

Informationen zu z39.50

Ein Vortrag gehalten beim SISIS-Systemverwaltertreffen am 17. Januar 2001 in Heidelberg

Grundlage :

Z39.50 Grundlage

- **Kommunikationsprotokoll**
 - Entwicklung seit 1984
 - OSI-Anwendungsprotokoll
 - Heterogene Systeme
- **Anwendungen**
 - Verbund-OPACs
 - Fachinformationssysteme
 - Bibliothekssysteme
- **Projekt DBV-OSI-II**

Die Wurzeln des Protokolls Z39.50 reichen zurück bis ins Jahr 1984. Die derzeit gültige Definition Z39.50 V3 wurde 1995 verabschiedet. Die Vorgängerversionen 1 und 2 stammen aus den Jahren 1988 und 1992. Oft wird auch als Bezeichnung die Jahreszahl angehängt. Z39.50 wurde als OSI-Anwendungsprotokoll, d.h. als Protokoll der Ebene 7 des OSI-Schichtenmodells, definiert. Federführend als Maintenance Agency für Z39.50 ist die Library of Congress. Daneben gibt es noch die ZIG, die Z39.50 Implementors Group, die sich mit der Weiterentwicklung von Z39.50 beschäftigt und mit der Realisierung von Z39.50-Anwendungen.

Z39.50 ist eine Schnittstelle bzw. ein Protokoll zur Online-Kommunikation zwischen heterogenen Bibliothekssystemen.

Z39.50 ist heute die gebräuchlichste Schnittstelle, die für die Kommunikation zwischen unterschiedlichen bibliographischen Datenbanken verwendet wird. Also überall dort wo auf unterschiedliche bibliographische Datenquellen zugegriffen wird geschieht dies meist über Z39.50.

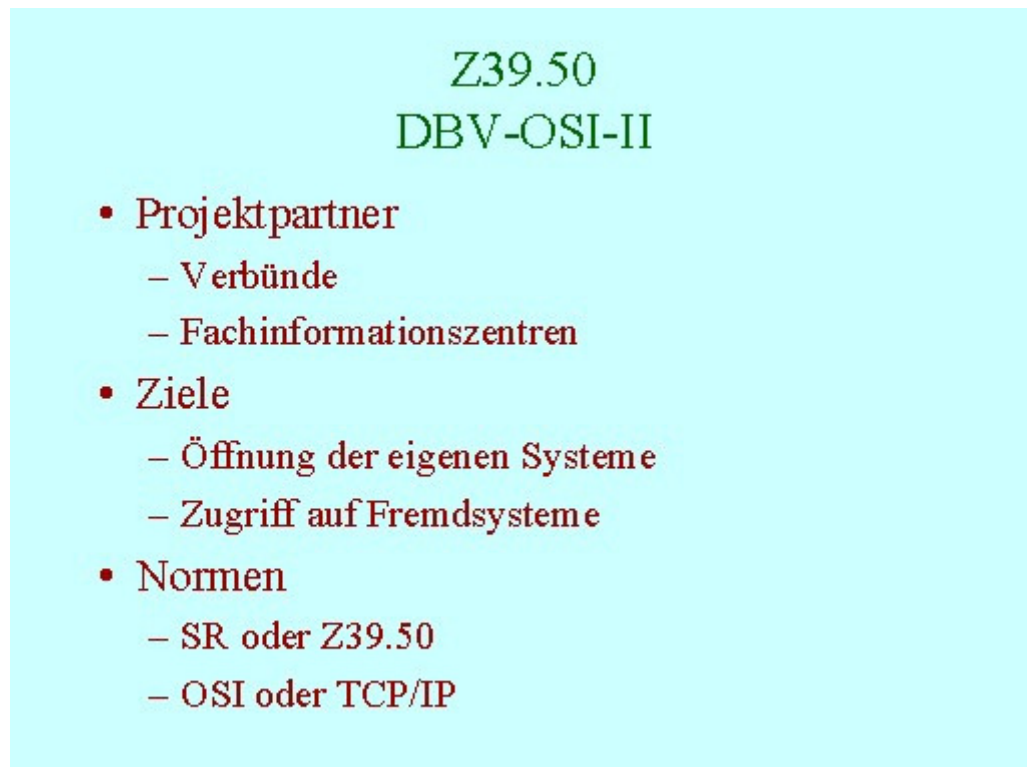
Über einen Z39.50-Client hat man die Möglichkeit auf die Daten anderer Verbünde zuzugreifen oder auf andere Datenbanken wie die ZDB (Zeitschriftendatenbank) oder auf Fachinformationsdatenbanken. Z39.50 dient primär dazu auf fremde

Datenquellen zuzugreifen und, wenn dies erlaubt wird, die Daten nicht nur anzuzeigen, sondern auch in einem strukturierten Format zu übernehmen.

