scheiterten immer wieder an den Kosten für den Kauf der Titeldaten und am Personalaufwand der Konvertierung der Daten in das jeweilige Bibliothekssystem. Die seit 2010 am Markt angebotenen „Discovery und Delivery Services“ (z.B. Primo Central, EBSCO Discovery Services, Summon) haben den Anspruch, genau diese Lücke zu füllen. Sie ermöglichen eine nahtlose Integration von Millionen von Metadaten zu unselbständiger Literatur und zwar in einem weit größeren Umfang als dies durch Eigenentwicklungen möglich wäre.

Für die am Österreichischen Bibliothekenverbund teilnehmenden Einrichtungen, die Primo als Bibliothekssuchmaschine bereits einsetzen oder einzusetzen planen, ist besonders das in Primo integrierbare Produkt Primo Central von Interesse. Daher beauftragte die Vollversammlung des Österreichischen Bibliothekenverbundes die AG PrimoImplementierung¹, zu untersuchen, inwieweit Primo Central zum aktuellen Zeitpunkt (Winter 2010) bereits in der Lage ist, die von den Primo-Partnern innerhalb des Österreichischen Bibliothekenverbundes abonnierte bzw. lizenzierte unselbständige Literatur abzubilden.

Die hier vorgelegte Publikation stellt den gekürzten und leicht veränderten Bericht der AG PrimoImplementierung an die Vollversammlung des Österreichischen Bibliothekenverbundes vom 12. Mai 2011 dar². Der Bericht richtet sich in erster Linie an an Primo Central interessierte Einrichtungen. Wir glauben jedoch, dass die im Rahmen der Evaluierung von Primo Central gemachten Beobachtungen auch bei der Evaluierung anderer Discovery and Delivery Services hilfreich sein können.

2. Produktbeschreibung

2.1 Übersicht

Primo Central ist ein relativ neues Produkt des Software-Unternehmens Ex Libris. Auf Basis der Software Primo stellt Ex Libris seinen Kunden einen aus unterschiedlichsten Quellen gespeisten, zentralen Index zur Verfügung. Nach eigenen Aussagen hat Ex Libris derzeit bereits mehrere hundert Millionen Datensätze geladen und indiziert³.

Ex Libris wirbt damit, dass bereits die Mehrheit der internationalen Verleger und Aggregatoren wissenschaftlicher Literatur erlaubt, ihre Plattformen für elektronische Zeitschriften bzw. elektronische Bücher für Primo Central zu indexieren. Über Ingenta, PubMed, Teile von Web of Science und die fachübergreifenden Volltextdatenbanken von Gale - Cengage Learning konnte eine Vielzahl kleiner Verlage und akademischer Gesellschaften in einem Schritt, ohne zeitraubende Einzelabkommen, in Primo Central aufgenommen werden.

Für Primo Central werden nicht nur bibliographische Zitate, Abstracts und allfällige Schlagwörter, sondern teilweise auch die Volltexte der Zeitschriftenaufsätze bzw. Buchbeiträge indexiert. Angezeigt wird immer das bibliographische Zitat, teilweise auch Abstract und Schlagwörter.

Primo Central kann nur mit den von den Anbietern gelieferten Datensätzen arbeiten. Qualität und Aktualität der Einträge hängen somit unmittelbar von den ursprünglichen Quellen und der Häufigkeit der Aktualisierung ab. Es fehlen in Fachdatenbanken typische Hintergrunddaten (wie Normindizes oder Thesauri). AutorInnennamen werden wie im Datensatz eingetragen, ohne Verweise und oft als reine Buchstabenfolge übernommen.

2.2 Funktionsweise

- Die Suche in Primo Central liefert prinzipiell sowohl lizenzierte, als auch nicht-lizenzierte Inhalte einer Einrichtung. Auf Basis der automatisierten Meldung der eigenen Bestände (= Lizenzen einer Institution) an Ex Libris kann den EndnutzerInnen die Wahlmöglichkeit angeboten werden, die Menge der Treffer auf die der Einrichtung verfügbaren Dokumente einzuschränken.
- Eine Suche im Primo Central-Index ist genau so schnell wie eine Suche im lokalen Katalog.
- Die Suchergebnisse aus dem zentralen Index können auf Wunsch mit den lokalen Treffern aus Katalog und anderen Quellen vermischt werden („Blended Search“). Falls es gewünscht ist, lokale Daten gegenüber der enormen Menge von Treffern in Primo Central zu bevorzugen, können diese mit unterschiedlichen, aber derzeit beschränkten Mitteln auf die erste Ergebnisseite gezwungen werden. Betreiben mehrere Einrichtungen eine gemeinsame Primo-Instanz, so müssen solche, die Suchmaschine betreffende Einstellungen unter den teilnehmenden Einrichtungen abgestimmt werden.
- Für den direkten Zugriff auf die verfügbaren Dokumente via Verlinkung wird ein OpenURL-fähiger Link-Resolver benötigt.

