

Coli-conc

Mapping Knowledge Organization System

Morsheda Akter

Content

- Objectives
- Software Features
- Software Architecture
- Software Development Process
- Technical Specifications
- Conclusion

Objectives

- Effective Creation and Management of Mappings
- Improvement the Quality of the Mappings
- Facilitate Use and Exchange of KOS and their Mappings

Software Features

Software Features

- Data Conversion
 - Marc21 Xml into JSKOS format
 - Filter and cleanse the data
 - Upload the data into the database server
- JSKOS API:
 - Restful API which ensures reusability
 - Provide third party access
 - Display KOS information in JSKOS format

Software Features

- Coli-Conc Web (CCWeb)
 - Display KOS information (notation/caption)
 - Display notation hierarchy
 - Display scope notes and linked relative index terms

Software Feature

- Coli-Conc Web (CCWeb) cont. ...
 - Enhanced by displaying “Top Concepts”.
 - Displays:
 - DDC-GND mapping and relevance
 - Co-occurrences from the title data
 - Mappings from Coli-conc Concordance database
- KOS-Registry Interface
 - Retrieval and display of metadata
 - Equipped with download function

Software Architecture

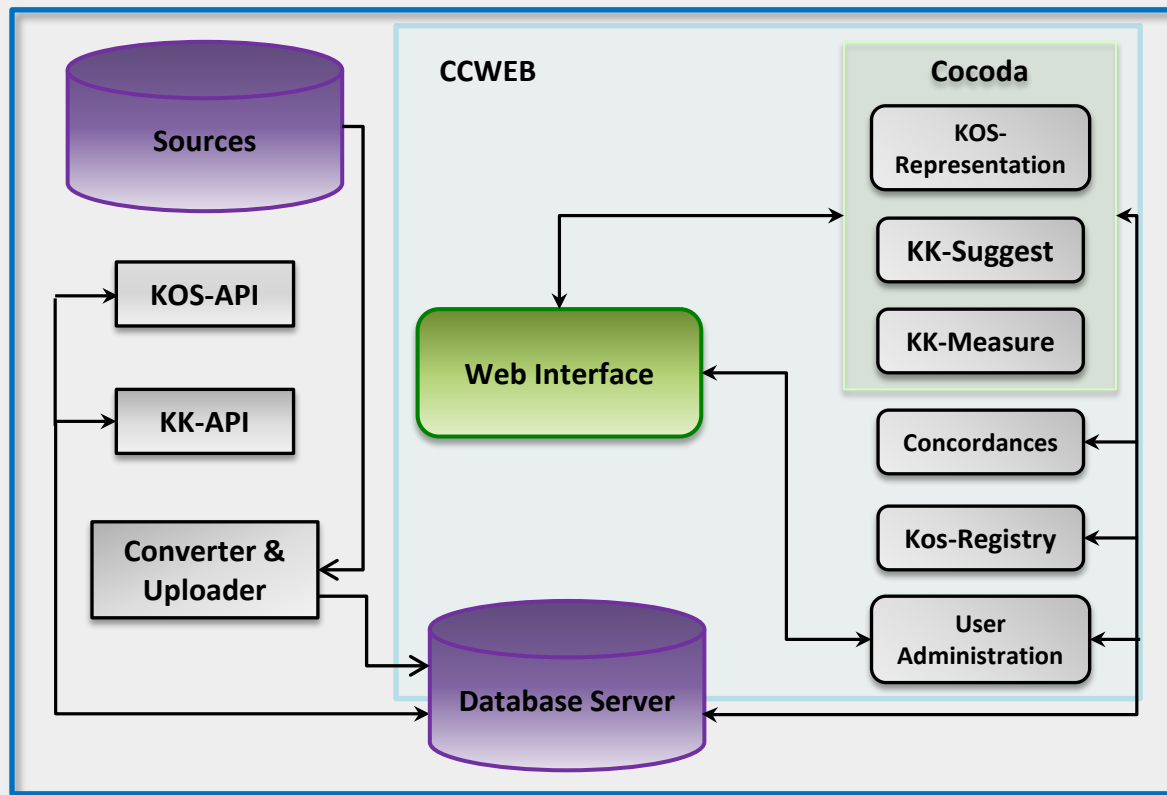


Fig. 1: Software Architecture of the CCWeb

Software Development Process

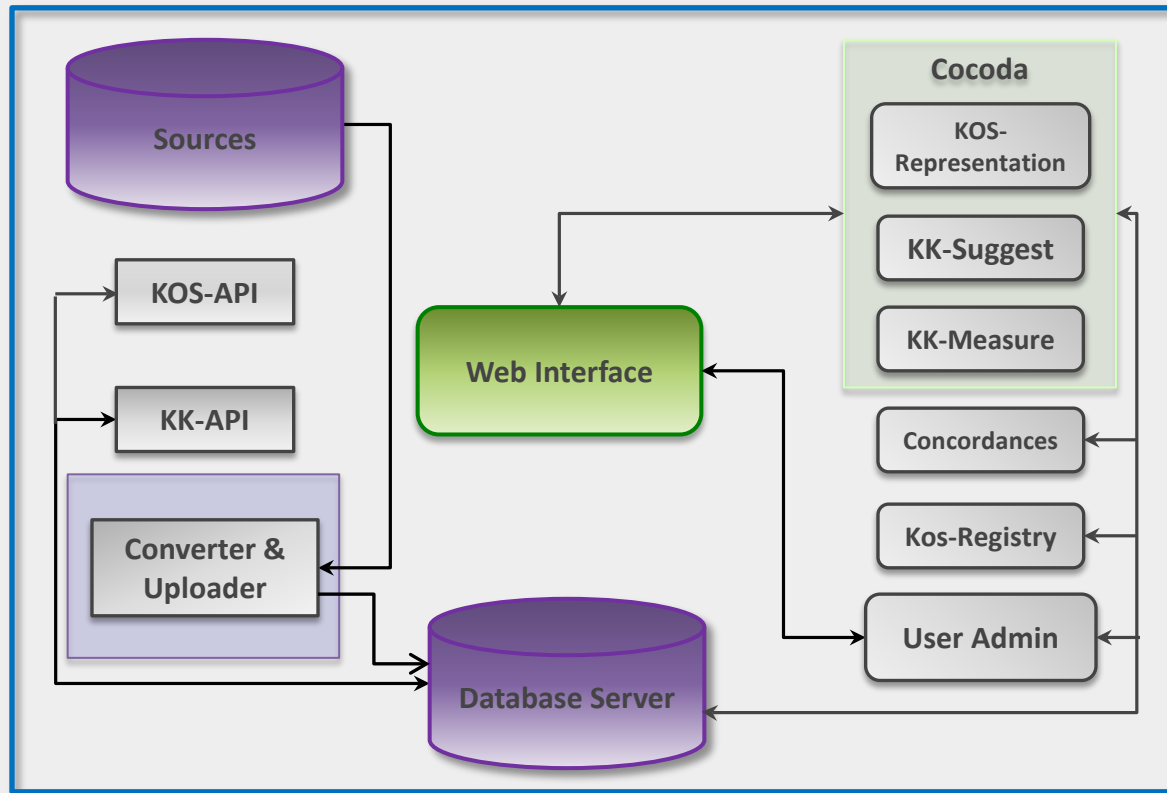


Fig. 1: Software Architecture of the CCWeb

Data Converter

- DDC Marc21 Xml:

```
<mx:datafield tag="084" ind2=" " ind1="0">↓
  <mx:subfield code="a">ddc</mx:subfield>↓
  <mx:subfield code="c">23de</mx:subfield>↓
  <mx:subfield code="e">ger</mx:subfield>↓
</mx:datafield>↓
<mx:datafield tag="153" ind2=" " ind1=" ">↓
  <mx:subfield code="a">000</mx:subfield>↓
  <mx:subfield code="e">00</mx:subfield>↓
  <mx:subfield code="j">Informatik, Informationswissenschaft, allgemeine Werke</mx:subfield>↓
  <mx:subfield code="9">ess=en</mx:subfield>↓
  <mx:subfield code="9">ess=eh</mx:subfield>↓
</mx:datafield>↓
```

