

LIVIVO: Neue Herausforderungen an das ZB MED-Suchportal für Lebenswissenschaften

LIVIVO: new challenges for ZB MED's search portal for life sciences

Abstract

The German National Library of Medicine (ZB MED) has a long track record providing search portals for life sciences. In 2015 LIVIVO was introduced as a new product, which database covers ZB MED's entire range of subjects: medicine, health, nutrition, environment and agriculture.

Initially LIVIVO's development aimed at realizing a modern search engine combining the functionalities of the predecessor portals. An entirely new web interface was developed, driven by a highly improved usability and a responsive web design.

LIVIVO has a great potential in providing search services based on more than 55 million metadata records. Current work of ZB MED focuses on automatic interfaces. Concurrently setting up ZB MED Knowledge Environment creates an indispensable data base for research work at ZB MED.

This article will address the current challenges of a scientific portal using the example of LIVIVO. It will outline the approaches to solutions and development lines.

Keywords: medicine, health, nutrition, environment, agriculture, life sciences, LIVIVO, MEDPILOT, GREENPILOT, ZB MED, search, research, challenges, knowledge environment, SRU, Z39:50, interfaces

Zusammenfassung

Die Deutsche Zentralbibliothek für Medizin (ZB MED) hat als Anbieterin von Suchportalen in den Lebenswissenschaften eine lange Tradition. Mit LIVIVO steht seit 2015 ein neues Produkt zur Verfügung, das erstmals das gesamte Fächerspektrum von ZB MED abdeckt: Medizin, Gesundheit, Ernährungs-, Umwelt- und Agrarwissenschaften.

In der Anfangsphase von LIVIVO stand der Aufbau eines modernen Fachportals mit einer neuen Suchmaschine im Vordergrund, das die Funktionalitäten der Vorgängerportale miteinander vereinigt. Dabei wurde eine neue Weboberfläche entwickelt, die sich durch eine hohe Benutzerfreundlichkeit und ein responsives Webdesign auszeichnet.

Das große Potential für die Entwicklung von LIVIVO liegt im Bereitstellen von Suchdiensten basierend auf den mehr als 55 Millionen Metadaten-sätzen. Aktuelle Arbeiten von ZB MED beschäftigen sich nun damit, automatische Schnittstellen für Suchservices anzubieten. Gleichzeitig wird mit dem Aufbau des ZB MED-Knowledge-Environment eine unverzichtbare Datenbasis für Forschungsarbeiten an ZB MED geschaffen. Dieser Aufsatz wird auf die aktuellen Herausforderungen eines wissenschaftlichen Portals am Beispiel von LIVIVO eingehen, Lösungsansätze skizzieren und davon ausgehend die Entwicklungslinien vorzeichnen.

Schlüsselwörter: Medizin, Gesundheit, Ernährung, Umwelt, Agrar, Lebenswissenschaften, LIVIVO, MEDPILOT, GREENPILOT, ZB MED, Recherche, Forschung, Herausforderungen, Knowledge Environment, SRU, Z39:50, Schnittstellen

Christoph Poley¹

1 ZB MED – Leibniz-
Informationszentrum
Lebenswissenschaften,
Bereitstellung von
Informationsdiensten /
Leitung LIVIVO-Entwicklung,
Köln, Deutschland



Abbildung 1: LIVIVO – ZB MED-Suchportal Lebenswissenschaften

Einleitung

LIVIVO – so heißt das ZB MED-Suchportal Lebenswissenschaften (Abbildung 1, <https://www.livivo.de>). Es bietet seinen Kundinnen und Kunden als Nachfolgeprodukt von MEDPILOT [1] und GREENPILOT umfassende und kostenfreie Recherchewerkzeuge für die interdisziplinäre Literatursuche in den Fächern Medizin und Gesundheit, Ernährung, Umwelt und Agrar. Eine intuitive Oberfläche bietet den bequemen Einstieg dazu. Für komplexe Anfragen steht eine frei kombinierbare Expertensuche zur Verfügung.

