

SUPERTHES

# SUPERTHES

Enhanced Software Tools for  
Construction, Maintenance  
and Web Visualisation of  
Multilingual Thesauri

umweltbundesamt<sup>U</sup>



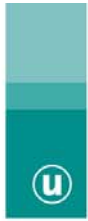
*Consiglio Nazionale delle Ricerche*

Umwelt  
Bundes  
Amt   
Für Mensch und Umwelt



ASO BRNO  
AUSTRIAN SCIENCE AND RESEARCH  
LIAISON OFFICE BRNO

umweltbundesamt<sup>U</sup>  
[www.umweltbundesamt.at](http://www.umweltbundesamt.at)



SUPERTHES

# Construction, Maintenance and Visualisation of Multilingual Thesauri “SuperThes”

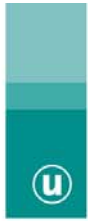
SuperThes Web Visualizer

Wolf–Dieter Batschi  
Rudolf Legat  
Paolo Plini  
Hermann Stallbaumer

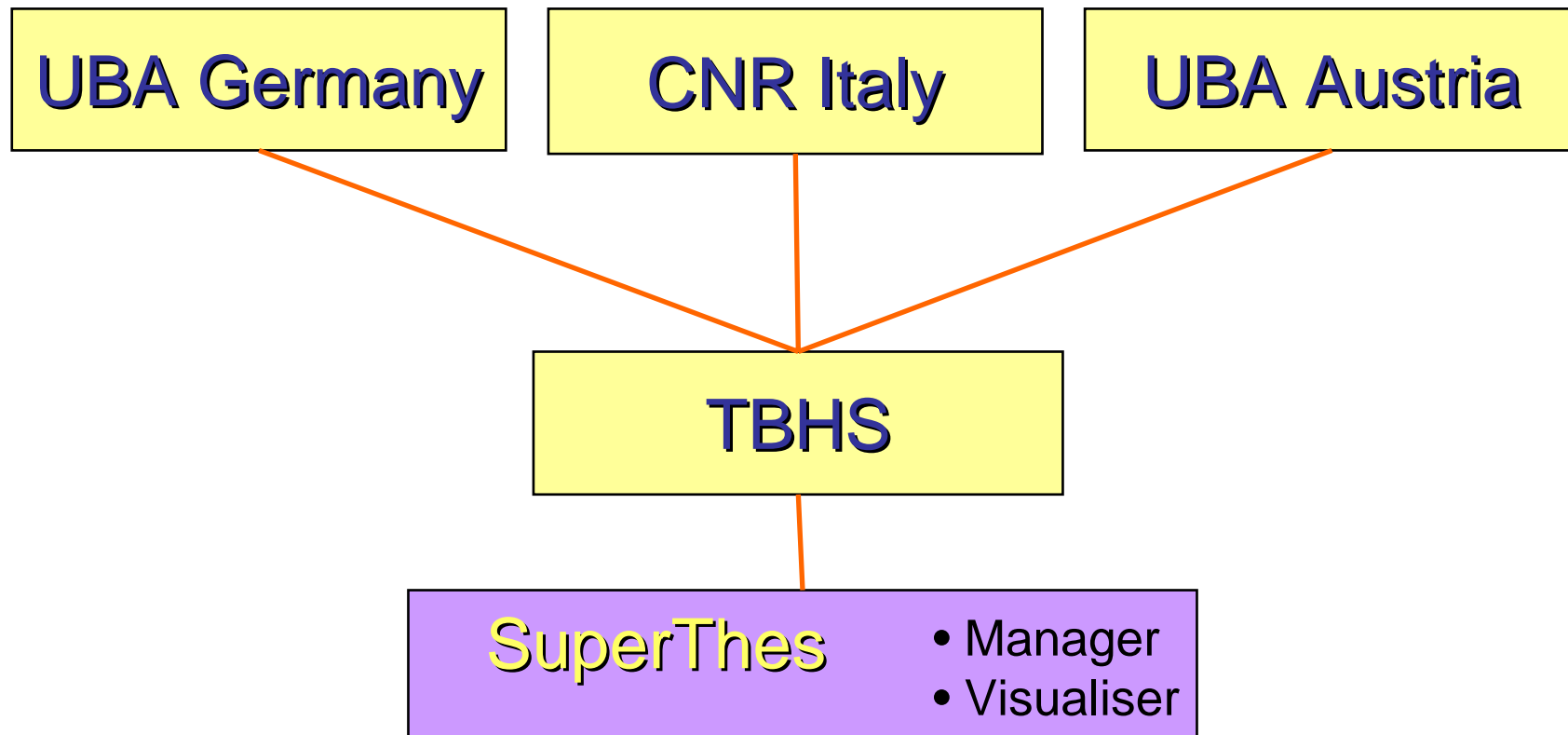


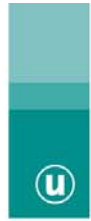
ASO BRNO  
AUSTRIAN SCIENCE AND RESEARCH  
LIAISON OFFICE BRNO

umweltbundesamt<sup>U</sup>  
[www.umweltbundesamt.at](http://www.umweltbundesamt.at)



