

B-WIN4M Baden-Württemberg

Teilprojekt:

Aufbau eines elektronischen wissenschaftlichen  
Publikationsverbunds

---

#### 4. Meilensteinbericht

## **Erste Verfügbarkeitsanalyse elektronischer Volltexte an der Universität Stuttgart und DB-Spezifikation**

Annette Maile<sup>1</sup> und Frank Scholze<sup>2</sup>

11.9.1997

---

<sup>1</sup> Rechenzentrum Universität Stuttgart, Abt. Informationsdienste

<sup>2</sup> Universitätsbibliothek Stuttgart

# Inhaltsverzeichnis

<b>1 EINLEITUNG</b>	<b>1</b>
<b>2 DOKUMENTARTEN</b>	<b>1</b>
<b>3 VERFÜGBARKEIT DER DOKUMENTE</b>	<b>2</b>
3.1 Fakultät 3 (Chemie)	2
3.2 Andere Fakultäten	2
3.3 Rechtliche Fragen hinsichtlich Prüfungsarbeiten	4
3.4 Fazit	5
<b>4 WEITERENTWICKLUNG DES ORGANISATORISCHEN UND TECHNISCHEN KONZEPTS</b>	<b>5</b>
4.1 Vorbereitung der Dokumente	5
4.1.1 Problem bei der Konvertierung von DVI in PDF	5
4.1.2 Lösung bei der Konvertierung von DVI in PDF	6
4.2 Erfassen und Speichern der Metadaten	6
4.2.1 Nachweis der Metadaten im regionalen und lokalen Verbund- system	6
4.2.2 Datenbankauswahl	8
4.2.2.1 Projektspezifische Anforderungen an das DBMS	8
4.2.2.2 Vergleich verschiedener DBMS auf Grundlage des Anforderungskatalogs	10
4.2.3 Datenbank-Struktur	10
<b>5 FAZIT</b>	<b>14</b>

# 1 Einleitung

In Forschung und Lehre werden eine Vielzahl von Dokumenten produziert. Diese sind wissenschaftliche Publikationen (Studien-, Diplom-, Doktorarbeiten), Forschungsberichte, Konferenzproceedings, Vorlesungsmanuskripte, Lehrunterlagen, Fachartikel usw. Ziel des Projekts ist der Aufbau eines elektronischen Publikationsverbunds, der Universitätsangehörige, seien es Lehrende, Lernende oder Forschende, in die Lage versetzt, für sie relevante, veröffentlichte Dokumente aus Forschung und Lehre innerhalb kürzester Zeit auf ihrem Bildschirm einzusehen und auszudrucken.

Im ersten bzw. zweiten Meilensteinbericht wurden zunächst die Online-Publikationskonzepte von Verlagen und Universitäten recherchiert. Daraufhin erfolgte die Untersuchung und Analyse der im WWW vorherrschenden Formate. Aus der Recherche der gängigen DTP-Formate im wissenschaftlichen Bereich resultierte die Untersuchung und Analyse der existierenden Konverter vom DTP-Format zum WWW-Format. Aus diesen Ergebnissen erfolgte eine erste grobe Konzeption über das Einbringen, Abspeichern und Retrieval der Dokumente. Im ersten Teil des dritten Berichts wurden Art, Auswahl und Zahl der möglichen Dokumente in einem elektronischen Publikationsverbund analysiert und beschrieben. Daran anschließend folgte die organisatorische und technische Verfeinerung des Konzepts zur Behandlung dieser Dokumente.

Im vorliegenden Bericht wurde die Verfügbarkeit von Dokumenten an der Universität Stuttgart näher untersucht. Detailspezifikationen (z.B. Datenbankauswahl, Festlegung der Datenbankstruktur) und Lösungen einzelner Problemfelder (z.B. Konvertieren von DVI in PDF) wurden innerhalb des bereits entwickelten Konzepts erarbeitet.