Die Suchtechnologie von LIVIVO basiert auf Apache Solr. Darüber hinaus werden Recherchen durch linguistische Verfahren aufbereitet und semantisch mit sprachunabhängigen Konzepten annotiert. Als Fachthesauri werden für die Fächer Medizin und Gesundheit die Medical Subject Headings [2], für die Ernährungs-, Umwelt- und Agrarwissenschaften AGROVOC [3] und UMHES [4] verwendet. Durch das Abbilden der Fachbegriffe in unterschiedlichen Sprachen auf ihre linguistischen Repräsentationen können sprachübergreifend Suchergebnisse gefunden werden. Gleichzeitig wird die Suche nach Wortvarianten und Synonymen ermöglicht.

Der folgende Abschnitt behandelt zunächst das neue Webfrontend als Grundlage für die weitere Produktentwicklung von LIVIVO. Darauf aufbauend werden anschließend neue Nutzungsszenarien beschrieben. Außerdem wird abschließend eine neue zentrale Architekturkomponente zur Anreicherung und Vernetzung von Metadaten vorgestellt.

Das LIVIVO-Suchportal

Eine wichtige Herausforderung beim Anbieten von Suchportalen, so auch in LIVIVO, liegt zunächst in der Darstellung komplexer Inhalte. Dabei bildet die Gestaltung der Trefferlisten ein zentrales Element für das Aussehen und

die Funktionalität der Weboberfläche. Es ist deshalb essentiell, dass der Benutzer auf möglichst intuitive Art und Weise durch sie geführt wird – Voraussetzung für die Akzeptanz eines jeden Webangebots.

Deshalb wurde bei der Entwicklung der LIVIVO-Weboberfläche großes Augenmerk auf die Usability und die User Experience gelegt. Dank des umfangreichen Feedbacks der Anwender und der Ergebnisse von Usability-Untersuchungen [5] konnten bereits kurz nach dem Start der Testphase von LIVIVO eine ganze Reihe von Verbesserungen durchgeführt werden. Diese betrafen vor allem die Trefferqualität, die Filtermöglichkeiten, die Verknüpfung mit Forschungsdaten und die Nutzung von LIVIVO auf mobilen Endgeräten.

Das Ranking der Suchergebnisse basiert auf dem tf.idf-Mechanismus [6] und weiteren Parametern. Als Alternative ist die Sortierung nach dem Erscheinungsdatum möglich. Durch das Anbinden von semantisch-linguistischen Komponenten an die Suchmaschine werden für jede Trefferliste Terme extrahiert, mit denen sich Suchen sinnvoll ergänzen und die Ergebnismenge mit einem Klick weiter einschränken lassen. Durch das Nutzen von Filtern kann jede Ergebnismenge weiter verfeinert werden, beispielsweise nach Dokumenttyp, Datenquelle oder Erscheinungsjahr.

Für den Großteil der Treffer werden Verlinkungen zu den Verlagsangeboten und vielen Open-Access-Repositoryn bereitgestellt. Sie werden je nach zur Verfügung stehender Datenlage in verschiedenen Qualitäten angeboten. Im Optimum führen sie direkt zum passenden Volltext, alternativ zu Abstracts, Verlagsangaben oder den Homepages der Zeitschriften.

Als Grundlage für die Verlinkungen existiert eine Reihe von Quellen, die sich unabhängig vom Standort der Benutzer einbinden lassen. Sie bestehen aus Links zu Open-Access-Publikationen, freien Zeitschriften und den speziell für ZB MED-Kunden lizenzierten Volltexten. Für die standortabhängige Verfügbarkeit lizenzierter Angebote von Zeitschriften und Artikeln wird der Linking-Dienst der