## Memorandum of Understanding (2000)



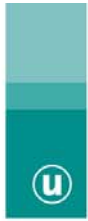


# Requirements



1. Contemporary thesaurus data are targeted towards a world-wide audience requiring Unicode capabilities
2. Thesaurus constructors demand flexible thesaurus structures, the freedom to create thesauri from scratch, using various data types and tables
3. Convenient data exchange with standard office applications using state-of-the-art technologies like drag and drop
4. Data storage based on proven client server technology
5. Powerful report generator
6. Bulk data extraction using either standard file formats or XML





# Language Definition

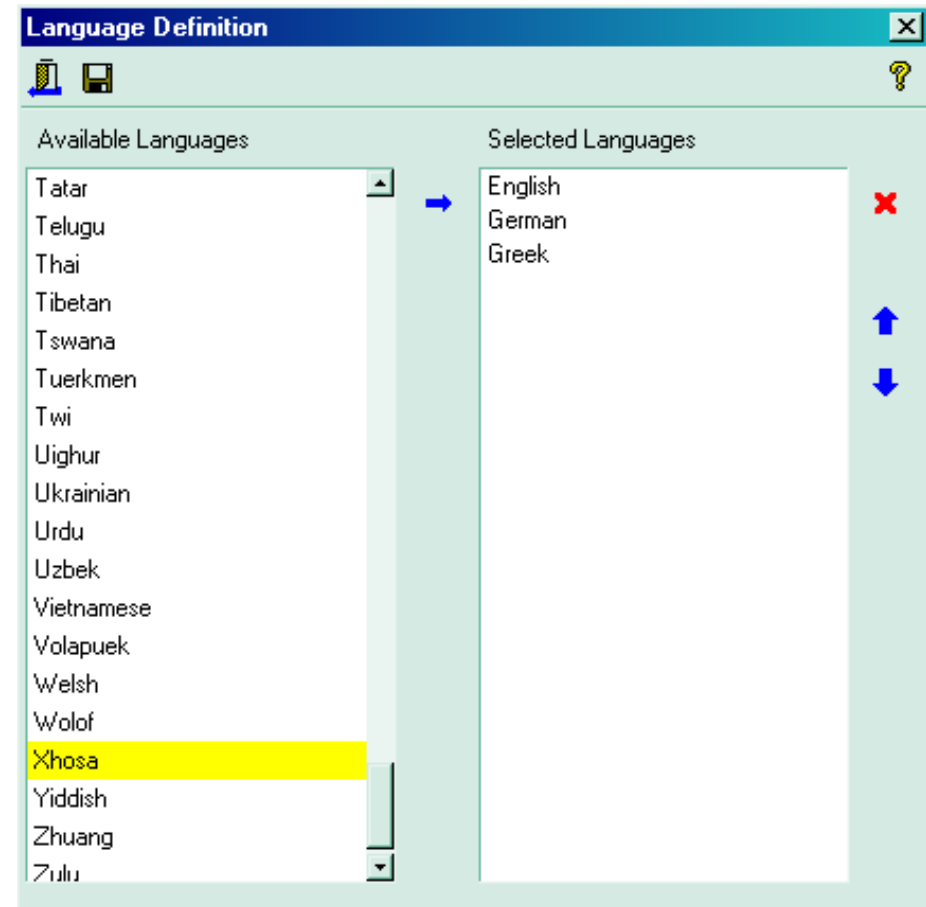
SUPERThES

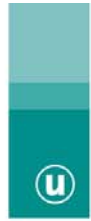
Contemporary thesaurus data are targeted towards a worldwide audience requiring true Unicode capabilities

SuperThes is Unicode compliant and stores all data in UCS-2 formats

All languages defined in ISO 639-1 are pre-defined

The availability of a particular language depends on the host systems capabilities like installed fonts, input method editors and right to left facilities





# SuperThes background: THESmain

SUPERTHES

## Origins:

THESmain was created 1995 due to the lack of commercial software with sufficient multilingual capabilities

