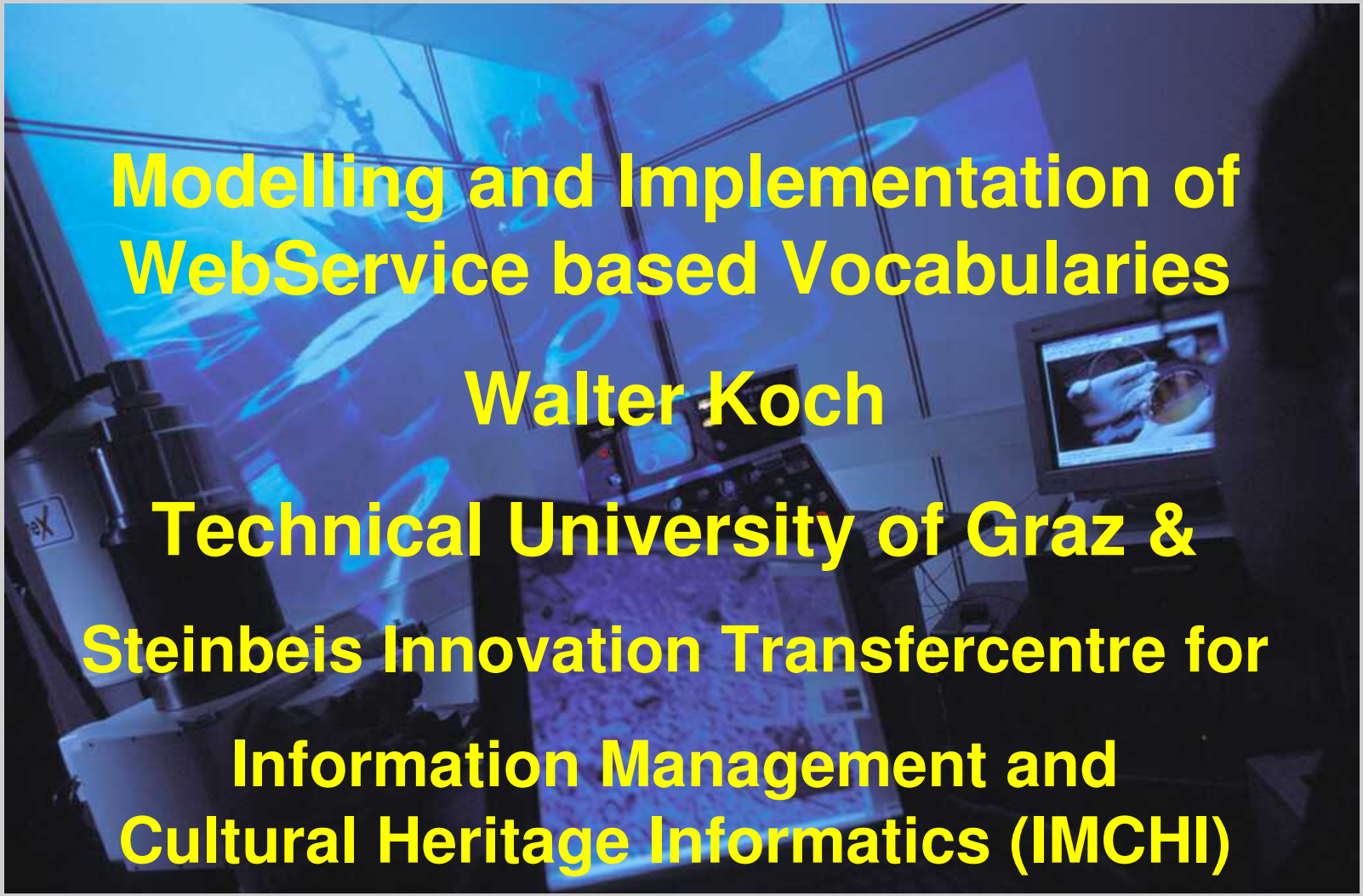
STEINBEIS-STIFTUNG



Modelling and Implementation of WebService based Vocabularies

Walter Koch

Technical University of Graz &

Steinbeis Innovation Transfercentre for

Information Management and Cultural Heritage Informatics (IMCHI)

Wir machen Innovationen schneller

Steinbeis Foundation

Service Areas

Consultancy, Research & Development, Expert Assessments,
International Technology Transfer, Education, Steinbeis University Berlin

Network

765 Steinbeis enterprises formed in 2008 the „**Steinbeis Network**“

Turnover 2008

124 Million Euro

Employees (2008)

1.383 Staff Members, **3.338** Freelancers, **801** Professors

Headquarter

Stuttgart, Germany

“Modelling and Implementation of WebService Based Vocabularies“

Content:

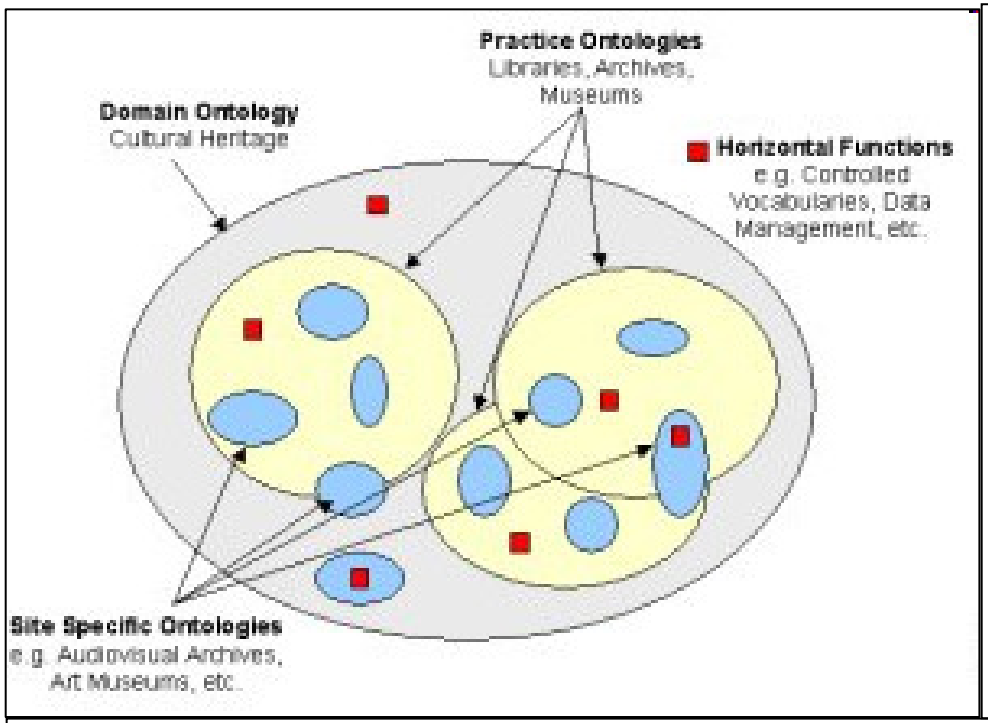
- Methods and Technologies
- Implementation Issues
- Workflows
- Business Model
- Examples

