



EUROPEAN TOPIC CENTRE ON
CATALOGUE OF DATA SOURCES (ETC/CDS)
EUROPEAN ENVIRONMENT AGENCY
NIEDERSÄCHSISCHES UMWELTMINISTERIUM



WebCDS 2.5

Retrieving European environmental meta-data through the Web

Version 1.0.8

Wassili Kazakos
E-mail: kazakos@fzi.de
Tel: +49 721 9654 712

Forschungszentrum Informatik (FZI)
Haid-und-Neu-Straße 10-14
D-76131 Karlsruhe
Germany



Table of Contents

1	INTRODUCTION.....	6
1.1	Overview.....	6
1.2	Scope and intended audience.....	6
1.3	Motivation.....	6
1.4	Outline	7
2	CORNER STONES	8
2.1	Catalogues	8
2.2	Thesauri	8
2.3	CDS	9
3	REQUIREMENTS AND APPROACH.....	11
3.1	Functionality	11
3.2	Multilingual Aspects.....	11
3.2.1	Multilingual User Interface.....	11
3.2.2	Multilingual Retrieval.....	12
3.3	Portability.....	12
3.4	Extensibility.....	13
3.5	Performance.....	13
4	OVERALL ARCHITECTURE	15
5	WEBCDS CLIENTS.....	17
5.1	HTML Client	17
5.2	Java Client (J-CDS).....	19
6	TECHNICAL SPECIFICATION.....	22
6.1	Server Software.....	22
6.2	Server Hardware.....	23
6.3	Client Software & Hardware	23
6.4	Communication.....	23

7	INSTALLING WEBCDS 2.5.....	25
7.1	Preconditions	25
7.2	WebCDS distribution archive	25
7.3	Import of the CDS schema and the data.....	26
7.4	Configuration of the DB Server.....	26
7.4.1	Oracle.....	26
7.4.2	Other DB management systems.....	27
7.5	Configuration of the Servlet Client and HTTP server.....	27
7.5.1	Servletrunner	27
7.5.2	Jserv	29
7.6	Access WebCDS HTML	30
7.7	Customizing WebCDS 2.5	31
7.7.1	Adding new languages.....	31
7.7.2	Replacing pictures and static HTML pages	31
7.7.3	replacing introduction, explanation and help text	32
7.7.4	Changing colours	32
7.7.5	Changing feedback E-mail address.....	33
7.8	Search fields	33
7.8.1	Quick Search for Addresses.....	33
7.8.2	Quick Search for Data Sources.....	33
7.8.3	Expert Search for Addresses.....	33
7.8.4	Expert Search for Data Sources.....	34
7.9	Troubleshooting WebCDS 2.5.....	34
7.9.1	Instability after some days of use.....	34
8	INSTALLING J-CDS.....	35
8.1	Configuration of J-CDS Client and Server	35
8.2	Installing the Java Plug-In on the client.....	37
8.3	Troubleshooting J-CDS.....	38
8.3.1	SecurityException.....	38
8.3.2	Detailed View does not work.....	38
9	CURRENT INSTALLATIONS OF WEBCDS 2.5	39
10	WEBCDS DATA MODEL.....	40
10.1	Description.....	40
10.2	Diagram	62
	REFERENCES	64

1 Introduction

1.1 Overview

The Catalogue of Data Sources (CDS) of the European Environment Agency is an environmental catalogue system to direct the user towards information about the state of the environment in Europe. The CDS comprises metadata (data about data sources of European interest), addresses and a multilingual thesaurus.

To provide global access to the CDS data via the Web, WebCDS has been developed. Major requirements were multilingual search and result forms, multilingual search by use of the multilingual thesaurus, independence of platforms, database management system (DBMS), and Web servers, user friendliness, and easy extensibility.

The Web application incorporates Java (on the server site and optionally on the client side), JDBC, RMI, servlets, and HTML. Servlets were utilized in order to overcome the problems of the classical HTML/CGI approach, e.g., servlets allow the communication of a client with an initialized service. We combined Java and HTML in WebCDS in order to gain nearly all advantages of Java without requiring Java capable Web browsers. In addition to the HTML based retrieval module (WebCDS) a Java based retrieval module (J-CDS) was designed and implemented. For further information about CDS and WebCDS please refer to the URL <http://www.mu.niedersachsen.de/cds> or [Kazakos et al., 1999] or contact the CDS Project Leader at ETC/CDS Stefan Jensen (stefan.jensen@mu.niedersachsen.de).

1.2 Scope and intended audience

In this document the overall WebCDS approach and the two WebCDS retrieval tools are presented in order to give an general overview about the developments for the interested audience. The second part describes an WebCDS installation and is intended mainly for WebCDS administrators.

1.3 Motivation

In 1993 the European Environment Agency (EEA) was launched by the European Union (EU) with a mandate to orchestrate, cross-check and put to strategic use information of relevance to the protection and improvement of Europe's environment and to ensure that the

public is properly informed about the state of the environment (<http://www.eea.eu.int/>). For the development of a versatile tool that directs the user towards information about the state of the environment in Europe the European Topic Centre on Catalogue of Data Sources and Thesauri (ETC/CDS) was founded by the EEA. To fulfil its task the ETC/CDS developed an environmental catalogue system, the Catalogue of Data Sources (CDS) and - to cope with the variety of European languages - the General European Multilingual Environmental Thesaurus (GEMET).

Public access to environmental information is an important right and an essential support to environmental policy making. Therefore, the primary goal of the CDS is to direct the users towards environmental information relevant on a European scale.

The importance attached to this tool will of course increase simultaneously with an increase of the volume of the available meta-data. The requirements identified for WebCDS include multilingual search and result forms, multilingual search by use of the multilingual thesaurus, portability by independence of platforms (operating system, database management system (DBMS), and Web server), user friendliness, extensibility, and performance. The WebCDS approach considering these requirements as well as the installation instruction of the tool are presented in this document.

1.4 Outline

This document is organized as follows. In Section 2, the corner stones are introduced. Section 3 focuses on the requirements for WebCDS. In Section 4, the general architecture of WebCDS are presented. Section 6 describes the technical specification of the software. Section 7 gives the installation instructions of WebCDS HTML version and finally Section 8 describes the installation of the J-CDS.

2 Corner Stones

In this section, we provide the background necessary to understand catalogues and catalogue systems and thesauri as an important means to search for information in catalogues. Furthermore, we introduce the Catalogue of Data Sources (CDS) as the starting point for WebCDS.

2.1 Catalogues

In order to react upon complex environmental problems, it is necessary to have information of various application fields at hand which in most cases will be located at different sites. Only if the respective information is sufficiently available it is possible to find complex interrelations, to effectively survey environmental laws, but also to mutually use available information. Both the recollection of data that already has been collected and keeping data without any further use because of missing knowledge about the data can be avoided.

The answer to a certain question requires to learn *which* information is available, *where* the information is managed, *how* this information can be obtained, and *how* to interpret the information correctly. The information that is necessary to obtain above mentioned information is called *meta-information*. This information can be compared to the information of classical index cards in library catalogues, which describe books but are not the books themselves [Crossley, 1996]. Due to the present explosion of the volume of data, it becomes even more important for the user to rely on information about existing data in order to find what he or she needs. Equivalent to the library catalogues which contain index cards, meta-information systems or *catalogue systems* are the more general electronic form.

2.2 Thesauri

Thesauri are a proven means to provide a uniform and consistent vocabulary for indexing and retrieval of information bearing objects (IBOs [Smith, 1997]), such as catalogue entries and addresses) and to supply users with a certain vocabulary for the retrieval of

IBOs in such systems. A *thesaurus* is a set of terms selected from a natural language representing the vocabulary of a certain field; it is used for the indexing, saving, and retrieval of IBOs [Wersig, 1985].

Thesaurus terms can be classified by *descriptors* and *non-descriptors*. Descriptors are terms used directly for the indexing and retrieval of objects. Non-descriptors are not used for indexing; they are supplementary entry points for the user, providing a wider range of terms for the retrieval of objects. Information about descriptors, e.g., an explanation of the meaning of a term, can be stored in scope notes. The terms in a thesaurus are related to each other through diverse relationships such as hierarchical relationship, equivalence relationship, and associative relationship.

A *multilingual thesaurus* is a regular thesaurus, with an equivalent term for every descriptor (possibly for non-descriptors as well) in each of the languages covered. The advantage of multilingual thesauri is that users do not have to be familiar with a particular language. They can use several languages for the retrieval from an information system, even though terms of just one language were used for indexing. Therefore multilingual thesauri are the basis for multilingual access to IBOs.

2.3 CDS

In Europe environmental metadata are stored and managed at different levels, e.g., regional metadata are managed by the environmental ministries of the states, national metadata are managed by federal environment agencies, and metadata of European-wide interest are managed by the EEA. For the latter case, the Catalogue of Data Sources (CDS) was developed.

The CDS is an environmental catalogue system to direct the user towards information about the state of the environment in Europe. It comprises metadata (data about data sources of European interest), addresses, and a multilingual thesaurus, the General European Multilingual Thesaurus GEMET [CNR, Rome, Umweltbundesamt, Berlin, 1997]. The CDS is designed for the collection, management, and dissemination of detailed descriptions of information resources. Currently for the collection and maintenance of metadata mainly a

Microsoft Windows based application (WinCDS) is used. WebCDS as presented in this document is the retrieval tool which allows global access via the Web.

CDS records are classified into several classes (documents, datasets, ...), indexed by subject by using GEMET. The records can be flexibly linked (parent/children relationship) in order to interconnect resources.

The data model for the metadata itself is split into three levels:

- The *core level* contains core attributes for European and EEA requirements. Attributes included in the core level are recommended to be used in all national CDS for compatibility reasons. These attributes are the basis for interoperability between environmental catalogues.
- *Level 1* attributes contain core level attributes and additional attributes needed by the EEA.
- *Level 2* attributes contain core level attributes and additional attributes which are defined by national authorities for national CDSs.

The CDS is a focal point for interoperability among environmental data catalogues [Kramer et al, 1997b], because data from national catalogues is used to fill parts of the CDS. For compatibility with other international metadata standards and with other environmental meta-information systems, the CDS data model draws upon existing international standards like the Global Information Locator Service (GILS).

3 Requirements and Approach

The requirements for WebCDS can be grouped into five categories: functionality, multilinguality, portability, extensibility, and of course performance. We describe these requirements in turn and indicate how they are addressed.