## 2 Dokumentarten

Zusätzlich zur inhaltlichen Typologie, auf die im letzten Bericht näher eingegangen wurde, kann man bereits eine formale Typologie elektronischer Dokumente feststellen.<sup>3</sup>

- Primäre elektronische Publikationen (Erstveröffentlichungen in elektronischer Form)
- Parallele elektronische Publikationen (Publikationen sowohl in gedruckter als auch in elektronischer Form)

---

<sup>3</sup> Vgl. Hans Hehl: Das Internet als Quelle bibliographischer Ermittlung und elektronischer Beschaffung, in: *Bibliotheksdienst* Heft 7, 1997 ([http://www.dbi-berlin.de/dbi\\_pub/bd\\_art/97\\_07\\_00.htm](http://www.dbi-berlin.de/dbi_pub/bd_art/97_07_00.htm))

- Elektronische Preprints (der Physik, Mathematik, Kongreßpreprints)
- Elektronische Reprints (Digitalisierung gedruckter Texte)

Für den Publikationsverbund kommen zunächst nur die ersten drei Formen in Betracht, da im Rahmen des Projekts keine retrospektive Konvertierung bzw. Scannen von konventionell vorliegenden Dokumenten vorgenommen werden soll<sup>4</sup>.

## 3 Verfügbarkeit der Dokumente

In einer Umfrage zur elektronischen Fachinformation, konzipiert von Herrn Kämper (UB) und Frau Bassen (Institut für physikalische Chemie) vom Arbeitskreis Fachinformation, wurden alle Institute der Universität u.a. auch zum gegenwärtigen Stand elektronischer Publikationen befragt. Die Institute der Fakultät Chemie wurden dabei ausführlicher auch nach der Bereitschaft zur zukünftigen Ablieferung elektronischer Volltexte befragt. Die Umfrage bildet die Grundlage einer detaillierteren Verfügbarkeitsanalyse.

### 3.1 Fakultät 3 (Chemie)

Alle 11 Institute der Fakultät antworteten (zum Teil mußte gemahnt werden), davon sind 9 grundsätzlich zur Mitarbeit bereit. Das Institut für Technische Chemie I (0307) und das Institut für Nichtmetallische Anorganische Materialien<sup>5</sup> (0311) sind nicht bereit, elektronische Volltexte zur Verfügung zu stellen.

Format für die Erstellung elektronische Volltexte ist überwiegend Word (10). Ebenfalls genannt wurden LaTeX (2) Word-Perfect (1) und Star (1).

Als Archivierungsformat genannt wurden Word (5), Postscript (2), LaTeX (2), Word-Perfect (1)

Nur das Institut für Theoretische Chemie (0304) betreibt einen Dokumentenserver mit Institutsberichten, Reports, Hard- und Softwareanleitungen etc.

### 3.2 Andere Fakultäten

Die übrigen Fakultäten der Universität Stuttgart wurden danach befragt, ob elektronische Volltexte zugänglich gemacht werden und wenn ja in welchem Format die Texte angeboten werden.

---

<sup>4</sup> Einen derartigen Ansatz verfolgt beispielsweise das Projekt Digitale Bibliothek Frankfurt. Vgl. <http://www.tm.informatik.uni-frankfurt.de/diglib/>

<sup>5</sup> Abstracts würden hier jedoch zur Verfügung gestellt werden, wohingegen das Institut für Technische Chemie I keinerlei Zusammenarbeit wünscht.

Fakultät		antwortende Institute	davon bieten elektronische Dokumente an
1	Architektur und Stadtplanung	7	-
2	Bauingenieur- und Vermessungswesen	20	3 (203, 206, 219)
4	Elektrotechnik	7	2 (403, 407)
5	Energietechnik	9	4 (501-03, 505)
6	Konstruktions- und Fertigungstechnik	12	3 (601, 607, 612)
7	Geo- und Biowissenschaften	8	1 (701 nur Abstracts)
8	Geschichts-, Sozial- und Wirtschaftswissenschaften	1	-
9	Luft- und Raumfahrttechnik	4	1 (903)
10	Mathematik	1	Fakultät
11	Philosophie	8	1 (1105)
12	Physik	9	5 (1203-07)
13	Verfahrenstechnik	keine Antwort	
14	Informatik	2	Fakultät

(Tab. 3.1: Umfrage zu elektronischen Volltexten an der Universität Stuttgart)

Als Dokumentformate für die Archivierung von Texten genannt wurden:

- HTML - 34x
- PostScript - 26x
- TeX/LaTeX - 14x
- Word - 14x
- ASCII - 7x
- DVI - 7x
- Word Perfekt - 6x
- PDF - 2x

Wobei auch von Instituten, die keine Dokumente im Volltext oder als Abstract anbieten, Formate genannt wurden. Die Frage nach einem Dokumentformat wurde wahrscheinlich in einigen Fällen auf das WWW generell und nicht auf den Dokumenten-Server des Instituts bezogen. Dies erklärt mit, warum HTML am häufigsten genannt wurde, obwohl es in der Praxis nicht das Format ist, in dem die Mehrzahl der Volltext-Dokumente der Universität Stuttgart angeboten wird.