2.3 Implementierung

In der neuesten Version 3.0.2 von Primo kann die Aktivierung des Dienstes und der verfügbaren Ressourcen unkompliziert über eine Auswahlliste erfolgen. Die Administrationsoberfläche bietet Informationen, von welchen Anbietern (Verlage, Datenbanken, Aggregatoren) Daten in Primo Central enthalten sind. Primo-AdministratorInnen können so entscheiden, ob Daten eines bestimmten Anbieters durchsucht werden sollen. Die Auswahl erfolgt aber auf Anbieterebene und nicht auf der Ebene einzelner Zeitschriften oder Bücher. Es erscheint daher sinnvoll, alle Quellen auszuwählen – mit Ausnahme der nicht von einer Einrichtung lizenzierten Datenbanken⁴.

Die Aktivierung von Primo Central erfolgt auf Ebene der einzelnen Institution, muss also nicht zwischen den teilnehmenden Institutionen abgesprochen werden.

Für die Verlinkung von Treffern in Primo Central zum lizenzierten Volltext ist ein Link-Resolver notwendig (siehe auch Kap. 5.1). Die Pflege der anzuzeigenden Lizenzen erfolgt in der Administrationsoberfläche des Link-Resolvers. Sie obliegt der jeweiligen Einrichtung und kann nicht zentral durchgeführt werden.

2.4 Unterschiede zur Federated Search (z.B. MetaLib)

Primo Central ist in seiner Funktionsweise nicht mit der bekannten Federated Search-Technologie (z.B. MetaLib) vergleichbar. Während MetaLib die Originaloberflächen in Echtzeit durchsucht und somit erst auf die Ergebnisse der einzelnen Anbieter warten muss, werden in Primo Central alle Inhalte vorab in die zentrale Primo-Instanz⁵ geladen und dort indiziert. Es werden also nicht die Originaldatensätze, sondern ein Speicher regelmäßig aktualisierter und ergänzter Kopien durchsucht. Das darauf zugreifende lokale Primo kann auf den bereits vorbereiteten Index genauso schnell zugreifen wie auf den lokalen Index. Im Gegensatz zu Primo Central erlaubt MetaLib keine Einschränkung der durchsuchten Inhalte auf die konkrete Verfügbarkeit an einer Institution.

3. Vorgangsweise bei der Evaluierung von Primo Central

Für die Evaluierung von Primo Central stand der Kontakt mit der Herstellerfirma Ex Libris im Vordergrund. Gleich zu Beginn wurde daher ein detaillierter Bericht der in Primo Central indizierten Quellen angefordert und in

der AG besprochen. Da die Auflistung in diesem Fall sehr umfangreich ausfiel, wurde auch eine Liste der Anbieter angefordert und auch geliefert. Ex Libris bietet an Primo Central interessierten Einrichtungen die Möglichkeit eines direkten Abgleichs der indexierten Quellen mit den eigenen Lizenzen, von der die Universitäts- und Landesbibliothek Tirol Gebrauch machte (siehe Kap. 4.1).

Zur Klärung allgemeiner Fragen, die nicht die indexierten Quellen betreffen, wurde eine Fragenliste für Ex Libris vorbereitet, deren Beantwortung auf schriftlichem Weg im Herbst 2010 erfolgte.

Ab Herbst 2010 stand der AG ein zeitlich auf drei Monate beschränkter, aber in seinem Umfang uneingeschränkter Test-Zugang zu Primo Central zur Verfügung. Primo Central wurde in den Primo Test-Sichten der Universitäten Innsbruck und Wien implementiert. Im Zuge dieses Tests konnten hinsichtlich Datenumfang, Erschließungstiefe sowie Filterung der indexierten Inhalte viele Fragen geklärt werden. Die während der Tests aufgetretenen Unregelmäßigkeiten und Fragen wurden abschließend in einer Telefonkonferenz im Jänner 2011 mit dem Primo Product Management besprochen. Die AG hatte dabei Gelegenheit, ihre Erwartungen und Wünsche an eine zukünftige Entwicklung von Primo Central direkt gegenüber den zuständigen Personen zu formulieren.

4. Testergebnisse

Die Darstellung der Testergebnisse ist bewusst ausführlich gehalten. Die AG möchte der Tatsache Rechnung tragen, dass beim Test aufgefallene Mängel für manche an Primo Central interessierte Einrichtungen möglicherweise ein entscheidendes Kriterium für dessen Erwerb/Implementierung darstellen.

4.1 Erfahrungen der ULB Tirol mit dem Vorab-Abgleich ihrer Lizenzen mit Primo Central

Für einen Abgleich der von einer Einrichtung lizenzierten oder gedruckt abonnierten Titel mit den in Primo Central indexierten Ressourcen benötigt Ex Libris den so genannten Google-Scholar-Export aus SFX. Wenn eine Einrichtung nicht SFX-Kunde ist, kann ein tab-delimited Textfile mit den Feldern ISSN, Startjahr (der Lizenz) und Endjahr, wie z.B. die „Titelliste mit Lizenzzeiträumen“ aus der EZB-Administration, für den Datenabgleich herangezogen werden.