SOOM[1)] - Reference Implementation: The Vocserv[2)]-System

1) Service Oriented Object Management:
   http://scrum.ait.co.at/projects/soom/
2) Vocabulary Services:
   http://scrum.ait.co.at/projects/soom/wiki/Vocabulary-Issues

Methods and Technologies

- Ontology Engineering: IDEF5 (Integrated Definition 5)
  http://www.idef.com/IDEF5.html



The **Domain Ontology** "Cultural Heritage" covers different sections, namely the **Practice Ontologies** "Libraries", "Archives" and "Museums". These are the application areas within the Domain Ontology.

Site Specific Ontologies at the lowest level of generality are special forms of Practice Ontologies. For the Practice Ontology "Archives" this for example could be Audiovisual Archives.

Vocabulary WebServices are supporting business functions and can be domain, practice, or site specific or domain independent.

Methods and Technologies

- Construction of Controlled Vocabularies

B.1. Properties of Lists, Synonym rings, Taxonomies, and Thesauri

Table 3: Controlled vocabulary properties

Property	List	Synonym Ring	Taxonomy	Thesaurus
Types of Terms				
Preferred terms	Yes	No	Yes	Yes
Entry terms	No	Yes	No	Yes
Candidate terms	No	No	No	Optional
Provisional terms	No	No	No	Optional
Deleted terms	No	No	No	Optional
Relationships	No	Yes	Yes	Yes
Equivalence		Yes	No	Yes
Hierarchy		No	Yes	Yes
Part/Whole		No	Yes	Yes
IsA		No	Yes	Yes
HasA		No	Yes	Yes
Classification		No	Optional	Optional
Related terms		No	No	Yes
Facet		No	No	Optional
Notes	No	No	Optional	Optional
Scope note			No	Optional
History note			No	Optional
Other notes			No	Optional

ANSI/NISO Z39.19-2005
Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies

Vocabulary independent (metadata)

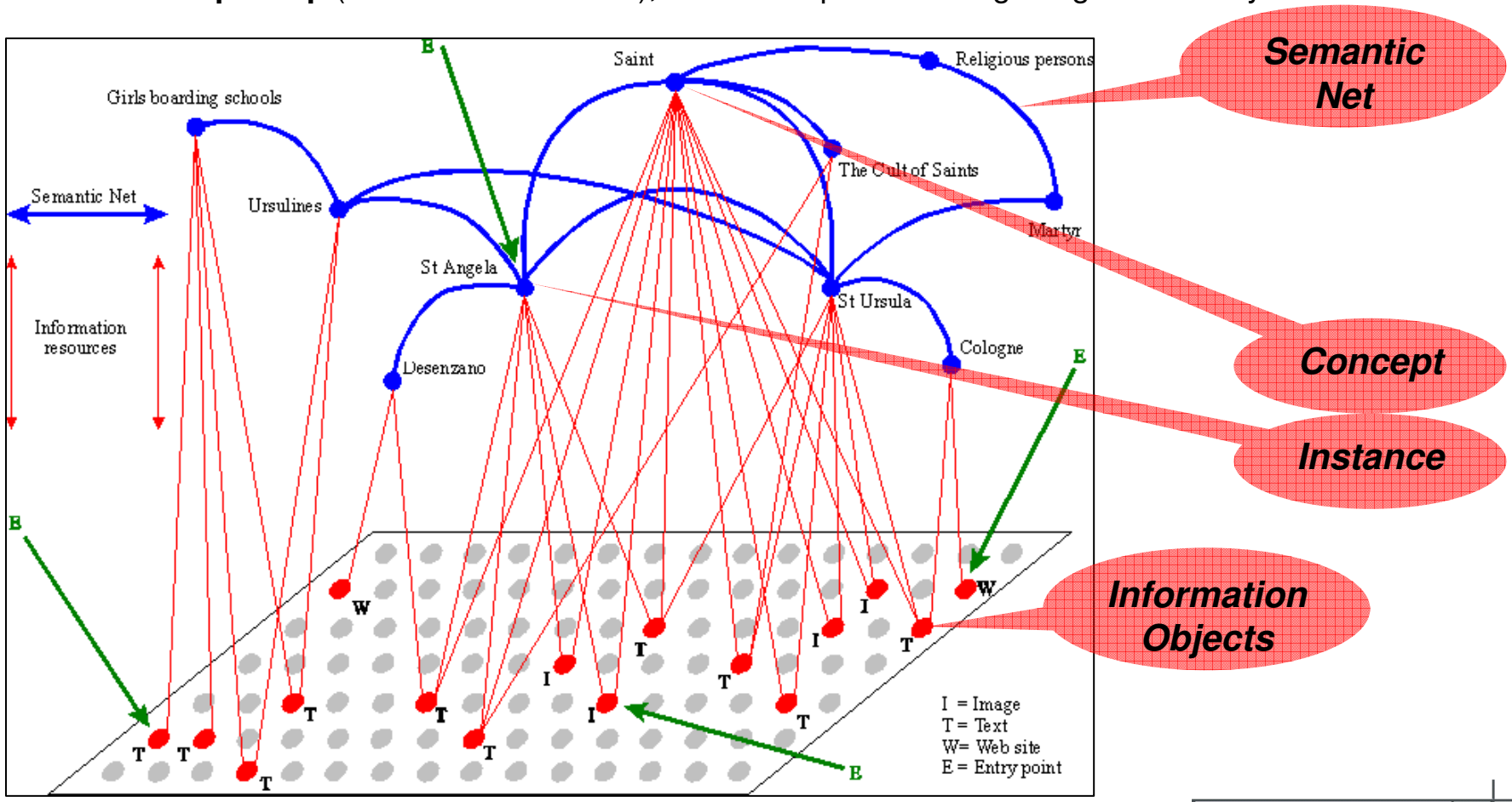
Vocabulary independent (structural elements)