3.1 Functionality

From the user's point of view, the basic retrieval functionality is the most important requirement WebCDS has to fulfil. For this purpose, and with respect to the different data stored in the catalogue, the catalogue system supports two different search types: a data source search and an address search. For both of them we introduced two search forms: a simple search form and an expert search form. This gives an experienced user the opportunity to refine the search and to add more constraints to the request. Independent of the selected search form, the user should have the possibility to use terms from a thesaurus which the data sources and addresses are indexed with for the retrieval.

3.2 Multilingual Aspects

The CDS is an European catalogue that is used all over Europe. Hence, support for multilinguality was one of the major requirements. In contrast to other settings such as local area networks, the Internet is a challenge for multilingual applications because of its global character. Usually, English is used to support access to local data for foreign people. However, this approach is insufficient, if -as in the case of WebCDS- the user should have the opportunity to search data as well in his native language, and especially if the data itself is multilingual as well.

3.2.1 Multilingual User Interface

The problem of a multilingual user interface covers two major aspects: First, the search forms and the result pages should be displayed in the language the user has chosen. This implies that all visible fields should be language sensitive. For this purpose we use the

resource bundle functionality as suggested by the internationalisation specification for Java [JavaSoft, 1997a]. The abstract class `ResourceBundle` gives us the possibility to build locale sensitive resources. The resource bundle approach gives us the possibility to switch between the different languages, with the special advantage of an easy extensibility of the catalogue system.

The second aspect, which cannot be solved by this approach, is the problem to represent multilingual data. In an ideal case, all data retrieved from the catalogue should be represented also in the language of the user interface. Because this is not possible without translating all data in all languages supported, we decided to support multilingual representations as far as possible. Attributes with restricted sets of values can be translated in all supported languages. By doing so, most information retrieved is represented in the native language of the user.

3.2.2 Multilingual Retrieval

The CDS supports the indexing of data sources and addresses with terms from one or more multilingual thesauri. Usually, a new data source will be indexed with terms from one or maximal two languages. This results in two more aspects which should be considered by the catalogue system:

- Users searching for proper thesaurus descriptors and data sources indexed with these descriptors, do not need to know in which languages data was indexed. This should be transparent to the user and he or she should find all indexed data sources, independent of the selected language.
- the thesaurus descriptors associated with a data source or address should be displayed in the native language of the user and not in the language that was initially used for indexing.

3.3 Portability

Due to the fact that the WebCDS server will be installed on different platforms all over Europe one of the basic requirements was to achieve portability. The independence from

the operating system was achieved using Java to implement WebCDS on server side as well as optionally on client side.

3.4 Extensibility

The catalogue system is a fairly dynamic system. A system like the WebCDS should quite easily be adaptable to user demands. Moreover we had also some additional requirements which should be considered during the first design and implementation phase: altering the schema of the database and even replacing the DBMS itself should be possible; furthermore future releases should be able to support several thesauri simultaneously.

3.5 Performance

The choice of Java for the server site application has many advantages with respect to portability, multilinguality, and distributed computing. As Java is interpreted Java applications are slower than compiled applications. To solve potential performance problems, just-in-time compilers have been introduced, which compile the byte code on the fly, so that the execution is speed up. Java compilers that generate native code have been suggested as well.

In case of accessing a database from the Word Wide Web, two other bottlenecks are even more run-time consuming. Both cannot be influenced by the choice of the programming language - the first is the Web access itself, and the other the database access. The first factor can only be influenced by the amount of data transmitted through the net, which is more or less application- or even query-dependent, and therefore fixed.

The most powerfull performance optimization can be reached by optimizing the database queries and the overall database access mechanisms. Whenever possible, a user query should result in exactly one database query. The reuse of results should be encouraged. The result set should be cached, so that the user can navigate through them without the need of new database queries, and it should be possible for other users to re-use the results. The optimization of database queries can result in an important performance improvement. We cannot rely on the optimization capabilities of the underlying database management system

(DBMS) only, because they change from DBMS to DBMS and there is no single best optimization of a specific application (or query).

4 Overall Architecture

WebCDS retrieval is designed as a 3-tier architecture as shown in Figure 1. The application server builds the core of the WebCDS architecture. It receives the requests from the clients and forwards the queries to the database. Its task is to transform the request objects of the client to proper SQL statements and to send the statements to the database. After receiving the results, the application server encapsulates them into result objects and sends them back to the clients.

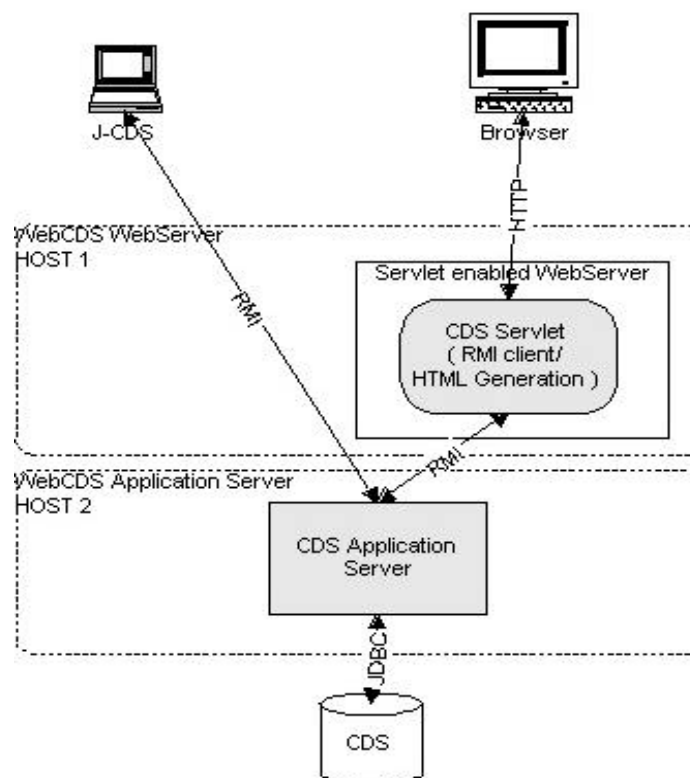


Figure 1 WebCDS Overall Architecture

WebCDS application server was implemented as an RMI server. We chose RMI instead of middleware solutions, such as CORBA, for several reasons. The clients and the server are written entirely in Java, we did not have the need for an intermediate object model and the RMI technology is distributed with the JDK and therefore free of charge. In addition, we avoid the installation and administration of an object request broker (ORB) at every

location at which WebCDS 2.5 is installed. Last but not least, we did not encounter any performance drawbacks.

WebCDS currently support to types of client; an HTML based and a Java based client. For the presentation of the results in the HTML client - any web browser with HTML 3.2 support- the results are converted into HTML in the HTML generation engine of the CDS Servlet. On the other hand experienced users can use the J-CDS client which accesses the application directly and displays the results within an integrated desktop, allowing more interoperability. Although the general architecture supports a distribution of the separate components for load balancing, all components can be installed on one computer to reduce communication overheads. Current installations of WebCDS are accessing an Oracle Databases at the ETC/CDS in Hannover, the Canton of Genva and the BUWAL in Bern. Another installation at the EEA in Copenhagen with access to an MS SQL Server 7.0 is also now available.

5 WebCDS Clients

5.1 HTML Client

WebCDS 2.5 HTML was the first web client developed for the CDS. The look and feel of the tool is comparable to standard search engines with several simple and more complex search forms and the search results presented in a list (Figure 2). With this tool users have fast access to CDS and they can search for data sources and addresses with a standard browser supporting HTML 3.2 and above.

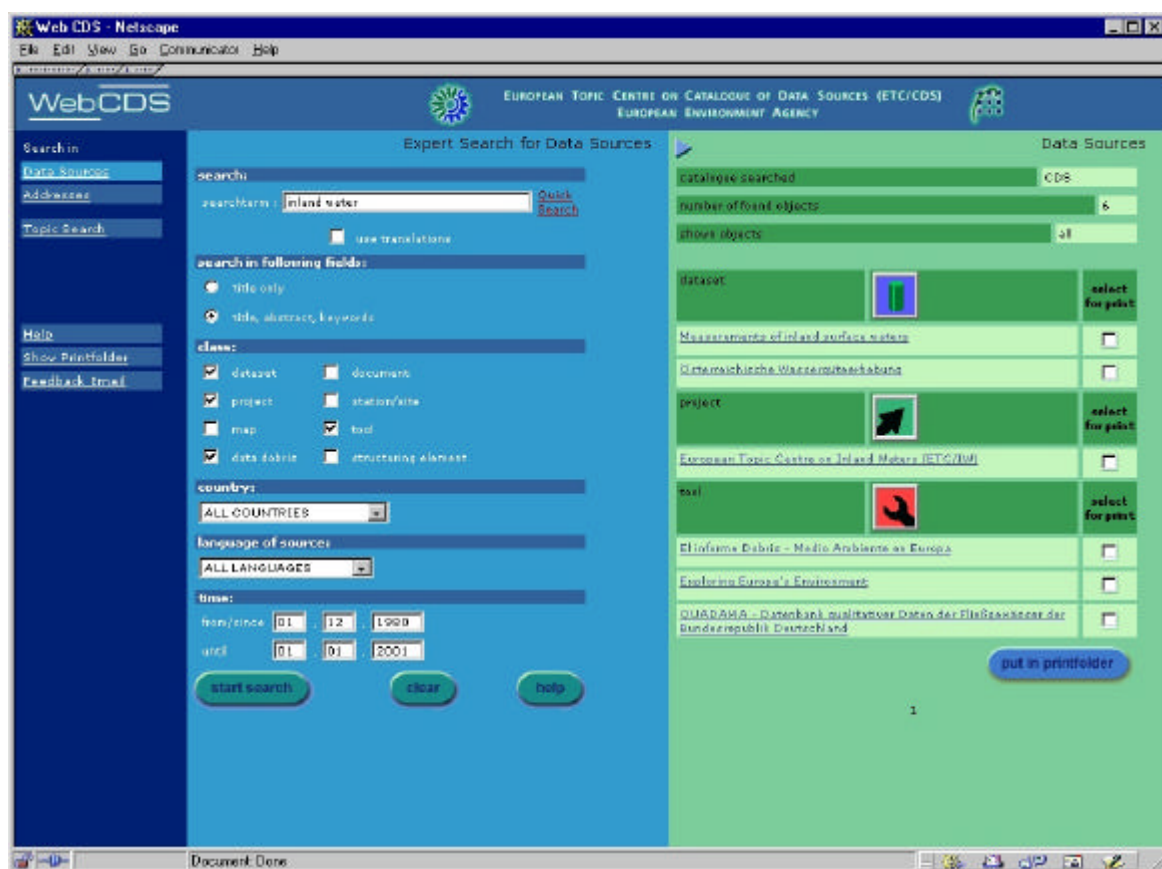


Figure 2 Expert Search for Data Sources of WebCDS HTML

The left frame, or menu frame, displays all search possibilities, like address search, data source search and topic search. In future additional searches can be added. In the address search mode as well as in the data sources search mode the user can switch between simple search (the default) which supports a simple keyword search only, and expert search with a more sophisticated search interface.

