Vom Zitat zum Volltext: Der ReDI-Lösungsansatz im Vergleich zu anderen Linkingsystemen

Bernd Oberknapp, UB Freiburg

Die Regionale Datenbank-Information Baden-Württemberg (ReDI)[1] bietet den teilnehmenden Einrichtungen neben dem Hosting von Datenbanken und der Einbindung der Datenbanken in einen einheitlichen Rahmen auch so genannte Mehrwertdienste an. Der wichtigste dieser Dienste ist die ReDI-Bestandsabfrage, mit der im Anschluss an die Recherche festgestellt werden kann, ob der gewünschte Titel in elektronischer oder gedruckter Form verfügbar ist oder gegebenenfalls bestellt werden kann. Die ReDI-Bestandsabfrage ermöglicht damit eine Integration des Datenbankangebots mit anderen Dienstleistungen der Bibliotheken.

Die ReDI-Bestandsabfrage ist Kontext-sensitiv und OpenURL-kompatibel, die Volltextverlinkung erfolgt bisher aber nur über die Elektronische Zeitschriftenbibliothek (EZB)[2] und damit nur auf Zeitschriftenebene. Mit der Einbindung von CrossRef[3] und den inzwischen um die Jahrgangs- bzw. Heftinformationen erweiterten Bestandsangaben aus der EZB sollte es möglich sein, auch eine Volltextverlinkung auf Artikelebene über die ReDI-Bestandsabfrage zu realisieren. Dieser Lösungsansatz hat im Vergleich zu Linkingsystemen wie SFX[4], EBSCO LinkSource[5] oder Ovid's LinkSolver[6] den großen Vorteil, dass die aufwendige Pflege von Bestandsdaten im Linkingsystem entfällt. Erste Tests zur Einbindung von CrossRef sind viel versprechend verlaufen.

1. Kontext-sensitives Linking

Die Anzahl der von den Bibliotheken lizensierten oder frei im Internet verfügbaren elektronischen Ressourcen ist in den letzten Jahren stark angestiegen. Den Wissenschaftlern und Studierenden stehen heute eine Vielzahl von

- Datenbanken,
- elektronischen Zeitschriften,
- Dokumentenservern (Dissertationen, Preprints, ...),
- Katalogen,
- (elektronischen) Lieferdiensten und
- sonstigen Internet-Ressourcen

wie zum Beispiel Suchmaschinen zur Verfügung, deren Nutzung aber dadurch erschwert wird, dass sie über verschiedenste, weltweit verteilte Systeme verstreut sind. Ein Ziel der Bibliotheken muss daher sein, die verfügbaren Ressourcen soweit wie möglich durch *Reference Linking* zu verknüpfen und den Nutzern so eine einfache Navigation in den für sie relevanten Ressourcen zu ermöglichen.

Unter dem Begriff Reference Linking werden dabei alle Dienste zusammengefasst, die den Nutzer ausgehend vom Zitat (bibliographischen Daten) zu weiteren, im Zusammenhang mit dem Zitat relevanten Informationen leiten. Dazu zählen insbesondere die Links “vom Zitat zum Volltext”, die im Folgenden hauptsächlich betrachtet werden, aber auch Dienste wie “Wo wird dieser Autor oder dieser Aufsatz zitiert?” oder “Was hat dieser Autor noch veröffentlicht?”, die zum Beispiel über Links, die Suchanfragen in Web of Science oder anderen geeigneten Datenbanken starten, implementiert werden können. Welche Links dem Nutzer zu einem Zitat angeboten werden können, ist vom *Kontext* abhängig:

- Daten: Welche bibliographischen Daten stehen zur Verfügung? Ein Link zur Elektronischen Zeitschriftenbibliothek (EZB)[2] ist zum Beispiel nur dann sinnvoll, wenn es sich bei dem Zitat um einen Aufsatz handelt.
- Rechte: Welche lizenzpflichtigen Ressourcen stehen für den Nutzer zur Verfügung? Dies hängt üblicherweise davon ab, zu welcher Benutzergruppe der Nutzer gehört.
- Technik: Auf welche Ressourcen kann der Nutzer zugreifen? Auf einige Datenbanken wie Beck-Online und Juris kann der Nutzer zum Beispiel nur zugreifen, wenn er sich auf dem Campus befindet, da eine Nutzung über Einwahlzugänge nicht gestattet ist.

Werden diese Abhängigkeiten und Beschränkungen beim Linkangebot berücksichtigt (zumindest teilweise, die technischen Voraussetzungen für die Nutzung der Ressourcen zu prüfen ist unter Umständen schwierig oder sogar unmöglich), so spricht man von *Kontext-sensitivem Linking*.

Voraussetzung dafür, dass zwei Ressourcen durch Reference Linking verknüpft werden können, sind entsprechende Schnittstellen: Das Ausgangssystem (Source) muss zumindest die Definition eines Links pro Zitat erlauben, über den die bibliographischen Daten oder Teile davon an das Zielsystem (Target) weitergeben werden können. Das Zielsystem muss eine Schnittstelle anbieten, über die die bibliographischen Daten übergeben werden können und die gewünschte Aktion angestoßen werden kann. Wenn die Systeme OpenURL unterstützen – OpenURL ist *der* Standard zum Austausch von bibliographischen Daten über URLs[7] –, vereinfacht das die Verknüpfung deutlich, in diesem Fall muss im Idealfall nur die Basis-OpenURL des Zielsystems im Ausgangssystem angegeben werden.

2. ReDI-Bestandsabfrage

Die meisten in ReDI eingebundenen Datenbanksysteme mit Weboberfläche bieten inzwischen Schnittstellen an, die ein Reference Linking ermöglichen (zumindest mit den Datenbanken als Ausgangssystem), wobei seit etwa zwei Jahren in zunehmendem Maße OpenURL unterstützt wird. Die vielen Windows basierten Datenbanken, die über ReDI angeboten werden (Stand Februar 2004: etwa 250), bieten aber leider bis auf ganz wenige Ausnahmen keine entsprechenden Schnittstellen an, so dass eine Einbindung in das Reference Linking praktisch nicht möglich ist.