Vocabulary dependent (metadata)

Methods and Technologies

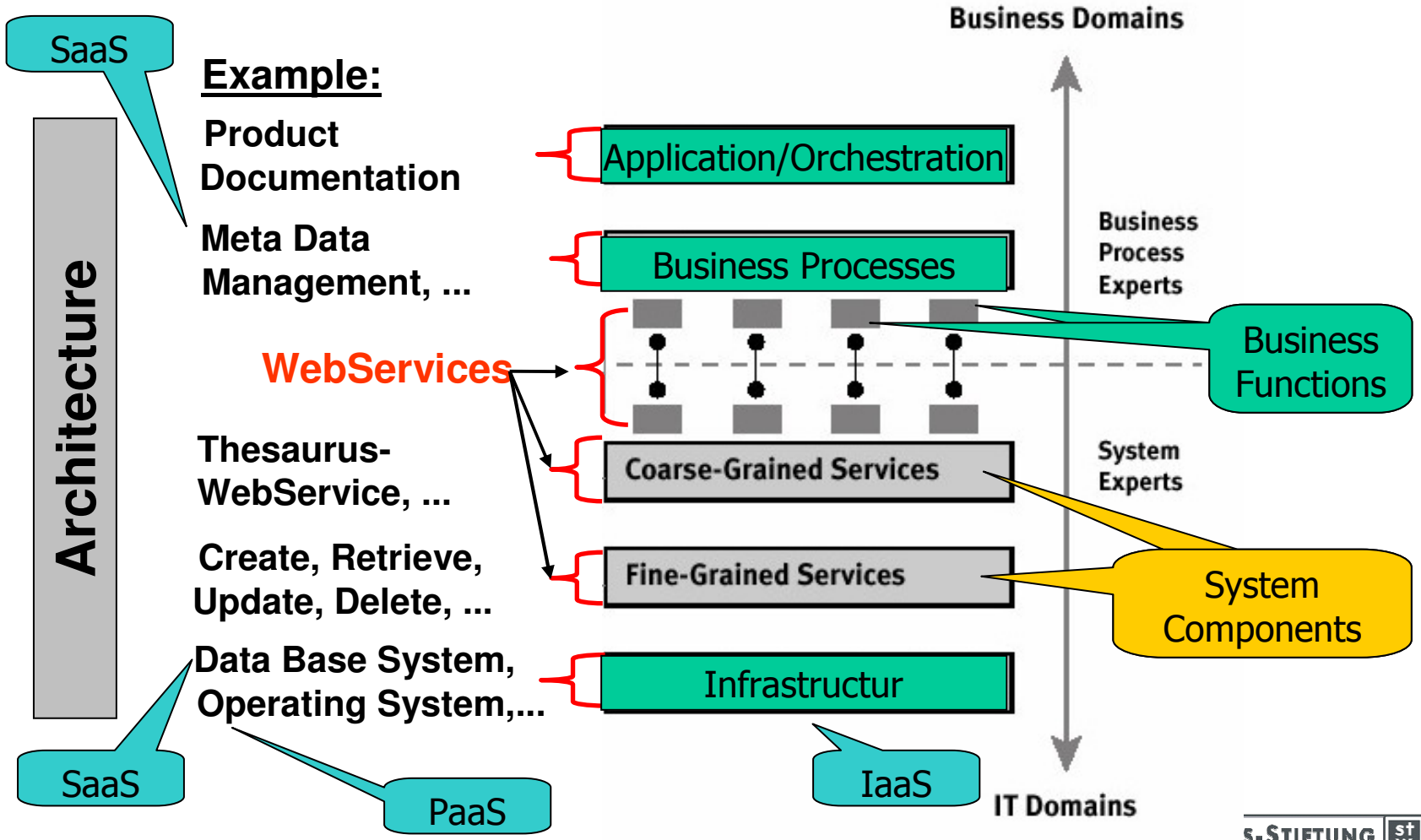
Semantic Technologies

XTM-Topicmap (ISO/IEC 13250:2000), **SKOS** Simple Knowledge Organisation System



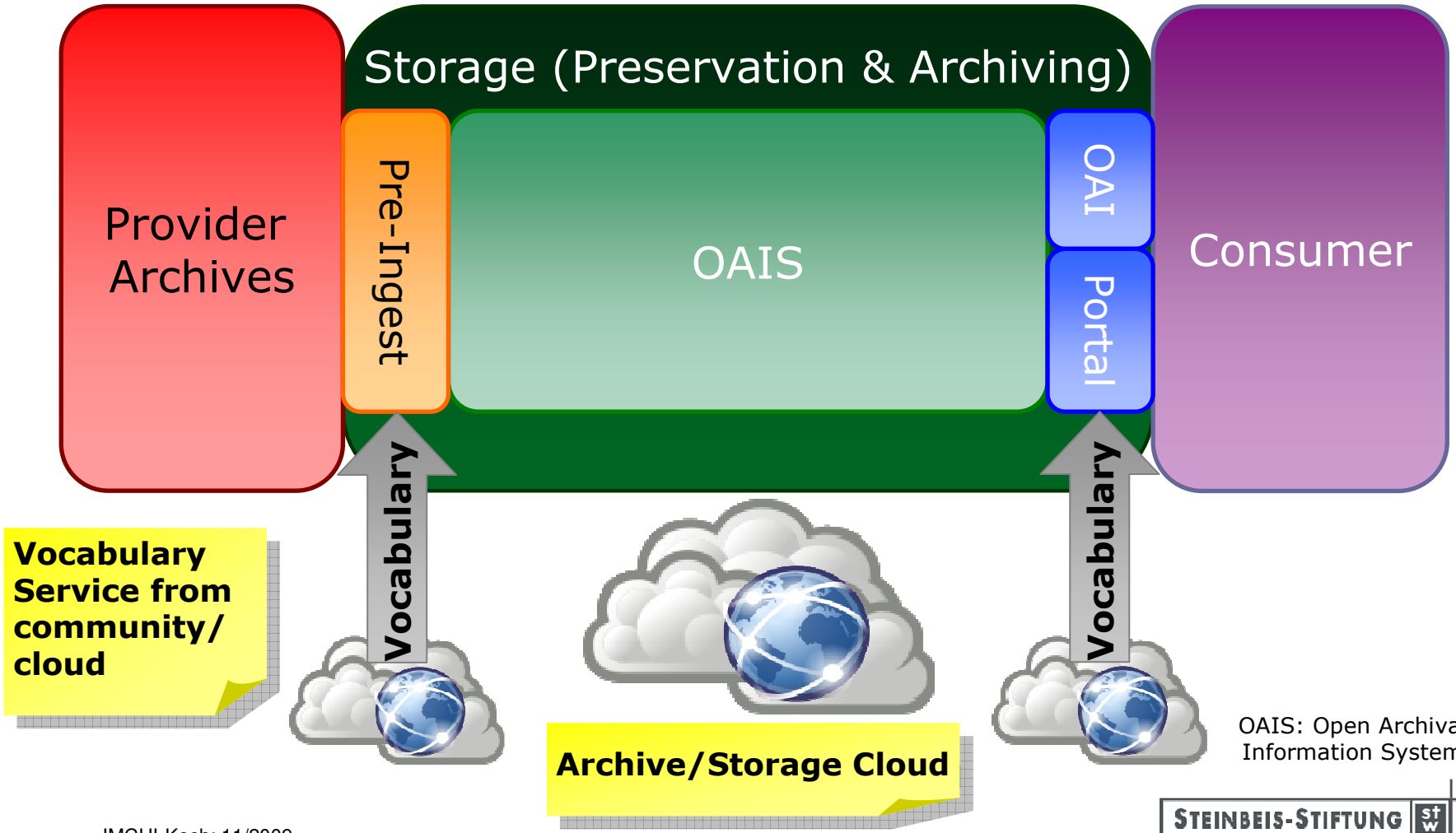
Implementation Issues

- Service Orientied Architecture (SOA) - WebServices



Implementation Issues

- Cloud Computing/Storage



OAIS: Open Archival Information System

Implementation Issues

- Data Structure
  Follows ANSI/NISO Z39.19-2005 Standard with extensions to provide multilinguality
- Structur of a Node

<Thesaurus>	(1)	Name of Controlled Vocabulary
ID	(1)	Unique Identifier (Numeric)
LANG	(1)	Language
TRM	(1)	Name of Term
BT	(0-n)	Broader Term (Parent Node)
NT	(0-n)	Narrower Term (Child Node)
USE	(0-n)	Use
UF	(0-n)	Use for
