

Literatur für interdisziplinäre Wissenschaft – eine Anforderungsanalyse aus Anwendersicht

Literature for interdisciplinary scientists – a requirement engineering approach

Abstract

Medical research is subject to continuous change, and scientific literature must be timely assessable for research, diagnosis, and teaching. However, biomedical sciences have become interdisciplinary, collective, and translational. Furthermore, novel technologies such as the Internet and smart phones have been established and are available almost everywhere. Medical libraries have to follow these trends offering novel integrated services to optimize the support of their customers. In this paper, we analyse the domain from the users' perspective and give a categorization of future bibliographic services.

Keywords: requirement engineering, literature access, interdisciplinary work

Zusammenfassung

Literatur ist die Kernkomponente der Wissenschaft, aber die medizinische Forschung ist im ständigen Wandel. Die Metapher des Elfenbeinturmes für den Arbeitsplatz des Wissenschaftlers ist längst nicht mehr gültig. Erfolgreiche Wissenschaft im biomedizinischen Bereich ist heute interdisziplinär, kollektiv und translational. Dieser Prozess wird begleitet von technologischen Innovationen wie dem Smart Phone und der Informationsressource Internet, die überall und immer verfügbar ist. Das medizinische Bibliothekswesen muss diesen Wandel mitvollziehen und mit innovativen, prozessintegrierten Diensten seine Kunden optimal unterstützen. Dieser Beitrag beleuchtet das Problemfeld aus Sicht des Anwenders und versucht, die Anforderungen an bibliothekarische Dienste zu kategorisieren und zu priorisieren.

Schlüsselwörter: Anforderungsanalyse, Literaturzugriff, Interdisziplinäres Arbeiten

1 Einleitung

Die Bibliothek und speziell die medizinische Bibliothek sammelt medizinische Fachliteratur, um sie ihren „Kunden“ möglichst einfach und schnell zugänglich zu machen. Unter Literatur wird dabei gemeinhin der Bereich aller mündlich oder schriftlich fixierten sprachlichen Zeugnisse verstanden [1]. Die Bedeutung der Literatur, also die ihr beigemessene Wichtigkeit, ist schon schwieriger zu fassen. Aus dem Inhalt eines derzeit laufenden DFG-Projektes stammt das folgende Zitat: „Die kognitive Funktion der Literatur kann ohne eine grundsätzliche Modellierung der komplexen Interaktion von Bedeutungs-, Wahrheits- und Wissensfragen nicht verstanden werden...“ [2], [3]. Allgemein kann Literatur als fundamentale Grundlage allen Wissens und jeder Wissenschaft aufge-

fasst werden. In den medizinischen Anwendungsbereichen der Forschung, Lehre/Ausbildung und Patientenversorgung ist sie gleichermaßen wichtig und es gelten dieselben Zielvorstellungen hinsichtlich des Literaturzugriffs: nämlich zeitlich sofort und inhaltlich vollständig. War vor 25 Jahren der Literaturzugriff ausschließlich über tabellarische und alphabetisch sortierte Karteikästen organisiert, über die das gesuchte Werk aus dem Magazin angefordert werden konnte, sind heute diese Einträge digital im Internet verfügbar. Der Ort des Zugriffs hat sich damit von zentral nach dezentral verschoben (Tabelle 1). Dieser Trend begann bereits vor 20 Jahren, als die ersten Literaturdatenbanken über BTX-Terminals abgefragt werden konnten. Die Abstracts einer solchen Recherche, die zentral am Terminal der medizinischen Bibliothek durchgeführt wurde, erhielt man drei Wochen später als

Thomas M. Deserno¹

¹ Universitätsklinikum der RWTH Aachen, Institut für Medizinische Informatik, Aachen, Deutschland

Tabelle 1: Literaturzugangssysteme im Wandel

Zeitpunkt	Beispiel	Ort des Zugangs	Ort der Ressource
1987	Karteikasten	zentral	zentral
1992	BTX-Terminal	zentral	dezentral
2002	Internet Datenbank (PubMed)	dezentral	dezentral
heute	Internet Portal (Ovid)	dezentral	virtuell
morgen	Annotierte Relevanz	virtuell + kooperativ	virtuell

Ausdruck, und dann begann erst der zeit- und arbeitsaufwändige Prozess der eigentlichen (Fernleih-)Bestellung der als relevant vermuteten Aufsätze. Mittlerweile bieten Portale, wie Ovid, den Zugriff auf mehrere Datenbanken gleichzeitig, und – je nach Ausstattung der Universität oder Bibliothek – gleich den Durchgriff auf die im PDF-Format verfügbaren Fachbeiträge. In 20 Jahren ist aus einem zentralen Bestellprozess über mehrere Wochen ein dezentraler Vorgang mit wenigen Mausclicks geworden.