Die Basis für alle diese Lösungen wurde in Deutschland durch das Projekt DBV-OSI-II gelegt.

Das Projekt DBV-OSI-II wurde 1993 begonnen und wurde vom Bund und der DFG gefördert. Das Projekt wurde 1997 erfolgreich abgeschlossen.

DBV-OSI-II :



Z39.50
DBV-OSI-II

- **Projektpartner**
 - **Verbünde**
 - **Fachinformationszentren**
- **Ziele**
 - **Öffnung der eigenen Systeme**
 - **Zugriff auf Fremdsysteme**
- **Normen**
 - **SR oder Z39.50**
 - **OSI oder TCP/IP**

Es waren fast alle Verbünde beteiligt und verschiedene Fachinformationszentren. Zunächst sollten auf Basis der ISO-Norm SR (Search and Retrieve) und den ISO-Kommunikationsprotokollen entsprechende Anwendungen realisiert werden und die einzelnen Systeme geöffnet werden.

Die Projektpartner waren :

- Die Deutsche Bibliothek (DDB), bei der auch die Projektleitung lag,
- Deutsches Bibliotheksinstitut (DBI),
- Fachinformationszentrum Karlsruhe (FIZ),
- Deutsche Institut für Medizinische Dokumentation und Information (DIMDI),
- Bibliotheksverbund Bayern (BVB),
- Südwestdeutscher Bibliotheksverbund (SWB),
- Gemeinsamer Bibliotheksverbund (GBV),
- PICA.

Nicht beteiligt waren z.B. HBZ für NRW und auch Hessen. Zentral sollte die Kommunikationsmaschine mit dem Protokoll-Stack realisiert werden, für den Einsatz bei den Projektpartnern.

Die Verbund- und Fachinformationssysteme sollten ihre OPACs um die SR-Schnittstelle erweitern. Eine Besonderheit in Bayern war, dass auch die SISIS-Lokalsysteme mit eingebunden waren.

Im Laufe der Konzeptionsphase von DBV-OSI-II wurde dann beschlossen statt der ISO-Norm SR und der OSI-Kommunikationsprotokolle das Projekt auf der Basis von Z39.50 und TCP/IP zu realisieren.

Z39.50 war zwar nur eine ANSI-Norm und TCP/IP nur ein Quasi-Standard, aber das waren die Protokolle, die in der Realität, vor allem in Amerika, bereits weit verbreitet waren.

Alle Dienste, die es in SR gab, waren auch in Z39.50 in gleicher Weise enthalten, Z39.50 enthielt aber wichtige weitere Dienste. Das TCP/IP-Protokoll ist mit der Entwicklung des Internet das Standardkommunikationsprotokoll geworden.

Services :



**Z39.50
Services**

- **Einheitliche Festlegungen**
- **Bereiche**
 - **Dienste**
 - **Attribute**
 - **Datenformate**
- **Dienste**
 - **Basisdienste**
 - **Erweiterte Dienste**

Z39.50 läßt wie alle Kommunikationsprotokolle einen Interpretationsspielraum, d.h. selbst wenn zwei Systeme Z39.50 unterstützen müssen sich diese noch nicht verstehen.

Eine der Hauptaufgaben von DBV-OSI-II war Festlegungen zu finden, die eine einwandfreie Kommunikation zwischen den Systemen der beteiligten Projektpartner erlauben sollten.

Die Grundfunktionen von Z39.50 beschreiben, wie eine Suchanfrage zu formulieren ist und wie das Ergebnis der Suche zurückgeliefert wird.

Drei Bereiche sind hierbei besonders wichtig:

- Die unterstützten Dienste,
- das Attribut-Set, also welche Suchmöglichkeiten werden unterstützt,
- das Datenformat, also wie werden die Daten übertragen

Es galt zunächst festzulegen welche Dienste sollen unterstützt werden.

Basisdienste:

- Init-Service
 - Authentifizierung
 - Protokollspezifische Festlegungen für die Sitzung, wie Protokollversion, Anwendbare Dienste etc.
- Search-Service
 - Suchanfragen
 - Namen der Result-Sets
 - Format für Rechercheergebnisse
 - Übertragung der ersten Datensätze
- Present-Service für den Transport der Daten
 - Namen des Result-Set
 - Bevorzugtes Übertragungsformat
 - Übertragung der Datensätze
- Delete-Service
 - Löschen von Result-Sets beim Targetsystem
- Resource-Report-Service
 - Austausch von Abrechnungsinformationen
- Scan-Service
 - Browsinglisten
 - Register

Diese Basisdienste sollten von den Projektpartnern verbindlich realisiert werden.