RT	(0-n)	Related Term
GS	(1)	Generic Structure (Path)
GS_LEVEL	(1)	Depth within GS (Sub Path)
NOTATION	(0-1)	Alternate Unique ID (Text/Label)
SCOPE	(0-1)	Container for Meta Data (vocabulary dependent)
<ATTR.>	(1-n)	Metadata Element

Implementation Issues

- AIT Vocabulary Service Description V1 (API)

  http://demo.ait.co.at/thesaurus/AIT Vocabulary Service Description V1.pdf

  http://demo.ait.co.at/thesaurus/Thesauri.wsdl

wsdl:message (30 rows)

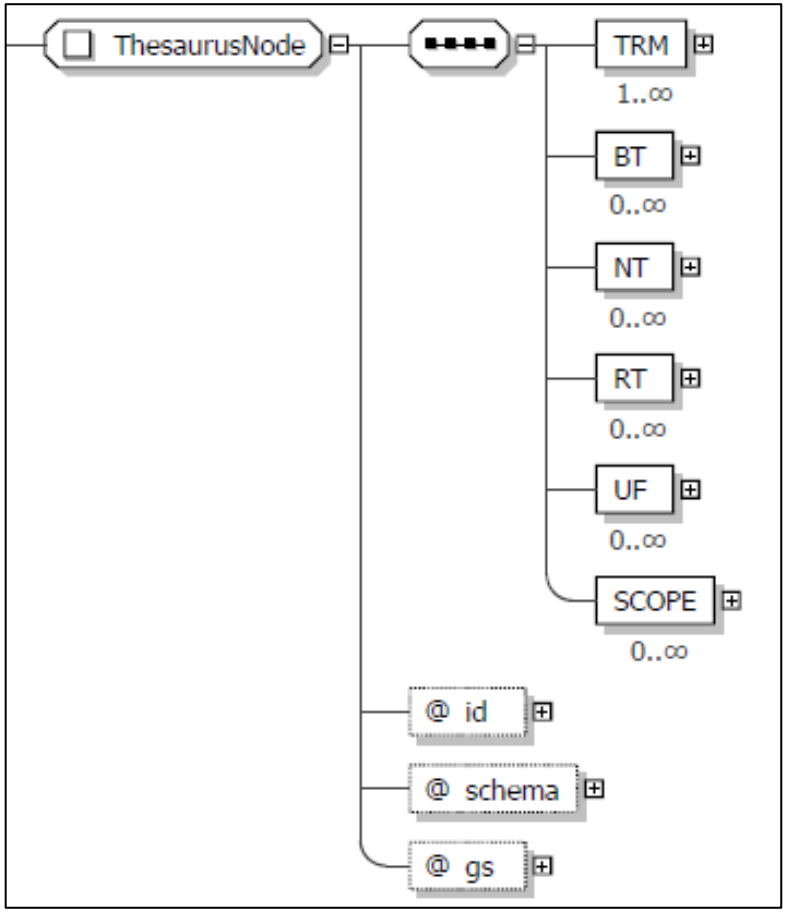
	@name
1	getAllThesauriRequest
2	getAllThesauriResponse
3	getLanguagesRequest
4	getLanguagesResponse
5	queryRequest
6	queryResponse
7	getNodesRequest
8	getNodesResponse
9	getMetaAttributeRequest
10	getMetaAttributeResponse
11	loginRequest
12	loginResponse
13	logoutRequest
14	logoutResponse
15	createNodeRequest