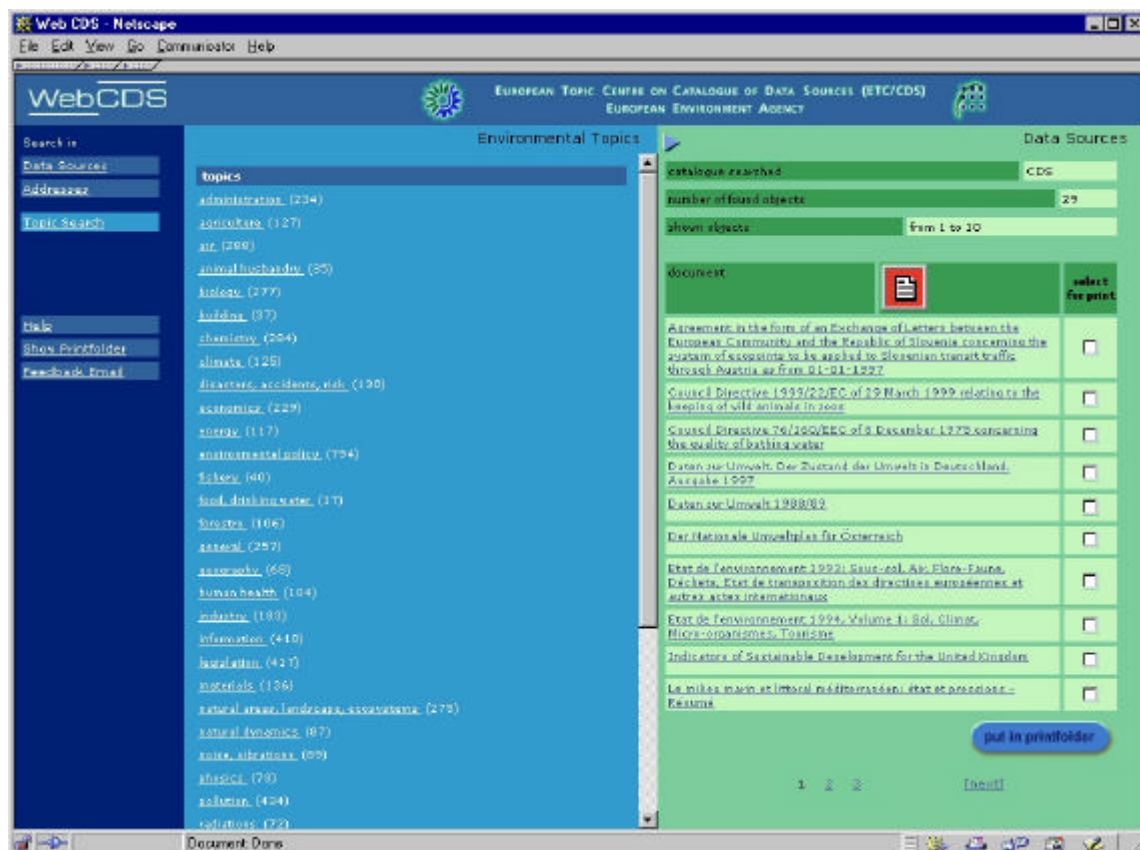


Figure 3 WebCDS HTML Topic Search

In both cases a list of results is displayed in the right frame. By selecting one of the results a detailed description of the address or data source is displayed in the centre frame. Besides the metadata describing the source, retrieval information is displayed, like URLs or person and organisations from where to purchase the real data.

Additionally to the classical access through search terms, the 'Topic Search' (Figure 3) allows access to data sources through the 40 environmental topics of the GEMET (General Multilingual Environmental Thesaurus).

5.2 Java Client (J-CDS)

J-CDS focuses on more interactivity. It integrates several tools within one desktop and allows the user a comfortable access to CDS data and addresses:

- **GeoView (Location)** is a easy to use tool for spatial access. It allows zooming, moving and selection of countries and bounding boxes (Figure 4).

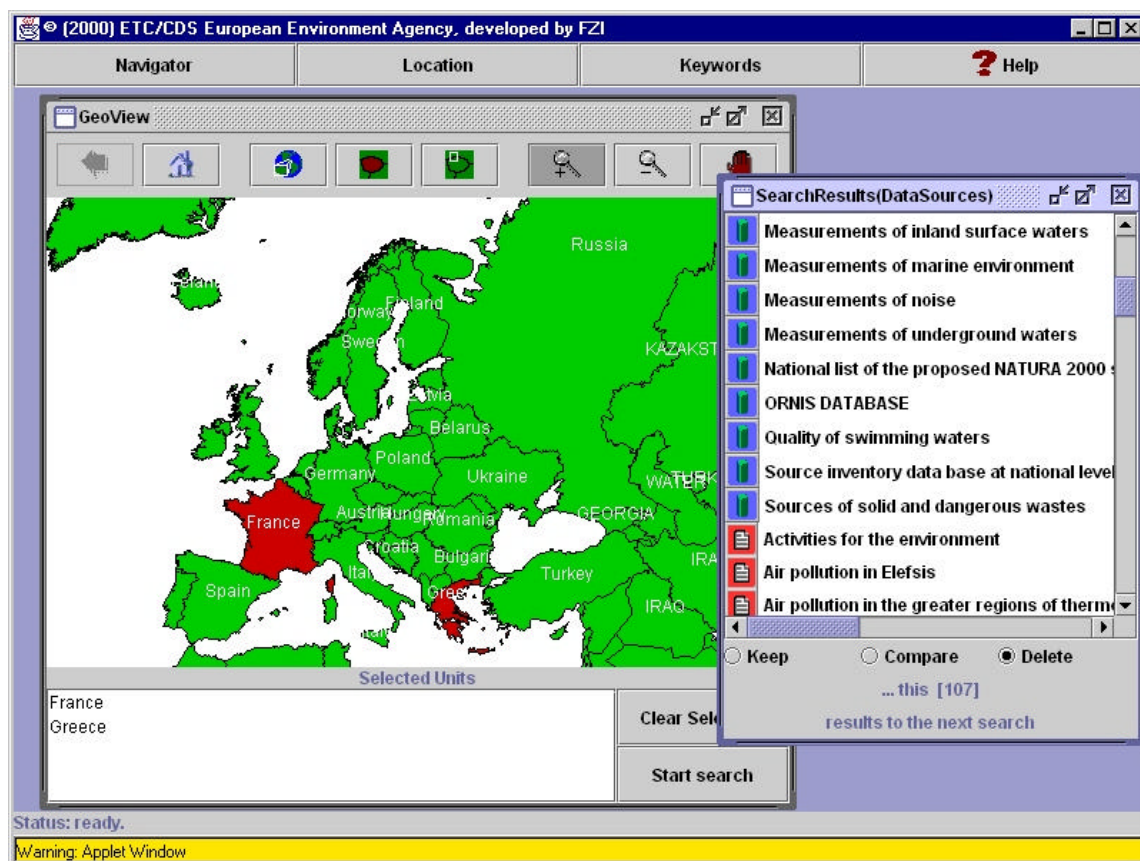


Figure 4 J-CDS with GeoView (Location) as the selected application

- **GenThes (Keywords)** allows navigation within the terms and hierarchies of the GEMET thesaurus (Figure 5). Besides a search for thesaurus descriptors and synonyms, GenThes supports browsing of the thesaurus hierarchy, detailed description of descriptors, and translation of a selected term. User can select one or more terms and access the CDS data sources. To search for addresses within GenThes the EEA Roles and topics hierarchy can be accessed and displayed. The GenThes architecture is described in more detail in [Nikolai et al., 1999].

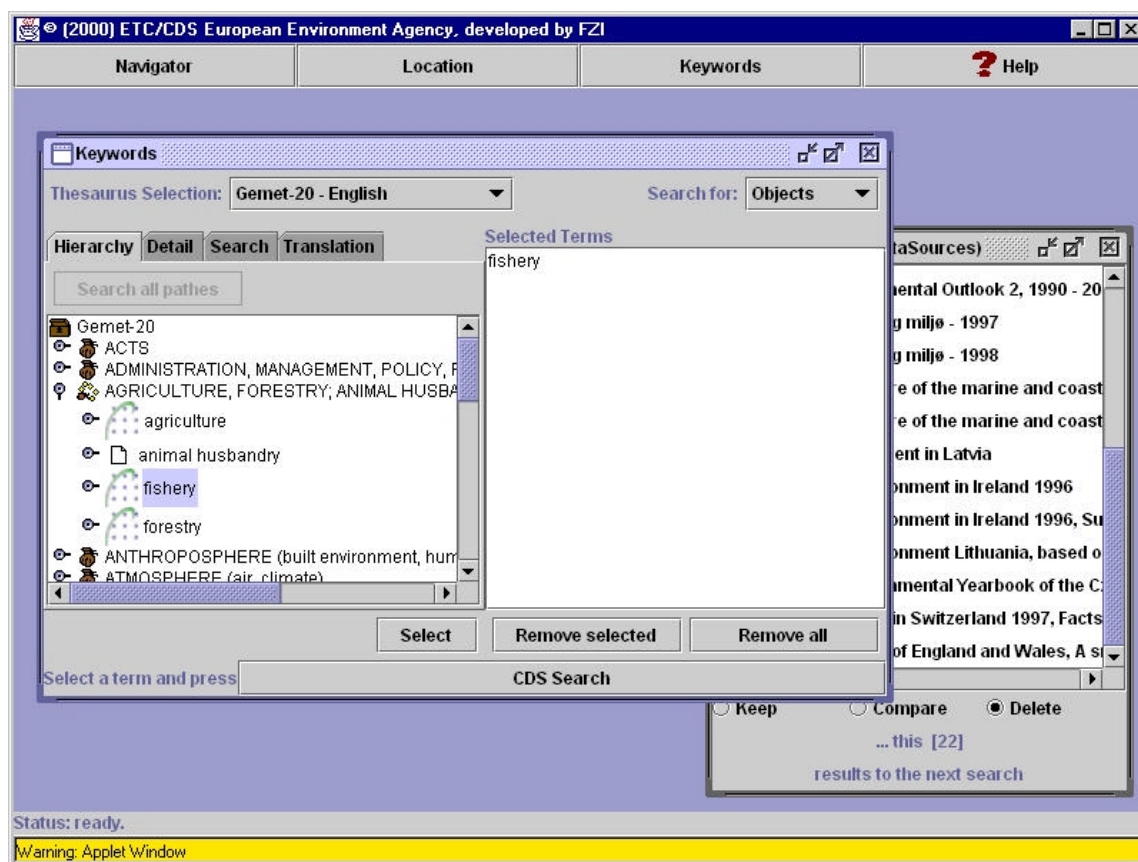


Figure 5 J-CDS with GenThes (Keywords) selected

- **The Navigator** allows to browse addresses and data sources. The user can navigate through the hierarchy and select an object of interest to view detailed information.