Erweiterte Dienste:

- Explain-Service für den Abgleich der Systeme
- Segment-Service für große Treffermengen
- Extended-Services z.B. für Update oder Item Order
- Access-Control-Service
- Ressource-Control-Service
- Sort-Service für Sortierung von Treffermengen

Retrieval:

Z39.50 Retrieval

- **Suchmöglichkeiten**
 - **Indexierte Felder**
 - **Registerfunktionen**
 - **Arten der Indexierung**
 - **Trunkierung**
 - **Proximity**
- **Suchsyntax**
 - **RPN Reverse Polish Notation**
 - **CCL Common Command Language**

Bei den Suchmöglichkeiten ist man immer abhängig vom jeweiligen Zielsystem.

- Welche Felder sind indexiert,
- gibt es Registerfunktionen fürs Browsing,
- welche Arten der Indexierung werden angewandt,
- welche Möglichkeiten der Trunkierung gibt es,
- gibt es Proximityfunktionen.

Diese Angaben definieren das Attribut-Set. Hierzu Festlegungen zu treffen und Mindestanforderungen zu definieren war sehr aufwendig. Das Hauptproblem dabei ist, dass nicht jedes System alle Attribute kennt bzw. nicht alle Attribute in gleicher Weise interpretiert und unterstützt werden.

Für die Formulierung der Suchanfrage gibt es zwei gebräuchliche Formate:

- RPN Reverse Polish Notation
- CCL Common Command Language

Diese Unterschiede sind auch der Grund dafür, dass die Suchergebnisse, je nachdem welches Zielsystem abgesucht wird, oft sehr unterschiedlich ausfallen und auch davon abhängig ist welcher Client verwendet wird.

Indexierung:

Z39.50 Indexierung

- **AG Indexierung**
 - **Begriffsdefinitionen**
 - **Indextypen**
 - **Feldzuordnungen**
 - **Mindestanforderungen**
- **Begriffe**
 - **Wort- oder Stringindex**
 - **Register**
 - **Phrasensuche**

Es gab eine Arbeitsgruppe der KfR die sich mit der unterschiedlichen Indexierung in bibliographischen Datenbanken beschäftigt hat. Diese wird im Rahmen des Standardisierungsausschusses wohl weitergeführt.

Die Hauptaufgaben sind:

- Begriffsdefinitionen
- Indextypen
- Feldzuordnungen
- Mindestanforderungen.

Es soll eine Grundlage geschaffen werden, für eine möglichst einheitliche Indexierung in den verschiedenen Bibliothekssystemen. Eine offene Frage wird sein, ob sich die kommerziellen Anbieter an diese Definitionen halten werden. Die Papiere für die String- und Wortindexierung wurden von mir in der jetzigen Version verteilt. Es handelt sich aber noch nicht um eine offiziell verabschiedete Version.

Einige Begriffe die festgelegt wurden:

- Wortindex
- Stringindex
- Register
- Phrasensuche

Die Phrasensuche basiert auf einem Wortindex, einen Phrasenindex gibt es nicht. In den Mindestanforderungen wurden den wichtigsten Datenfelder die verschiedenen Indexierungstypen zugeordnet.

Datenformate:

Z39.50
Datenaustausch

- **Datenformate**
 - **Unstructured Text**
 - **MAB**
 - **MARC**
- **Nutzungsmöglichkeiten**
 - **Anzeige in OPACs**
 - **Datenübernahme**

Bei den Datenformaten sieht die Sache etwas einfacher aus.

Folgende Formate stehen primär zur Auswahl:

- Unstructured Text
- MAB-Formate
- MARC-Formate

Für die Anzeige von Suchergebnissen ist bereits Unstructured Text als Übertragungsformat ausreichend. Die Daten werden vom Target aufbereitet (z.B. als ISBD) und so wie sie geliefert werden vom Client angezeigt. Daneben gibt es die strukturierten Formate.

MAB ist nur im deutschsprachigen Raum gebräuchlich, während MARC international verwendet wird. Die Formate sind an die jeweiligen geltenden Regelwerke angelehnt.

Es gibt aber auch noch mal Unterschiede wie MAB1 und MAB2 bzw. UNI-

MARC und US-MARC usw. Nur bei strukturierten Formaten ist auch die Datenübernahme für eine Weiterverarbeitung, z.B. im Rahmen der Katalogisierung ins eigene System, möglich.

Für die Anzeige von Mehrfachtrefferlisten stellt der Target ein definiertes Kurzformat zur Verfügung.

Realisierung:

Z39.50 Realisierung

- **DBV-OSI-II**
 - Festlegungen der Projektpartner
 - Datenformate
- **Stufen der Realisierung**
 - Client-Funktionalität
 - Target-Funktionalität
 - Update-Funktionalität
- **Bibliothekssysteme**

Im Projekt DBV-OSI II wurden Festlegungen getroffen damit zumindest für die Projektpartner die Kommunikation funktioniert.