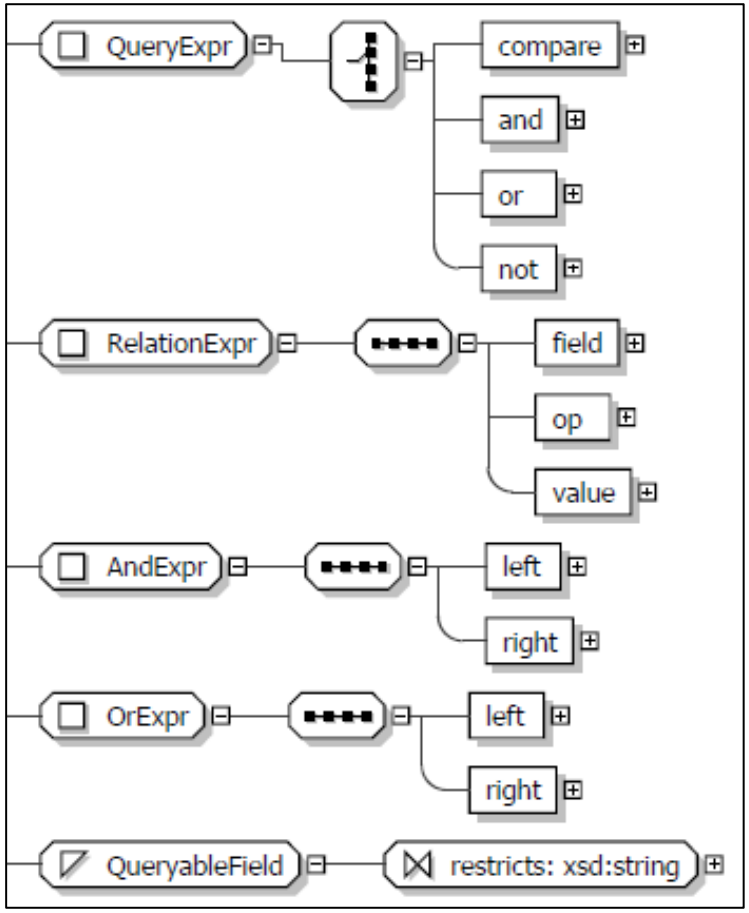
xsd:element (14 r...

	@name	xsd:complexType
1	getLanguages	xsd:complexType
2	getLanguagesResponse	xsd:complexType
3	query	xsd:complexType
4	queryResponse	xsd:complexType
5	getNodes	xsd:complexType
6	getMetaAttribute	xsd:complexType
7	getMetaAttributeResponse	xsd:complexType
8	login	xsd:complexType
9	loginResponse	xsd:complexType
10	logout	xsd:complexType
11	logoutResponse	xsd:complexType
12	thesaurusFault	xsd:complexType
13	query_fault1	xsd:complexType
14	getLanguages_fault	xsd:complexType