Rückblick

Beim Betriebsstart von ReDI 1999 war es nur bei zwei der eingesetzten Systeme (OvidWeb und WebSPIRS) möglich, eigene Links zu definieren, wobei lediglich auf die ISSN oder ISBN der Zitate zurückgegriffen werden konnte. Zunächst wurden den Nutzern nur einige Links zu den wichtigsten Katalogen angeboten, wobei das Linkangebot für alle an ReDI teilnehmenden Einrichtungen einheitlich war.

Der Versuch, die gewünschten Links für jede an ReDI teilnehmende Einrichtung in jedem Datenbanksystem individuell zu konfigurieren, hätte zu einer sehr komplexen Konfiguration geführt: Bei im Schnitt fünf Katalogen pro Einrichtung (lokaler Katalog, SWB-Katalog, KVK, ZDB und EZB) und mehr als 40 Einrichtungen hätten pro Datenbanksystem mehr als 200 Links konfiguriert werden müssen! Und mit jedem weiteren Datenbanksystem (Web of Science, EBSCO, CSA, ...) und jedem zusätzlichen Link hätte sich die Gesamtzahl der Links drastisch erhöht. Abgesehen vom Pflegeaufwand schied eine solche Lösung auch schon deshalb aus, weil in einigen Datenbanksystemen (zum Beispiel Web of Science) überhaupt nur ein einziger Link definiert werden kann.

Mit der ReDI-Bestandsabfrage, die ab Mitte 2000 in einer ersten Version zur Verfügung stand, konnten diese Probleme weitgehend gelöst werden. Die ReDI-Bestandsabfrage ist Kontext-sensitiv, sie

- zeigt automatisch das Ergebnis der Suche im für die Einrichtung des Nutzers voreingestellten Katalog an,
- bietet dem Nutzer weitere, individuell pro Einrichtung konfigurierbare Kataloge, Lieferdienste und andere Ressourcen an und
- zeigt bei Zeitschriften automatisch das Ergebnis der Suche in der EZB für die Einrichtung des Nutzers an und vermittelt so den Zugang zu den elektronischen Zeitschriften

(siehe Abbildung 1). In den Datenbanksystemen selbst muss nur einmalig pro Einrichtung ein Link definiert werden, der die ReDI-Bestandsabfrage aufruft und dabei die bibliographischen Daten des Zitats übergibt.

Dieser Lösungsansatz hat mehrere entscheidende Vorteile:

- Für den Nutzer steht damit in allen Datenbanken, in denen ein Link zum Aufruf der ReDI-Bestandsabfrage definiert werden kann, ein einheitlicher Weg “vom Zitat zum Volltext” zur Verfügung.
- Die Bibliotheken können ihren Nutzern für diese Datenbanken jeweils alle gewünschten Dienste anbieten – unabhängig davon, ob das Datenbanksystem diese Dienste direkt unterstützt oder nicht.
- Die Administration der Links zu den Ressourcen, die den Nutzern angeboten werden sollen, wird drastisch vereinfacht – die Konfiguration und Pflege erfolgt zentral an nur einer Stelle, nämlich in der ReDI-Bestandsabfrage.

Das Konzept, die Links zu den Ressourcen (Zielsysteme) unabhängig von den Datenbanksystemen (Ausgangssysteme) in einer zentralen Komponente anzubieten und zu pflegen, entspricht dem eines zentralen Linkresolvers (siehe Abschnitt 4).