## Design goals were:

- Separation of tasks “thesaurus maintenance” and “thesaurus visualisation”
- Graphical user interface
- Compliance to DIN 1462 /ISO 2788 and DIN /ISO 5964
- Must cover all spoken European languages
- Data storage in standard database format
- Standardized data exchange via SGML

Wide spread use: UDK–Thesaurus, GEMET, ENVOC, AWG, Forstglossar, Scandinavian Health Thesaurus, Ethik–Thesaurus





# Thesaurus Definition



Thesaurus constructors demand flexible thesaurus structures (freedom to create thesaurus fields, data types, relations, additional tables, due to the properties of the supplied data and design goals of the thesaurus)

The screenshot shows a 'Table Definition' window with two main sections. The top section, titled 'Available Tables', contains a table with two columns: 'Tablename' and 'Description'. The 'TestA' row is highlighted in yellow. The bottom section, titled 'Field properties of selected table', contains a table with five columns: 'Fieldname', 'Datatype', 'Index', 'LDF', and 'Description'. The 'Term' row is highlighted in yellow.

Tablename	Description
THESAURUS	
TestA	

Fieldname	Datatype	Index	LDF	Description
ID	ID	Unique		
Sort	Numeric		<input type="checkbox"/>	A single field
Term	Date	Unique	<input checked="" type="checkbox"/>	new desc
Percentage	Scientific	Ascend	<input type="checkbox"/>	

Thesauri consisting of several table may be constructed from scratch using the table definition editor





# Ease of Use



A rich graphical user interface with drag and drop features and context menus allow quick and efficient data handling

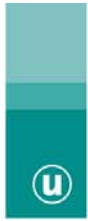
The screenshot displays the SuperThes application interface. The main window, titled "SuperThes - Thesaurus: SuperThes", features a menu bar (File, View, Windows, Tools, Reports, Help) and a toolbar with various icons. Below the toolbar is a "Tabular View" window containing a table of terms. The table has columns for ID, DescNr, Print, Term.English, Term.German, Term.Greek, and Term.Russian. The first few rows are visible, showing terms like "abandoned industrial site", "abandoned vehicle", and "abiotic factor".

ID	DescNr	Print	Term.English	Term.German	Term.Greek	Term.Russian
1	7	1	abandoned industrial site	Altstandort	εγκαταλειμμένος βιομηχαν...	Изо ставен индустриален обект
2	8	1	abandoned vehicle	Abgestelltes Fahrzeug	εγκαταλειμμένο όχημα	Изо ставено превозно средство
3	11	1	abiotic factor	Abiotischer Faktor	αβιοτικός παράγοντας	Абиотичен фактор
4	13	1	absorption (exposure)	Absorption (biologisch)		
5	20	1	acceptable risk level	Zulässiges Risikoniveau		
6	21	1	agreement (administrative)	Verwaltungsabkommen		
7	22	1	access road	Zufahrtsstraße		
8	23	1	access to culture	Zugang zur Kultur		
9	24	1	access to the sea	Meereszugang		
10	25	1	accident	Unfall		
11	28	1	accidental release of organis...	Unfallbedingtes Freiset		
12	30	1	accident source	Unfallsursache		
13	36	1	accumulation in body tissues	Anreicherung im Körper		
14	38	1	accumulator	Akkumulator		
15	42	1	acid deposition	Saure Deposition		
16	44	1	acidification	Versauerung		
17	47	1	acidity	Acidität		
18	48	1	acidity degree	Säuregrad		
19	51	1	acid rain	Saurer Regen		
20	52	1	acid	Säure		

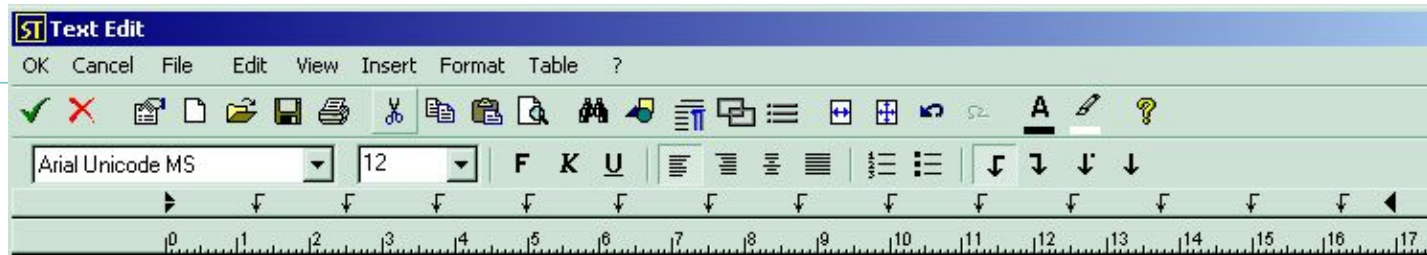
An inset window, also titled "Tabular View", shows a detailed view of terms. It has a toolbar and a search bar. The table in this window has columns for Term.English and Term.Greek. The term "environmental protection officer" is highlighted in yellow, with its Greek translation "(αξιωματούχος) αρμόδιος (επάλληλη)" visible next to it.