Als Ergebnis des Datenabgleichs erhielt die ULB Tirol die Information, dass 94 % der zur Überprüfung eingereichten Titel in Primo Central enthalten seien. Erst nach Rückfrage und Unterzeichnung eines Non-Disclosure-Agreements wurde uns eine Liste der in Primo Central enthaltenen Einzeltitel übersandt. Übermittelt wurde ISSN, E-ISSN und Titel, keine Angabe des Verlages und/oder der Ressource, aus der diese Zeitschrift übernommen wird. Es handelte sich um eine – 277 Seiten umfassende – PDF-Datei, deren Inhalt nicht in Excel übertragbar und daher nicht automatisiert weiterverarbeitbar ist. Eine (viel kürzere und dadurch leichter überblickbare) Liste der nicht enthaltenen Titel wurde nicht zur Verfügung gestellt.

Diese Liste ist sicherlich wertvoll, um nachzusehen, ob einzelne Titel von regionalem oder sehr fachspezifischem Interesse in Primo Central enthalten sind. Sie wurde auch für diesen Bericht bei der stichprobenartigen Erhebung des Abdeckungsgrades einzelner Ressourcen in Kap. 4.2.3 verwendet.

4.2 Inhaltliche Abdeckung

4.2.1 Fachbereiche, Sprachen, Dokumententypen

Beispielhafte Suchen aus den Fachbereichen Medizin, Umwelt, Wirtschaft, Psychologie, Pädagogik, Literatur, bildende Kunst, Musik und Recht haben Folgendes gezeigt:

- Zu Fragen aus allen Fachbereichen werden mehr als 100, oftmals mehr als 1000 Treffer gefunden – Primo Central deckt also auch die Geistes- und Sozialwissenschaften ab.
- Auch zu Kombinationen aus mehreren Suchwörtern werden mehrere 1000 Treffer gefunden, wobei allerdings bereits innerhalb der ersten 30, teils sogar bereits innerhalb der ersten 10 Treffer nicht mehr in allen Treffern alle Stichwörter in den angezeigten Metadaten vorkommen – d.h., dass die großen Treffermengen durch die Volltextindexierung und / oder durch Anzeige von Treffern, die nicht alle Suchbegriffe enthalten, entstehen.
- Suchen nach nicht-englischsprachigen Suchbegriffen erbringen deutlich weniger oder keine Treffer, was prinzipiell das Sprachenverhältnis von online-verfügbarem Material widerspiegelt. Fachdatenbanken verwenden darüber hinaus meist englische Verschlagwortung und Abstracts.

- Primo Central enthält mehrheitlich Zeitschriftenaufsätze, wesentlich seltener Buchbeiträge. Andere Dokumententypen desselben Anbieters (z.B. Statistiken, Wirtschaftsanalysen, Normen, Entscheidungen, Patente) konnten nicht gefunden werden.

4.2.2 Von Primo Central indexierte Informationsquellen (Stand 22.02.2011)

Die Mehrzahl der für Primo Central indextierten Ressourcen betrifft Verleger von wissenschaftlichen Zeitschriften. Dabei kommen die Metadaten nur zu einem geringen Teil vom Verlag selbst (26 Verlage). Die Metadaten der großen Verlage kommen von CrossRef (62 Verlage), über IngentaConnect kommen vor allem akademische Gesellschaften und kleine Verlage (228 Quellen). Hinsichtlich der von CrossRef gelieferten Metadaten werden 62 Verleger namentlich aufgezählt und dann wird noch „Rest of CrossRef collections, not specified“ genannt. Die bislang enthaltenen E-Book-Plattformen weisen fast ausschließlich Bücher als Ganzes nach.

Folgende Datenbanken wurden im Februar 2011 in Primo Central integriert: PubMed, ERIC, Academic One File, ECCO, General One File, Legal Trac, The Sabin Collection, Index to Hebrew Periodicals und Web of Science 1989 ff. PubMed und ERIC sind frei zugänglich. Gale stellt die Metadaten seiner Volltextdatenbanken in Primo Central frei zur Verfügung, nur für den Zugriff auf die enthaltenen Volltexte ist eine Lizenz notwendig. Der Index to Hebrew Periodicals und das Web of Science dürfen nur von Institutionen über Primo Central durchsuchbar gemacht werden, die diese Datenbanken lizenziert haben.

Bibliographische Datenbanken, die vorwiegend nicht-englischsprachige Zeitschriften nachweisen (z.B. deutsche und österreichische Rechtsdatenbanken, PSYINDEX, RSWB, wiso) sind nicht in Primo Central enthalten. Die von EBSCO Publishing hergestellten Volltextdatenbanken können seit 1. Jänner 2011 nicht mehr angeboten werden (Überlegungen zu Vor- und Nachteilen der Integration von Datenbanken in Discovery und Delivery Services siehe Kap. 5.2.3).

4.2.3 Abdeckungsgrad ausgewählter Ressourcen auf Einzeltitelebene

Aus der Ressourcenliste in der Administrationsoberfläche von Primo Central kann nur dann auf den Abdeckungsgrad lokaler Lizenzen in Primo Central geschlossen werden, wenn davon ausgegangen werden kann, dass von den in der Ressourcenliste von Primo Central genannten Anbietern jeweils sämtliche Zeitschriften (und zwar alle Jahrgänge) bzw. Bücher in Primo

Central enthalten sind. Dies könnte an Hand einer in Excel oder Access importierbaren Titelliste mit ISSNn bzw. ISBNn sehr leicht festgestellt werden. Eine derartige Liste wurde der AG trotz mehrfacher Nachfragen nicht zur Verfügung gestellt.