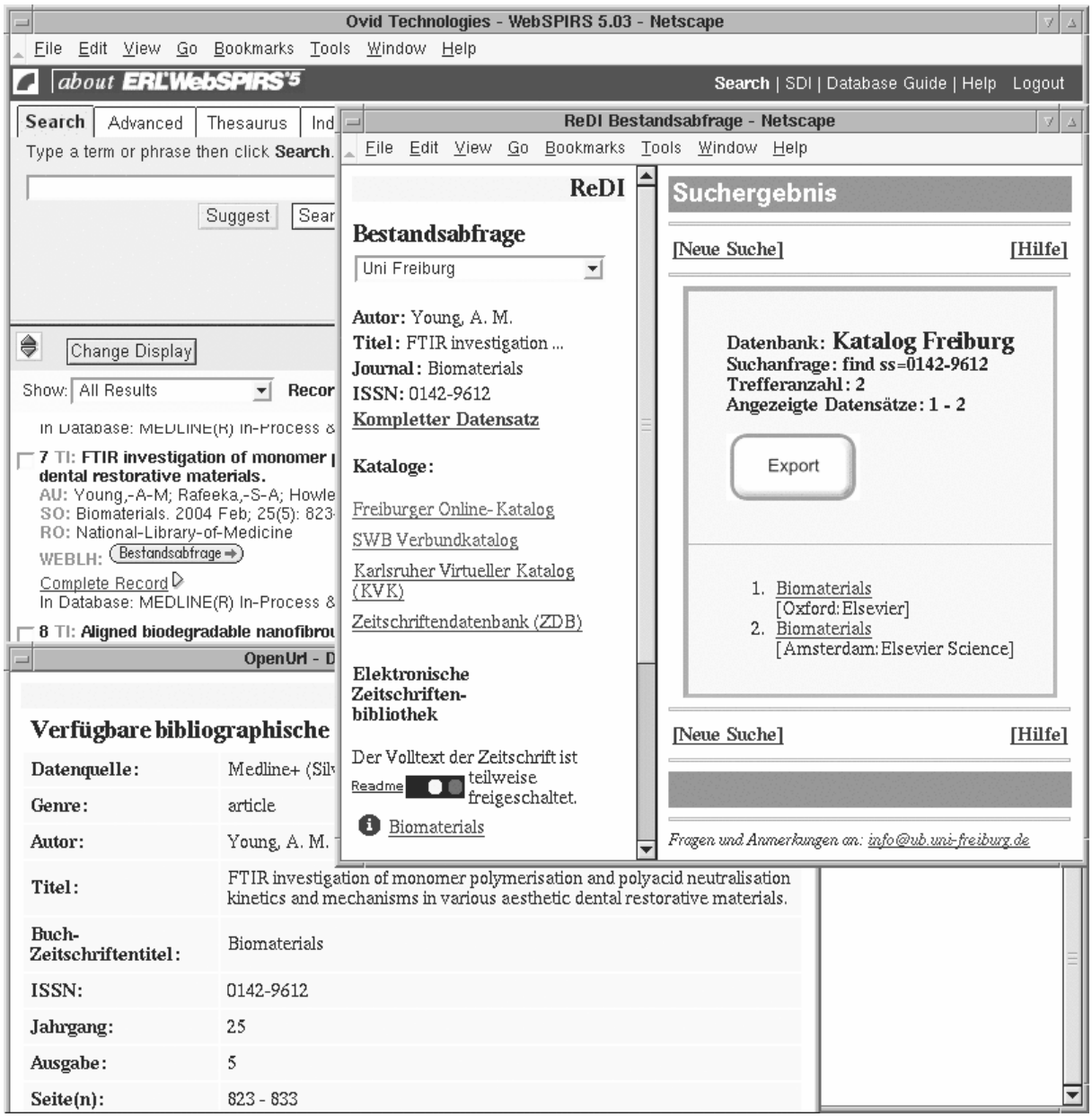

Abbildung 1:
Die ReDI-Bestandsabfrage wird über den entsprechenden Link in der Datenbank (im Hintergrund) aufgerufen, dabei werden die bibliographischen Daten möglichst vollständig aus der Datenbank in die Bestandsabfrage übernommen (unten links). Es wird das Ergebnis einer Suche im für die Einrichtung voreingestellten Katalog angezeigt, bei Zeitschriften auch das Ergebnis der Suche in der EZB, außerdem werden weitere Kataloge angeboten (oben rechts).

Die ReDI-Bestandsabfrage konnte zunächst nur die ISSNs und ISBNs der Zitate aus den Datenbanken übernehmen – mehr konnten die meisten Datenbanksysteme am Anfang aber auch nicht liefern – und damit nur einen Teil der gewünschten Dienste anbieten. Nachdem sich eine immer breitere Unterstützung von OpenURL[7] abzeichnete und die Datenbankanbieter begannen, OpenURL-kompatible Schnittstellen in ihren Produkten anzubieten, war eine entsprechende Erweiterung der ReDI-Bestandsabfrage die logische Konsequenz.

Aktueller Stand

Um ein Kontext-sensitives Reference Linking auf Basis von OpenURL über die ReDI-Bestandsabfrage realisieren zu können, waren nicht nur eine OpenURL-Schnittstelle zur Übernahme der bibliographischen Daten erforderlich, sondern vor allem auch Komponenten zur Aufbereitung der Daten und zur Bereitstellung von neuen Diensten auf Basis dieser Daten. Während die Implementation einer Schnittstelle zur Übernahme von Daten im OpenURL-Format eine eher triviale Aufgabe ist, kann die Aufbereitung der Daten sehr aufwendig sein: Der Aufwand variiert je nach Qualität der OpenURL-Unterstützung durch das Ausgangssystem von nicht vorhanden (strikte Einhaltung des OpenURL-Standards), über gering bis mäßig (einzelne Felder müssen nachbearbeitet werden, typischerweise die Autorenfelder) bis hoch (Ausgangssystem kann die Quelldaten nicht aufgeschlüsselt in Titel, Jahrgang, Heft und Seiten liefern). Auch in letzterem Fall können die Daten per OpenURL übergeben werden, und zwar im Private Identifier bzw. als Private Data. Die ReDI-Bestandsabfrage als Zielsystem muss diese Daten dann entsprechend interpretieren und aufbereiten, was besonders aufwendig ist, wenn die Quellangaben im Ausgangssystem nicht einheitlich erfasst sind.

Ab Ende 2002 wurde nach und nach die Anbindung der ReDI-Bestandsabfrage an die einzelnen Datenbanksysteme auf OpenURL umgestellt, und es wurden die notwendigen Komponenten zur Aufbereitung der Daten erstellt. Bis auf OvidWeb – die Umstellung hat sich wegen der Migration der Ovid-Datenbanken von den ReDI-eigenen Ovid-Servern auf den Online-Server und der Einführung des neuen Ovid-Linkingsystems Links@Ovid verzögert – sind inzwischen alle Datenbanksysteme, die eine entsprechende Schnittstelle zur Definition von Links bieten (Ovid/ERL-WebSPIRS, CSA, EBSCOhost, ISI/Web of Knowledge, FIZ-Technik, GBI/WISO-Net, ABC-CLIO, ...) auf OpenURL umgestellt.

Die aktuelle Version der ReDI-Bestandsabfrage übernimmt jetzt alle bibliographischen Daten, die die Datenbanksysteme zu einem Zitat per OpenURL bereitstellen können (siehe Abbildung 1). Damit können im Vergleich zur alten ISSN/ISBN-basierten ReDI-Bestandsabfrage erweiterte und neue Dienste angeboten werden wie

- Volltextverlinkung auf Artikelebene (siehe unten)
- Übergabe der bibliographischen Daten an Bestellsysteme: Damit kann das Bestellformular automatisch ausgefüllt werden, was dem Nutzer das Übertragen der Daten per Hand oder per Copy&Paste erspart. Die Anbindung der vorhandenen lokalen Bestellsysteme ist in Arbeit, die Einbindung der im Aufbau befindlichen elektronischen Fernleihe ist geplant.
- Erweiterte Recherchemöglichkeiten in Katalogen: Bisher erfolgt die Suche nur auf Basis von ISSN/ISBN. Steht keine ISSN/ISBN zur Verfügung oder ergibt die Suche keine Treffer, so kann zum Beispiel auch nach dem Zeitschriftentitel oder nach Titel und/oder Autor gesucht werden. Eine erste Testversion einer erweiterten Suche ist bereits für einige Kataloge implementiert.
- “Wo wird dieser Autor oder dieser Aufsatz zitiert?”: Suche in Web of Science.
- “Was hat dieser Autor noch veröffentlicht?”: Suche in geeigneten Datenbanken – wobei vom Zitat bzw. vom Ausgangssystem (Kontext) abhängig ist, welche Datenbanken geeignet sind.
- “Elektronisches Regal”: Die Zitate, für die die ReDI-Bestandsabfrage aufgerufen wird, könnten für eine spätere Weiterverarbeitung – Datenbank übergreifend – in einer Merkliste gespeichert werden.