Implementation Issues

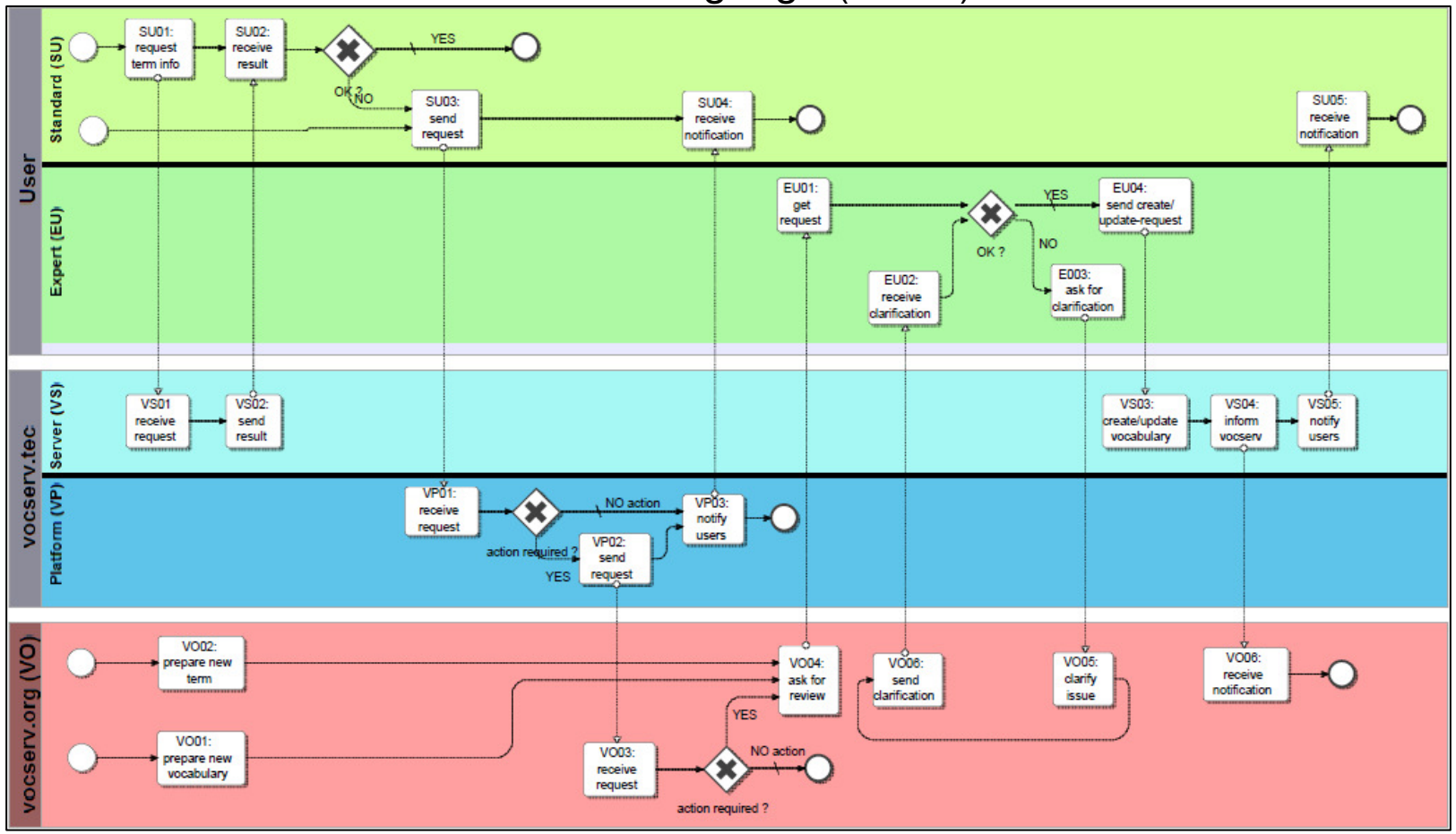
AIT Vocabulary Service - Schema





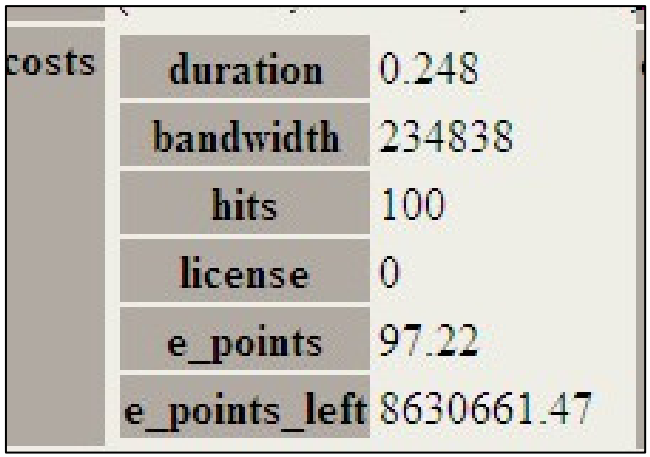
Workflows

- Business Process Modelling Notation (BPMN)
- Business Process Execution Language (BPEL)



Business Model

- As WebService (Software as a Service): Pay Per View (ePoints):

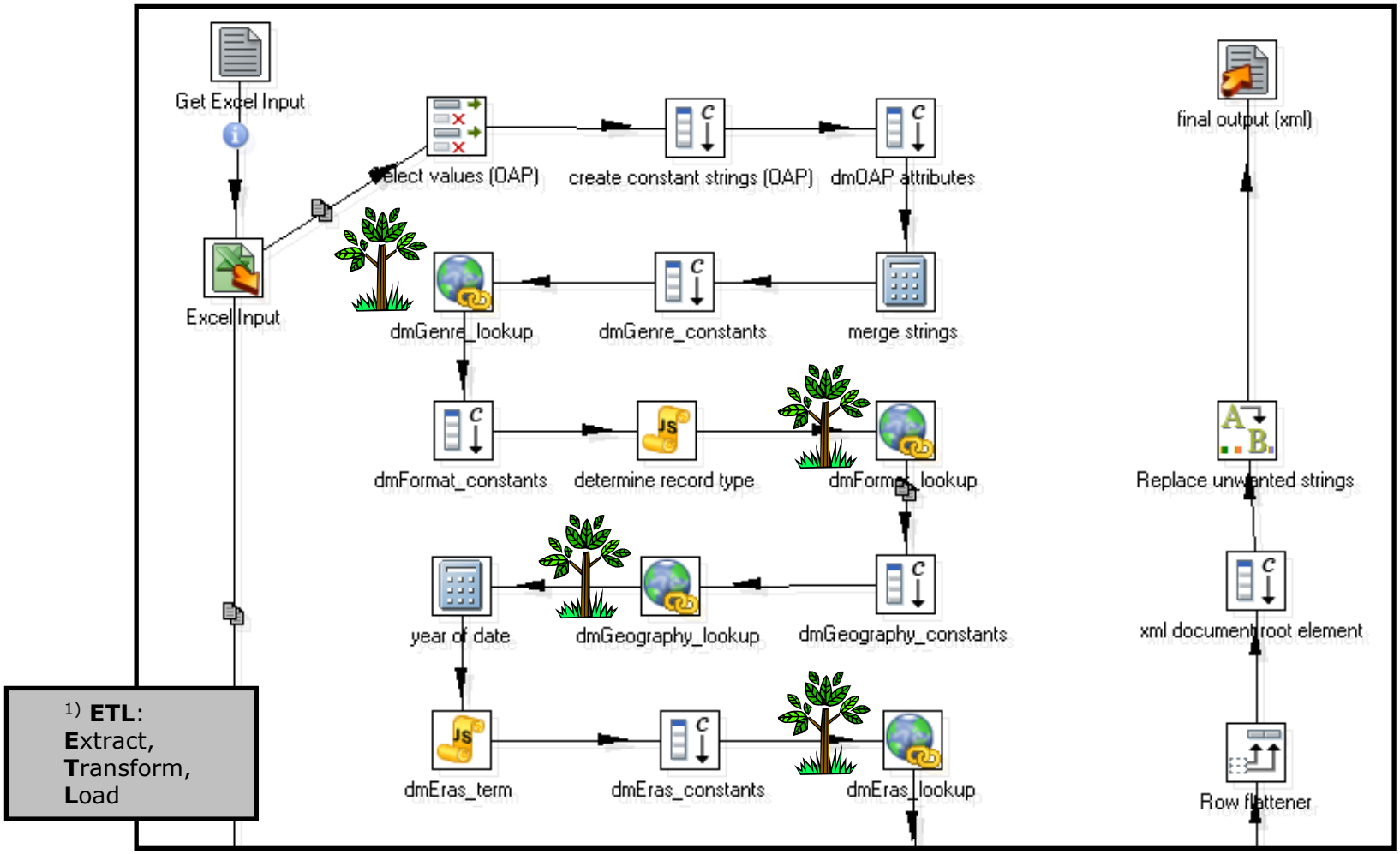


- Virtual Machine Image (TGN[1)]-example: 16GB including database and all system components)
- Infrastructure as a Service (integrated in a Virtual Machine in a Cloud)

[1)] Thesaurus of Geographic Names (2008 ~ 1 mio place names)

Example: Data Aggregation (semantic enrichment)

- Usage of ETL[1)]-Technology: DISMARC case (http://www.dismarc.eu)