Independent of the selected tool the results are collected in a result list window. The user can now combine the results of one search with the results of the next search, either by just

keeping the old results in the list or by comparing them with the new ones. To display the detailed information a mini browser was implemented which displays the results by accessing the HTML engine of WebCDS.

6 Technical Specification

One of the major requirements of the WebCDS was as much platform independence as possible (see also Section 3.3). To achieve this all modules were developed in Java (JDK 1.1.x) and no proprietary database or operating systems specific development tools and modules were used. This way we were able to reduce the installation and configuration overhead of WebCDS on different platforms to a minimum. Nevertheless it is not possible to test all possible combinations of operating system platforms, database systems and web server. Especially the database support is currently limited to Oracle 7.3.4 and above and to Microsoft SQL Server 7.0. This section describes the Hardware and Software requirements to install and run the WebCDS 2.5 software and is intended to assist the administrator to select a proper platform for the installation.

6.1 Server Software

Software (server side)	Installed and tested	On demand
Operating System	UNIX Solaris HP-UX 10.20 Suse Linux 6.2 Windows NT 4.0	Further UNIX-Systems
Java	JDK 1.1.8 & Swing 1.0.2	
Database System	Oracle 7.3.4 Oracle 8 MS SQLServer 7.0	Further relational Databases
Database Connectivity	For all supported Oracle databases: Oracle-Thin-JDBC-driver 8.0.0.5 For MS SQL Server 7.0: TurboJ JDBC driver	Further JDBC-driver
Web Server	Servletrunner Sun Microsystems Java Web Server Apache with jserv-Module	All Web Servers supporting Servlets (like Jigsaw, IIS,...)

Table 1 Software requirements server side

Table 1 gives an overview about the software needed to install WebCDS on a server. The 'Installed and tested' column describes the platforms on which we have already installed WebCDS. The 'on demand' column gives an overview about possible software

configuration which may substitute the installed and tested software components if need. If you need WebCDS running on a specific platform, please contact additionally ETC/CDS (etc/cds@mu.niedersachsen.de) or FZI (kazakos@fzi.de).

6.2 Server Hardware

The hardware requirements vary, depending on the number of parallel accesses expected and on the performance requirements. Currently WebCDS is installed on a Sun UltraSparc and a HP workstation with >128MB main memory at the ETC/CDS in Hannover. Additional Installation at the EEA, the BUWAL and the Canton Geneva are using a Windows NT PC with 128 to 256 MB

Currently all software modules are installed on the same workstation. They can also be installed in distributed environment, due to the client-server architecture of WebCDS. One server for each of the components web server, database server and WebCDS server is possible. This way the hardware requirements for each server can be reduced or/and the performance can be improved.

6.3 Client Software & Hardware

On client side we have to distinguish between the HTML client and the Java client. For the HTML Client there are no special software or hardware requirements, except an HTML 3.2 capable Web Browser like Netscape Navigator or Internet Explorer and an Internet connection. The Java-Client (J-CDS) needs the Java Plug-in (see section 8.2)

6.4 Communication

There are two possibilities to run WebCDS. Either as one application or as a client server application. In the second case WebCDS components communicate with each other via Java Remote Method Invocation (RMI). To allow this communication a firewall between the components has to be opened on some well known ports. Additionally problems where encountered when the server on which WebCDS Server runs has two (2) IP addresses.

The first configuration is the default one.

7 Installing WebCDS 2.5

As mentioned in Section 6 WebCDS can be installed on different platforms with different configurations. In this section we describe what an administrator needs to do to install WebCDS. We exemplify the installation and configuration for an installation under UNIX with a Oracle database and as a single service (no client server). The servletrunner handles the servlet requests. Additionally we assume that a web server is available (no servlet support is needed). The installation and configuration on other combination is possible and comparable to this one.

7.1 Preconditions

In order to install WebCDS 2.5 following preconditions have to be fulfilled:

- Installed database management system (e.g. Oracle 7.3.4) with
 - ? a database user (e.g. cds25)
 - ? a password (e.g. cds25pwd), and
 - ? a SQL*Net address (e.g. `www.oraclehost.com:1521:orcl`).
- Installed jdk1.1.x and swing 1.0.2.
- Installed JDBC driver (e.g. JDBC-THIN Driver 8.0.0.5)
- Installed servletrunner (part of the freely available JSDK 2.0)

7.2 WebCDS distribution archive

All files and classes needed to run WebCDS 2.5 are packaged in the archive WebCDS25_DATE.zip where DATE is the distribution date of the archive. To start installation of WebCDS create a home directory for the installation and extract the archive to that directory. We refer to this directory from now on as \$WebCDS. Following subdirectory structure is created:

- catalogue, containing all files needed for the HTML generation
- dmp, containing the Oracle dumps and schema
- html_server, containing the application server specific classes and files

- `system`, containing all static pages, buttons etc..

7.3 Import of the CDS schema and the data

To run WebCDS a CDS 2.5 database schema and the data have to be imported into the Oracle user (e.g. `cds2.5`). This can be done by Oracle specific import tool (`imp`). Please use the `cds25.dmp` on the CDS for a initial version of the database user. Afterwards you can replace data with your data.

Type in

```
> imp user/passwd
```

and follow the instructions of the import tool.

7.4 Configuration of the DB Server

7.4.1 Oracle

In directory `$WebCDS/html_server/classes` edit the file `DB_common.properties`. Change the following entries to match to your environment:

- `database_type`, e.g. `oracle7`, `oracle8`, `sqlserver`
- `url`, the SQL*Net address. It has always the following syntax
`jdbc:oracle:thin:@"server_on_which_oracle_is_installed":`
`port":"listener"`, e.g.
`jdbc:oracle:thin:@www.oraclehost.com:1521:orcl`
- `jdbc_driver`, the JDBC driver class
- `login`, the user name of your Oracle user, and
- `passwd`, the password of your Oracle user.

A sample configuration for Oracle 8 looks like this (changes in the files are marked **bold**):

```
database_type=oracle8  
jdbc_driver=oracle.jdbc.driver.OracleDriver
```

```
url=jdbc:oracle:thin:@www.oracle.com:1521:orcl  
login=cds25  
passwd=cds25
```

7.4.2 Other DB management systems

WebCDS is build to easily support other database servers. Nevertheless, due to incompatibilities between the database management systems, some changes of the software have still to be done, to support other databases. This has already achieved for MS SQL Server. The following description exemplifies the installation on SQL Server.

A sample configuration for MS SQL Server 7 looks like this (changes in the files are marked **bold underlined**):

```
database_type=sqlserver  
jdbc_driver= com.ashna.jturbo.driver.Driver  
url= jdbc:JTurbo:// www.sqlserverhost.com/cds25  
login=cds25  
passwd=cds25
```

In this sample we use a commercial driver from ashna corp. (<http://www.ashna.com>). This is needed since MS SQL Server does not ship it's own driver.

7.5 Configuration of the Servlet Client and HTTP server

7.5.1 Servletrunner

The WebCDS servlet (CatalogueServlet.class) has to be registered within the servlet capable HTTP server. The following description is based on a servletrunner installation. If an other HTTP servlet capable web server will be used, please refer additionally to the servlet registration instructions of this webserver. Following steps have to be done (changes in the files are marked **bold underlined**):

- Create a directory 'system' in the public html directory of your webserver
- Create a soft link from the public_html/system/ directory to \$WebCDS/system/cds25.
- In file \$WebCDS/catalogue/servlet/properties/catalogue.properties change the URLs to match to your environment, following values have to be set (servlethost and webserverhost are usually the same):

```
mainServlet.host=www.servlethost.com
mainServlet.port=12345
main.index=http://www.webserverhost.com/system/cds25/index.html
main.index2=http://www.webserverhost.com/system/cds25
[ ... ]
picture.dir=http://www.webserverhost.com/system/cds25/pictures/
bundle.dir=absolut path to
$WebCDS/system/catalogue/properties
```

- In file \$WebCDS/catalogue/servlet/servlet.properties set the paths and the URL to match to your environment, i.e.:

```
propertyDir=absolut path to the property-directory
pictureDir=URL to the picture-directory
configfile=absolut path to the catalogue.properties file
```

To start the servlet, a sample script is also in this directory. It sets the classpath and path environment variables and starts the servletrunner. Edit the file WebCDS25_HTML in the directory \$WebCDS/catalogue/servlet/, i.e. set the java path, classpath etc. to match to your environment. To start the database server execute WebCDS25_HTML (WebCDS25_HTML.bat in Windows NT environments).

If you are using an other servlet capable web server instead of the servletrunner, please be sure, that the following entries are in the classpath of the webserver:

./

\$JAVA/jdk1.1/lib/classes.zip

\$ORACLE/jdbc/lib/classes111.zip

\$JAVA/swing-1.0.2/swingall.jar

\$JAVA/JSDK/lib/jsdk.jar

\$WebCDS/html_server/classes

\$WebCDS/catalogue/classes

\$WebCDS/catalogue/servlet/properties/

\$WebCDS/catalogue/servlet/DE.fzi.dbs.webcdsplus.mediator.jar

Where \$ORACLE, \$JAVA, and \$WebCDS are directories depending on your environment.