Die Volltextverlinkung auf Artikelebene ist sicher der mit Abstand wichtigste der neuen Dienste, sie ist aber auch mit Abstand am schwierigsten zu realisieren. In den folgenden beiden Abschnitten wird ein Überblick über die dabei auftretenden Probleme und die verschiedenen Lösungsansätze gegeben.

3. Volltextverlinkung auf Artikelebene

Neben bibliographischen Daten sind weitere Informationen erforderlich, um eine Volltextverlinkung auf Artikelebene realisieren zu können. Benötigt werden:

- Aufsatzdaten: ISSN, Titel, Autor, Jahrgang, Heft, Seite. Die Aufsatzdaten müssen dabei nicht unbedingt vollständig vorliegen, der Aufsatz muss nur anhand der vorliegenden Daten eindeutig zu identifizieren sein.
- Bestandsangaben: Ist der Aufsatz elektronisch verfügbar und wenn ja, bei welchen Anbietern? Bei der Beantwortung dieser Frage müssen die Rechte des Nutzers berücksichtigt und das so genannte "appropriate copy"-Problem gelöst werden – die Volltextverlinkung muss also Kontext-sensitiv sein.
- Linkinginformationen: Wie lauten die Volltext-URLs für den Aufsatz?

Fehlen notwendige Aufsatzdaten oder Linkinginformationen, so lassen sich die Volltext-URLs nicht ermitteln. Fehlen Bestandsangaben, so kann zwar unter Umständen eine Volltext-URL ermittelt werden, es kann aber nicht festgestellt werden, ob der Nutzer Zugriff auf den Volltext hat. Die typischen Probleme sind:

1) Ein Teil der Aufsatzdaten fehlt oder die Aufsatzdaten sind nicht hinreichend aufgeschlüsselt.
2) Die Bestandsangaben sind nicht vollständig erfasst oder nicht bzw. nur teilweise elektronisch auswertbar.
3) Die Volltext-URLs sind je nach Anbieter unterschiedlich aufgebaut und teilweise nicht aus den Aufsatzdaten zu ermitteln.

Problem 1 lässt sich eventuell durch Ergänzung der fehlenden oder nicht aufgeschlüsselten Daten aus einer anderen Quelle lösen. Dies kann entweder automatisch oder unter Mithilfe des Nutzers geschehen, zum Beispiel durch Suchanfragen an andere Datenbanken, in denen der Aufsatz auch nachgewiesen sein könnte – wie beim Dienst "Was hat dieser Autor noch veröffentlicht?", in beiden Fällen wird dieselbe Funktionalität benötigt. Liegen die Daten nicht hinreichend aufgeschlüsselt vor, so kann man versuchen den Anbieter zu überzeugen, die Daten in einem geeigneteren Format zur Verfügung zu stellen, oder die Daten mit entsprechend hohem Aufwand selbst aufbereiten (siehe vorheriger Abschnitt).

Nicht oder nur teilweise elektronisch auswertbare Bestandsangaben fallen für ein Programm auch unter "nicht oder nur teilweise vorhanden", so dass bei Problem 2 nur diese beiden Fälle zu unterscheiden sind. Liegt nur ein Teil der Bestandsangaben vor – typischerweise fehlt die Angabe, für welchen Zeitraum eine Zeitschrift lizensiert ist –, so kann zumindest festgestellt werden, bei welchen Anbietern, für die Volltext-URLs für den gewünschten Aufsatz ermittelt werden konnten, der Nutzer Zugriff auf den Volltext haben könnte. Die entsprechenden Volltext-URLs können dem Nutzer dann angeboten werden, es sollte aber darauf hingewiesen werden, dass der Zugriff eventuell nicht funktioniert. Liegen gar keine Bestandsangaben vor, so bleibt nur die Möglichkeit, dem Nutzer alle ermittelten Volltext-URLs anzubieten, wobei aber darauf hinwiesen werden sollte, dass nicht festgestellt werden konnte, ob über eine der URLs der Zugriff auf den Volltext möglich ist.

In der Elektronischen Zeitschriftenbibliothek (EZB)[2] können mit dem erweiterten Datenformat inzwischen auch die Lizenzzeiträume erfasst werden, die Daten werden dabei so weit wie möglich in elektronisch auswertbarer Form gespeichert. Mit den Bestandsangaben aus der EZB sollte es zukünftig also möglich sein festzustellen, ob und wenn ja über welche Anbieter der Nutzer Zugriff auf den Volltext des gewünschten Aufsatzes hat (von der Frage der technischen Probleme beim Zugriff zum Beispiel durch IP-Beschränkungen einmal abgesehen).