Daher mussten wir uns darauf beschränken, für zwei große Verlage, einige mittelgroße Verlage und einige akademische Gesellschaften stichprobenartig zu überprüfen, ob die ausgewählten – bevorzugt nicht dem STM-Bereich zugehörigen bzw. nicht englischsprachigen – Zeitschriften in Primo Central nachgewiesen sind und ob die Anzahl der gefundenen Aufsätze der Anzahl laut Verlags- oder Aggregatorplattform entspricht.

Die Abdeckung war mehrheitlich sehr gut. Alle Zeitschriften konnten gefunden werden, meist waren auch die geprüften Hefte vollständig nachgewiesen.

Primo Central ist weniger aktuell als die Anbieterplattformen selbst – so fehlt wiederholt das aktuelle Heft einer Zeitschrift, während andererseits Zeitschriftenaufsätze aufscheinen, die in Primo Central noch keine Bandzählung aufweisen (online ahead of print), laut Verlagsseite aber bereits einem Band/Heft zugeordnet sind. Das Problem liegt hier also offensichtlich in verzögerter Datenlieferung an Primo Central bzw. am Zeitaufwand für die Normalisierung der Daten.

4.3 Suche und Darstellung der Daten

4.3.1 Suchbarkeit

Ex Libris wirbt damit, dass der Primo Central Index die gesamte, auf der Plattform des jeweiligen Anbieters verfügbare Information nachweist – auch die dort verfügbaren Volltexte. Angezeigt werden in Primo Central in der Detailanzeige der Treffer dann aber nur die im vorigen Kapitel aufgelisteten Metadatenfelder – das könnte erklären, weshalb es immer wieder zu Treffern kommt, bei denen nicht klar ist, warum sie angezeigt werden bzw. warum sie unter den erstgereihten Treffern sind, obwohl sie in den angezeigten Metadaten nicht alle Suchbegriffe enthalten⁶.

Da sehr viele Primo-Central-Einträge aus CrossRef stammen – einer Plattform, die keine Volltexte, ja nicht einmal Abstracts enthält, – erscheint es unwahrscheinlich, dass eine Volltextindexierung derzeit bereits für viele Zeitschriften verfügbar ist. Auch PubMed und das Web of Science enthalten keine Volltexte. Wohl werden aber im Falle des Web of Science sämtliche Informationen in den Detailinträgen indiziert, auch die Titel der zitierenden und der zitierten Arbeiten.

In der Administrationsoberfläche besteht derzeit keine Möglichkeit, einzustellen, dass Treffer, die einen Suchbegriff nur im Volltext (oder nur in den Titeln der zitierten und zitierenden Arbeiten) enthalten, nicht berücksichtigt werden sollen. Auch den Benutzenden steht diese Möglichkeit über die Facettierung nicht zur Verfügung. Bei der einzelnen Suchabfrage kann man nur in der einfachen Suche in allen Feldern suchen oder in der erweiterten Suche auf eine reine Titelsuche einschränken und schließt damit aber dann die Suche in anderen inhaltsrelevanten Feldern (Schlagwortfeld, Abstract) aus.

4.3.2 Anzeige und Einflussmöglichkeiten auf die Normalisierung

Für die korrekte Anzeige der Treffer aus Primo Central in der lokalen Primo-Sicht ist es für Primo-AdministratorInnen erforderlich, zu wissen:

- a) welche Felder im PNX-Satz (Primo Normalized XML, das Datenformat in Primo) befüllt sind und daher im Frontend angezeigt werden müssen.
- b) welche Inhalte via Normalisierungsregeln in welche Felder gemappt werden. Auf das Mapping von Datensätzen aus Primo Central kann von den Primo-AdministratorInnen kein Einfluss genommen werden.
- c) welche Standardwerte in statischen Facetten (d.h. Facetten, in denen nur Werte aus einer vordefinierten Liste angezeigt werden) vorkommen können.

Ex Libris stellt diese Informationen im sog. „Primo Central Index Configuration Guide“ zur Verfügung. Die in dieser Anleitung enthaltenen Informationen wurden jedoch als unzureichend erfahren. Ex Libris ist das Problem bewusst und es wird bereits über mögliche Wege nachgedacht, Primo-AdministratorInnen mit mehr Informationen zu den PNX-Daten zu versorgen.

Die Qualität der Einträge ist je nach Anbieter sehr unterschiedlich und nicht bei allen Zeitschriften eines (von CrossRef) gelieferten Anbieters gleich. In Primo Central sind mehrheitlich nicht Verlagsdaten sondern CrossRef-Daten enthalten. Dadurch, dass viele dieser Zeitschriften aber gleichzeitig von PubMed und/oder Web of Science nachgewiesen werden, können den Benutzenden in Primo Central dennoch mehrheitlich aussagekräftige Einträge angezeigt werden.

4.3.3 Deduplizierung

Es gibt eine große Menge an Dubletten in Primo Central, z.B. Treffer aus CrossRef bzw. direkt vom Verleger und Treffer aus PubMed und/oder Web

of Science. Die Deduplizierung von Primo Central-Daten (in sich bzw. mit lokalen Daten) ist technisch nicht möglich.