7.5.2 Jserv

Alternatively to servletrunner, some installation may use Apache web server and jserv module. If this is the case, the following files have to be edited instead:

Jserv.properties in directory /etc/httpd/jserv (\$JSERV):

- Insert servlet Zone: cdsservlets
- Define path: cdsservlets.properties = \$WebCDS/catalogue/servlet/ servlet.properties
- Define classpath as:
 - wrapper.classpath=\$WebCDS/catalogue/classes
 - wrapper.classpath=\$WebCDS/WebCDS/html_server/classes
 - wrapper.classpath=\$WebCDS/catalogue/servlet/properties
 - wrapper.classpath=\$WebCDS/catalogue/servlet/DE.fzi.dbs.webcdsplus.mediator.jar
 - wrapper.classpath=\$JAVA/lib/classes.zip
 - wrapper.classpath=\$JAVA/swingall.jar
 - wrapper.classpath= \$JSDK/lib/jsdk.jar
 - wrapper.classpath= \$ORACLE/jdbc/lib/classes111.zip

\$JAVA, \$WebCDS, \$ORACLE depend on your environment

Please be sure that you use one line for each classpath part!

In the file `servlet.properties`:

- update path to property directory; picture directory and config file
- path: `$WebCDS/catalogue/servlet/servlet.properties`

In the file `jserv.conf` (directory `$JSERV`):

- insert: `ApJServMount/cdsservlets/cdsservlets`

7.6 Access WebCDS HTML

To access WebCDS via HTTP open your WebBrowser and type in the following URL:

`http://www.myhost.com:12345/servlet/CatalogueServlet?page=CCDSFrameSet&language=en`

This starts the English version of WebCDS. To start WebCDS in an other language use the following 2 letter language code:

en	English
de	German
fr	French
gr	Greek
pt	Portuguese
es	Spanish
it	Italian
no	Norwegian

This languages are currently available. To add further languages please read section 7.7.1.

7.7 Customizing WebCDS 2.5

7.7.1 Adding new languages

WebCDS 2.5 has been designed to easily add new languages (see also section 3.2). By using so called PropertyResourceBundle it is possible to extend the WebCDS language support by just translating the text files. Currently WebCDS supports eight languages as mentioned in section 7.6. To add a new language copy the file catalogueBundle.properties in the directory \$WebCDS/catalogue/servlet/properties to catalogueBundle_LANG.properties (LANG corresponds to the 2-letter ISO 639 language code of the new language) and translate the tags on the right of the equation mark. E.g. :

Original Text in file catalogueBundle.properties:

```
detail.top.text.ds=Detailed Information
```

French translation in file catalogueBundle_fr.properties:

```
detail.top.text.ds=information détaillée
```

After translating the file WebCDS 2.5 is available in the new language as described in section 7.6.

If you want to translate the help and introduction page, you can use a standard HTML editor, write in the text, and save it as help_LANG.properties and intro_LANG.properties in the \$WebCDS/catalogue/servlet/properties/ directory.

7.7.2 Replacing pictures and static HTML pages

All pictures and static HTML pages used within WebCDS are located in the system subdirectory of the installation directory. In order to adapt WebCDS further to the corporate identity of the installing institution pictures in the following directories may be replaced (without changing the names):

- `$WebCDS/system/cds25/pictures` containing general purpose pictures, e.g. the WebCDS headbar
- `$WebCDS/system/cds25/pictures/languages` containing the language buttons, e.g. english
- `$WebCDS/system/cds25/pictures/classes` containing the data classes, e.g. data set, map etc.
- `$WebCDS/system/cds25/pictures/flaggen` containing the country flags

The starting page for the language selection is located in directory `cds25/` and is called `index.html`. This page can be replaced by any other starting page pointing to the URL:

<http://x.y.z/servlet/CatalogueServlet?page=CCDSFrameSet&language=en>

7.7.3 replacing introduction, explanation and help text

If you want to replace the help and introduction page, you can use a standard HTML editor, write in the text, and save it as `help_LANG.properties` and `intro_LANG.properties` in the `$WebCDS/catalogue/servlet/properties/` directory. `LANG` is a substitute for the language of the text (e.g. `help_en.properties` for the English help text).

To change the explanation text you have to edit additionally the files `explad_LANG.properties` (explanation for addresses) and `explds_LANG.properties` (explanation form data sources).

7.7.4 Changing colours

If you want to change the colours of WebCDS 2.5 you have to adopt the `catalogue.properties` file in the `$WebCDS/catalogue/servlet/properties` directory. E.g.:

```
detail.bottom.bgcolor=#7DCD9F
```

```
detail.bottom.text=#000000
```

```
detail.bottom.link=#50509B
```

```
detail.bottom.vlink=#B0B0B0
```

The colour information is coded in standard hexadecimal RGB values.

7.7.5 Changing feedback E-mail address

If you want to change the colours of WebCDS 2.5 you have to adopt the catalogue.properties file in the \$WebCDS/catalogue/servlet/properties directory. E.g.:

```
main.feedback=abc@company.com
```

7.8 Search fields

This chapter describes the mappings between the field in the search forms and the attributes in the database schema

7.8.1 Quick Search for Addresses

Quick search for addresses searches in the following fields of the CDS:

- First and last name of a person, and name of an organisation, i.e. addresscore.adr_name and addresscore.adr_ps_first_name
- Acronyms of an organisation, i.e. addresslocal.adr_loc_acro

7.8.2 Quick Search for Data Sources

Quick search for data sources searches in the following fields of the CDS

- Title, English title, and abstract of a data source, i.e. datacore.dat_title, datacore.dat_engl_title, datacore.dat_abstract.
- GEMET terms associated with a data source, i.e. thesdesc.th_term joined with datthesrel.

The search in the GEMET terms is always multilingual, e.g. by searching for the german word 'Wasser' you will find all data sources indexed with the corresponding English word 'water'.

7.8.3 Expert Search for Addresses

First you have to select if you are searching for a person, an organisation, or an address being a part of the EIONET (EEA roles and topics). In addition you can restrict the results by restricting the country and the city of the address. The corresponding fields are:

- addresscore.adr_name, addresscore.adr_ps_first_name, addresslocal.adr_loc_acro when organisation or person is selected.
- thesdesc.th_term for the thesaurus search when EEA Roles and Topic is selected.
- addresscore.adr_ps_city
- addresscore.adr_cc (country code of the searched country)

7.8.4 Expert Search for Data Sources

The mapping between the search fields and the database attribute is as follows:

- Search term is mapped to datacore.dat_title, datacore.dat_ababstract, datafreeterms.dat_terms and the thesaurus terms within thesdesc.th_term
- Use translations indicates that the term is translated into all available GEMET languages before send to the search.
- Class is mapped to datacore.dat_class
- Country is mapped to dataextloc.loc_code
- Language is mapped to datacore-dat_lg_meta
- Time is mapped to dataext.tm_from and dataext.tm_until

7.9 Troubleshooting WebCDS 2.5

On problems installing or running WebCDS 2.5 please inform the ETC/CDS (matthias.menger@mu.niedesachsen.de) or/and the developers (kazakos@fzi.de).

7.9.1 Instability after some days of use

Problem: WebCDS 2.5 is up and running. Nevertheless on some installations encounter instabilities after running some days.

Reported platform: Oracle 7, JDBC driver thin 7.3.4

Reason: The reason seems to be in the communication between Application and Oracle Server.

Resolution: Install Oracle JDBC thin driver 8.0.0.5 or/and upgrade to Oracle 8.

8 Installing J-CDS

8.1 Configuration of J-CDS Client and Server

To install JCDS unzip the delivered archive, i.e. jcds_DATE.zip (DATE corresponds to the distribution date) in a directory of your choice, e.g. \$WebCDS/jcdis.

The properties which have to be adopted to the local environment are located in the following directories:

- `./serverProperties` contains the properties for the server. The access to this directory should be restricted to administrators, since it contains security information about the database .
- `./` (working directory) contains the client side configuration information for J-CDS.
- `./DE/fzi/catalogue/components/genthes` contains the client side configuration information for GenThes.

The JCDS Server has to be installed on the same host as the WebCDS server, since applets can connect only to the computer they are downloaded from.

To configure J-CDS following steps have to be taken (changes in the files are marked **bold underlined**):

- In the file `./JCDS.properties` (in the working directory) change the IP address to the IP address of the computer installed and change the target URLs for the detailed view:

```
DBAccessURL = rmi://www.rmihost.com:3000/JCDSServer
[ ... ]
DetailServlet_OBJECT =
http://www.servlethost.com:12345/servlet/CatalogueServ
let?page=CDetailedDataSourcesBottomFrame&language=en&u
serID=JAVAxzy&id=
```

```
DetailServlet_ADDRESS =  
http://www.servlethost.com:12345/servlet/CatalogueServlet?page=CDetailedAddressBottomFrame&language=en&userI  
D=JAVAAxyz&id=  
HTMLDetailRedirect =  
http://www.servlethost.com:12345/servlet/CatalogueServlet?
```

- In the file `serverProperties/JCDSserver.properties` change the IP address to the IP address of the computer installed. Change additionally the JDBC properties, the username and the password:

```
driverName = oracle.jdbc.driver.OracleDriver  
jdbcURL =  
jdbc:oracle:thin:@www.oraclehost.com:1521:orcl  
user = login  
password = passwd  
RMIRegistryHost = 111.222.111.222
```

(please use the IP address and not the domain name of the rmihost)

- In the file `GenThes_JCDS.properties` change the IP address to the IP address of the computer installed. Change additionally the JDBC properties, the username and the password. The file `GenThes_JCDS.properties` appears twice (for the client in directory `./DE/fzi/catalogue/components/genthes/` and the server in directory `./serverProperties/`). Please change the properties in both files. The login and password should be added only at the server properties, i.e.:

- In file `./serverProperties/GenThes_JCDS.properties` only:

```
cds.database.login=login  
cds.database.password=passwd
```

- In both `GenThes_JCDS.properties` files:

```
cds.database.genthes.url=jdbc:oracle:thin:@www.oraclehost.com:1521:orcl
```

cds.genthes.rmiurl = //111.222.111.222:3000

(please use the IP address and not the domain name of the rmihost)

- Create a soft link from the HTTP-Root to the directory \$WebCDS/jcds.
- To start J-CDS servers you can use the startServer_Win32_JCDS.bat (Windows) or startServer_Unix_JCDS (UNIX) script, by changing the paths and classpath according to your environment.
- The installation of the server is now complete, and the Applet is available under the URL, e.g. <http://www.xyz.de/jcds/index.html> . *Please note: you have to install the Java Plug-in to run the client.(see section 8.2)*

8.2 Installing the Java Plug-In on the client

As long the Java 1.1 support of the Browser does not include 100% Java compliance, J-CDS needs a Java Plug-In to be installed on the clients. The Version number needed is 1.1.2 and can be downloaded from the SUN sites. To install the plug-in follow the following steps:

- Download the Plug-in from <http://java.sun.com/products/plugin/1.1.2/index-1.1.2.html>
Please note: Do not use any other version of the plug-in!
- Uninstall older or other versions of the Plug-in
- Install Java Plug-in 1.1.2: Execute the exe-file and follow the instructions
- Configure the Java Plug-In 1.1.2
 - ? Start->Program files-> Java Plug-in ControlPanel
 - ? In the tabbed folder "Basic" set:
 - ? Enable Java Plug-in: Yes
 - ? Show Java Console: Yes
 - ? Cache JARs in memory: No
 - ? Network access: Unrestricted

8.3 Troubleshooting J-CDS

On problems installing or running J-CDS please inform the ETC/CDS (matthias.menger@mu.niedesachsen.de) or/and the developers (kazakos@fzi.de).