Alle möglichen Volltext-URLs zu einem Aufsatz zu ermitteln, ist ein äußerst schwieriges Problem. Die Volltext-URLs sind je nach Anbieter unterschiedlich aufgebaut und lassen sich teilweise nicht zuverlässig oder sogar gar nicht aus den Aufsatzdaten ermitteln. Besonders problematisch ist, wenn der direkte Zugriff auf den Volltext nur über eine interne Identifikationsnummer der Anbieters, die nicht oder nur durch ein aufwendiges Verfahren aus den Aufsatzdaten ermittelt werden kann, möglich ist. Das Problem 3 lässt sich daher praktisch nur mit Hilfe einer *Link-*

datenbank lösen, wie dies bei den am Markt verfügbaren Linkingtools und Linkresolver geschieht (siehe nächster Abschnitt). Die Linkdatenbank enthält dabei sowohl die bibliographischen Daten der Aufsätze als auch die zugehörigen Volltext-URLs, so dass die Volltext-URLs über eine Datenbankabfrage mit den bibliographischen Daten ermittelt werden können. Vollständigkeit zu erreichen, also zu einem Aufsatz jeweils alle Volltext-URLs zu ermitteln, ist aber auch mit einer Linkdatenbank praktisch nicht möglich. Kann für einen Anbieter, über den der Nutzer Zugriff auf den Volltext hat, keine Volltext-URL ermittelt werden, so sollte der Nutzer darauf hingewiesen und statt dessen ein Link zum Heft, in der sich der Aufsatz befindet, zur Homepage der Zeitschrift oder zumindest zur Webseite des Anbieters angeboten werden.

4. Linkingsysteme

Bei den Linkingsystemen, die eine Volltextverlinkung auf Artikelebene bieten und so ein Kontext-sensitives Reference Linking ermöglichen, kann grundsätzlich zwischen zwei Typen unterschieden werden: In die Datenbanksysteme der Anbieter integrierte Linkingtools wie zum Beispiel ISI-Links, EBSCO Custom/SmartLinks oder Links@Ovid (siehe Abbildung 2) und zentrale Linkresolver wie SFX von Ex Libris[4], die per OpenURL in die Datenbanksysteme – und natürlich auch andere Ausgangssysteme – eingebunden werden (siehe Abbildung 3). Inzwischen gibt es mit Linkingsystemen wie EBSCO LinkSource[5] und Ovid's LinkSolver[6] auch Mischformen, d.h. einerseits in Datenbanksysteme (EBSCOhost bzw. OvidWeb und WebSPIRS) integrierte, andererseits aber auch unabhängig davon als zentraler Linkresolver nutzbare Linkingsysteme, die entsprechend auch als separates Produkt lizensiert werden können.

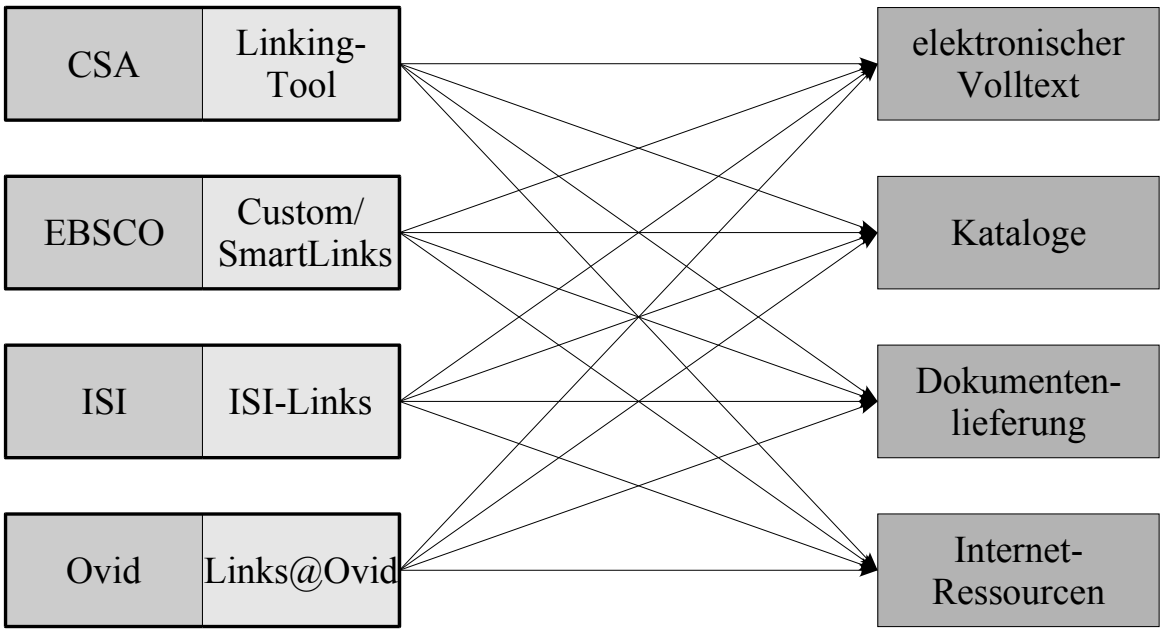

Abbildung 2:
In Datenbanksysteme integrierten Linkingtools bieten den Nutzern zwar die “direktesten” Links, die Konfiguration ist aber im Vergleich zu zentralen Linkresolvern (siehe Abbildung 3) überaus aufwendig und komplex (die Anzahl der Pfeile kann als Maß für die Komplexität genommen werden).

SFX, ein ursprünglich an der Universität Ghent entwickeltes System (siehe “Publications” auf der SFX-Homepage[4]), war Ende 2000 mit als erster Linkresolver am Markt verfügbar und hat neue Maßstäbe für das Kontext-sensitive Reference Linking gesetzt. Unter anderem wurde im Rahmen von SFX auch OpenURL[7], *der* Standard zum Austausch von bibliographischen Informationen über URLs, entwickelt.

Praktisch alle Linkingsysteme verwenden eine Linkdatenbank, um die Volltext-URLs zu den Zitaten zu ermitteln (siehe vorheriger Abschnitt). Die Basis hierfür bildet bei vielen Linkingsystemen die CrossRef[3] Metadaten-Datenbank, die momentan (Stand: Februar 2004) Linkinginformationen für mehr als 9000 Zeitschriften von mehr als 275 Verlagen umfasst (Anmerkung: Die im Vortrag genannte Zahl von etwa 5000 Zeitschriften war falsch, sie beruhte auf veralteten Angaben.). Um die Menge der Zeitschriften, für die Volltextlinks angeboten werden können, weiter zu erhöhen, schließen die Anbieter der Linkingsysteme üblicherweise so genannte Linkingabkommen mit weiteren, nicht an CrossRef teilnehmenden Verlagen ab.