Die Dokumentserver der Institute sind zum Teil nicht öffentlich bzw. über WWW zugänglich.

### 3.3 Rechtliche Fragen hinsichtlich Prüfungsarbeiten

Vom Institut für Anorganische Chemie (0301) und vom Institut für Technische Biochemie (0310) wurde auf die persönliche Einverständniserklärung von Diplomanden und Doktoranden als Voraussetzung zu einer elektronischen Veröffentlichung hingewiesen. Auf Diplom-, Studien- und Magisterarbeiten trifft dies auch uneingeschränkt zu, da in diesen Fällen das Sach- bzw. Eigentumsrecht beim Verfasser bzw. Urheber verbleibt. Eine Veröffentlichung kann somit juristisch nicht verpflichtend verlangt oder verboten werden.<sup>6</sup> Sie kann sogar im Hinblick auf eine mögliche anschließende Promotion aus Sicht des Verfassers in manchen Fällen unerwünscht sein.

Im Fall der Dissertation geht das Sacheigentum jedoch in das Eigentum der Universität über, die Promotionsordnung regelt dabei Form und Art der möglichen Veröffentlichung. Insoweit ist das Urheberrecht des Verfassers eingeschränkt.<sup>7</sup>

Ein Entwurf zur Änderung der Promotionsordnung der Universität Stuttgart (§ 9, Abs. 2) ist vorbereitet worden. (Anlage) Dieser sieht die zusätzliche Möglichkeit der Ablieferung einer Dissertation in elektronischer Form vor. Das der Promotionsordnung anhängende „Merkblatt für Doktoranden“ soll entsprechend erweitert werden. Hier werden formale Spezifikationen zur elektronischen Dissertation festgelegt, wie

- Ablieferungsart (voraussichtlich File-Upload, Diskette)
- Ablieferungsformat (PDF)
- Verpflichtung des Doktoranden, Metadaten (einschließlich Schlagwörtern und Zusammenfassung) abzuliefern
- Authentizitätsgarantie des Doktoranden (Übereinstimmung der elektronischen Dissertation mit den abgegebenen Papierexemplaren in Form und Inhalt)

Auch an anderen Universitäten sind derartige Änderungen der Promotionsordnungen geplant, in den meisten Fällen bislang jedoch nicht realisiert. Gegenwärtig ist daher die Ablieferung der Dissertation in elektronischer Form eine zusätzlich zur bislang vorgeschriebenen Ablieferungspflicht erbrachte Leistung des Doktoranden (Karlsruhe, Augsburg, Konstanz u.a.).

An der Universität Marburg ist eine „Vorläufige Regelung zur Ablieferung von digitalen Dissertationen“ seitens der UB erstellt worden, die jedoch noch nicht in die bestehende Promotionsordnung Eingang gefunden hat.<sup>8</sup> Die Universität Hannover ermöglicht auf-

---

<sup>6</sup> Vgl. Rudolf Kraßer, Gerhard Schrickler: *Patent- und Urheberrecht an Hochschulen: die Verwertung von Erfindungen und urheberrechtlich geschützten Werken an Hochschulen wissenschaftlich Tätiger*. Baden-Baden, 1988, S. 150.

<sup>7</sup> Auskunft des Dezernat Recht der Universität Stuttgart vom 16.06.97 (Az. V 0521.3)

<sup>8</sup> <http://www.ub.uni-marburg.de/meta/sta/sta4/1997-04.html>

grund eines Senatsbeschlusses die Ablieferung elektronischer Dissertationen seit dem 9. Juli 97.<sup>9</sup> Hier sind sechs maschinenschriftliche Exemplare zusätzlich zur elektronischen Form abzugeben.

### **3.4 Fazit**

Aufgrund der Umfrageergebnisse innerhalb der Fakultät Chemie und weiteren Gesprächen, die mit der Fakultät Mathematik geführt wurden, läßt sich feststellen, daß die generelle Bereitschaft zur Ablieferung elektronischer Dokumente vorhanden ist. Fragen der Qualitätssicherung und der Vollständigkeit müssen dabei weiter diskutiert werden. Eine spezifische Umfrage zum Online Publikationsverbund wurde vorbereitet und soll in Abstimmung mit dem Arbeitskreis Fachinformation im Oktober 1997 durchgeführt werden.