If you encounter problems running J-CDS please be sure to enable the java console by setting "Show Java Console" in Start/Program Files/Java Control Panel. You need to restart your Browser afterwards. On Problems the Java Console will display some exception messages. If none of the following troubleshooting tips helps solving the problem, please send the content of the console and your specific platform information (browser, version, operating system, plug-in etc.) to the ETC/CDS.

8.3.1 SecurityException

This Exception occurs when a firewall is installed between the client (the computer where the browser is installed) and the server. Please inform your system administrator to open the J-CDS ports. The default ports are 3000,3001,3002.

8.3.2 Detailed View does not work

This problem occurs when Java Plug-in 1.1.3 is installed. Please be sure to install Java Plug-in 1.1.2.

9 Current Installations of WebCDS 2.5

WebCDS 2.5 is up and running. It is currently installed in the following locations:

- European Topic Centre on Catalogue of Data Sources (ETC/CDS), Hannover
Public access possible: <http://www.mu.niedersachsen.de/cds>
- European Environmental Agency (EEA), Kopenhagen
Currently Intranet only
- Bundesamt für Umwelt, Wald und Landschaft (BUWAL), Bern
Currently Intranet only
- Kanton Geneva, Geneva
Currently Intranet only

10 WebCDS Data Model

10.1 Description

Datamodel Description WebCDS 2.5

Table Name	Field Name	Description
------------	------------	-------------

AccessConstrains contains the entries for the access constraints pick list

dat_access	unique identifier for the access constraint
acc_constrain	Name of the access constraint to be selected from the Data form
acc_sequence	sequence for the entries in the picklist
Old16	identifier for the access constraint name in CDS Version 1.6

AddressCore contains address information of the CDS core datamodell

adr_id	Unique Identification of Address; generated
adr_type	Type of Address (Institutions etc =2, Persons = 1); mandatory; coded
adr_name	Name of institution or person; mandatory
adr_ps_form	Title of addressee (Mr, Mrs, Miss, Doctor, His Holiness); optional
adr_ps_title	Job title of addressee; optional
adr_ps_first_name	Given name or nickname (John, Michael, Edward the Confessor); optional
adr_tel	Telephone number; mandatory (if no e-mail or fax or...)
adr_fax	Telefax number; mandatory (if no e-mail or phone or...)
adr_email	E-Mail address; mandatory (if no phone or fax or ...)

adr_url	URL address; mandatory (if no phone or fax or ...)
adr_cc	Country code; mandatory; picklist
adr_ml_zip	Postcode; mandatory (if no email or phone or fax or URL)
adr_ml_street	Name of street; mandatory (if no email or phone or fax or URL)
adr_ml_city	Name of village; mandatory (if no email or ...); picklist form database
adr_pb_zip	Postbox: postcode; mandatory (if no email or phone or fax or URL)
adr_pb	Postbox: number; mandatory (if no email or phone or fax or URL)
adr_date	Date of creation or last revision of Address; generated; mandatory
adr_creation_date	Date of creation of Address; generated; mandatory
adr_id_orig	Id of original record
adr_valid	Yes if all mandatory fields are filled
adr_has_parent	Performance enhancement: Has Parent = 1..., No Parent = 0; generated
adr_mp_flag	flag for export and search conditions

AddressLanguage languages of addressee

adr_id	Unique Identification of Address; generated
adr_language	preferred languages; mandatory; default

AddressLocal local name/acronym of organisation

adr_id	Unique Identification of Address; generated
adr_language	selected languages; picklist; default

adr_loc_name local address name (different organisation names based on the selected language)

adr_loc_acro local address acronym (different organisation acronyms based on the selected language)

AddressType [pick list for the address type selection \(organisation, person\)](#)

adr_type Type number of Address (Organisations etc =2, Persons = 1)

adt_name Type description

adt_sequence sequence for the entries in the picklist

AdrDatRelation [reference to data sources](#)

adr_id Unique Identification of Address; generated

dat_id Unique Identification of Data Source; generated

adt_type Function or Type of Relation; mandatory

AdrDatRelationType [pick list of reference types](#)

adt_type unique identifier for the relation type

adt_text Name of the relation type to be selected from the Address or Data form

adt_sequence sequence for the entries in the picklist

Old16 identifier for the relation type in CDS Version 1.6

AdrHasChildren [relation between addresses and they children](#)

adr_id Unique Identification of Address

adr_has_children Unique Identification of the child Address

AdrRel [reference to addresses \(parent-child relation\)](#)

adr_parent_id Unique Identification of Parent Address; generated

adr_id Unique Identification of Address; generated

adr_name Name of Parent Address; redundant

adr_parent_name Name of Address; redundant

AdrThesRel[reference to thesaurus descriptors](#)

th_thes_no	No of term list or Thesaurus (see table: ThesThes); generated
th_lang_no	No of language (see table: Language); generated
th_desc_no	No of descriptor; generated
th_no	Compiled internal hash-codeing
adr_id	Unique Identification of Address; generated

DataClass[pick list for the data class selection](#)

dat_class	Class of Data Source (Dataset, Activity,...); mandatory; coded
dtc_name	Description
dtc_dialog_box	Name of the extension form which will be opened by clicking the next button on the form Data
dtc_sequence	sequence for the entries in the picklist
Old16	identifier for the data class in CDS Version 1.6

DataCore[contains data source information of the CDS core datamodell](#)

dat_id	Unique Identification of Data Source; generated
dat_title	Title of Data Source; mandatory
dat_title_engl	English title of Data Source; optional
dat_abstract	Abstract; optional
dat_abstract_engl	English Abstract, optional
dat_ordering	Ordering information; optional
dat_access	Accessibility of data, e.g. public, for official use only, not available; optional
dat_use	Usage restrictions; optional
dat_lg_meta	Language of meta information; mandatory; default; picklist

dat_date	Date of last revision of Data Source; generated; mandatory
dat_creation_date	Date of creation of Data Source; generated; mandatory
dat_class	Class of Data Source (Dataset, Activity,...); mandatory; coded
dat_id_orig	Id of Origin
dat_valid	Yes if all mandatory fields are filled
dat_has_parent	Perfomence enhancement: Has Parent = 1..., No Parent = 0; generated
dat_date_source	Date of creation/modification of the resource (no locator); mandatory
dat_sel_criteria	Indicates if the locator has to be exported (i.e. if it fulfills special conditions for export)
dat_mp_flag	flag for export and search conditions

DataDistribution**Formats in which the data source is available**

dat_id	Unique Identification of Data Source; generated
dat_distrib	Distribution medium (CD-Rom, paper, online); optional

DataDobris**contains the data source information for the requirements of the Dobris+3 report**

dat_id	Unique Identification of Data Source; generated
Data_source	Data Source description
Definitions	Definitions
Remarks	Remarks
Geo_coverage	Geographical coverage
Geo_detail	Geographical level of detail
Time_series	Time series
Units	Units

Presentation	Probable presentation
Problems	Aggregation problems
Target	Target/projections/scenario
Compilation	Data compilation
Contact	Contact
Action	Action required
OLE	Embedded OLE object

DataExt

contains data source information of the level 1 extension

dat_id	Unique Identification of Data Source; generated
ext_subclass	Name of the group; Subclass Information (1.7-)
ext_notes	Special Notes and additional information on Data Source; optional
ext_status	State, e.g. planned, closed, in progress, n/a; optional; picklist
su_descr	General description of technical coverage; regular
su_sta_type	Type of Station (1.7-)
su_sta_equip	Equipment of Station (1.7-)
su_prod_usage	Scope of Tool, e.g. Statistical calculations, test and regression analysis; regular (-1.6 purp; 1.7-)
su_proj_type	Type of project(1.7-)
su_map_data	Type of datacollection underlying the map; regular
su_map_type	Type of map, e.g. topographic or thematic; mandatory; picklist
su_map_gis	GIS-System; (1.7-)
su_map_res	Resolution of Map; (1.7-)
su_map_res_unit	Unit of Resolution of Map, (dpi, l/mm); picklist (1.7-)

su_map_coord	Coordinate system used for the map, e.g. UTM, Gauss-Krüger,...; regular; picklist (1.7-)
su_doc_publisher	Publisher of Document (1.7-)
su_doc_publ_place	Place of Publishing (1.7-)
su_doc_year	Year of Publishing (1.7-)
su_doc_isbn	ISBN/ISSN Number (1.7-)
su_doc_source	Published in / Source (1.7-)
su_doc_type	Type of Document (1.7-)
lo_bb_auto	Automatically generated BB?; (1.7-)
lo_bb_x_coord1	Coordinate of bounding box, lower left, x-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_bb_y_coord1	Coordinate of bounding box, lower-left, y-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_bb_z_coord1	Coordinate of bounding box, lower-left, z-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_bb_x_coord2	Coordinate of bounding box, upper-right, x-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_bb_y_coord2	Coordinate of bounding box, upper-right, y-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_bb_z_coord2	Coordinate of bounding box, upper-right, z-coordinate; mandatory (if no dat_lo_id); (1.7-)
lo_geoloc	Additional geographic identification; compatibility
lo_descr	Additional free text description of spatial coverage; regular
tm_from	Starting date; optional
tm_until	Ending date; optional
tm_step	Empirical Data: Time step (value) of update intervall (see tm_interv for unit), e.g. 5; optional

tm_interv	Empirical Data: Type of update intervall, e.g. second, minute, hour, day, month, year; optional; piclist
tm_period	Empirical Data: Indicates if the data collections if updated updated regularly, not regularly or never; optional; piclist
tm_descr	General description of time coverage; regular
su_ext_ole	Embedded OLE object