Laut Auskunft von Ex Libris wird es eine Deduplizierung auf Basis des FRBR-Mechanismus⁷ geben, wobei die Daten des Publishers standardmäßig am höchsten gewichtet werden. Die Umsetzung ist für das 2. Quartal 2011 geplant.⁸

5. Primo Central und andere Dienste

5.1 Notwendigkeit von SFX bzw. eines beliebigen Link-Resolvers in Primo Central

Information über die Volltextlizenzen der jeweiligen Institution wird in Primo Central benötigt:

- a) Um mit einem Mausklick zwischen der Anzeige aller in Primo Central gefundenen Treffer und der Einschränkung auf die von der Einrichtung lizenzierten Zeitschriften bzw. Bücher wechseln zu können.

Damit im Scope „Lokaler Bestand“ auf die von einer Einrichtung lizenzierten Treffer eingeschränkt werden kann, muss Ex Libris eine URL genannt werden, von der eine Liste der lizenzierten Titel mit Angabe des Lizenzzeitraumes in regelmäßigen Abständen abgeholt werden kann. Diese Liste muss in einem bestimmten Format gehalten sein, damit sie von Primo Central interpretiert werden kann.

Wenn eine Einrichtung SFX als Link-Resolver verwendet, so kann diese Liste aus SFX generiert werden. Es handelt sich um den so genannten Google-Scholar XML-Export, der auf dem SFX-Server in einem bestimmten Verzeichnis abgelegt wird und Primo Central muss dann Zugriff auf dieses Verzeichnis des SFX-Servers erhalten. Diese Liste wird laut Angabe von Ex Libris wöchentlich abgeholt.

Auf die Frage, ob eine Einrichtung auch einen anderen Link-Resolver verwenden könne, oder, ob diese Liste auch in der EZB-Administration erzeugt werden könnte, wurde uns wiederholt beteuert, dass es nicht zwingend eine mittels SFX erzeugte Liste sein muss. Eine Einrichtung, die ohne SFX eine Einschränkung von Primo Central auf die lokalen Bestände durchführen möchte, möge die ihr vorliegende Holdingsliste an Ex Libris senden und man werde sich bemühen, eine Lösung zu finden. Eine derartige Lösung besteht aber wohl noch nicht; sie würde im Bedarfsfall erst erstellt.

- b) Um beim einzelnen Treffer eine Verfügbarkeitsanzeige bieten zu können – also möglichst direkt zum lizenzierten elektronischen Volltext verlinken zu können oder alternative Wege zum Originaldokument anzeigen zu können (Printbestand, Fernleihe).

Für die Verfügbarkeitsanzeige wird ein Link-Resolver benötigt. Hier erscheint es durchaus plausibel, dass auch andere Link-Resolver in der Lage sein sollten, aus den von Primo Central übermittelten Metadaten einen direkten Link zum Volltextdokument zu generieren.

Ein effizienter Einsatz von Primo Central bedarf eines aktuell und präzise gehaltenen Nachweises der elektronischen Holdings, idealerweise auch der Print-Bestände. Diese Arbeit muss von den einzelnen Einrichtungen geleistet werden.

5.2 Primo Central und wissenschaftliche Datenbanken

5.2.1 In Primo Central integrierte Datenbanken

PubMed und Web of Science sind in Primo Central integriert. Dadurch, dass diese Datenbankeinträge eine größere Erschließungstiefe aufweisen als CrossRef-Einträge der selben Zeitschrift, erhöhen sie die Trefferanzahl und die für die einzelnen Treffer verfügbaren Metadaten ganz wesentlich.

5.2.2 In Primo Central fehlende Datenbankinhalte

Nicht in Primo Central nachgewiesen sind die deutschsprachigen Zeitschriften aus wiso, RSWB und PSYNDEX, sowie alle Titel, die in beck Online, Rechtsdatenbank und LexisNexis Österreich enthalten sind. Recherchen zu österreichischem und deutschem Recht sind daher in Primo Central nicht möglich. Bei Recherchen aus Wirtschaft, Sozialwissenschaften, Bauingenieurwesen und Psychologie fehlen bei einer Suche in Primo Central die deutschsprachigen Veröffentlichungen. Da wiso und die Rechtsdatenbanken Volltextdatenbanken sind, kann somit auch ein Teil des Volltextbestandes österreichischer und deutscher Institutionen nicht über Primo Central nachgewiesen werden.

Primo Central weist ausschließlich Zeitschriftenaufsätze und Buchbeiträge nach. Faktendatenbanken (z.B. Reaxys / Beilstein, Firmen- und Finanzdatenbanken, Statistikdatenbanken, Bildarchive, Normen, Patente) müssen weiterhin separat durchsucht werden⁹.

5.2.3 Einträge aus Volltextdatenbanken und Fachbibliographien in Primo Central wünschenswert?

Die in (internationalen) Volltextdatenbanken enthaltenen Zeitschriften sind zu einem überwiegenden Teil wissenschaftliche Zeitschriften und daher zumindest als CrossRef-Eintrag, evtl. sogar als vom Verlag bereitgestellter Eintrag in Primo Central enthalten. Abonnenten dieser Datenbanken können aus einem von CrossRef oder vom Verlag stammenden Primo-Central-Treffer zum Volltext in der jeweiligen Volltextdatenbank verlinken.