DataField notation

prefLabel

Fig. 2: Marc21 Xml data

- Converts the raw Marc21 xml KOS data
- Extracts the fine grained classification data

Data Converter

- Converts it into project Format : JSKOS
- Stores data in the database server

Data Converter

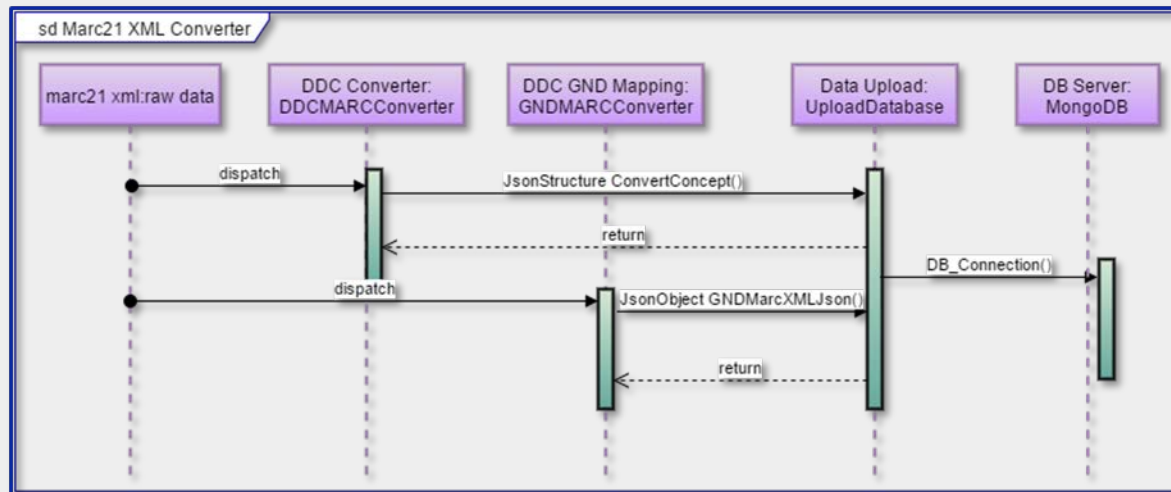


Fig. 3: DDC Marc21 Xml Data Converter Sequence Diagram

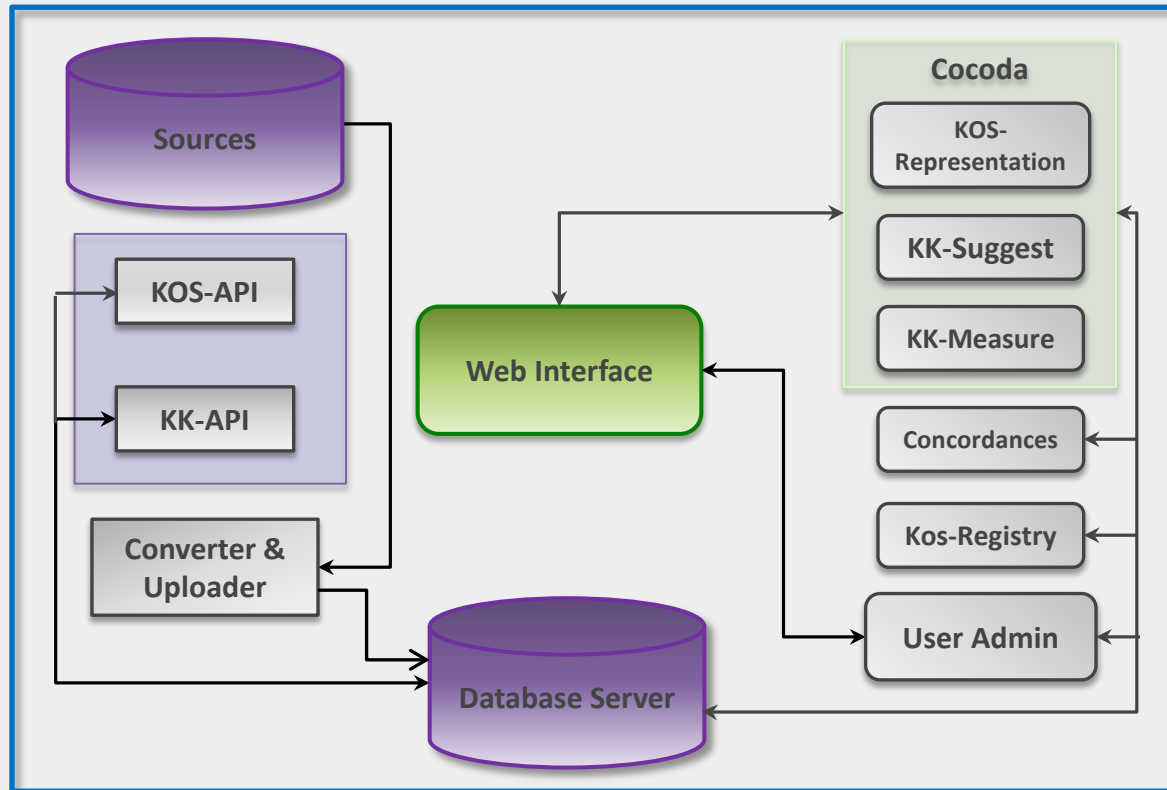


Fig. 1: Software Architecture of the CCWeb

JSKOS API (KOS API)

- http://191.***.***.***:8080/jskos-webservice/rest/DDC/?notation=001
- http://191.***.***.***:8080/jskos-webservice/rest/DDC/uri/?uri=http://dewey.info/class/174/e23



The screenshot shows a web browser window with the address bar containing the URL `8080/jskos-webservice/rest/DDC/?notation=001`. The browser's developer tools are open, displaying a JSON response. The JSON data is as follows:

```
[{
  "uri" : "http://dewey.info/class/001/e23",
  "prefLabel" : { "de" : "Wissen"},
  "notation" : ["001"],
  "type" : ["http://www.w3.org/2004/02/skos/core#Concept"],
  "broader" : [{ "notation" : ["000"], "uri" : "http://dewey.info/class/00/e23"}],
  "inScheme" : [{ "uri" : "http://dewey.info/scheme/edition/e23"}, { "uri" : "http://dewey.info/scheme/ddc"}]
}]
```

Fig. 4: JSKOS data view for DDC notation-“001”