## **4 Weiterentwicklung des organisatorischen und technischen Konzepts**

Die Struktur des organisatorischen Konzepts zur Behandlung der Dokumente und der sie beschreibenden Metadaten im elektronischen Publikationsverbund wurde in den vorangegangenen Berichten prinzipiell beschrieben und festgelegt. Davon ausgehend erfolgten Detailspezifikationen und Lösungen einzelner Problemfelder innerhalb des Konzepts.

### **4.1 Vorbereitung der Dokumente**

Ein großer Teil der technisch-naturwissenschaftlichen Dokumente wird mit Hilfe von TeX bzw. LaTeX erstellt. Deren Konvertierung in PDF gestaltet sich komplexer als die Konvertierung aus anderen DTP-Systemen.

#### **4.1.1 Problem bei der Konvertierung von DVI in PDF**

Dokumente, die von TeX oder LaTeX in PS oder PDF konvertiert werden, enthalten oft nicht skalierbare Fonts, die am Bildschirm oder beim Ausdruck zu schlechten Resultaten führen. Dies liegt daran, daß PostScript die Fonts in unterschiedlichen Formaten darstellen kann. PostScript Typ 3 Fonts können entweder als fixed-resolution bitmaps oder als skalierbare Vektor-Fonts beschrieben werden. PostScript Typ 1 Fonts sind auflösungsunabhängige Vektor-Fonts, die an geringe Auflösungen und kleine Fontgrößen angepaßt werden können.

Die „Computer Modern Fonts“, die üblicherweise mit Donald Knuth's TeX verwendet werden, werden als fixed-resolution bitmaps ausgegeben und im DVI-Format (Device

---

<sup>9</sup> <http://www.tib.uni-hannover.de/diss/disshome.htm>

independent) gespeichert. Das Programm DVIPS, welches aus DVI PostScript erstellt, konvertiert normalerweise diese Fonts in Typ 3 fixed-resolution bitmap fonts.

Wenn ein solches PS-File in PDF konvertiert wird, so ist das Resultat am Bildschirm sehr schlecht. Da PDF dafür konzipiert wurde, daß es an unterschiedlichen Monitoren und Druckern mit unterschiedlichen Auflösungen angesehen bzw. ausgedruckt wird, müssen skalierbare Vektor-Fonts benutzt werden.

#### **4.1.2 Lösung bei der Konvertierung von DVI in PDF**

Um die Typ 3 Fonts mit fester Auflösung durch skalierbare PS-Typ 1 Fonts zu ersetzen, werden die entsprechenden Fonts incl. Fontmap-Datei in dem allgemein zugänglichen Verzeichnis

```
/sw/share/ps_type1_fonts/pfb
```

zur Verfügung gestellt.

Nun muß nur noch dieses Verzeichnis durch den System Administrator zugänglich gemacht werden und die Konfigurationsdatei für DVIPS

```
config.distillembed
```

in das eigene Home-Directory kopiert werden.

Das Konvertieren des DVI-Files mit gleichzeitigem Ersetzen der Fonts in ein PS-File erfolgt mit dem Befehl

```
dvips -P distillembed <Name des dvi-Files>
```

Beispiel: Konvertieren von test.dvi in test.ps

```
dvips -P distillembed test.dvi
```

Falls die Quelldatei, d.h. TeX- oder DVI-File, nicht verfügbar ist, so ist dieses Verfahren nicht anwendbar. Es gibt eine Fixfont-Utility, die versucht im PS-File die Typ 3 Fonts durch Typ 1 Fonts zu ersetzen, jedoch waren die Resultate nicht sichtbar besser.