[1)] **ETL**:
Extract,
Transform,
Load

Example: Page Level Indexing (1)

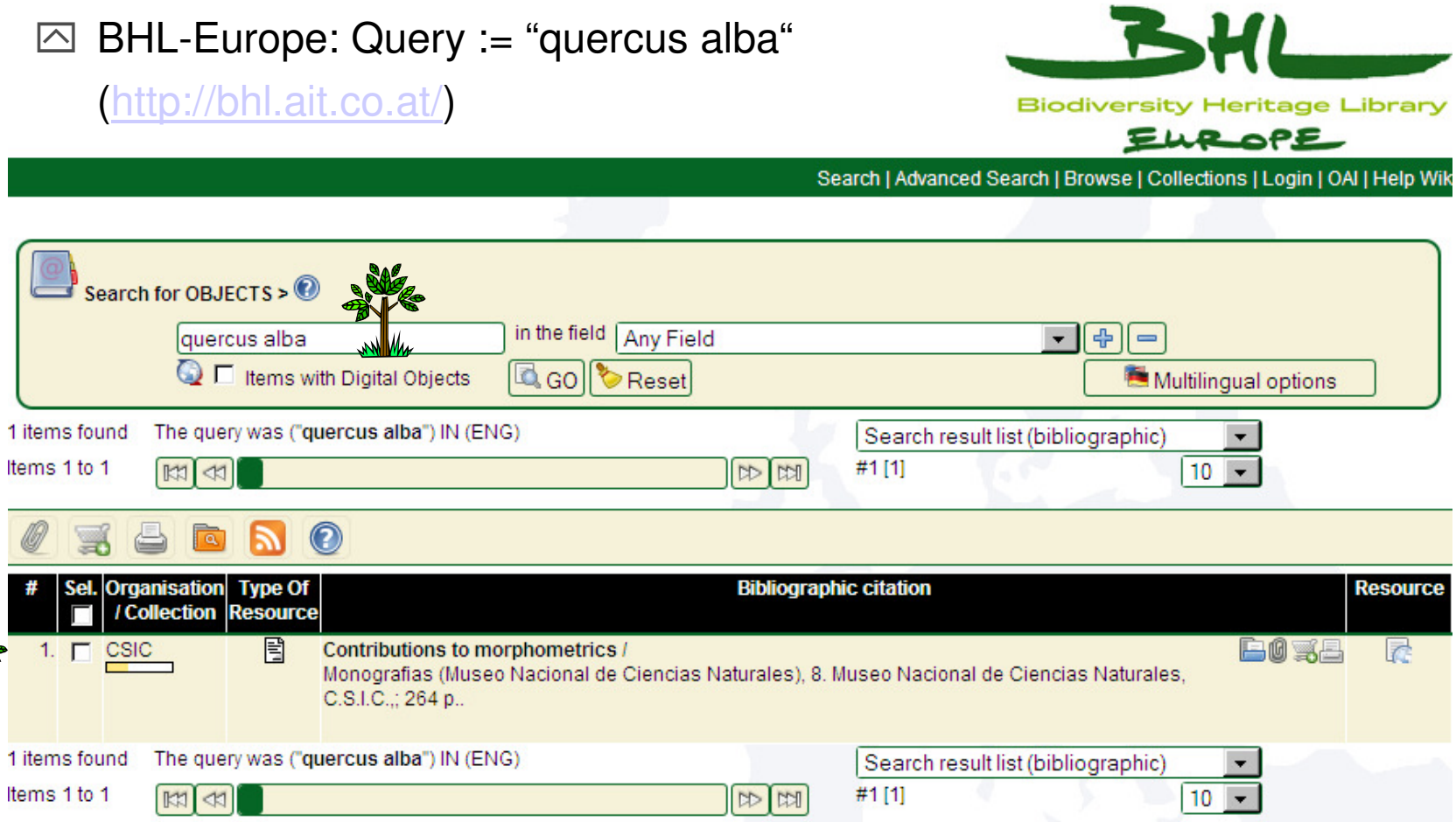
- BHL-Europe: TaxonFinder lookup on page level.

- (This function is an attempt to merge the TaxonGrab and FindIT algorithms.) It will identify **taxonomic name strings** within documents.
- **input arguments:**
  - url: URL of document to be searched
  - freeText: block of text to be searched. (Will only be looked at if URL is empty.)
  - includeLinks: returns external URLs associated with returned name strings if set to 1
- **Output:** Taxon/Taxa

- **WSDL**: http://names.ubio.org/soap/

Example: Page Level Indexing (2)