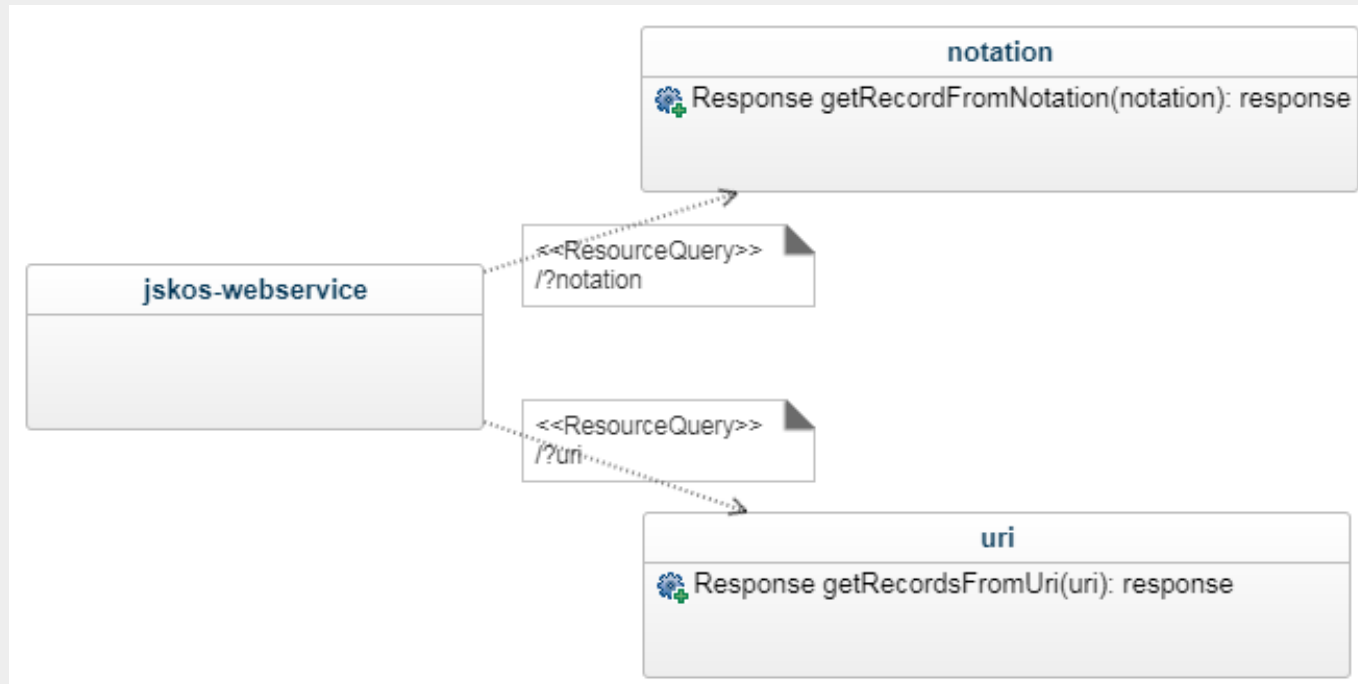


Fig. 5: JSKOS Web Service Model (RESTful)

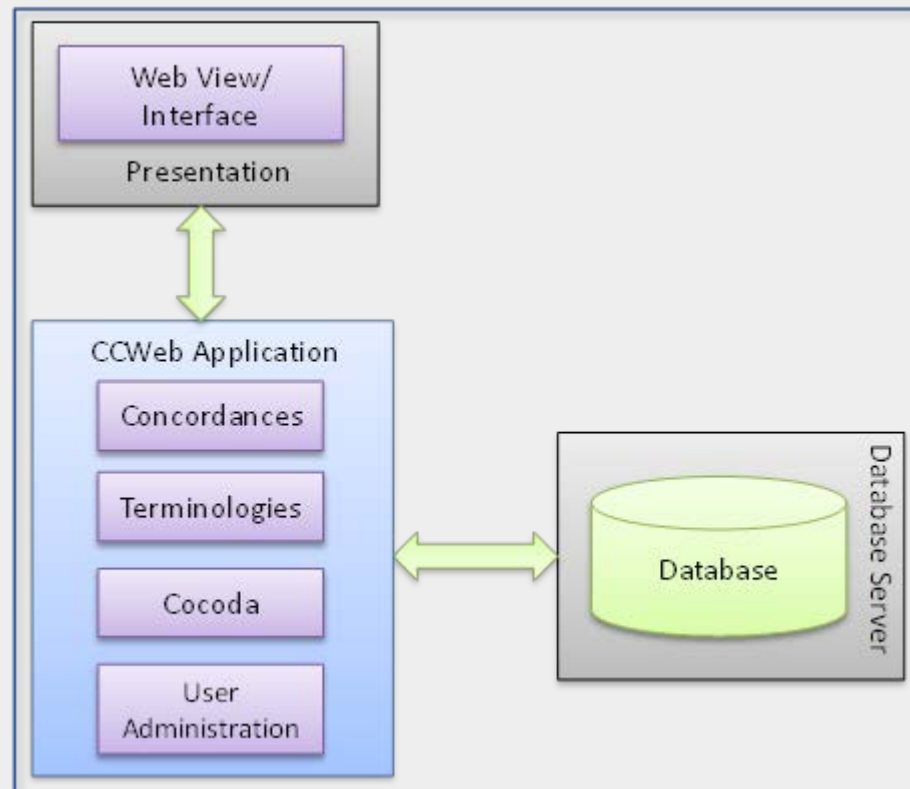


Fig. 6: CCWeb Application Block Diagram

KOS Registry

- Selection of a KOS
- Expansion feature for the display of metadata
- Download option of the metadata in different formats:
 - JSKOS,
 - XML,
 - CSV etc.