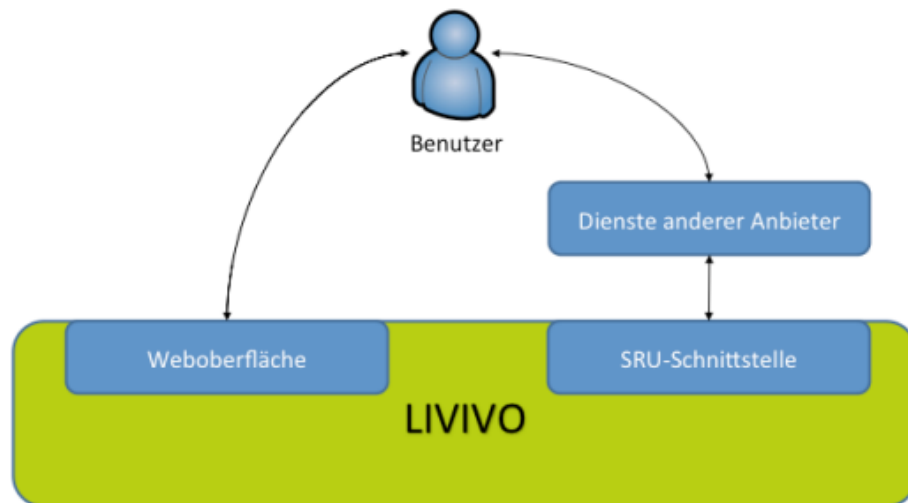


Abbildung 2: Suchdienste von LIVIVO

Elektronischen Zeitschriftenbibliothek [7] herangezogen und nach dem bekannten Ampelsystem eingebunden. Gleichzeitig wurde mit dem Online-Mietervice DeepDyve ein neuer Dienst für den Zugang zu Volltexten integriert. Zudem erlaubt ein weiterer Service die bedarfsorientierte Erwerbung von Literatur durch Patron Driven Acquisition. Als Besonderheit existieren bei einer Reihe von Publikationen Verknüpfungen zu assoziierten Forschungsdaten, zum Beispiel aus dem Dryad Digital Repository (<http://datadryad.org/>).

Neben der Bereitstellung der gewünschten Literatur werden auch Links zu weiteren Dienstleistungen angeboten, wie beispielsweise die ZB MED-Dokumentlieferung. Für den Export von Suchergebnissen in Literaturverwaltungsprogramme existiert eine Z39:50-Schnittstelle.

Die Weiterentwicklung des LIVIVO-Suchportals bleibt auch in Zukunft ein wichtiger Bestandteil. Parallel dazu werden seit Anfang 2016 neue Herausforderungen angegangen. Sie sind Inhalt der folgenden beiden Abschnitte.

Neue Nutzungsszenarien

Eine gut benutzbare Weboberfläche ist eine notwendige Bedingung für die weitere Entwicklung von LIVIVO. Denn gerade im Bereich der Lebenswissenschaften existiert ein hochkompetitiver Markt für Suchportale. Allein in den beiden Fachgebieten „Medizin“ und „Gesundheit“ gibt es einen sehr starken Wettbewerb durch das Rechercheangebot PubMed der National Library of Medicine (<http://www.ncbi.nlm.nih.gov/pubmed/>). Darüber hinaus haben sich fachübergreifende Datenbanken am Markt platziert, die anders als LIVIVO in vielen Fällen kostenpflichtig sind. Für LIVIVO heißt das, sich einerseits als Portal noch stärker zu profilieren und andererseits für neue Nutzungsszenarien zu öffnen.

Als Grundlage dafür wird das LIVIVO-Portal um eine neu entwickelte SRU(Search/Retrieve via URL)-Schnittstelle erweitert (Abbildung 2). Dank einer zugrunde liegenden serviceorientierten Architektur kann die LIVIVO-Suche gekapselt und rein datenbasiert in Produkte anderer

Dienstleister eingebunden werden. Somit lässt sich LIVIVO auf einfache Art als Suchdienst für die Recherche in den Lebenswissenschaften in andere Produkte integrieren. Ein wesentlicher Aspekt für den Betrieb einer solchen SRU-Schnittstelle ist, dass sich damit nahezu sämtliche LIVIVO-Suchscenarien nachstellen lassen. Neben den kompletten Metadaten aus der Trefferliste lassen sich auch alle Facetten aus dem Portal ausgeben, um entsprechende Filterqueries realisieren zu können. Die Facetten selbst umfassen beispielsweise die Fachgebiete, das Erscheinungsjahr, die Sprache, den Dokumenttyp oder die zugrunde liegende Datenquelle. Zusätzlich sind Filter über ISSN und ISBN möglich, um die Suchanfragen auf bestimmte Monographien sowie Zeitschriften und damit verknüpfte Artikel einzuschränken.

Das ZB MED Knowledge Environment

Ein Suchportal wie LIVIVO lebt von den Daten und dem damit verknüpften Wissen, das über Schnittstellen als Suchdienste zur Verfügung gestellt wird. Als Quelle dienen zunächst Metadaten verschiedener Anbieter, die normalisiert in einen Datenindex überführt werden. Für das Anbieten von umfangreichen und vernetzten Recherchemöglichkeiten als Anwenderfall reicht dies im Allgemeinen jedoch nicht mehr aus.