Datenbankeinträge enthalten im Gegensatz zu CrossRef-Einträgen Abstracts und eine vom Datenbankhersteller durchgeführte Beschlagwortung. Sie sind somit besser auffindbar. Unter diesem Gesichtspunkt erscheint die Einbindung von Datenbanken in Primo Central erstrebenswert.

Andererseits ist die Erstellung und Beschlagwortung von Datenbankeinträgen zeitaufwändig, sodass Datenbanken weniger aktuell sind als die Verlagsplattformen.

Ex Libris sollte daher angeregt werden, mit den Verlagen und Aggregatoren direkt die Indexierung der von ihnen angebotenen Zeitschriften und E-Books zu verhandeln und diese nicht aus Datenbanken zu übernehmen.

5.3 Primo Central im Vergleich zu Google Scholar

Google Scholar erfasst ebenfalls wissenschaftliche Literatur. In den Trefferlisten sind deutlich mehr Einträge von Verlagsplattformen (auch für Elsevier, Wiley, Springer) zu finden als in Primo Central. Google Scholar enthält zusätzlich Einträge aus institutionellen Repositorien und Fachrepositorien, Primo Central nur arxiv.org. Zu PubMed-Einträgen wird verlinkt, Einträge aus kostenpflichtigen Datenbanken sind selbstverständlich nicht vorhanden.

Google Scholar erreicht eine gute Abdeckung bei STM-Verlagen. Bei Suchanfragen aus dem geistes- und sozialwissenschaftlichen Bereich werden trotz der inzwischen erfolgten Integration von JSTOR immer noch mehrheitlich Zitierungen von Büchern, Buchbesprechungen und Webseiten angezeigt.

Die Suchmöglichkeiten von Primo Central entsprechen denen von Google Scholar. Die Facettierung und die Querverweise sind zusätzliche Features von Primo Central. Auch in Google Scholar kann ein Link-Resolver eingepflegt werden. Es ist aber nicht möglich, nur die Treffer anzeigen zu lassen, die zum elektronischen Bestand einer Einrichtung gehören.

Google Scholar fasst jetzt bereits Dubletten zu einem Eintrag mit Links zu den als Versionen bezeichneten Einzeleinträgen zusammen¹⁰. Google weist zitierende Literatur nach, Primo Central nicht.

6. Schlussfolgerungen

6.1 Zusammenfassende Bewertung von Primo Central und Einordnung in das Rechercheangebot wissenschaftlicher Bibliotheken

Primo Central bildet für die BenutzerInnen wissenschaftlicher Bibliotheken einen One-Stop-Shop für die Recherche in den Beständen und lizenzierten Inhalten ihrer Einrichtungen. Fachbereiche, in denen vorwiegend in internationalen, englischsprachigen Publikationsorganen publiziert wird, sind gut abgedeckt. Fachbereiche, für die nationale – im Falle österreichischer Bibliotheken deutschsprachige – Publikationen wichtig sind (primär Rechtswissenschaften, aber auch Bauingenieurwesen und Teilbereiche der Wirtschaftskunde), müssen für eine Recherche nach unselbständiger Literatur weitere Quellen heranziehen.

Primo Central ist gut geeignet für eine Einstiegsrecherche und, wenn die sofortige Verfügbarkeit der gefundenen Literatur wichtig ist. Im Vergleich zu Fachdatenbanken fehlen die wissenschaftlichen, fachspezifischen Suchwerkzeuge. Im Vergleich zu einer Federated Search (z.B. MetaLib) fehlt die Möglichkeit, die Suche auf Kombinationen bestimmter Quellen einzuschränken.

Primo Central durchsucht die Quellen nicht in Echtzeit. Die in Primo Central indexierten Inhalte unterscheiden sich in ihrer Aktualität und in ihrem Umfang teilweise stark von den jeweiligen Quellen.

Trotz dieser Einschränkungen ist Primo Central eine wertvolle Erweiterung des Informationsangebotes wissenschaftlicher Bibliotheken. Der Unterschied zwischen den von einer Einrichtung angebotenen Rechercheplattformen muss den BenutzerInnen allerdings kommuniziert werden. Zu unterscheiden sind:

- Individuelle Fachdatenbanken: Sie bieten hochspezialisierte Suchmöglichkeiten über die Literatur eines Faches. Besonders bei interdisziplinären Fragestellungen müssen mehrere Fachdatenbanken durchsucht werden. Dies bedeutet einen hohen Arbeitsaufwand, eventuell einen hohen Anteil an Dubletten, aber maximale Qualität des Ergebnisses. Die Trefferliste kann nicht auf die an einer Einrichtung verfügbaren Volltexte eingeschränkt werden.