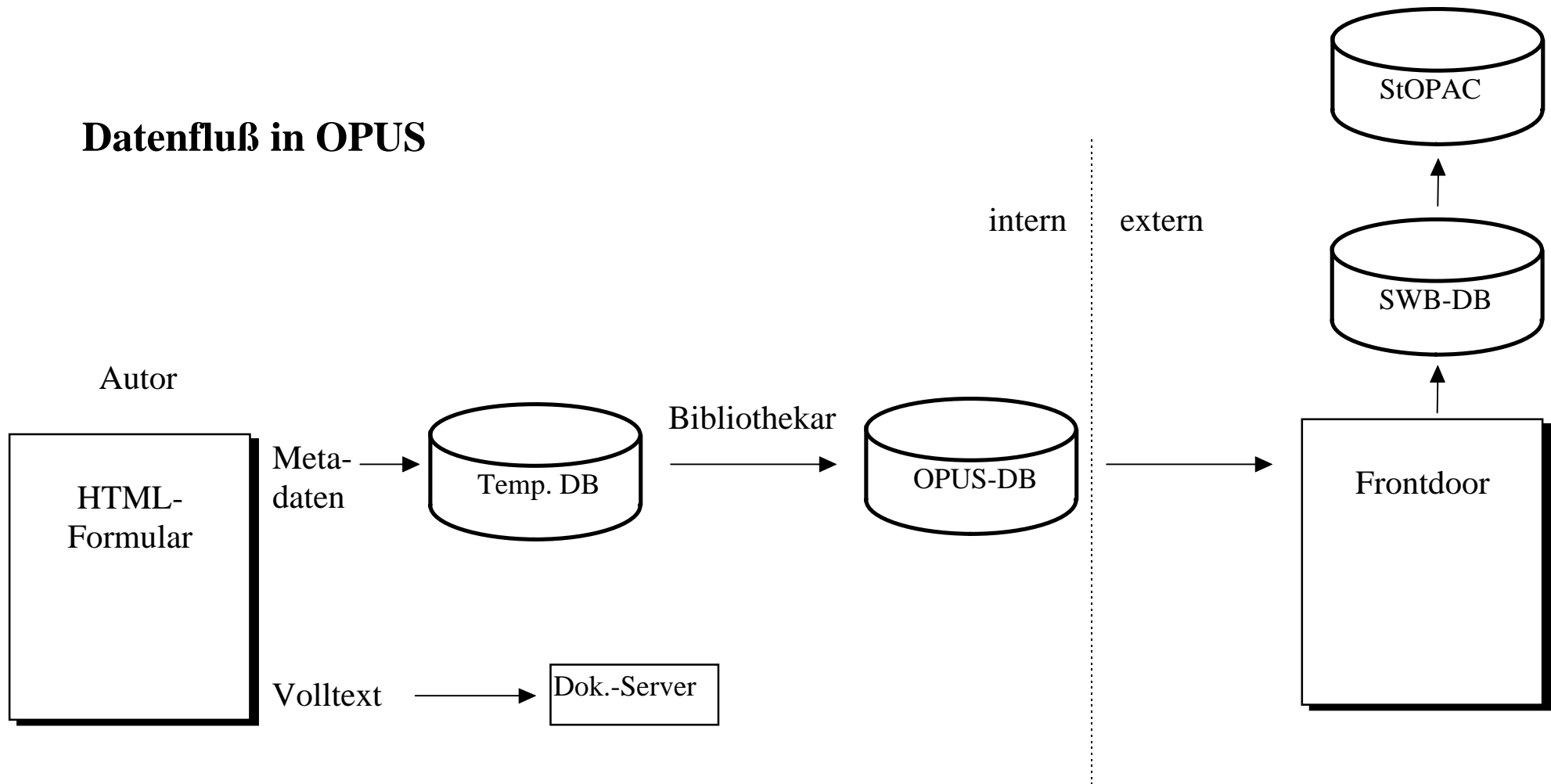
## **4.2 Erfassen und Speichern der Metadaten**

Bei diesem Punkt des Konzepts werden sowohl technische wie auch organisatorische Detailprobleme behandelt.

### **4.2.1 Nachweis der Metadaten im regionalen und lokalen Verbundsystem**

Bei einem Treffen mit Vertretern des Bibliotheksservicezentrums am 7. Juli 1997 in Konstanz wurde eine enge Zusammenarbeit vereinbart, um den Austausch der Metadaten technisch und organisatorisch zu gewährleisten (Abb. 4.1).

# Datenfluß in OPUS



Temp. DB:      Metadaten, temp. URL Volltext  
OPUS - DB:    Metadaten, URL Volltext

(Abb. 4.1: Datenfluß in OPUS)

Der Nachweis elektronischer Volltexte im lokalen bibliographischen Datenpool StOPAC gestaltet sich schwierig, da die Daten aus der Verbunddatenbank des SWB übernommen werden. Für die Übernahme der IBAS-Kategorien (SWB), die die URL des Volltextservers enthalten, sind in der OLIX-Datenbank (StOPAC) jedoch keine korrespondierenden MAB-Kategorien definiert, da zum Zeitpunkt der Einführung dieser neuen Kategorien die Entwicklung des OLIX-Systems bereits abgeschlossen bzw. abgebrochen war. Die Informationen dieser Kategorien gehen derzeit bei der Übernahme der Daten in den StOPAC vollständig verloren. Um diesem gravierenden Informationsverlust zu begegnen, hat die UB Mannheim eine eigene Lösung des Problems erarbeitet. Dort wird die URL beim Umsetzen der Daten in Kategorie 501 geschrieben. Mit Hilfe eines selbst entwickelten CGI-Interface, kann diese Kategorie bei der Abfrage im WWW als Link zum Volltextdokument angezeigt werden.<sup>10</sup> Eine gleichwertige Lösung ist auch für Stuttgart anzustreben.