Als Datenformat wurde primär MAB2 festgelegt, aber auch Unstructured Text und UNI-MARC sollten möglich sein. Die Daten sollten, soweit möglich, konvertiert werden. Schließlich wurden Unstructured Text und UNI-MARC von allen Projektpartnern unterstützt, MAB jedoch nicht von allen. Es gibt verschiedene Stufen der Realisierung von Z39.50 zu unterscheiden.

Stufe 1 : Clientfunktionalität

Zugriff aus dem eigenen System auf fremde Datenquellen im Rahmen von OPACs oder in Katalogisierungsanwendungen

Stufe 2 : Targetfunktionalität

Öffnung der eigenen Datenbank für den Zugriff von außen

Stufe 3 : Updatefunktionalität

Katalogisierung mit "Fremd"-Clients in eine Datenbank

Diese Stufen der Realisierung sind in den verschiedenen Bibliothekssystemen unterschiedlich weit entwickelt.

Man sollte also immer genau prüfen welche der Z39.50-Funktionalitäten realisiert sind, wenn davon die Rede ist, dass ein System eine Z39.50-Schnittstelle hat.

Zumindest die Z39.50-Clientfunktionalität für den OPAC und die Katalogisierung, für den Zugriff auf fremde Datenquellen, ist für jedes Bibliothekssystem mittlerweile ein Muss.

Aktuelle Anwendungen:



Z39.50
Aktuelle Anwendungen

- **Realisierung in SISIS**
 - **Z39.50-Client**
 - OPAC, Web-OPAC, JOPAC
 - SIKIS, Katalog-Client
 - **Z39.50-Target**
- **Verbundsysteme**
 - OPACs
 - KOBV

Wie sieht es nun bei SISIS damit aus ?

- Es gibt seit längerer Zeit ein Add-On für den OPAC, den SR-OPAC, für die Recherche über Z39.50 in unterschiedlichen Datenbanken und es gibt die Möglichkeit in SIKIS über Z39.50 Daten aus unterschiedlichen Datenquellen für die Katalogisierung zu übernehmen. In den neuen SunRise-Komponenten Web-OPAC, Java-OPAC und Katalog-Client ist von Anfang an eine Z39.50-Schnittstelle integriert worden.
- Es gibt daneben eine Z39.50-Targetanwendung, um den Zugriff über Z39.50 auf die eigene Datenbank zuzulassen. Dieser Z39.50-Target wurde im Rahmen des Projekts DBV-OSI-II entwickelt und steht daher allen SISIS-Anwendern kostenfrei zur Verfügung.
- Die Z39.50-Update-Funktionalität ist, wie auch in allem mir sonst bekannten Systemen, nicht implementiert.

Seit 1996 gibt es in den Verbund-OPACs, Z39.50-Schnittstellen die vom Ursprung her auf das DBV-OSI-II-Projekt zurückgehen. Es gibt auch viele Anwendungsbeispiele aus jüngerer Zeit. Im KOBV z.B. können nur Systeme eingesetzt werden, in denen eine Z39.50-Schnittstelle realisiert ist. Die Portale oder Information Broker wie Elektra oder Gateway Bayern verwenden an vielen Stellen ebenfalls Z39.50 als Schnittstelle.

Durch DBV-OSI-II hat sich Z39.50 in Deutschland als Standard für die Vernetzung von Bibliotheks- und Fachinformationssystemen durchgesetzt. Bei der Entwicklung von Z39.50-Anwendungen in Deutschland wird meist auf die Festlegungen aus dem DBV-OSI-II-Projekt aufgesetzt.

Interoperabilität:

Z39.50 Interoperabilität

- **Voraussetzungen**
 - **Gleicher OSI-Protokoll-Stack**
 - **Kompatible Anwendungsprotokolle**
 - **Gleiche unterstützte Dienste**
 - **Gemeinsames Attribut- und Diagnostic-Set**
 - **Gemeinsame Datenformate**
 - **Gleicher Zeichensatz für die Suchanfragen**
 - **Authentifizierungs- und Abrechnungsstrukturen müssen gegenseitig bekannt sein**

Nicht jeder Z39.50-Client muss sich mit jedem Z39.50-Target verstehen.

Voraussetzung für eine Interoperabilität sind:

- Gleicher OSI-/Protokoll-Stack
- Kompatible Anwendungsprotokolle
- Gleiche unterstützte Dienste
- Gemeinsames Attribute- und Diagnostic-Set
- Gemeinsame Datenformate
- Gleicher Zeichensatz für die Suchanfragen
- Authentifizierung- und Abrechnungsstrukturen müssen gegenseitig bekannt sein.

Heidelberg, 17.01.2001

Robert Scheuerl