Integrierte Linkingtools

Mit der Integration eines Linkingtools in ein Datenbanksystem ist es möglich, dem Nutzer ein hohes Maß an Komfort zu bieten: Die Links für die verschiedenen Dienste inklusive der Links zum Volltext werden üblicherweise bereits bei den Zitaten in der Trefferliste angezeigt. Alle Dienste können so “mit nur einem Klick” aufgerufen werden, außerdem ist für den Nutzer unmittelbar ersichtlich (zumindest scheint es so, siehe unten), für welche Zitate ein elektronischer Volltext vorhanden ist. Für die Bibliotheken haben die integrierten Linkingsysteme den Vorteil, dass sie meistens als Mehrwertdienst zu den Datenbanken kostenlos zur Verfügung gestellt werden.

Das größte Problem der integrierten Linkingtools ist, dass die gewünschten Links in jedem Linkingtool einzeln konfiguriert werden müssen. Damit steht man genau vor dem schon beim Rückblick auf die Entwicklung der ReDI-Bestandsabfrage beschriebenen Komplexitätsproblem: Die Anzahl der zu konfigurierenden Links wächst mit jedem zusätzlichen Datenbanksystem und jedem zusätzlichen Link enorm an (siehe Abbildung 2). Insbesondere müssen jedem Linkingtool die Bestandsdaten zur Verfügung gestellt werden, damit es dem Nutzer Kontext-sensitive Volltextlinks anbieten kann. Da eine mehrfache Pflege der Bestandsdaten nicht akzeptabel ist, ist dies nur machbar, wenn die Bestandsdaten ohne großen Aufwand aus einer zentralen Datenbank wie zum Beispiel der EZB exportiert, für die Linkingtools aufbereitet und dann geladen werden können – was regelmäßig erfolgen muss, um die Bestandsdaten aktuell zu halten.

Ein zusätzliches Problem ist, dass das Linkangebot je nach Funktionalität des Linkingsystems und Umfang der Linkdatenbank variiert. Beispielsweise bieten die meisten Datenbanksysteme mit integriertem Linkingtool die Möglichkeit, die Treffermenge auf die Zitate zu beschränken, für die elektronische Volltexte verfügbar sind (Volltextlimit). Diese Funktion wird von den Anbietern gerne als Argument für die integrierten Linkingtools angeführt, sie ist aber problematisch, da jedes Linkingtool nur für eine – vom Linkingtool abhängige – Teilmenge aller Volltexte die entsprechenden URLs kennt. So kann es vorkommen, dass für das gleiche Zitat in einer Datenbank ein Volltextlink angeboten wird, in einer anderen Datenbank aber nicht, da nur einer der Anbieter ein Linkingabkommen mit dem Verlag abgeschlossen hat. Erfolgt die Suche mit Volltextlimit, so ist das Zitat bei einer der Datenbanken nicht in der Treffermenge enthalten, obwohl ein Volltext zur Verfügung steht! Den Nutzern ist dieses Problem häufig nicht bewusst. Beheben lässt sich das Problem nur, wenn das Linkingsystem die Definition eigener Volltextkollektionen ermöglicht (was natürlich auch wieder pro Linkingsystem gemacht werden muss und den Aufwand weiter erhöht) und diese beim Volltextlimit mit berücksichtigt.

Zentrale Linkresolver

Die Vorteile eines zentralen Linkresolvers gegenüber den in Datenbanksysteme integrierten Linkingtools sind offensichtlich, sie wurden im Wesentlichen bereits beim Rückblick auf die Entwicklung der ReDI-Bestandsabfrage als Argumente für einen zentralen Lösungsansatz aufgezählt:

- Mit einem Linkresolver kann den Nutzern für alle Ausgangssysteme, die eine entsprechende Schnittstelle bieten, ein einheitliches Dienst- und Volltextlinkangebot zur Verfügung gestellt werden.
- Im Ausgangssystem muss zur Einbindung des Linkresolvers nur ein OpenURL-Link definiert werden. Insbesondere ist damit auch für Datenbanksysteme ohne integriertes Linkingtool eine Volltextverlinkung möglich.

- Die gesamte Konfiguration der Dienste und der Volltextlinks erfolgt zentral an nur einer Stelle – im Linkresolver.

Für den Nutzer bedeutet ein zentraler Linkresolver zwar "einen Klick mehr" bis zum gewünschten Dienst, dieser Nachteil kann aber durch ein verbessertes Dienstangebot mehr als wettgemacht werden.