Select a Terminology

Search:

Expand for Metadata

- MSC (Mathematics Subject Classification)
- ▼ PACS (Physics and Astronomy Classification Scheme)

Title: PACS (Physics and Astronomy Classification Scheme)
Alternative Title: PACS
Abstract: The Physics and Astronomy Classification Scheme® (PACS) was developed by the American Institute of Physics (AIP) and has been used in Physical Review since 1975 to identify fields and sub-fields of physics. It is used in a variety of ways, for example, in the online journals as a tool in searching for articles by subject. PACS is arranged hierarchically, by subdivision of the whole spectrum of subject matter in physics- and astronomy-related sciences into segments and then repeating the process of subdivision down to four levels. The latest edition of PACS is the 2010 edition.
Author: American Institute of Physics (AIP)
Type: classification scheme
Additional information: http://en.wikipedia.org/wiki/Physics_and_Astronomy_Classification_Scheme
Link: http://publish.aps.org/PACS
Topic: physical sciences, astronomy
DDC: 520, 530

- FKDigBib (Fachklassifikation digitale Bibliothek)

Fig. 7: KOS Registry web view

- Selection of source scheme/target scheme for the mapping process
- Search field for notation/term entry for a selected KOS
- Display of a hierarchical view of the given notation/caption
- Possibility to extend with details
- Facilitates display of top concepts of the selected KOS
- The “Concordance Database” retrieves existing mappings.