Dennoch gibt es Potential für weitere Verbesserungen, dem die medizinischen Bibliotheken nachkommen sollten. Diese Anforderungen beinhalten die Art des Zugangs: individuell vs. kooperativ, die Art der Suche: text bzw. bildbasiert, und die Art der Ressource: PDF-Volltext vs. Abstract. Diese Punkte werden im Folgenden mit Beispielen illustriert.

2 Material und Methode

Medizinische Fachliteratur enthält Wissen, das jedoch in den bibliographischen Tabellen nicht repräsentiert wird. Die allgemein bekannte Informationspyramide (Abbildung 1) zeigt die schrittweise Abstraktion von Zeichen über Daten, Information und Wissen zur Weisheit. Gemäß dieser Aufteilung sind tabellarische Listen eher Daten als Information; die medizinische Bibliothek – trotz optisch ansprechender Suchmasken im Web – betreibt heute also Daten- statt Informations- oder Wissensmanagement. Erst kürzlich wurde begonnen, die vielfältigen Informationen, die in den Listen der Bibliotheken schlummern, für den Nutzer verfügbar zu machen. Um dies an strukturierten Beispielen zu illustrieren, muss der gesamte Lebenszyklus von Daten betrachtet werden. Dieser wurde in dem Open Archival Information System (OAIS) Modell, das vom Consultative Committee for Space Data Systems (CCSDS) und der National Aeronautics and Space Administration (NASA) entwickelt wurde, als sechsstufiger Zyklus in acht Schritten modelliert [4]. Abbildung 2 zeigt den Zyklus, der neben der Generierung von Daten auch deren Nutzung, Modifikation und Aktualisierung behandelt [5]. Medizinische Literatur wird hier also nicht mehr als statisches Objekt, sondern als sich stetig ändernde dynamische Einheit aufgefasst.

Eine (medizinische) Bibliothek sollte den gesamten Prozess unterstützen. Im folgenden werden jedoch nur die Bereiche Verteilung, Benutzung und Verbesserung exemplarisch betrachtet.

2.1 Verteilung gesammelter Daten (Share)

Der Zugriff auf medizinische Literatur erfolgt immer noch alphanumerisch, d.h. in den entsprechenden Textfeldern für Autorennamen, Titel oder Abstract können Suchbegriffe angegeben werden. Es gibt bereits viele Aussagen darüber, dass ein Information Retrieval hinsichtlich seiner Query Completion – also der Antwortvollständigkeit von Suchanfragen – stark verbessert werden kann, wenn die Suche auf Texturen und Bildmuster ausgeweitet wird, die insbesondere in der Medizin durch die medizinische Bildgebung eine besondere Relevanz haben [6]. Die sich derzeit durchsetzenden Open Access Publikationswege unterstützen eine Auswertung bildlicher Information bereits generisch, denn Bilder und Graphiken werden im PDF als individuell extrahierbare Objekte eingebunden und werden in den Webinterfaces von z.B. Pubmed bereits gesondert angezeigt.

Es ist also naheliegend, die Suche auf Bilder auszuweiten. Ein entsprechendes Interface ist in Abbildung 3 visualisiert [6]. Neben Autorennamen und Stichwörtern können Bildmuster zur Suche angegeben werden.

Auch die listenartige Anzeige der Suchergebnisse ist nicht mehr up to date. In Internet-Portalen wie Amazon können Ergebnislisten nach „Relevanz“ sortiert werden, die z.B. aus der kollaborativen Bewertung anderer ermittelt wird. Google Scholar hat hier bereits erste Ansätze umgesetzt. Weiterhin sollten Filtermöglichkeiten nach Verfügbarkeit des (PDF-)Artikels integriert werden, um nur solche Ergebnisse anzuzeigen, die auch für den jeweiligen Nutzer direkt verfügbar sind. Auf Internet-Plattformen zu Hotelbuchung ist dies bereits Standard – ausgebuchte Hotels werden gar nicht erst angezeigt. Die heutige Pubmed-Funktion alleine reicht nicht aus, denn je nach Universität sind die Zugriffserlaubnisse mit den einzelnen Verlagen und Copyright-Haltern unterschiedlich.