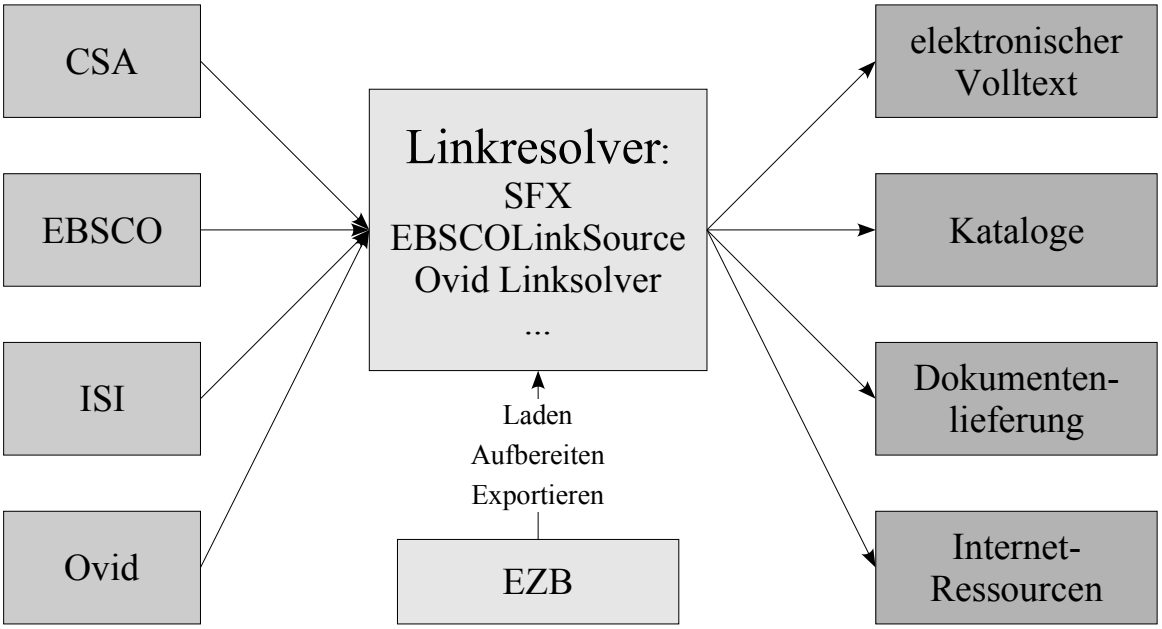

Abbildung 3:
Die Konfiguration eines zentralen Linkresolvers ist im Vergleich zu der mehrerer integrierter Linkingtools deutlich einfacher (wie schon die deutlich geringere Anzahl der Pfeile im Vergleich zu Abbildung 2 nahe legt). Mit dem zentralen Linkresolver ist ein einheitliches Linkangebot für alle Datenbanken möglich, der Weg zum gewünschten Dienst ist aber "einen Klick länger".

Auch bei einem zentralen Linkresolver müssen die Bestandsdaten hinterlegt werden, um eine Kontext-sensitive Volltextverlinkung realisieren zu können. Wie bei den Linkingtools bietet es sich an, die Bestandsdaten regelmäßig aus der EZB zu exportieren und sie nach entsprechender Aufbereitung in den Linkresolver zu laden (siehe Abbildung 3). Der Aufwand hierfür sollte nicht unterschätzt werden, er kann je nachdem welche Schnittstellen der Linkresolver für das Laden der Daten zur Verfügung stellt (bzw. nicht zur Verfügung stellt) ganz erheblich sein.

Ein Problem sind auch die unter Umständen hohen Kosten für einen zentralen Linkresolver. Bei inzwischen 50 an ReDI teilnehmenden Einrichtungen würden bei einer landesweiten Lösung – und nur die wäre sinnvoll, sonst müsste für einen Teil der Einrichtungen zusätzlich eine andere Lösung entwickelt werden – Lizenzkosten in erheblicher Höhe anfallen. Bei einer Lösung wie SFX, die im Gegensatz zu EBSCO LinkSource und Ovid's LinkSolver den Betrieb eines eigenen Linkingservers erfordert, ist auch dies bei der Ermittlung der Gesamtkosten zu berücksichtigen.

5. Der ReDI-Lösungsansatz

Mit der aktuellen Version der ReDI-Bestandsabfrage steht im ReDI-System bereits eine zentrale Linkinglösung zur Verfügung, die bis auf die Volltextverlinkung auf Artikelebene alle wesentlichen Funktionen eines zentralen Linkresolvers bietet (siehe Abschnitt 2). Es stellt sich daher die Frage, ob und wenn ja mit welchem Aufwand die ReDI-Bestandsabfrage um eine Volltextverlinkung auf Artikelebene erweitert werden kann.

Im Wesentlichen sind hierfür Erweiterungen in zwei Bereichen notwendig:

- Bestandsangaben: Die Prüfung der Verfügbarkeit muss auf Heftebene statt wie bisher auf Zeitschriftenebene erfolgen.
- Linkinginformationen: Es müssen die Volltext-URLs für die gewünschten Aufsätze ermittelt werden.

Die Erweiterung der Verfügbarkeitsprüfung von Zeitschriften- auf Heftebene sollte für diejenigen Einrichtungen, die ihre Bestände in der EZB nachweisen, auf jeden Fall möglich sein – mit dem erweiterten Datenformat der EZB stehen alle notwendigen Daten zur Verfügung (siehe Abschnitt 3). Für Einrichtungen, die ihre Bestände nicht in der EZB sondern in einem anderen Systemen nachweisen (wie zum Beispiel die Universität Heidelberg im HZV), muss eine mit der Lösung für die EZB vergleichbare Lösung zum Abruf der Bestandsangaben durch die ReDI-Bestandsabfrage gefunden werden – auch das sollte eigentlich problemlos möglich sein. Falls nur Bestandsangaben auf Zeitschriftenebene oder gar keine Bestandsangaben vorliegen, kann wie in Abschnitt 3 beschrieben vorgegangen werden.

Die Ermittlung der Volltext-URLs sollte ebenfalls möglich sein, und zwar wie bei anderen Linkingsystemen auch mit CrossRef[3] als Linkdatenbank (siehe Abschnitt 4)! Die CrossRef-Mitgliedschaft ist für Bibliotheken seit Mitte 2003 kostenlos.

Anhand der bibliographischen Daten kann durch Abfrage der CrossRef Metadaten-Datenbank der DOI (Digital Object Identifier[8]) des gewünschten Aufsatzes ermittelt und der DOI über das Handle-System (siehe http://dx.doi.org/ oder http://hdl.handle.net/) wiederum in die entsprechenden Volltext-URLs aufgelöst werden. Eine Ermittlung der DOIs in Echtzeit ist möglich, CrossRef bietet hierfür eine Telnet basierte Schnittstelle an, deren Einbindung in die ReDI-Bestandsabfrage kein Problem sein sollte. Mit Hilfe der CrossRef Metadaten-Datenbank ist es gegebenenfalls auch möglich, fehlende bibliographische Daten zu ergänzen. Die CrossRef Metadaten-Datenbank deckt momentan wie gesagt bereits mehr als 9000 Zeitschriften ab (Stand: Februar 2004). Ob der Nutzer über eine der ermittelten Volltext-URLs Zugriff auf den Volltext hat, kann anhand der Bestandsangaben festgestellt werden.