- Federated Search: Eine Federated Search Engine (z.B. MetaLib) bietet nur einfache Suchmöglichkeiten, aber Informationsquellen können zu fachlich relevanten (oder fachübergreifenden) Subsets gebündelt und mit einer Suchanfrage durchsucht werden. Die Suche selbst ist langsam. Die Quellen werden aber in Echtzeit durchsucht, MetaLib ist daher genauso aktuell wie die Quellen selbst. Die Trefferliste kann nicht auf die an einer Einrichtung verfügbaren Volltexte eingeschränkt werden.
- Primo Central: Vermöge der Vorab-Indexierung der Quellen ist die Suche in Primo Central so schnell wie die Suche in den Datenbanken selbst, jedoch sind die Suchergebnisse unter Umständen weniger aktuell. Es gibt nur sehr einfache Suchmöglichkeiten und keine Möglichkeit der Einschränkung der Suche auf fachlich relevante Quellen. Die Trefferliste kann auf die an einer Einrichtung verfügbaren Volltexte eingeschränkt werden. Allerdings ist es für eine fundierte wissenschaftliche Literaturstudie nicht ausreichend, nur das zu berücksichtigen, was an der eigenen Einrichtung (gedruckt oder in elektronischer Form) vorhanden ist.

6.2 Offene Fragen

Folgende Fragen zu Primo Central konnten im Rahmen der Tätigkeit der AG nicht geklärt werden:

6.2.1 Kosten von Primo Central

Primo Central ist ein kostenpflichtiger Dienst. An Ex Libris ist eine einmalige Freischaltungsgebühr (für den Ex Libris entstehenden Arbeitsaufwand) und eine jährliche Nutzungsgebühr zu entrichten. Für diese Gebühr erwirbt die Einrichtung das Recht, in ihrer Primo- oder MetaLib-Instanz den Primo Central Index zu durchsuchen. Es werden keinerlei Datensätze aus Primo Central dauerhaft auf den zentralen Primo-Server der jeweiligen Bibliothek bzw. des jeweiligen Bibliothekenverbundes übertragen.

Die Einholung von Preisangeboten für Primo Central gehörte nicht zu den Aufgaben der AG, daher können hiezu keine genaueren Angaben gemacht werden.

6.2.2 Nachhaltige Verfügbarkeit der derzeit in Primo Central indexierten Quellen

Die Rücknahme der EBSCO-Datenbanken aus Primo Central mit 1. 1. 2011 hat aufgezeigt, dass die nachhaltige Verfügbarkeit der derzeit in Primo

Central indexierten Quellen nicht selbstverständlich ist (zur Notwendigkeit/Verzichtbarkeit von Volltextdatenbanken und bibliographischen Datenbanken in Primo Central siehe Kap. 5.2.3).

Weder die AG PrimoImplementierung noch Ex Libris können abschätzen, welche Anbieter in Zukunft Metadaten für Primo Central zur Verfügung stellen werden. Das selbe gilt für vergleichbare Discovery und Delivery Services anderer Hersteller.

6.2.3 Sichtbarkeit der Katalogdaten angesichts der Treffermengen aus Primo Central

Die via Primo Central verfügbare Menge an Datensätzen wird sofort jeden herkömmlichen Katalog fluten. Der am häufigsten im Katalog gesuchte Bestand (Lehrbücher, Monographien, Nachschlagewerke) wird zwischen hochwissenschaftlicher Literatur versteckt. Es muss an jeder Institution geklärt werden, für welche primäre BenutzerInnengruppe Primo Central angeboten wird: Werden FachwissenschaftlerInnen Primo Central nutzen? Welchen Wert hat hochwissenschaftliche Literatur für Studierende? Derzeit müssen sich alle an einer zentralen Primo-Instanz teilnehmenden Einrichtungen gemeinsam auf eine Strategie einigen, falls sie den herkömmlichen Katalogbestand verstärkt sichtbar machen möchten. Das kann zu Interessenskonflikten führen, wenn Einrichtungen unterschiedlichen Typs (Universaluniversitäten, Fachuniversitäten, National- und Landesbibliotheken) sich eine zentrale Instanz teilen.

6.3 Wünschenswerte Erweiterungen / Verbesserungen von Primo Central

6.3.1 Inhaltliche Erweiterungen

1. Einbinden neuer Quellen, insbesondere Buch-Quellen, Reihenwerke, deutschsprachige Verlage.
2. Die Erweiterung der Inhalte durch weiter zurückreichende Quellen.
3. Einbinden von Ressourcen mit inhaltlicher Erschließung (Keywords, Thesauri, Abstracts). Insbesondere Wechsel von CrossRef-Datensätzen zu originären Verlagsquellen.
4. Durch schnelle, häufige Aktualisierung müssen rezente Artikel (insbes. online early u.dgl.) bzw. auf Oberflächen neu verfügbare ältere Inhalte schnell und standardisiert verfügbar sein.
5. Bessere und tiefere Indizierung von Reihen- und Buchpublikationen. So ist z.B. ist ein Großteil der Buchpublikationen mangelhaft bzw. nicht auf Beitragsebene indiziert.

6. Es bedarf auch einer nachvollziehbaren und maschinell verarbeitbaren Dokumentation aller nachgewiesenen Bestände und ihrer Aktualisierungen.

6.3.2 Verbesserung und Standardisierung der Datenqualität

1. Wenn unterschiedliche Datensätze auf der gleichen Plattform mit identischer Qualität vorhanden sind, sollten diese auch in Primo Central so wiedergegeben werden.
2. Bessere Mechanismen zur Deduplizierung müssen implementiert werden. Identische Artikel sollten unter einem Treffer angezeigt werden.
3. Die Unterstützung von Unicode ist noch mangelhaft, weshalb es häufig zu falscher Darstellung von Sonderzeichen, sichtbaren HTML-Tags u.Ä. kommt.