DataExtAuthor [list of all authors, subtable for the document class](#)

dat_id	Unique Identification of Data Source; generated
su_doc_author	Name of the Author (Data Class: Document)

DataExtDataType [data format subtable for the dataset class](#)

dat_id	Unique Identification of Data Source; generated
ext_datatype	Form of data, e.g. digital, analog, digital and analog; regular; (1.7-)

DataExtEditor [editor subtable for the document class](#)

dat_id	Unique Identification of Data Source; generated
su_doc_editor	Name of the Editor (Data Class: Document)

DataExtLoc [reference to the location tables](#)

dat_id	Unique Identification of Data Source; generated
loc_code	location code (NUTS); mandatory
loc_code_name	selected location
loc_region	Region Name coded via Location0 table (determines the loc_code_name if loc_country is empty)
loc_country	Country Name coded via Location1 table (determines the loc_code_name if loc_state is empty)
loc_state	State Name coded via Location2 table (determines the loc_code_name if loc_county is empty)

loc_county	County Name coded via Location3 table (determines the loc_code_name if loc_municipal is empty)
loc_municipal	Municipal Name coded via Location4 table (determines the loc_code_name if not empty)
dlc_add_code	additional, optional location code

DataExtOrigScale [original scale, subtable for the map class](#)

dat_id	Unique Identification of Data Source; generated
mp_prod_scale	Basic scale used in the production of the map; regular; (1.7-)

DataExtPar [parameters associated with data source, subtable for datasets, maps and stations/sites](#)

dat_id	Unique Identification of Data Source; generated
ext_parameter	Datasets, Maps: physical, chemical or biological quantity - Stations and Sites: Type of station;
ext_unit	Datasets: Unit of measurements, e.g. m ³ , days, µm; regular;
ext_method	Datasets: Standards or Methods of data acquisition, e.g. gas chromatography, field investigation - Maps: Standards or Methods of data acquisition, e.g. satellite photography; regular;
ext_quality	Datasets: Quality ensuring Norms or Procedures, e.g. according to ISO 9000; regular; (-1.6)
ext_precision	Datasets: Precision of measurements or calculations, e.g. 0.5-1mg; regular; (-1.6)
ext_limit	Datasets: Detectable limit of substances measured according to the equipment used; optional; (-1.6)
ext_evaluation	Dataset, Map: Description of the evaluation; optional; (-1.6)

DataExtPublScale [publication scale, subtable for the map class](#)

dat_id	Unique Identification of Data Source; generated
mp_publ_scale	Scale used for the publication of the map; regular; (1.7-)

DataFreeTerms Descriptive terms not drawn from the thesaurus

dat_id	Unique Identification of Data Source; generated
dat_distrib	Free Terms; optional

DataHasChildren Relation between a data source and the children

dat_id	Unique Identification of Data Source
dat_has_children	Unique Identification of Data Source Child

DataIndex Performance optimising index *(not yet supported)*

idx_word	
idx_dat_id	
idx_dat_title	

DataLanguage languages in which the resource is available

dat_id	Unique Identification of Data Source; generated
dat_lg_source	Language of DataSource; mandatory; default; picklist

DataLegislation legislation, the data source refers to or satisfies

dat_id	Unique Identification of Data Source; generated
leg_abbr	Abbreviation
leg_descr	Name
leg_country	Country of the Legislation; picklist
leg_no	Official Number
leg_part	Article number

DataSubClass pick list for the group selection

dat_class	Class identifier, where the subclasses are belonging to
-----------	---

ext_subclass	identifier for the subclass/group
ett_name	Name of the subclass
ett_sequence	sequence for the entries in the picklist

DataSubClassFields [list of all visible fields for the corresponding subclass](#)

dat_class	Class identifier, where the subclasses are belonging to
ext_subclass	identifier for the subclass/group
dsf_section	identifier for the extension type (1=Technical, 2=Spatial, 3=Temporal)
dsf_field	Field name on the extension form
dsf_visible	indicates if the field is visible for the affiliated subclass
dsf_enabled	unused (indicates if the field is enabled for the affiliated subclass)
dsf_locked	unused (indicates if the field is locked for the affiliated subclass)

DataUrl [urls associated with data source](#)

dat_id	Unique Identification of Data Source; generated
dat_url	URL address; optional

DatRel [reference to data sources \(parent-child relation\)](#)

dat_parent_id	Unique Identification of Parent Data Source; generated
dat_id	Unique Identification of Data Source; generated
dat_parent_title	Title of Parent Data Source; redundant
dat_title	Title of Data Source; redundant

DatThesRel [reference to thesaurus descriptors](#)

th_thes_no	No of term list or Thesaurus (see table: ThesThes); generated
th_lang_no	No of language (see table: ThesLang); generated
th_desc_no	No of descriptor (see table ThesDesc); generated
th_no	unused (unique identifier for the descriptor)
dat_id	Unique Identification of Data Source; generated

ISO3166**ISO 3166 entries (countries)**

ctr_cc	two letter abbreviation of the country
ctr_cc3	three letter abbreviation of the country
ctr_name	Countryname
ctr_no	unique index of the country
ctr_priority	indicates the sequence in the listboxes (where the countries are sorted by the priority)

ISO639**ISO 639 entries (languages)**

lg_code	three letter abbreviation of the language
lg_no	unique index of the language
lg_name	Languagename
lg_priority	indicates the sequence in the listboxes (where the languages are sorted by the priority)

Location0**top level locations list**

loc_code	unique identifier for the region
loc_sequence	sequence for the entries in the picklist
loc_name	name of region; picklist
loc_bb_x_coord1	x,y,z: minimum coordinates of the bounding box enclosing the region
loc_bb_y_coord1	
loc_bb_z_coord1	

loc_bb_x_coord2 x,y,z: maximum coordinates of the bounding
box enclosing the region

loc_bb_y_coord2

loc_bb_z_coord2

Location1

second level locations list (first nuts codes level)

loc_code unique identifier for the country

loc_parent identifier of the affiliated region (from table
Location0)

loc_name name of country; picklist

loc_bb_x_coord1 x,y,z: minimum coordinates of the bounding
box enclosing the country

loc_bb_y_coord1

loc_bb_z_coord1

loc_bb_x_coord2 x,y,z: maximum coordinates of the bounding
box enclosing the country

loc_bb_y_coord2

loc_bb_z_coord2

Location2

third level locations list (second nuts codes level)

loc_code unique identifier for the state

loc_parent identifier of the affiliated country (from table
Location1)

loc_name name of state, picklist

loc_bb_x_coord1 x,y,z: minimum coordinates of the bounding
box enclosing the state

loc_bb_y_coord1

loc_bb_z_coord1

loc_bb_x_coord2 x,y,z: maximum coordinates of the bounding
box enclosing the state

loc_bb_y_coord2

loc_bb_z_coord2

Location3**fourth level location list (third nuts codes level)**

loc_code unique identifier for the county

loc_parent identifier of the affiliated state (from table
Location2)

loc_name name of county; picklist

loc_bb_x_coord1 x,y,z: minimum coordinates of the bounding
box enclosing the county

loc_bb_y_coord1

loc_bb_z_coord1

loc_bb_x_coord2 x,y,z: maximum coordinates of the bounding
box enclosing the county

loc_bb_y_coord2

loc_bb_z_coord2

Location4**fifth level location list**

loc_code unique identifier for the municipal

loc_parent identifier of the affiliated county (from table
Location3)

loc_name name of municipal; picklist

loc_bb_x_coord1 x,y,z: minimum coordinates of the bounding
box enclosing the municipal

loc_bb_y_coord1

loc_bb_z_coord1

loc_bb_x_coord2 x,y,z: maximum coordinates of the bounding
box enclosing the municipal

loc_bb_y_coord2

loc_bb_z_coord2

MapResolutionUnit picklist with resolution units, subtable for GIS System and GIS Layer, groups of data class map

su_map_res_unit Unit of Resolution of Map, picklist; (1.7-)

msu_map_res_unit Unit of Resolution of Map, picklist; (1.7-)

msu_sequence sequence for the entries in the picklist

MapTypes picklist with map types, subtable for GIS Systems and GIS Layers (groups of data class map)

su_map_type Type of map underlying the GIS; regular

mpt_map_type Type of map underlying the GIS; regular

mpt_sequence sequence for the entries in the picklist

StatusOfResource contains the entries for the status pick list

ext_status unique identifier for the status

str_status Name of the status to be selected from the Data form

str_sequence sequence for the entries in the picklist

OLD16 identifier for the status name in CDS Version 1.6

ThemeDatRel relation between GEMET Themes and Data Sources

tht_id unique identifier of the theme

tht_theme title of the theme

dat_id unique identifier of the data source

ThesDef thesaurus definitions to corresponding descriptors

th_thes_no Thesaurus number, coded via ThesThes table

th_lang_no Language number, coded via Language table

th_desc_no Descriptor number, unique number from descriptor ThesDesc

th_no unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)

definition Definition Note

ThesDesc thesaurus descriptors

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number for descriptor
th_no	Hashcode
th_type	1=Descriptor; 2=Used For; 3=Synonym
th_top_term	is 1 if no parent
th_term	Descriptor

ThesDescPicked [temporary thesaurus terms picked by the user](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc

ThesGrp [thesaurus descriptor to group reference](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
ths_id	Thesaurus super group number
thg_id	Thesaurus group number

ThesLang [thesaurus language table](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table

ThesNar [thesaurus narrower and broader term reference](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table

th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
th_nar_desc_no	Narrower Descriptor number, unique number from descriptor ThesDesc
th_nar_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)