In der biomedizinischen Literatur könnte die „Relevanz“ auch über die Referenzen in der Suchmenge angegeben werden. Viele der Referenzen eines in der Ergebnismenge enthaltenen Beitrages verweisen auf andere Elemente der Ergebnismenge: es wird ein Netz (Information) und keine Liste (Daten) erzeugt. Aus den Verknüpfungen innerhalb der Ergebnismenge lässt sich wiederum eine Relevanz bestimmen, um die Liste zu ordnen. Das interaktive Navigieren durch solche Ergebnisnetze ist in der wiederum fiktiven Abbildung 4 illustriert.

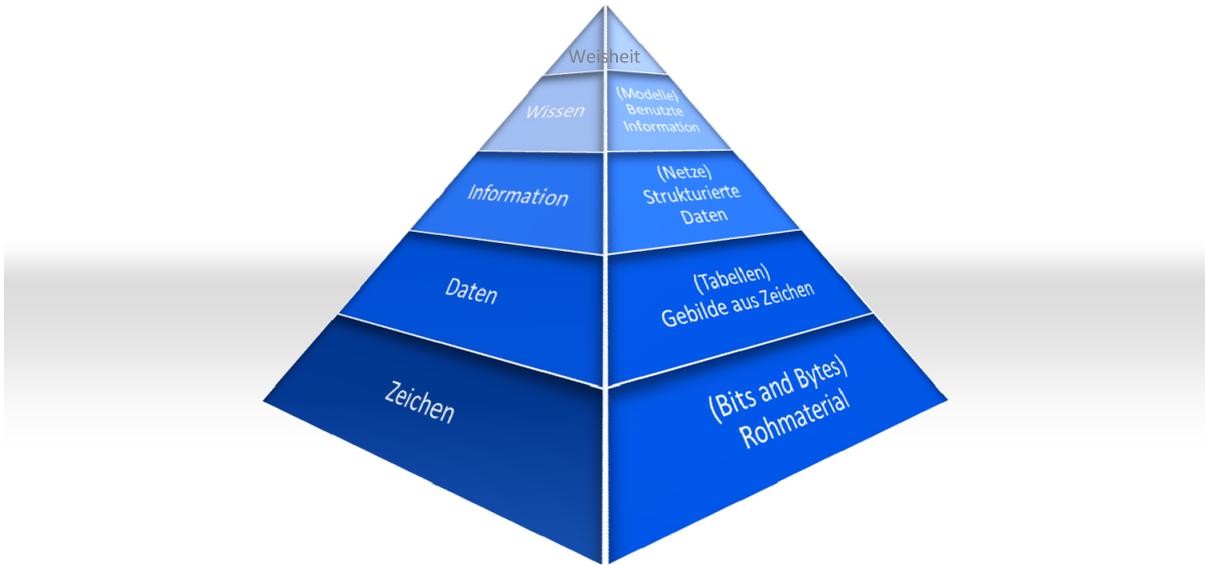


Abbildung 1: Informationspyramide

Der Bestandskatalog einer Bibliothek – sei er als Karteikasten oder als Tabelle im Internet organisiert – entspricht der Datenebene.

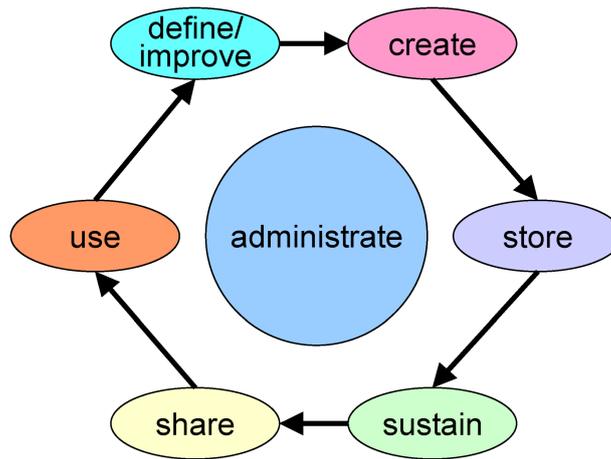


Abbildung 2: OAIS-Modell

Der Zyklus enthält die Schritte Definition, Erzeugung, Speicherung, Erhaltung, Verteilung, Verbesserung und Verwaltung, die alle anderen Bereiche zentral unterstützt [5].