## **4.2.2 Datenbankauswahl**

Die im letzten Bericht nur angedeutete Auswahl des DBMS soll hier noch einmal ausführlicher erläutert werden.

### **4.2.2.1 Projektspezifische Anforderungen an das DBMS**

Aufgrund des Projektziels und des daraus resultierenden Konzepts ergaben sich folgende Anforderungen an das DBMS:

- WWW-Gateway (als Bestandteil des DBMS)
  - Eingabe und Recherche über HTML-Formular
- Unterstützung verschiedener Datentypen:
  - „char“ von variabler Länge bis 2000 Zeichen
  - Datumsfelder, numerische Felder
- Boolesche Suche
- Stopwortlisten
- Sortiermöglichkeit
- Verwaltung und Speicherung von Volltexten in PDF
- Systemgrenzen
  - maximale Datenbankgröße (mind. 50 MB)
  - maximale Anzahl der möglichen gleichzeitigen Nutzer
- Zugriffskontrolle
- gesicherte Marktpräsenz des Produkts und Support
- Preis

---

<sup>10</sup> Die Mannheimer Lösung ist näher beschrieben in: *Bibliotheksdienst* Heft 4, 1997 ([http://www.dbi-berlin.de/dbi\\_pub/bd\\_art/97\\_04\\_06.htm](http://www.dbi-berlin.de/dbi_pub/bd_art/97_04_06.htm))

---

Anforderungen <sup>11</sup>	Access 97	BRS	Hyperwave	Informix	mSQL	Opentext	Oracle 7.3
WWW-Gateway	-	+	+	+	+	+	+
Eingabe und Recherche über HTML-Formular	-	+		+	+	+	+
„char“ von variabler Länge bis 2000 Zeichen	-	+	-	+	+/-	+	+
Datumfelder, numerische Felder	+	+	-	+	+	-	+
Boolsche Suche		+	-	+	+	+	+
Stopwortlisten		+	+	+	-	+	+
Sortiermöglichkeit		+	-	+	+	+	+
Verwaltung und Speicherung von Volltexten in PDF	-	-	-		-	+	+
maximale Datenbankgröße (mind. 50 MB)	+	+		+	+		+
maximale Anzahl der möglichen gleichzeitigen Nutzer		+	+		+		
Zugriffskontrolle		-	+		+	-	+
Marktpräsenz des Produkts und Support	+		-	+		+	+
Preis		38.000 DM / 8 Nutzer incl. Web- Server	frei		frei	40.000 DM excl. Web- Server	ca. 25.000 DM

(Tab. 4.1: Vergleich verschiedener DBMS auf Grundlage des Anforderungskatalogs)

---

<sup>11</sup> Großteil der Informationen aus: Jochen Meyer (Redaktion): *Entscheidungsgrundlage für die Auswahl von Volltextdatenbanksystemen für den MeDoc-Dienst*, <http://www-is.informatik.uni-oldenburg.de/~meyer/medoc/papers/vtdbms.ps>

---

#### 4.2.2.2 Vergleich verschiedener DBMS auf Grundlage des Anforderungskatalogs

Aufgrund des Anforderungsprofils wurden verschiedene DBMS vergleichend bewertet (Tab. 4.1). Dabei waren WWW-Gateway als Bestandteil des DBMS, Eingabe und Recherche über HTML-Formular mit die wichtigsten Anforderungen. Die Verwaltung und Speicherung von Volltexten in der Datenbank war kein zwingendes Kriterium für die DB-Auswahl, ist aber aufgrund der größeren Offenheit für zukünftige Entwicklungen ein wünschenswertes Merkmal hinsichtlich des Dokumentmanagements. Nach dem projektspezifischen Anforderungsprofil erwiesen sich Oracle und (mit kleineren Abstrichen) mSQL als die am besten geeigneten DBMS für den Publikationsverbund. Aufgrund der komplexen Administration und der zu erwartenden hohen Kosten von Oracle fiel die Entscheidung zunächst zu Gunsten von mSQL. Dieser Entschluß wurde dadurch erleichtert, daß es unter mSQL möglich ist, einen *Dump* der Datenbank im ASCII Format zu erstellen. Dieser *Dump* enthält alle notwendigen SQL-Befehle zur Rekonstruktion der Datenbank. Dadurch ist es möglich, die Datenbankinhalte in eine andere SQL-Datenbank (z.B. Oracle) zu portieren.