ThesOrigId [thesaurus compatibility table](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_orig_thes_name	Name of Original Thesaurus
th_orig_desc_no	Original Descriptor Number from different Thesaurus

ThesRel [thesaurus descriptor reference to related terms](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	nused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
th_rel_desc_no	related Descriptor number, unique number from descriptor
th_rel_no	unused (unique identifier for the related term of th_thes_no + th_lang_no + th_rel_desc_no)

ThesSco**thesaurus scope notes to corresponding descriptors**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
scope_note	Scope note

ThesSyn**thesaurus descriptor reference to synonyms**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the related term composed of th_thes_no + th_lang_no + th_rel_desc_no)
th_syn_desc_no	Synonym number, unique number from ThesDesc
th_syn_no	unused (unique identifier for the synonym composed of th_thes_no + th_lang_no + th_syn_desc_no)

ThesTheme**thesaurus descriptor to theme reference**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number for descriptor
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
tht_id	Theme number

ThesThes**list of thesauri**

th_thes_no	Thesaurus number, coded via ThesThes table
------------	--

th_name	Name of Thesaurus
th_url	URL or Thesaurus
th_description	Description; future use

ThesTop[thesaurus descriptor reference to top terms](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
th_top_desc_no	Description number, unique number from ThesDesc
th_top_no	unused (unique identifier for the top term composed of th_thes_no + th_lang_no + th_top_desc_no)

ThesTypes[thesaurus type reference](#)

th_type	index of the descriptor type
th_name	name of the descriptor type

ThesUse[thesaurus descriptor reference to "use termes"](#)

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
th_desc_no	Descriptor number, unique number from descriptor ThesDesc
th_no	unused (unique identifier for the descriptor composed of th_thes_no + th_lang_no + th_desc_no)
th_use_desc_no	used for Descriptor number, unique number from descriptor
th_use_no	unused (unique identifier for the use-for-term composed of th_thes_no + th_lang_no + th_use_no)

ThRefGrp**thesaurus group description**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
ths_id	Descriptor number, unique number from descriptor ThesDesc
thg_id	Thesaurus group number
thg_group_accro	Group acronym
thg_group	Group

ThRefSGrp**thesaurus super group description**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
ths_id	Descriptor number, unique number from descriptor ThesDesc
ths_super_group_accro	Super group acronym
ths_super_group	Super Group

ThRefTheme**thesaurus themes description**

th_thes_no	Thesaurus number, coded via ThesThes table
th_lang_no	Language number, coded via Language table
ths_id	Descriptor number, unique number from descriptor ThesDesc
tht_theme_accro	Theme acronym
tht_theme	Theme

TimeInterval**contains the entries for the units of the time interval pick
list**

tm_interv	Unit of the time interval; picklist
tt_interv	Unit of the time interval; picklist
tt_sequence	sequence for the entries in the picklist

TimePeriod contains the entries for the update period pick list

tm_period	Period of data collection or measurement; picklist
tp_period	Period of data collection or measurement; picklist
tp_sequence	sequence for the entries in the picklist

WebListAddress materialized view of addresscore and adrrel

adr_name	Name of institution or person
adr_inst_name	Name of parent institution or person
adr_id	Unique Identification of Address
adr_inst_id	Unique Identification of parent of Address
adr_ps_first_name	Given name or nickname (John, Michael, Edward the Confessor); optional
adr_ps_form	Title of addressee (Mr, Mrs, Miss, Doctor, His Holiness)
adr_type	Type of Address (Institutions etc =2, Persons = 1); mandatory; coded
adr_has_parent	Perfomence enhancement: Has Parent = 1..., No Parent = 0; generated
adr_ml_city	Name of village; mandatory (if no email or ...)
adr_cc	Country code; mandatory; picklist

XlastUpdate Last update of an data source or address

id	unique identifier of the data source or address
timestamp	time stamp, time of last update
xuser	user name who did the last update

XobjectType help table for the update module

id	unique identifier of an address or data sources
objecttype	objecttype of an address of data source

Xpermission access restrictions

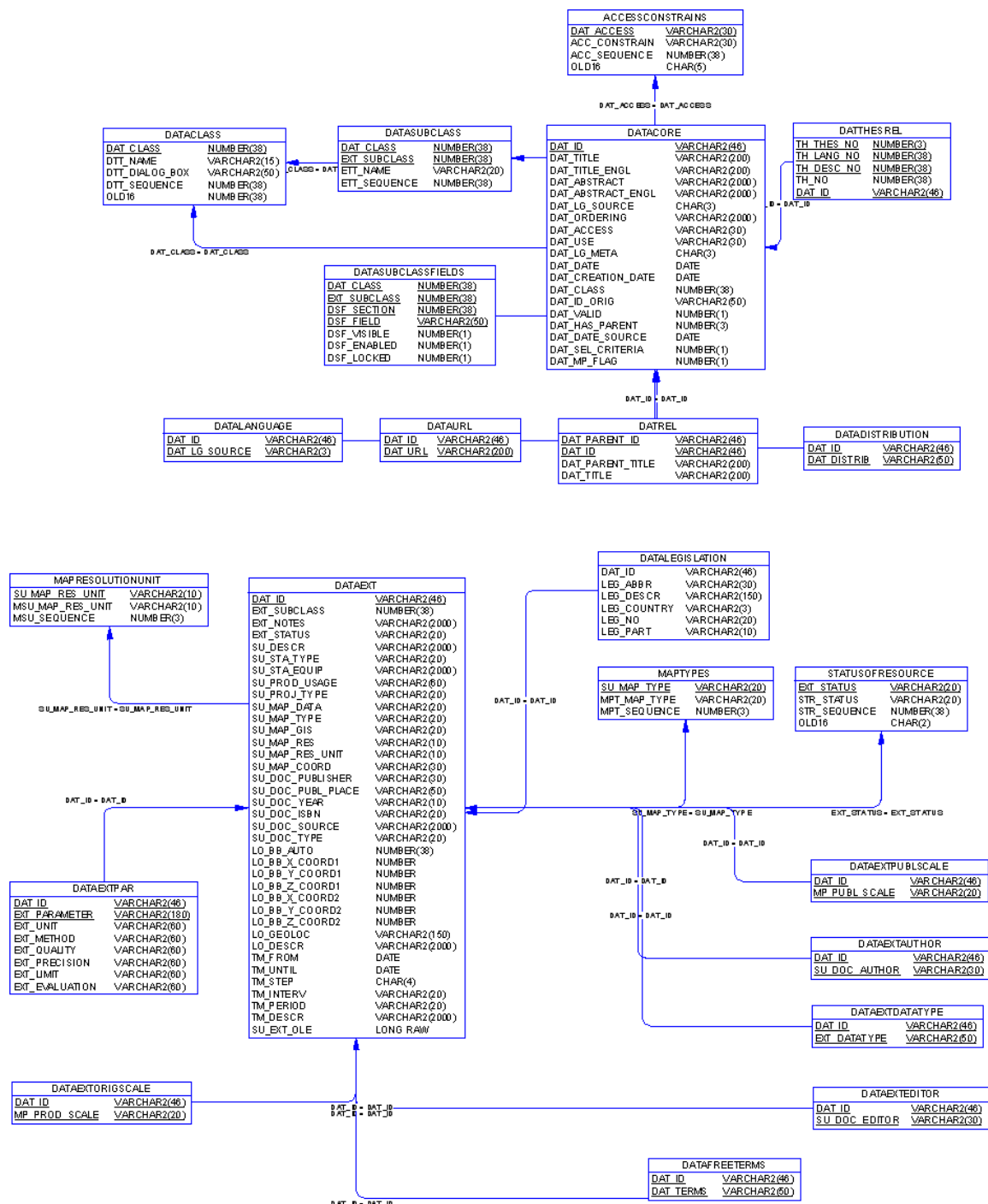
id	unique identifier of an address or data sources
xuser	user name having the permission to change the record

Xuser

basic authentication of users

xuser	user name
password	password

10.2 Diagram



References

CNR, Rome, Umweltbundesamt, Berlin, 1997

CNR, Rome, Umweltbundesamt, Berlin (1997). Gemet - general european multilingual environment thesaurus. Technical report, European Environment Agency. Version 1.0.

Crossley, 1996

Crossley, D. (1996). Wais through the web - discovering environmental information. <http://www.ncsa.uiuc.edu/SDG/IT94/Proceedings/Searching/crossley/paper.ps>.

JavaSoft, 1999a

JavaSoft (1997a). intlspec.doc1.html. <http://www.javasoft.com:80/products/jdk/1.1/intl/html/intlspec.doc1.html>.

JavaSoft, 1999b

JavaSoft (1997b). Java server documentation. <http://jserv.javasoft.com/products/java-server/documentation/index.html>.

JavaSoft, 1999c

JavaSoft (1997c). Java servlets. <http://java.sun.com/products/servlet/index.html>.

Kazakos et al., 1999

W. Kazakos, R. Kramer, R. Nikolai, C. Rolker. "WebCDS - A Java-based Catalogue System for European Environmental Data", Current Trends in Database Technology, Editors A. Dogac, M.T. Özsu, O. Ulusoy, Idea Group Publishing, Hershey USA, London UK, 1999.

Kramer et al., 1996

Kramer, R., Nikolai, R., and Rolker, C. (1996). World Wide Web Access to CDS: Software Design and Demonstration of First Prototype. In *Proc. 3rd Workshop on Catalogue of Data Sources (CDS) and Thesaurus*, pages 81-91, Copenhagen, Denmark. European Environment Agency, ETC/CDS.

Kramer et al., 1997b

Kramer, R., Nikolai, R., Rolker, C., Bjarnason, S., and Jensen, S. (1997b). Interoperability Issues of the European Catalogue of Data Sources (CDS). In *Proceedings of the 2nd IEEE Metadata Conference*.

<http://computer.org/conferen/proceed/meta97/papers/rkramer/rkramer.html>.

Nikolai et al., 1999-11-16

Ralf Nikolai, Ralf Kramer, Marc Steinhaus, Bruno Felluga, and Paolo Plini. Genthesis: A general thesaurus browser for web-based catalogue systems. In Proc.IEEE Meta-Data'99, Bethesda, Maryland, USA, April 1999. Available at <http://computer.org/proceedings/meta/1999/papers/49/rnikolai.html>.

Wersig, 1985

Wersig, G. (1985). *Thesaurus-Leitfaden*. K.G.Saur Verlag KG, München.