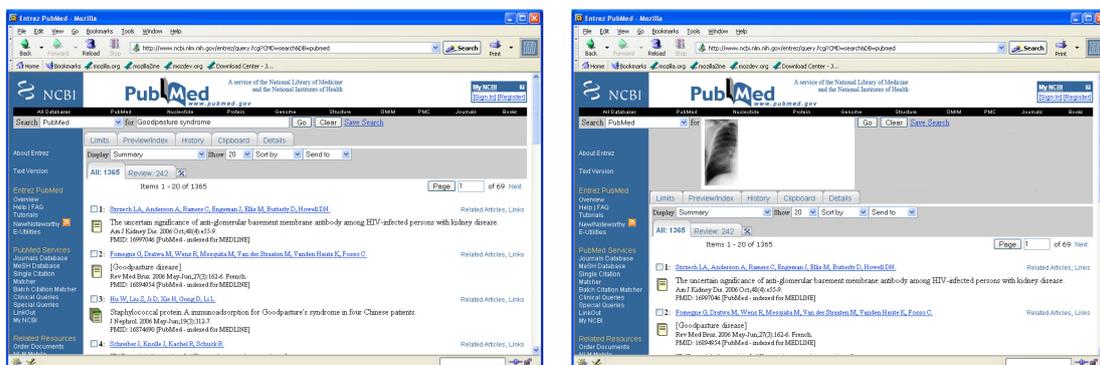


Abbildung 3: Pubmed Interface

Die Fotomontage zeigt das textbasierte Interface (links) von heute und die bildbasierte Version (rechts) von morgen, bei der Bilder via Copy/Paste zur Suche in Datenbanken verwendet werden können.



Abbildung 4: Netz-Visualisierung

Die Fotomontage veranschaulicht die Visualisierung der Ergebnislisten als Informationsnetzwerk, in dem die Verbindungen aus den Literaturreferenzen der einzelnen Elemente der Ergebnismenge ermittelt wurden. Die hiernach wichtigste Literaturstelle ist im Zentrum angeordnet.

2.2 Benutzung recherchierter Daten (Use)

Informationsservices zukünftiger medizinischer Bibliotheken sollten jedoch noch einen Schritt weitergehen, um die Anforderungen der Kunden zu erfüllen und optimal zu unterstützen. Die Literaturrecherche in der medizinischen Forschungsanwendung dient nicht zum Selbstzweck sondern ist integraler Bestandteil neuer Forschungsergebnisse, die wiederum publiziert werden. Diese Publikation kann von einer medizinischen Bibliothek in vielerlei Hinsicht unterstützt werden:

1. *Unterstützung beim Wiederfinden* einmal bezogener Literatur: In eStores wird diese Funktion bereits als selbstverständlich angenommen. Ein einmal bestellter Artikel bleibt in der zentralen Datenbank gespeichert. eProdukte wie elektronische Bücher und Musikdateien können erneut geladen werden, die „Bibliothek“ merkt sich, was der „Kunde“ bereits bestellt und geliefert bekommen hat. Ungeachtet etwaiger Lizenzprobleme wäre es genauso wünschenswert, wenn über diesen „Zwischenpuffer“ die Verfügbarkeitsquote für den Vollzugriff der Institution gesteigert werden könnte – Referenzen, die an andere Benutzer geliefert wurden, sind ebenfalls verfügbar.
2. *Unterstützung beim Weiterverwenden* einmal bezogener Literatur: Ebenso wäre es aus Anwendersicht wünschenswert, wenn in einer zentralen Datenbank (der Bibliothek) gesichert würde, was wann und wo vom Benutzer referenziert wurde. Mit entsprechenden Plugins zu Adobe Acrobat oder Microsoft Word ließe sich durch Rechtsklick ein entsprechendes Kontextmenü öffnen, in dem nicht nur auf den Eintrag mit

gleicher Nummer im Literaturverzeichnis desselben Dokuments verwiesen wird, sondern in dem das PDF direkt zum Öffnen verfügbar ist sowie die bereits erzeugten eigenen und fremden Werke, in denen selbiges Zitat vorkommt. Die hierzu erforderlichen technologischen Voraussetzungen sind vorhanden, aber es fehlt noch an entsprechenden Produkten.