#### 4.2.3 Datenbank-Struktur

Basierend auf den Spezifikationen zum Dublin Core Element Set, die gemeinsam mit dem Bibliotheksservice-Zentrum erarbeitet wurden, ergab sich folgende Datenbank-Struktur. Die vom Dokumentlieferanten obligatorisch auszufüllende Kategorien sind **fett** markiert.

	<b>Kategorie</b>	<b>Attributname</b>	<b>Datentyp</b>
1	<b>Titel</b> (DC.TITLE)	TITLE	char
2	<b>Verfasser</b> (DC.CREATOR.NAME)	CREATOR_NAME	char
3	<b>Urheber</b> (DC.CREATOR.CORPORATE)	CREATOR_CORPORATE	char
4	Inhaltliche Erschließung (DC.SUBJECT) Schlagwörter aus SWD (Scheme=SWD)	SUBJECT_SWD	char
5-20	Inhaltliche Erschließung (DC.SUBJECT) Fachspezifische Klassifikationen wie CR etc. (Scheme=CR...)	SUBJECT_CR etc.	char
21	Inhaltliche Beschreibung (DC.DESCRPTION)	DESCRIPTION	char
22	<b>Verleger bzw. Herausgeber</b> (DC.PUBLISHER) Universität oder Körperschaft	PUBLISHER_UNIVERSITY	char
23	<b>Verleger bzw. Herausgeber</b> (DC.PUBLISHER) Fakultät	PUBLISHER_FACULTY	char

	<b>Kategorie</b>	<b>Attributname</b>	<b>Datentyp</b>
24	Weitere beteiligte Personen und Körperschaften (DC.CONTRIBUTORS.NAME)	CONTRIBUTORS_NAME	char
25	Weitere beteiligte Personen und Körperschaften (DC.CONTRIBUTORS.CORPORATE)	CONTRIBUTORS_CORPORATE	char
26	<b>Datum</b> (DC.DATE) Erstellungsjahr der Ressource	DATE_YEAR	char
27	Veröffentlichungsdatum (DC.DATE.CREATION) intern	DATE_CREATION	date
28	Datum der letzten Änderung (DC.DATE.MODIFIED) intern SWB-Vorschlag: DC.DATE.CURRENT	DATE_MODIFIED	date
29	<b>Dokumentart</b> (DC.TYPE)	TYPE	char
30	Format (DC.FORMAT)	FORMAT	char
31	<b>Ressourcen-Identifikation</b> (DC.IDENTIFIER) Scheme=URL, ISBN etc.	IDENTIFIER	char
32	Quelle (DC.SOURCE) In diesem Element wird das gedruckte oder elektronische Werk, aus dem diese Ressource stammt, eingetragen.	SOURCE_TITLE	char
33	Verknüpfung mit SWB-Identnummer (Scheme=SWBidn?) SWB-Vorschlag: CONTENT=SWB-IDNR/1234567	SOURCE_SWB	char
34	<b>Sprache</b> (DC.LANGUAGE)	LANGUAGE	char
35	Zeitliche Maßangabe (DC.COVERAGE.TEMPORAL)	COVERAGE_TEMPORAL	date
36	Räumliche Maßangabe (DC.COVERAGE.SPATIAL)	COVERAGE_SPATIAL	char
37	Rechtliche Bedingungen (DC.RIGHTS)	RIGHTS	char
38	<b>Identifizierung</b> E-Mail-Adresse, Tel., etc. (nur temporär)	VERIFICATION	char
39	Dateigröße des Volltextes	FILESIZE	char
40	Beziehung zu anderen Ressourcen (DC.RELATION)	RELATION	char

(Tab. 4.2: Datenbank-Struktur der OPUS-DB)

Aus dieser Datenbank-Struktur ergibt sich folgende Struktur für das Eingabeformular der Metadaten.