Die CrossRef-Abfrage bzw. die Auflösung des DOI liefert momentan noch in vielen Fällen nur eine Volltext-URLs, auch wenn der gewünschte Volltext über mehrere Anbieter zur Verfügung steht. Die so genannte “multiple resolution”, die dieses Problem lösen sollte, befindet sich zwar in der Entwicklung, für einige wichtige Anbieter, für die sich die Volltext-URLs noch nicht über CrossRef ermitteln lassen, müsste aber zunächst eine andere Lösung gefunden werden. Dies trifft insbesondere auf wichtige Aggregatoren wie EBSCO zu, bei denen die an ReDI teilnehmenden Einrichtungen in größerem Umfang Volltexte lizensiert haben. Als Lösung bietet sich an, eine Suche nach dem gewünschten Aufsatz im Zielsystem zu starten – eine entsprechende “Jumpstart”-Funktion bieten die meisten Systeme an, im Idealfall sogar auf Basis von OpenURL. Ob dies tatsächlich eine brauchbare Lösung ist, müsste im Einzelfall geprüft werden, deshalb ist eine solche Vorgehensweise auch nur für große Anbieter praktikabel.

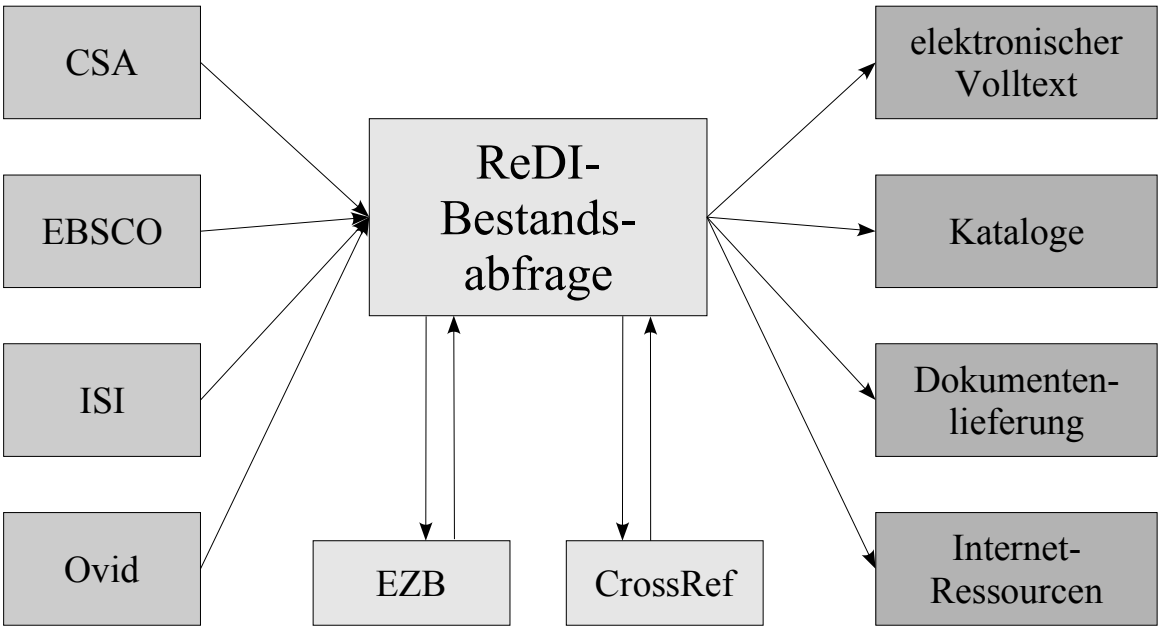

Abbildung 4:
ReDI-Bestandsabfrage als zentraler Linkresolver. Für die Volltextverlinkung auf Artikelebene werden die Bestandsangaben aus der EZB abgerufen und die Volltext-URLs über CrossRef ermittelt – jeweils in Echtzeit. Im Vergleich zu anderen Linkinglösungen entfällt so die aufwendige Pflege der Bestandsdaten im Linkingsystem.

Die Volltextverlinkung auf Artikelebene könnte damit wie folgt realisiert werden (siehe auch Abbildung 4): Anhand der per OpenURL übernommenen bibliographischen Daten prüft die ReDI-Bestandsabfrage über die EZB oder ein anderes Nachweissystem wie das HZV, ob und wenn ja über welche Anbieter der gewünschte Aufsatz für den Nutzer in elektronischer Form verfügbar ist. Ist für den Anbieter (zum Beispiel EBSCO) eine Sonderlösung implementiert, so wird dem Nutzer angeboten, eine Suche im Zielsystem zu starten, ansonsten wird versucht, die Volltext-URL über CrossRef zu ermitteln.

Dieser Lösungsansatz hat im Vergleich zu Linkingsystemen wie SFX, EBSCO LinkSource oder Ovid's LinkSolver den großen Vorteil, dass die aufwendige Pflege von Bestandsdaten im Linkingsystem entfällt. Erste Tests zur Einbindung von CrossRef sind viel versprechend verlaufen, eine erste Testversion der ReDI-Bestandsabfrage mit Volltextverlinkung auf Artikelebene sollte im 2. Quartal 2004 zur Verfügung stehen.

Literatur

[1] ReDI-Webseiten. http://www.redi-fr.belwue.de/
[2] Elektronische Zeitschriftenbibliothek. http://www.bibliothek.uni-regensburg.de/ezeit/
[3] CrossRef.org: The Reference Linking Backbone. http://www.crossref.org/
[4] SFX: Context-Sensitive Reference Linking. http://www.sfxit.com/
[5] EBSCO LinkSource. http://www.linkresolver.com/
[6] Ovid's LinkSolver. http://www.linksolver.com/
[7] Development of an OpenURL Standard. http://library.caltech.edu/openurl/
[8] The Digital Object Identifier System. http://www.doi.org/

Bernd Oberknapp
Universität Freiburg
- Universitätsbibliothek -
Postfach 16 29
D-79016 Freiburg
Tel.: +49(0)761/203-3852
E-Mail: bo@ub.uni-freiburg.de