2.3 Annotation recherchierter Daten (Improve)

Es gibt zwar (kommerzielle) Systeme, die die Verwaltung von Zitaten, Referenzen und PDF-Dokumenten unterstützen, aber diese Anwendungen fokussieren den einzelnen Benutzer und unterstützen Arbeitsgruppen nur unzureichend. Innerhalb einer interdisziplinären Arbeitsgruppe sollten alle einmal recherchierten Dokumente allen gleichzeitig verfügbar sein. Dies gilt auch für die Annotationen und Bemerkungen, die der Einzelne zu den Zitaten festhält. In den derzeitigen Single-User Systemen sind die Datenbanken auf dem Rechner der Einzelnen nicht austauschbar. Eine Bibliothek könnte hier die Arbeitsgruppe mit zentralen Diensten unterstützen, zum Beispiel eine rechtssichere Cloud zur Verfügung stellen, in der Teammitglieder copyrightgeschützte Literatur teilen dürfen. Darüber hinaus muss die Trennung zwischen Dokument und Annotation – wie sie in heutigen Systemen vorherrscht – hinterfragt werden. Sicher ist es besser und einfacher sowie genauer, die Anmerkung mit den gängigen PDF-Werkzeugen direkt in das PDF-Dokument an die entsprechende Stelle zu schreiben. Solche Möglichkeiten bietet bereits das Social Network Mendeley [7].