So lässt sich bei Artikeln ohne Auswerten der Metadaten kein eindeutiger Link zur korrespondierenden Zeitschrift einbinden. Ein weiterer Fall stellt die Verknüpfung von Forschungsdatensätzen mit den Metadaten aus den dazu gehörenden Publikationen dar. Bereits in der Vergangenheit wurden separate, den Anforderungen genügende Lösungen realisiert, die beispielsweise direkt bei der Datenkonvertierung mit eingebunden wurden. Bei Anwendungsfällen, in denen sich etwa verknüpfende Elemente regelmäßig ändern können (z.B. Generierung von Verfügbarkeiten) oder bei zeit- und rechenintensiven Prozeduren, wie der Annotation durch semantisch-linguistische Verfahren, bedarf es einer geeigneten Infrastruktur.

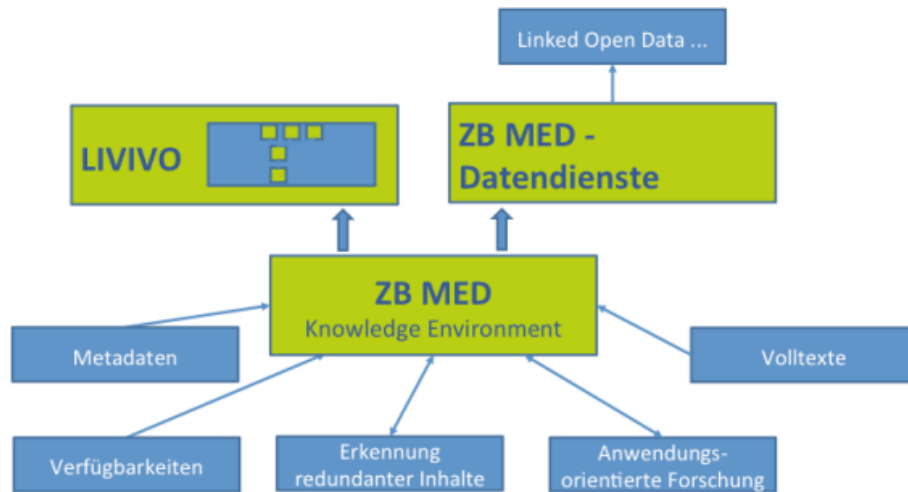


Abbildung 3: ZB MED Knowledge Environment

Diese Aufgabe übernimmt nun das ZB MED Knowledge Environment (ZB MED KE), das auf einer dokumentbasierten Datenbank aufbaut. Um sie herum lässt sich nun eine Vielzahl von Services andocken, die diese Datenbank mit Informationen füllen, anreichern oder diese zur weiteren Verarbeitung wieder ausgeben (Abbildung 3).

Im Gegensatz zum klassischen Einspeisen eines Datenindexes durch Metadaten können im ZB MED KE zusätzlich etwa freie Verfügbarkeiten und Querbezüge zu anderen Datensätzen gespeichert werden. Diese Informationen lassen sich nun durch neue Services anreichern, die bereits in Zusammenarbeit mit der ZB MED-Forschung in Arbeit sind [8]. Ein erster Prototyp ist bereits online und unter <http://labs.livivo.de> zu finden. Eine weitere Anwendung ist die Entwicklung von Algorithmen, die es erlauben, redundante Informationen zusammenzufassen und damit Dubletten zu beseitigen.

Für LIVIVO bedeutet das konkret, dass die Metadaten ab 2017 aus dem ZB MED KE gewonnen werden. Gleichzeitig eröffnet das ZB MED KE die Möglichkeit, neue beliebige Kollektionen für neue Datendienste und Produkte zusammenzustellen und diese zu verwalten. Beispielsweise können so Ergebnisse aus der ZB MED-Forschung mit den zugrunde liegenden Metadaten und den angereicherten Informationen als Linked Open Data automatisch zur weiteren Verarbeitung angeboten werden.

Fazit und Ausblick

Die gezeigten Ergänzungen der LIVIVO-Dienste und die Einführung des ZB MED KE als zentrale Wissensumgebung zum Speichern von Metadaten, Verfügbarkeiten, Volltexten und damit verknüpften Informationen bilden aktuell grundlegende Arbeiten an ZB MED, um als Anbieter von Such- und Datendiensten auch für die Zukunft gut gerüstet zu sein.