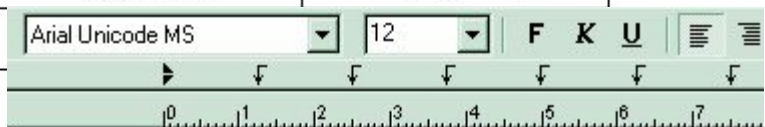




# Text Editors



Name	Text	Name	Text
Arabic	دو ك و ن و ي	Hindi	यूनिकोड
Armenian	Յունիկոդ	Japanese	ユニコード
Bengali	যুনিকোড		
Bopomofo	ㄅ ㄨ ㄣ ㄉ ㄛ ㄩ ㄨ ㄚ ㄛ		
Chinese	统一码		
Georgian	უნიკოდ		
Greek	Γιούνικοντ		



- Powerful word processor plug-in:
- supports tables and images
  - reads and writes RTF and HTML formats
  - reads and writes MS Words documents

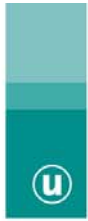
Chemical Formulas:

$$\begin{array}{c} \text{H} \\ | \\ \text{H}-\text{C}=\text{O} \end{array} \quad \text{H}_2\text{O} \quad \text{HCO}_3$$

$$\text{CrO}_4^{2-} \quad \text{Ba}(\text{NO}_3)_2$$

$$\text{Na}^+\text{Cl}^- \quad \text{Ba}^{2+} \quad \text{SO}_4^{2-}$$

$$\text{Na}_2\text{SO}_4 \quad \text{O} \cdot \quad \cdot \text{O}$$

# Multimedia Editors

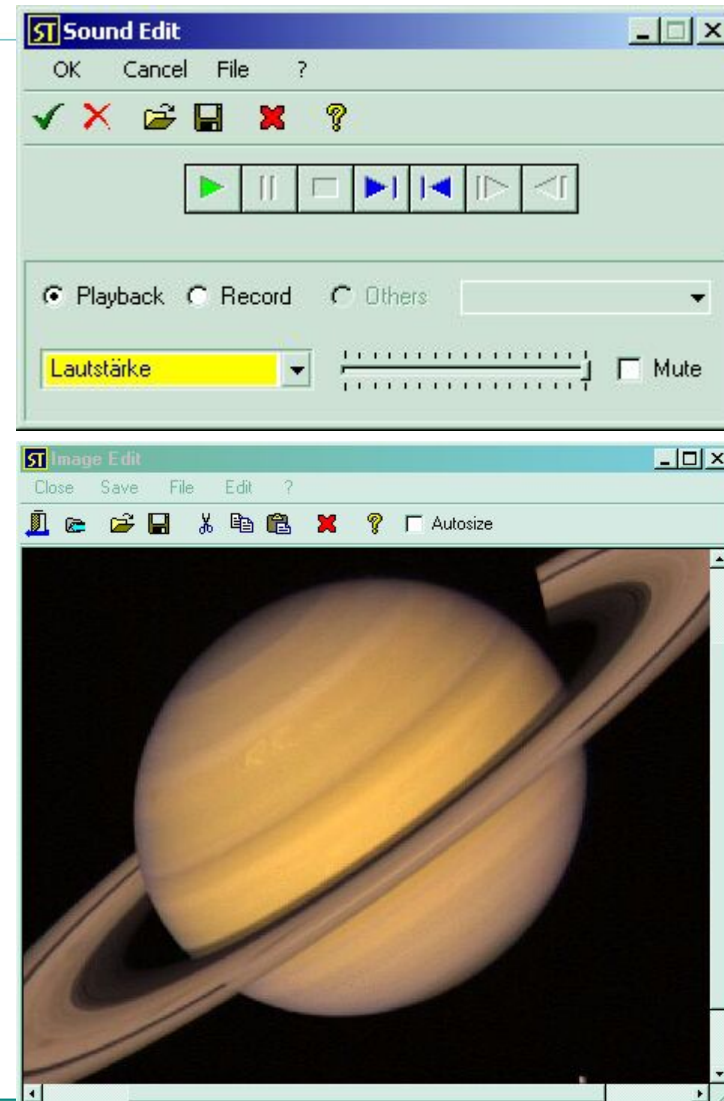


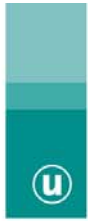
Multimedia editors for sounds and images:

- Supports various file formats (jpg, bmp, ico, emf, wmf)
- Data exchange with other applications via files, clipboard and drag and drop

SuperThes supports a wide range of additional data types:

- boolean, date, time, decimal, list, memo, numeric and scientific



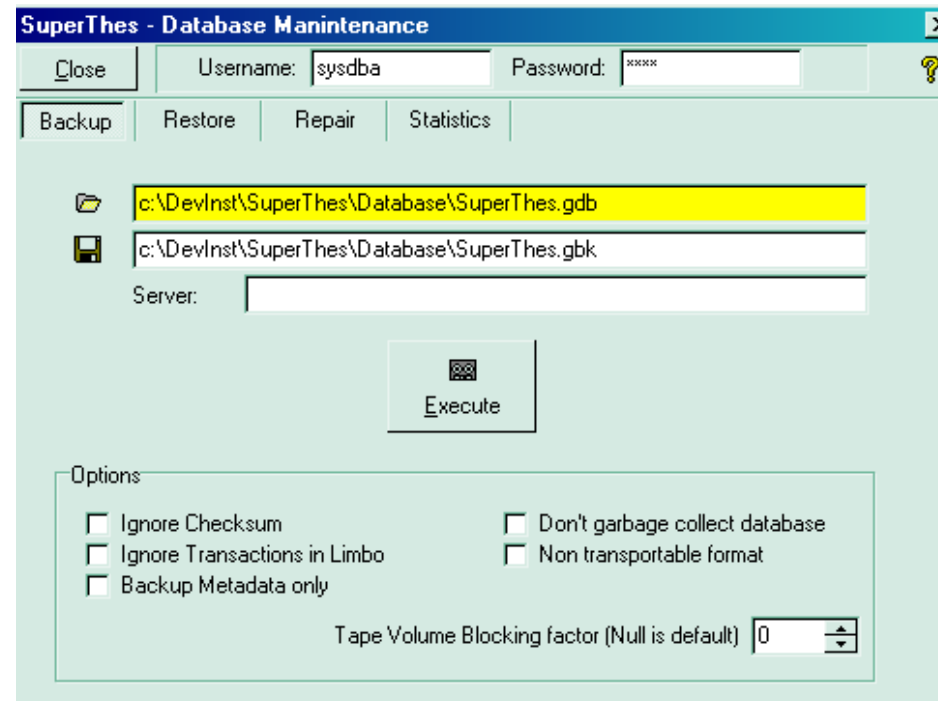


# Database Technologies



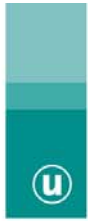
SuperThes relies on robust client server technology

For small installations, client and server may reside on the same computer



For those who do not want to learn the native database management environment, a comprehensive maintenance and backup utility is included





# SuperThes Web Visualizer

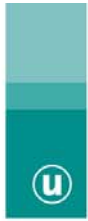
SUPERTHES



## Current developments:

The present work programme comprises a web-based tool called SuperThesVIZ which allows access to the SuperThes databases via the Internet. The main goal for the software development is to ensure the convenience of the user interface for the MS Windows-based application. SuperThesVIZ is platform independent, based on Java servlet technology.





# SuperThes Web Visualizer

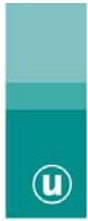


## Requirements for the SuperThes visualisation module:

Requirement that is different from standard web application

- Not the developer, but the user determines the thesaurus structure
- The visualisation module must therefore:
  - adapt to different database structures
  - be easy to configure





# SuperThes Web Visualizer

SUPERTHES

## Design objectives:

- System independence
- Compliance with standards
- Use of proven technologies
- Easy to configure
- Combination of several technologies
- Must be suitable for use by non programmers

it must be a type of self-configuring building block system

Configur

Home  
Back  
Contact  
Legal Notes

### About us

#### About the program

SuperThes VIZ is a client side agent used with SuperThes thesaurus construction package

SuperThes Visualizer was tested with Firefox

#### Origins

In late 2000, a Memorandum of Understanding was signed between:

- [EKOLab](#),
- [Umweltbundesamt Berlin](#),
- [Umweltbundesamt Vienna](#) and
- [Technisches Bureau Hermann Stallt](#)

The MoU was renewed for the period 2004-2006 and maintaining multi-lingual, poly-hierarchical thesaurus construction and maintenance packages.