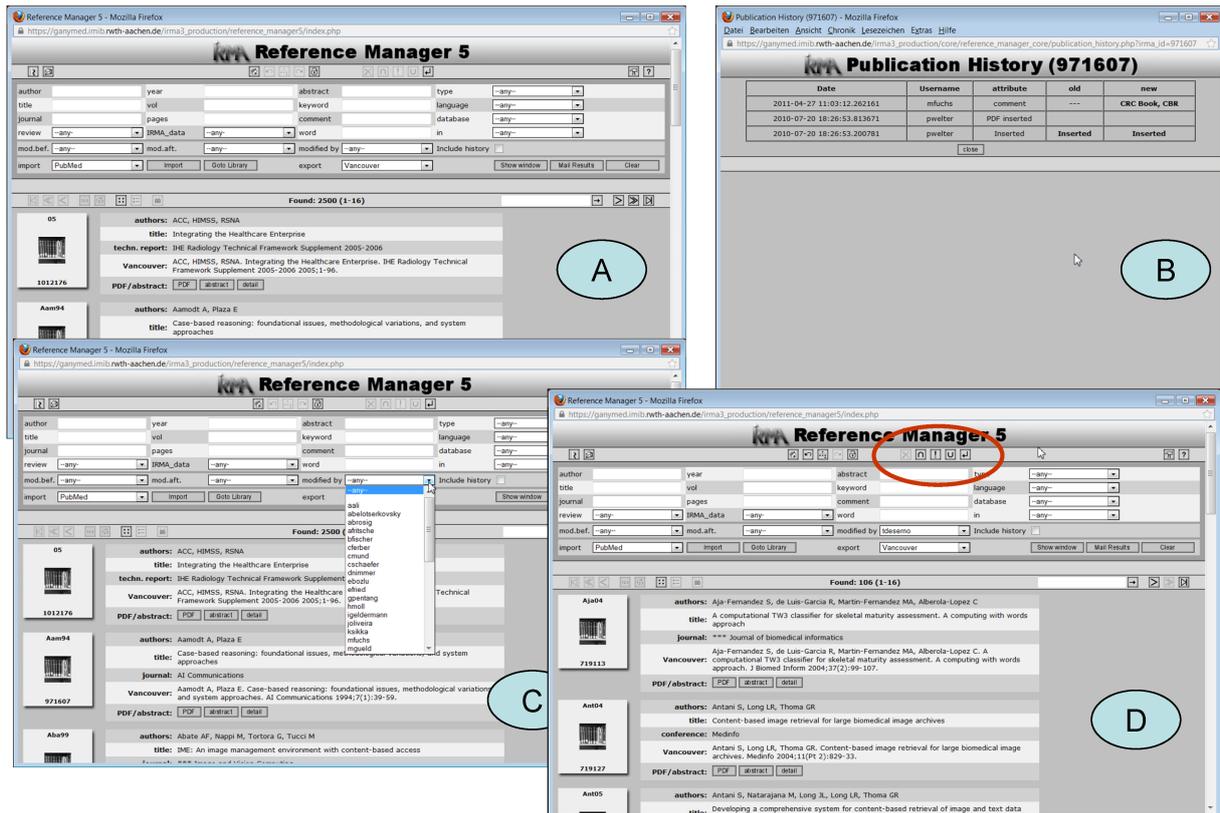


Abbildung 5: IRMA Reference Manager

Die aus dem webbasierten Verwalten großer Bilddatenmengen heraus entstandene Literaturlatenbank verwaltet literarische Metainformation und die als PDF-Dokument gespeicherten Referenzen einheitlich und User-bezogen in der kooperierenden Arbeitsgruppe (A). Annotationen werden personenbezogen erfasst (B) und können für vielfältige Filterungen eingesetzt werden (C). Das Webinterface erlaubt zudem logische Verknüpfungen einzelner Suchanfragen (D).

Dies zeigt anschaulich, dass der Weg von Literaturdaten zu Literaturinformation gehen wird, in dem die einzelnen Literaturstellen einem kontinuierlichen Änderungsprozess unterworfen sind, der gemäß der OAIS-Modellierung mit Bibliotheksdiensten unterstützt werden sollte. Einige Verlage haben dies bereits erkannt und publizieren in elektronischen Plattformen, in denen der Autor eines Beitrages eigenständig und online Änderungen an seinem Beitrag vornehmen kann (auch wenn dies Kommentare seiner Leser bislang noch nicht umfasst) [8]. Abbildung 5 veranschaulicht die Unterstützung von kooperativem Arbeiten am Beispiel des Image Retrieval in Medical Applications (IRMA) Projektes [9], in dessen Rahmen eine zunächst für die Verwaltung von großen annotierten Bildsammlungen konzipierte Datenbankanwendung auf die translationale Literaturverwaltung adaptiert wurde. In (A) ist die webbasierte Oberfläche dargestellt. Alle Änderungen an einem Eintrag werden in einer History automatisch festgehalten [10]. Hiermit kann jederzeit nachverfolgt werden, welcher Nutzer an welchen Einträgen wann welche Annotationen vorgenommen hat (B), was auch in entsprechende Sichten überführt werden kann (C). Darüber hinaus bietet das Interface weitere Möglichkeiten, um Ergebnismengen einzelner Suchanfrage miteinander zu verknüpfen. Diese logischen Mengenoperationen umfassen Und- und Oder-Operation sowie die Negation.

3 Resümee und Ausblick

Gemäß der Definition in Wikipedia ist die „Bibliothek eine Dienstleistungseinrichtung ... [zur] Beschaffung (Share) und Nutzung (Use) von gedruckten Publikationen“ [11]. Die Einschränkung „gedruckt“ sollte nach Auffassung des Autors entfallen. Die in diesem Beitrag durchgeführte Analyse mithilfe des OAIS-Modells hat deutlich gemacht, dass der Prozess der Literaturerzeugung, also der Annotation bestehender (Improve) und der Publikation neuer (Create) Literatur, ebenso von der medizinischen Bibliothek zu unterstützen ist.

Die künftigen Dienste sollten – geordnet nach den persönlichen Prioritäten des Autors – dabei folgendes umfassen:

1. Maximale Verfügbarkeit der elektronischen Volltexte
2. Kooperatives Arbeiten mit zentralen Diensten
3. Erweiterte Suche (Bilder)
4. Interaktive Suche (Ergebnis-Mining nach Relevanz)
5. Annotationen im Dokument (Improve)
6. Unterstützung beim Erzeugen neuer Literaturdokumente (Create)

Eine vollständige Verfügbarkeit der elektronischen Volltexte unterstützt auch die automatische Plagiatdetektion und wird somit helfen, Plagiate zu vermeiden. Automatische Volltextvergleiche sollten dabei als zentraler Service

von der Medizinischen Bibliothek angeboten werden und den Nutzer bei der Erstellung seiner Texte unterstützen. Damit wird das Datenmanagement von Literatur zum Informationsmanagement und dieses wird kooperativ, serviceorientiert und Multi-User tauglich. Die Medizinische Bibliothek wird somit dem Leitspruch der medizinischen Informatik folgen, und die für ihre Benutzer richtige Information zur richtigen Zeit am richtigen Ort bereitstellen [12].