Parallel dazu werden ab 2017 verstärkt Arbeiten am LIVIVO-Portal fortgesetzt. In Zusammenarbeit mit der ZB MED-Forschung werden Datensätze mit zusätzlichen Informationen angereichert und mit neuen Wissensdatenbanken und Ontologien verknüpft. Visualisierungen werden dazu

beitragen, komplexe Daten in einfacher Form darzustellen und in die Recherche einzubinden, zum Beispiel als Word Cloud.

Ein weiterer Baustein ist die kontinuierliche Verbesserung der Such- und Datenqualität und die Erweiterung des Suchraums sowie der Zugänge zu den freien und lizenzierten Volltexten. Flankiert durch Vermarktung und verstärkte Netzwerkaktivitäten sollen diese Maßnahmen dazu beitragen, den Bekanntheitsgrad von LIVIVO als Portal und als Suchdienst zu erhöhen und als Highlightprodukt von ZB MED fest in Wissenschaft und Forschung in den Lebenswissenschaften deutschlandweit und international zu etablieren.

Anmerkung

Interessenkonflikte

Der Autor leitet die technische Entwicklung von LIVIVO, ansonsten liegen keine Interessenkonflikte in Zusammenhang mit diesem Artikel vor.

Literatur

1. Poley C. Mit MEDPILOT auf dem Weg ins Semantic Web. *GMS Med Bibl Inf.* 2012;12(3):Doc22. DOI: 10.3205/mbi000258
2. National Library of Medicine (NLM). Medical Subject Headings. Verfügbar unter: <https://www.nlm.nih.gov/mesh/>
3. Food and Agriculture Organization of the United Nations (FAO). AGROVOC Multilingual agricultural thesaurus. Verfügbar unter: <http://aims.fao.org/vest-registry/vocabularies/agrovoc-multilingual-agricultural-thesaurus>
4. Umweltbundesamt. Umweltthesaurus UMTHEs. Verfügbar unter: <http://data.uba.de/umt/de.html>
5. Pössel J. Auf Nutzerbedürfnisse abgestimmtes Design: Usability-Untersuchung des neuen ZB MED-Suchportals LIVIVO. *GMS Med Bibl Inf.* 2016;16(1-2):Doc06. DOI: 10.3205/mbi000361
6. Kumar J. *Apache Solr Search Patterns*. Birmingham: Packt Publishing; 2015.

7. Friedrich G, Hutzler E, Scheuplein M, Schröder P. Schlussbericht zum Teilprojekt 4: „Integration der Elektronischen Zeitschriftenbibliothek (EZB)“ des Gesamtvorhabens: Realisierung einer fachübergreifenden Infrastruktur für elektronische Informationsdienstleistungen durch die Informationsverbünde; Teil I – Kurze Darstellung und Teil II – Eingehende Darstellung. Regensburg: Universitätsbibliothek Regensburg; 2007. URN: urn:nbn:de:bvb:355-opus-8530
8. Müller B, Hagelstein A. Beyond metadata enriching life science publications in livivo with semantic entities from the linked data cloud. In: Martin M, Cuquet M, Folmer E, editor. Joint Proceedings of the Posters and Demos Track of the 12th International Conference on Semantic Systems – SEMANTICS2016 and the 1st International Workshop on Semantic Change & Evolving Semantics (SuCESS'16) co-located with the 12th International Conference on Semantic Systems (SEMANTICS 2016), Leipzig, Germany, September 12-15, 2016. Available from: <http://ceur-ws.org/Vol-1695/paper1.pdf>

Korrespondenzadresse:

Christoph Poley
ZB MED – Leibniz-Informationszentrum
Lebenswissenschaften, Bereitstellung von
Informationsdiensten / Leitung LIVIVO-Entwicklung,
Gleueler Str. 60, 50931 Köln, Deutschland
poley@zbmed.de

Bitte zitieren als

Poley C. LIVIVO: Neue Herausforderungen an das ZB MED-Suchportal für Lebenswissenschaften. *GMS Med Bibl Inf.* 2016;16(3):Doc21. DOI: 10.3205/mbi000376, URN: urn:nbn:de:0183-mbi0003760

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/mbi/2016-16/mbi000376.shtml>

Veröffentlicht: 22.12.2016

Copyright

©2016 Poley. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.