6.3.3 Erweiterung der Suchmöglichkeiten und der Facettierung

1. Facettenumfang und -anzeige müssen erweitert werden.
2. Möglichkeit, die Suche in indizierten Volltexten, Literaturlisten u.Ä. zu deaktivieren. Im Idealfall können BenutzerInnen selbst die Volltextsuche an/abwählen.
3. Die Suchbarkeit von Zeitschriftentiteln, Reihen (= übergeordneten Datensätzen) v.a. um die Datenvollständigkeit zu überprüfen, aber auch als Service für BenutzerInnen.

7. Fazit

Primo Central erweitert die in Primo suchbare Literatur um ein Vielfaches. Vor allem durch die Einbindung unselbstständiger Zeitschriftenliteratur in eine gemeinsame Suchoberfläche ist eine enorme Service-Verbesserung für die BenutzerInnen zu erwarten. Eine signifikante Erhöhung der Nutzung elektronischer Ressourcen ist die zu erwartende Folge.

An Primo Central interessierten Einrichtungen wird empfohlen, vorab durch einen Abgleich ihrer Bestände mit den in Primo Central indizierten Inhalten zu klären, ob ein genügender Teil ihrer Bestände bereits in Primo Central enthalten ist.

Der Implementierungs- und Wartungsaufwand für Primo Central selbst ist gering. Voraussetzung für eine optimale Nutzung der Funktionalitäten von Primo Central ist jedoch die Integration eines Link-Resolvers und die

kontinuierliche Aktualisierung der in diesem Link-Resolver nachzuweisenden Lizenzen der jeweiligen Einrichtung – eine Tätigkeit, die an der Primo Central einsetzenden Einrichtung durchgeführt werden muss.

Dr.ⁱⁿ Eveline Pipp
Universitäts- und Landesbibliothek Tirol
Abteilung Datenbanken & Neue Medien
Innrain 52f
A-6020 Innsbruck
Telefon: +43 512 507 2494
E-Mail: eveline.pipp@uibk.ac.at

Mag. Markus Knitel
Die Österreichische Bibliothekenverbund und Service Gesellschaft m.b.H.
Brünnlbadgasse 17/2a (Zweigstelle Bennogasse 26/1)
A-1090 Wien
Telefon: +43 1 4035158-55
E-Mail: markus.knitel@obvsg.at

Wolfgang Mayer
Universitätsbibliothek Wien
eResource Management
Dr. Karl Lueger-Ring 1
A-1010 Wien
Telefon: +43 1 42 77-15020
E-Mail: wolfgang.mayer@univie.ac.at

Mag.^a(FH) Michaela Putz
Universitätsbibliothek Wien
Teinfaltstraße 8
A-1010 Wien
Telefon: +43-1-4277-153 13
E-Mail: michaela.putz@univie.ac.at

- 1 Die Mitglieder AG Primo-Implementierung: Viktor Babitchev (Österreichische Bibliothekenverbund und Service GmbH (OBVSG)), Wolfgang Hamedinger (OBVSG), Markus Knitel (OBVSG), Wolfgang Mayer (UB Wien), Peter Pilsl (UB Salzburg), Eveline Pipp (ULB Tirol), Michaela Putz (UB Wien), Albert Schnitzer (UB Graz), Georg Stern-Erlebach (ULB Tirol), Dirk Willinghöfer (ÖNB). Die AutorInnen danken den weiteren Mitgliedern der Arbeitsgruppe für Ihre Mitarbeit bei der Produkt-evaluierung und für Diskussionsbeiträge zum Bericht.
- 2 Die AutorInnen danken der Verbundvollversammlung für die Anregung zur Publikation dieses Berichtes.
- 3 http://www.exlibrisgroup.com/de/category/Primo_Central.
- 4 Bezüglich der in Primo Central enthaltenen Datenbanken siehe Kap. 4.2.2.
- 5 Wird von Ex Libris in einer Cloud-Computing-Umgebung gehostet.
- 6 In Primo Central werden die Treffer nach Relevanz gereiht (eine Umsortierung nach Erstautor, Titel und Publikationsjahr kann gewählt werden). Ex Libris hat bisher keine Beschreibung des Ranking-Algorithmus publiziert und rät davon ab, zu versuchen, für eine bestimmte Instanz diesen Algorithmus zu modifizieren.
- 7 = Gruppierung von Datensätzen. Für Dubletten erfolgt nur noch EIN Eintrag in der Trefferliste mit der Angabe „Es gibt x Versionen dieses Eintrages“. Klickt man auf diesen Link, so erscheinen die diesbezüglichen Dubletten in einem separaten Fenster, wie z.B. in Google Scholar.
- 8 Die Umsetzung erfolgte im Rahmen des wenige Tage vor Einreichung dieses Beitrages verfügbar gewordenen Service Packs.
- 9 Diese Datenbanken sind auch mehrheitlich nicht über Federated-Search-Engines, wie z.B. MetaLib durchsuchbar.
- 10 Mittels des von Ex Libris zwischenzeitlich umgesetzten FRBR-Mechanismus; siehe Kap. 4.3.3.