	<b>Beschreibung Kategorie</b>	<b>Formularfeldtyp</b>	<b>Formularfeldname</b>
1	<b>Titel</b> (DC.TITLE)	text	TITLE
2	<b>Verfasser</b> (DC.CREATOR.NAME)	text	CREATOR_NAME
3	<b>Urheber</b> (DC.CREATOR.CORPORATE)	text	CREATOR_CORPORATE
4	Inhaltliche Erschließung (DC.SUBJECT) Schlagwörter aus SWD (Scheme=SWD)	text	SUBJECT_SWD
5-20	Inhaltliche Erschließung (DC.SUBJECT) Fachspezifische Klassifikationen wie CR etc. (Scheme=CR...)	text	SUBJECT_CR etc.
21	Inhaltliche Beschreibung (DC.DESCRPTION)	text	DESCRIPTION
22	<b>Verleger bzw. Herausgeber</b> (DC.PUBLISHER) Universität oder Körperschaft	Liste default „Universität Stuttgart“	PUBLISHER_UNIVERSITY
23	<b>Verleger bzw. Herausgeber</b> (DC.PUBLISHER) Fakultät	Liste	PUBLISHER_FACULTY
24	Weitere beteiligte Personen und Kör- perschaften (DC.CONTRIBUTORS.NAME)	text	CONTRIBUTORS_NAME
25	Weitere beteiligte Personen und Kör- perschaften (DC.CONTRIBUTORS.CORPORATE)	text	CONTRIBUTORS_CORPORATE
26	<b>Datum</b> (DC.DATE) Erstellungsjahr der Ressource	text (4) default = aktuelles Jahr	DATE_YEAR
27	<b>Dokumentart</b> (DC.TYPE)	Liste	TYPE
28	Format (DC.FORMAT)	Liste default „application/pdf“	FORMAT
29	<b>Ressourcen-Identifikation</b> (DC.IDENTIFIER)	file	IDENTIFIER
30	Quelle (DC.SOURCE) In diesem Element wird das gedruckte oder elektronische Werk, aus dem diese	text	SOURCE_TITLE

	<b>Beschreibung Kategorie</b>	<b>Formularfeldtyp</b>	<b>Formularfeldname</b>
	Ressource stammt, eingetragen.		
31	<b>Sprache</b> (DC.LANGUAGE)	Liste default „deutsch“	LANGUAGE
32	<b>Identifizierung</b> E-Mail-Adresse, Tel., etc. (nur temporär)	text	VERIFICATION
33	Dateigröße des Volltextes	text	FILESIZE

(Tab. 4.3: Struktur des Eingabeformulars für OPUS)

## 5 Fazit

Die Verfügbarkeit elektronischer Dokumente wurde in einer ersten Umfrage untersucht. Es zeigte sich, daß die prinzipielle Bereitschaft zur Ablieferung vorhanden ist. Ein detailliertere Umfrage zu diesem Thema wird zu Beginn des Wintersemesters 1997/98 durchgeführt.

Das Problem der schlechten Darstellungsqualität bei der Konvertierung von DVI in PDF wurde gelöst. Auf der Grundlage der projektspezifischen Anforderungen an das DBMS erwiesen sich Oracle und mSQL als die am besten geeigneten DBMS für den Publikationsverbund. Die Entscheidung fiel zunächst zu Gunsten von mSQL.

Basierend auf dem Dublin Core Element Set, wurde die Datenbankstruktur und die daraus abgeleitete Struktur des Eingabeformulars festgelegt. Eingabe, Änderung und Recherche der Daten über ein HTML-Formular wurden während der Fertigstellung dieses Berichts prototypisch mit mSQL realisiert. Genauere Ergebnisse hierzu können deshalb erst im nächsten Bericht dargestellt werden.

Gestaltung und Umfang der Präsentation des Projekts im WWW sind weitgehend abgeschlossen. Zur Zeit ist eine Grundversion an Informationsseiten verfügbar (vgl. Abb. 5.1), die im weiteren Verlauf des Projekts sukzessive erweitert wird.



(Abb. 5.1: Basisversion der Einstiegsseite des Projekts im WWW)

## Anlage

### **Entwurf zur Änderung der Promotionsordnung der Universität Stuttgart**

#### §9 (2)

e) 6 Exemplare in kopierfähiger Maschenschrift zusammen mit einer Kopie in Form einer maschinenlesbaren Datei. In diesem Fall überträgt der Bewerber der Universität (ggf. auch der Deutschen Bibliothek in Frankfurt/Leipzig und der zuständigen Sondersammelgebietsbibliothek) das Recht, diese Form der Dissertation elektronisch zu speichern und in Datennetzen öffentlich zugänglich zu machen.

Im Fall einer Ablieferung nach §9 (2) a-d stellt der Bewerber ebenfalls eine Kopie der Dissertation in Form einer maschinenlesbaren Datei zur Verfügung.