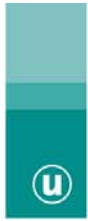
#### Features of SuperThes

SuperThes is used for visualization and multi-lingual, poly-hierarchical thesaurus construction and maintenance. SuperThes runs on Windows XP as an operating system. All thesauri managed by SuperThes as well as client computers resources. Besides the main thesaurus may be used to describe the main Thesaurus SuperThes supports all hierarchical relations number of relations per thesaurus and per thesaurus.

A thesaurus may be constructed from scratch







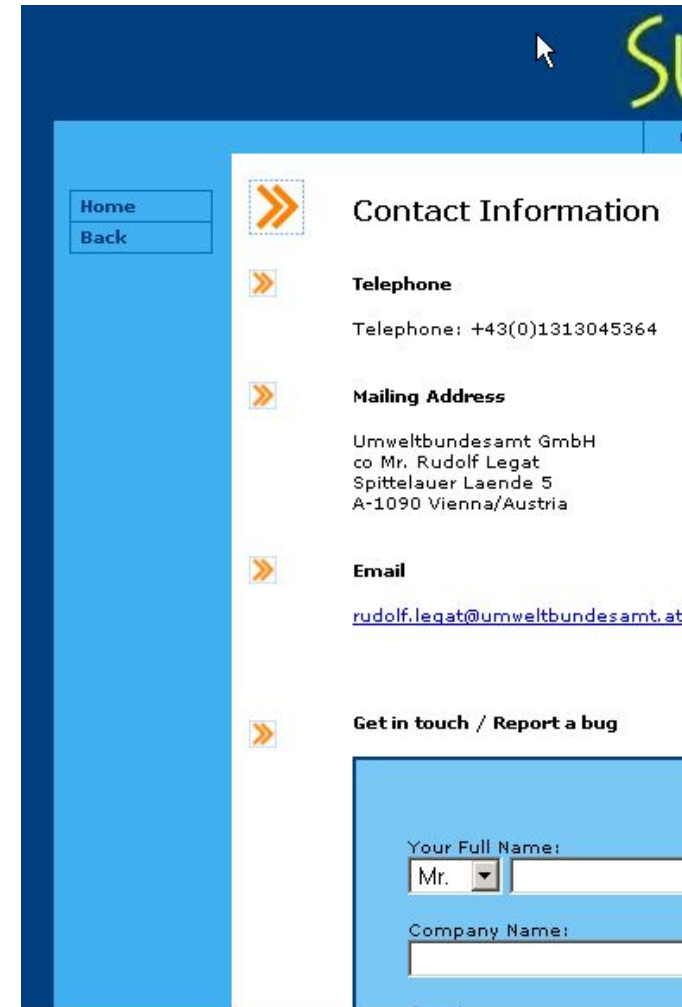
# SuperThes Web Visualizer

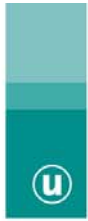
SUPERTHES

## Applied technologies:

**Important! Servers might be Windows or Unix/Linux based**

- XHTML 1.0
- Java Server Pages
- Java Servlets
- Firebird 1.5 (database)
- Runs on standard servlet containers, such as Tomcat 5.5





# SuperThes Web Visualizer

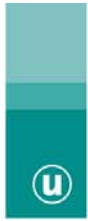
SUPERTHES

## Features:

- Simple way of deploying SuperThes databases on the Web.
- Runs on Windows computers as well as on Unix.
- Mixed environments possible (e.g Webserver Apache/Tomcat on Unix, Database on Windows 2003 Server)
- Supports user authentication via Database Server
- Supports https for secure working
- Feature set is determined automatically from attached database
- Thesaurus display is similar to THESshow, so users familiar with it will already know the user interface







# SuperThes Web Visualizer

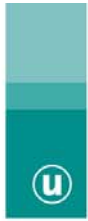
SUPERTHES

## Setup:

To deploy a SuperThes database on the web the following steps are required:

- Copy SuperThesVIZ.war into the web-apps directory of the desired servlet container
- Let the servlet engine expand the war-file
- Adapt settings in SuperThes.cfg and Mailer.ini (like location and name of database)
- Now the database should be accessible on the web





# SuperThes Web Visualizer

SUPERTHES

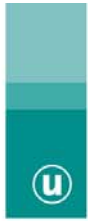
## Configuration of thesaurus presentation

- Setup is done by filling out a configuration file
- Database structure, available tables, fields, languages are extracted by the software automatically
- User related information (contact, terms of use, presentation of user organisation) is kept in simple html files.
- Html files are included into the JSP's at runtime.
- Layout of web pages is controlled from a central CSS-file, so adapting the appearance is an easy task.