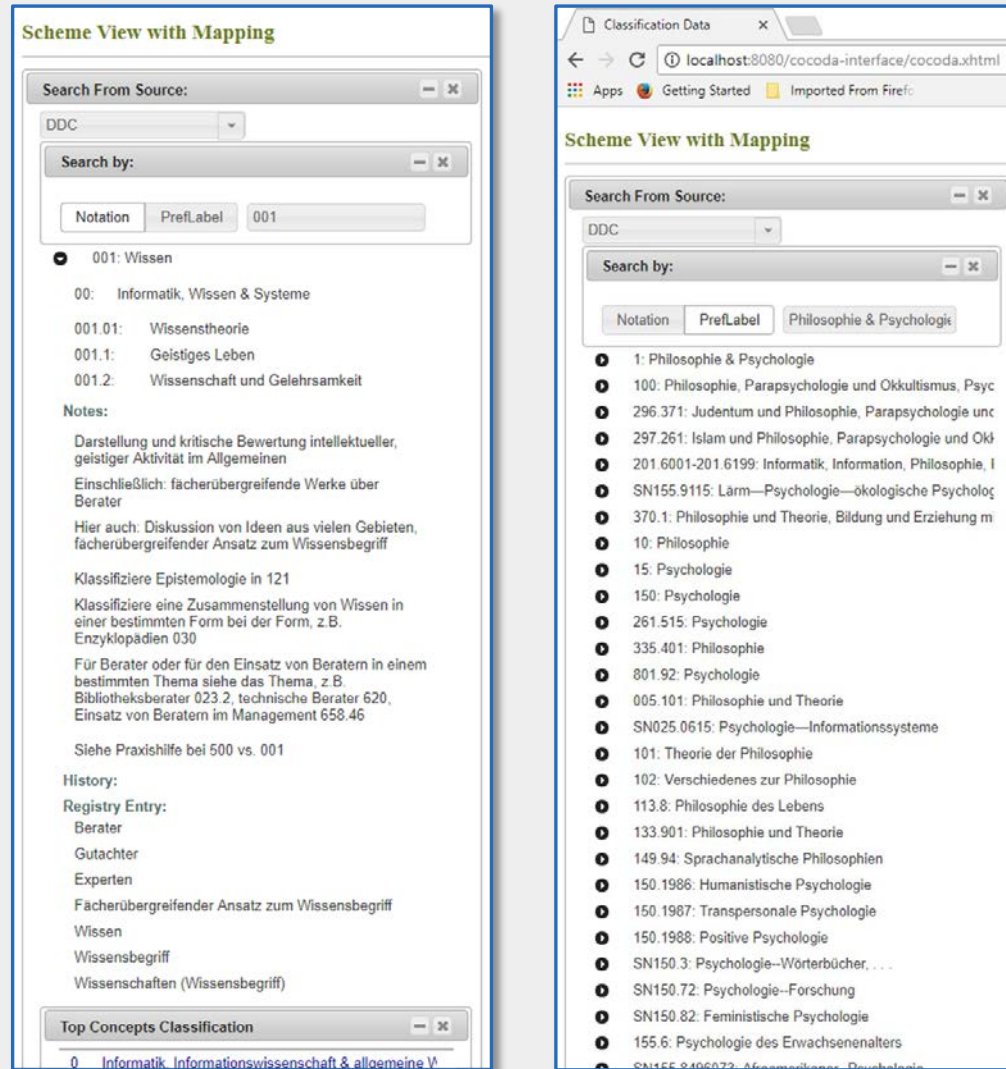


Fig. 8: Search view for notation/term

The screenshot displays the 'Scheme View with Mapping' interface of the CCWeb application. It is divided into several panels:

- Search From Source:** Shows a search for '100' in the DDC scheme. The result is '100. Philosophie, Parapsychologie und Okkultismus, Psyc'.
- Top Concepts Classification:** A hierarchical list of concepts including 'Informatik, Informationswissenschaft & allgemeine V', 'Philosophie & Psychologie', 'Religion', 'Sozialwissenschaften', 'Sprache', 'Naturwissenschaften', 'Technik, Medizin, angewandte Wissenschaften', 'Künste und Unterhaltung', 'Literatur', and 'Geschichte und Geografie'.
- DDC GND Mapping:** A table showing mappings between DDC Notation, GND PrefLabel, GND Broader, and Relevance.

DDC Notation	GND PrefLabel	GND Broader	Relevance
4045791-6	Philosophie	4045790-4	closeMatch
4185103-1	Theoretische Philosophie	4045791-6	closeMatch
4494545-0	Interkulturelle Philosophie	4045791-6	0.5
- Mapping:** A table showing mappings between Source Scheme, Source Concept, Target Scheme, Target Concept, and Creator.

Source Scheme	Source Concept	Target Scheme	Target Concept	Creator
DDC	100	RVK	CA-CK	VZG
DDC	100	RVK	CL-CZ	VZG
DDC	100	RVK	CP 9400	VZG
DDC	100	RVK	BE 9130	VZG
DDC	100	RVK	BF	VZG
- Library Catalogue:** A table showing mappings between DDC Notation, RVK Notation, and BK Notation.