Anmerkung

Danksagung

Ich danke Herrn Dr. Cord Spreckelsen, Aachen, für die kritische Durchsicht des ersten Manuskriptes und seine inhaltlichen Anregungen, die ich gerne aufgenommen habe.

Interessenkonflikte

Der Autor erklärt, dass er keine Interessenkonflikte in Zusammenhang mit diesem Artikel hat.

Literatur

- Literatur. In: Wikipedia. Die freie Enzyklopädie. Available from: <http://de.wikipedia.org/wiki/Literatur>
- Daiber J, Rott H. Wissen und Bedeutung in der Literatur. DFG-Forschungsprojekt. Universität Regensburg; 2009. Available from: <http://www.uni-regensburg.de/forschung/wissen-und-bedeutung-in-der-literatur/>
- Projekt Wissen und Bedeutung in der Literatur - Zwischen Wissenspoetologie und Analytischer Sprachphilosophie. GEPRIS-Suchergebnis. Deutsche Forschungsgemeinschaft (DFG). Available from: <http://gepris.dfg.de/gepris/OCTOPUS/?module=gepris&task=showDetail&context=projekt&id=165173840>
- Space data and information transfer systems – Open archival information system – Reference model. International Organization for Standardization; 2003. Available from: http://www.iso.org/iso/catalogue_detail.htm?csnumber=24683
- Deserno TM, Welter P, Horsch A. Toward a repository for standardized medical image and signal case data annotated with ground truth. *J Digit Imaging*. 2012;25(2):213-26. DOI: 10.1007/s10278-011-9428-4
- Deserno TM, Antani S, Long RL. Content-based image retrieval for scientific literature access. *Methods Inf Med*. 2009;48(4):371-80. DOI: 10.3414/ME0561
- Mendeley Ltd. [mendeley.com]. Available from: <http://www.mendeley.com>
- Springer-Verlag GmbH. Springer Reference [springerreference.com]. Available from: <http://www.springerreference.com/docs/index.html>
- Lehmann TM, Güld MO, Thies C, Fischer B, Spitzer K, Keyzers D, Ney H, Kohnen M, Schubert H, Wein BB. Content-based image retrieval in medical applications. *Methods Inf Med*. 2004;43(4):354-61.
- Deserno TM, Güld MO, Plodowski B, Spitzer K, Wein BB, Schubert H, Ney H, Seidl T. Extended query refinement for medical image retrieval. *J Digit Imaging*. 2008;21(3):280-9. DOI: 10.1007/s10278-007-9037-4
- Bibliothek. In: Wikipedia, Die freie Enzyklopädie. Available from: <http://de.wikipedia.org/wiki/Bibliothek>
- Medizinische Informatik. In: Wikipedia, Die freie Enzyklopädie. Available from: http://de.wikipedia.org/wiki/Medizinische_Informatik

Korrespondenzadresse:

Prof. Dr. Thomas M. Deserno
 Institut für Medizinische Informatik, Universitätsklinikum der RWTH Aachen, Pauwelsstraße 30, 52074 Aachen, Deutschland
deserno@ieee.org

Bitte zitieren als

Deserno TM. Literatur für interdisziplinäre Wissenschaft – eine Anforderungsanalyse aus Anwendersicht. *GMS Med Bibl Inf*. 2012;12(3):Doc20. DOI: 10.3205/mbi000256, URN: urn:nbn:de:0183-mbi0002564

Artikel online frei zugänglich unter

<http://www.egms.de/en/journals/mbi/2012-12/mbi000256.shtml>

Veröffentlicht: 20.12.2012

Copyright

©2012 Deserno. Dieser Artikel ist ein Open Access-Artikel und steht unter den Creative Commons Lizenzbedingungen (<http://creativecommons.org/licenses/by-nc-nd/3.0/deed.de>). Er darf vervielfältigt, verbreitet und öffentlich zugänglich gemacht werden, vorausgesetzt dass Autor und Quelle genannt werden.