## Example Configuration File

```
#location and name of database driver
db_drivename =
org.firebirdsql.jdbc.FB
#URL of web-server
db_serverurl = localhost
#location of SuperThes database
db_path =
c:/Programme/SuperThes/Dat
#name of SuperThes database
db_name = stsample.fdb
#URL of mail host
smtp.host = smtp.eunet.at
#client information to include in mail
valid.env = REMOTE_HOST,
REMOTE_ADDR, REMOTE_USER,
#list of allowed referrers
valid.referrers =
```





# SuperThes Web Visualizer

SUPERTHES

## Technical Requirements:

### Clientside:

- XHTML 1.0 compliant browser (IE, Netscape Firefox, Opera)

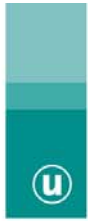
### Web-Server:

- J2EE container (compliant to Servlet 2.4 specs, JSP 2.0 specs)
- jdbc driver (Firebird XA-compliant JDBC driver Version 1.5)
- optional access to smtp mail service (for form mailer)

### Database-Server:

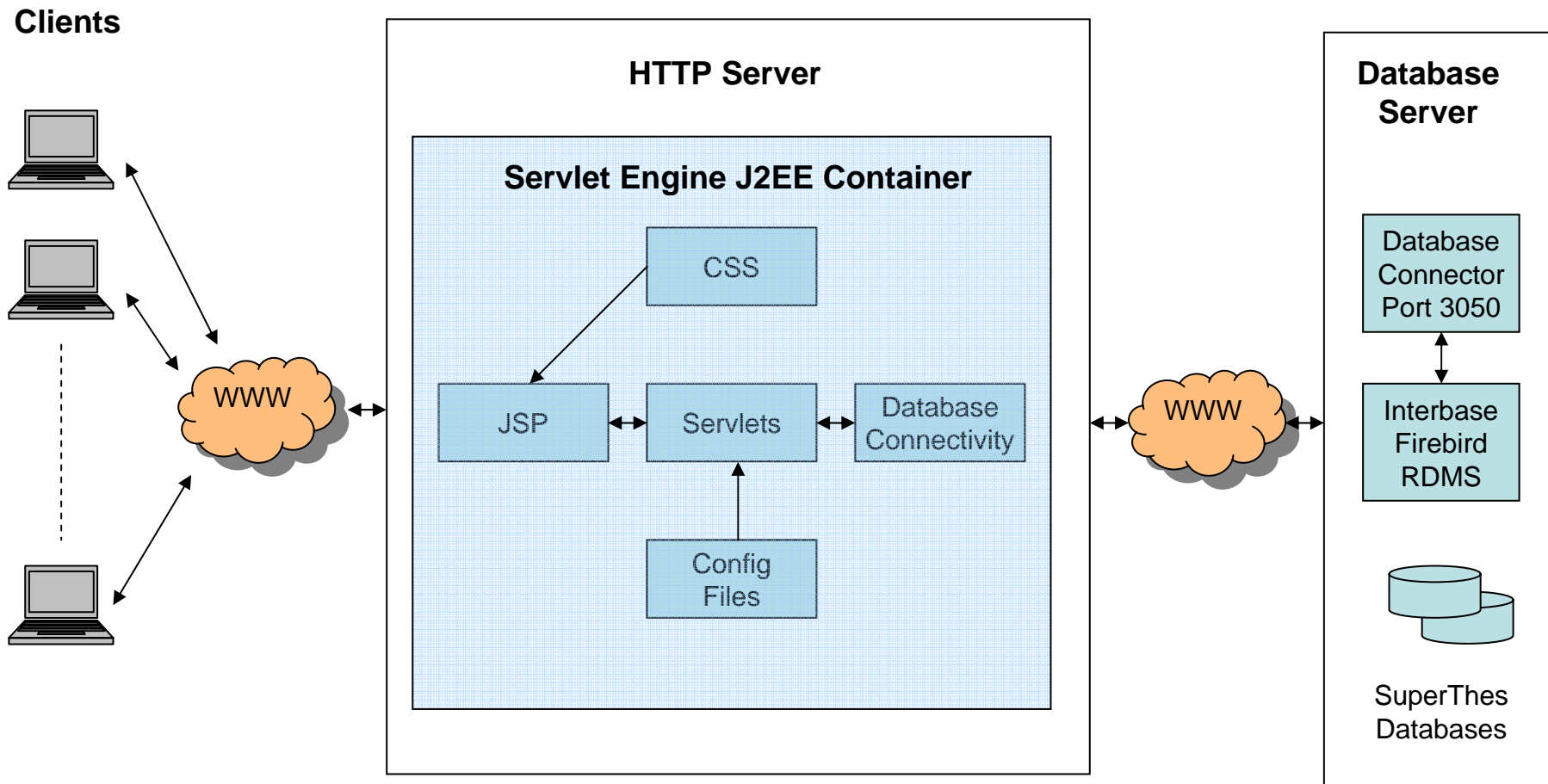
- Firebird Database engine (Version 1.5) (All Interbase engines newer than 6.0.1 should work also)