DDC Notation	RVK Notation	BK Notation
100	CD 3067	08.21
296 3092B100290	DD 5541	15 9611 2008.25
100	PC 4630	
- Search From Target:** Shows a search for 'CL 1000' in the RVK scheme. The result is 'CL 1000. Zeitschriften'.
- Top Concepts Classification:** A hierarchical list of concepts including 'Allgemeines', 'Theologie und Religionswissenschaften', 'Philosophie', 'Psychologie', 'Pädagogik', 'Allgemeine und vergleichende Sprach- und Klassische Philologie, Byzantinistik, Mittella', 'Germanistik, Niederlandistik, Skandinavistik', 'Anglistik, Amerikanistik', 'Romanistik', 'Statistik', 'Ethnologie', 'Klassische Archäologie', 'Kunstgeschichte', 'Musikwissenschaft', 'Politologie', 'Soziologie', 'Militärwissenschaft', 'Geschichte', 'Rechtswissenschaft', 'Wirtschaftswissenschaften', 'Geographie', 'Mathematik', 'Informatik', 'Allgemeine Naturwissenschaft', and 'Geologie und Paläontologie'.

Fig. 9: CCWeb Application view

Technical Specifications

Technologies:

Java

**Jersey
bundle**

**Web Servers:
Apache Tomcat
v 7.0 or higher**

**Framework:
JavaServer Faces
(JSF)
v 2.2**

JavaScript

**Application
Servers: J2EE
(Java 2 Platform,
Enterprise
Edition)**

IDE: eclipse

**Database
Servers:
Mongodb v
3.2.1**

**User Interface:
PrimeFaces v 5.0**

**Server
Request
Forms:
XHTML**

<https://si-it-workshop.gbv.de/>

Thank You