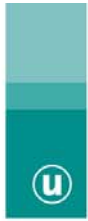




# System Configuration

SUPERTHES





# SuperThes Web Visualizer



## Configuration page

**Configuration** | Thesaurus | Support

Home  
Back

Table Selector: THESAURUS

Field Selector: Term.German

Detail Fields:  
ID  
Status  
Guide  
Term.German  
Term.English  
Term.Greek  
Definition.German  
Definition.English  
Print  
GemID

Clear [X]

Display:  
 Hierarchical  
 Tabular

Relations:  
Topterm  
Broader Term  
Narrower Term  
Related Term  
Related Term Antonym  
Nondescriptor  
Nondescriptor QS  
Nondescriptor Reg  
Nondescriptor Irreg  
Print

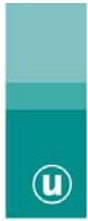
Clear [X]

Secondary Table: Themes

Detail Fields:  
ID  
Abbrev  
Sort  
Theme.German  
Theme.English  
Theme.Greek  
Image  
Document

Clear [X]





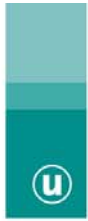
# SuperThes Web Visualizer

SUPERTHES

## Presentation page

The screenshot displays the SuperThes Web Visualizer interface. At the top, the title 'SUPERTHES V12' is shown in a stylized font. Below the title is a navigation bar with tabs for 'Configuration', 'Thesaurus', and 'Support'. A search bar with a 'Search' button is highlighted with an orange border. On the left side, there are 'Home' and 'Back' buttons. The main content area is divided into two columns. The left column lists terms under various headings: 'Abdichtung', 'Abdichtungssystem', 'Abdichtungsverfahren', 'Aberntung', 'Abfackelung', 'Abfälle (überwachungsbedürftig)', 'Abfälle (Vermeidung und Verwertung)', 'Abfälle zur Beseitigung', 'Abfälle zur Verwertung', 'Abfälle, geordnet abgelagert', 'Abfälle, ungeordnet abgelagert', and '[Abfall]'. The right column shows hierarchical relationships: 'Top Term' (Schadstoffe und Abfälle, Umweltverschmutzung), 'Broader Term' (Schadstoffe und Abfälle, Umweltverschmutzung), and 'Narrower Terms' (Abfallart, Abfallaufkommen, Abfallbehörde, Abfallbeschaffenheit, Abfallbeseitigung, Abfallerzeuger, Abfallminderung, Abfallrecht, Abfallwirtschaft, Altstoff (Abfall), Überlassungspflichtiger Abfall).





# The Team



## **Wolf-Dieter Batschi**

Umweltbundesamt Dessau, Wörlitzer Platz 1, D-06844 Dessau

email: [wolf-dieter.batschi@uba.de](mailto:wolf-dieter.batschi@uba.de)

Internet: [www.umweltbundesamt.de](http://www.umweltbundesamt.de)

## **Rudolf Legat**

Umweltbundesamt Wien, Spittelauer Lände 5, A-1090 Wien

email: [rudolf.legat@umweltbundesamt.at](mailto:rudolf.legat@umweltbundesamt.at)

Internet: [www.umweltbundesamt.at](http://www.umweltbundesamt.at)

## **Paolo Plini**

CNR – IIA – Environmental Knowledge Organisation Laboratory

Via Salaria Km 29,300 – CP 10, I-00016 Monterotondo Stazione, Rome – Italy,

email: [plini@iia.cnr.it](mailto:plini@iia.cnr.it)

Internet: [uta.iia.cnr.it/](http://uta.iia.cnr.it/)

## **Hermann Stallbaumer**

Fa. TBHS, Favoritenstraße 182, A-1100 Wien

email: [hermann@tbhs.co.at](mailto:hermann@tbhs.co.at)