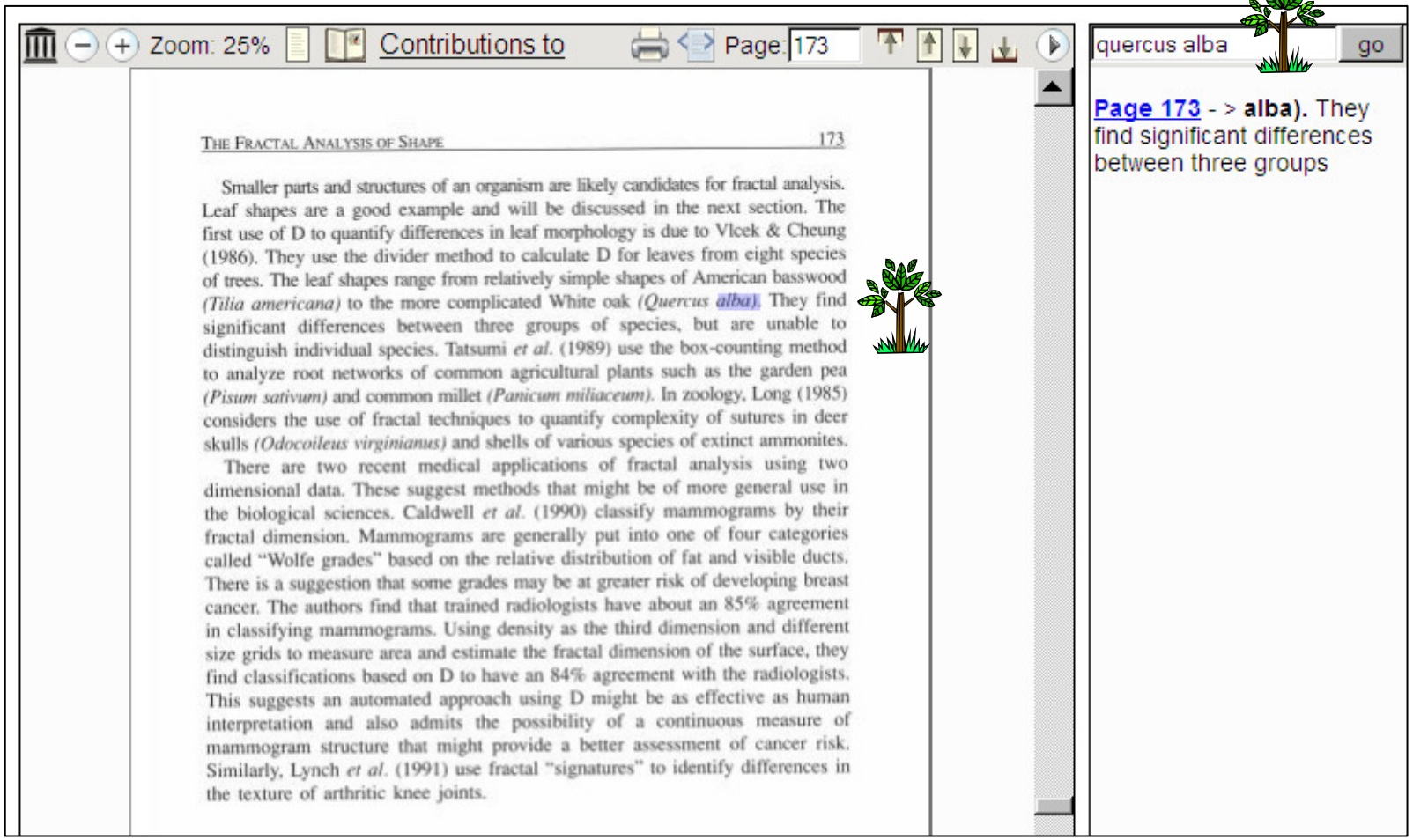
- BHL-Europe: Query := “quercus alba“ (http://bhl.ait.co.at/)



Example: Page Level Indexing (3)

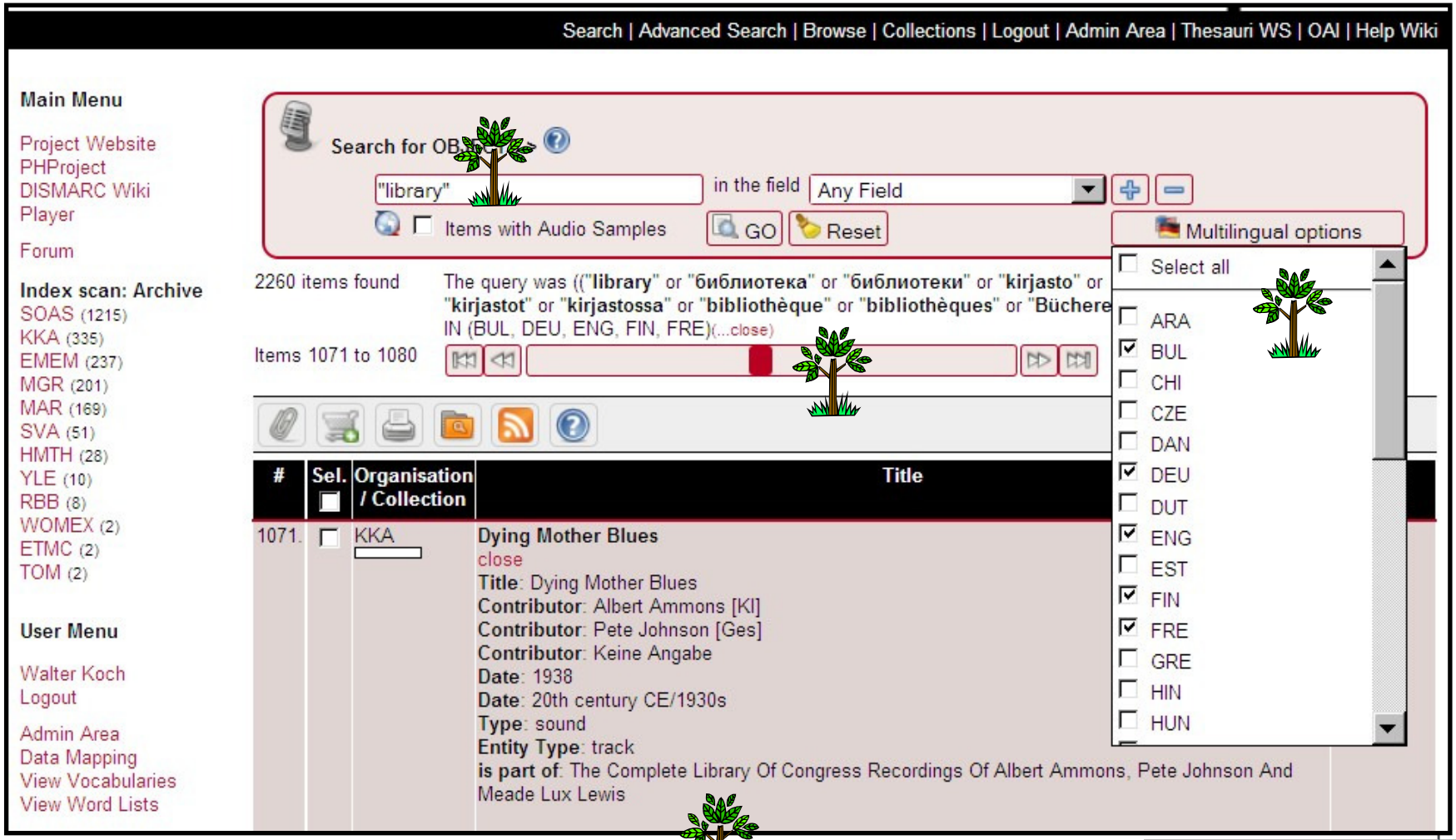
- Page Look Up

Subject : Quercus alba-- Link to page



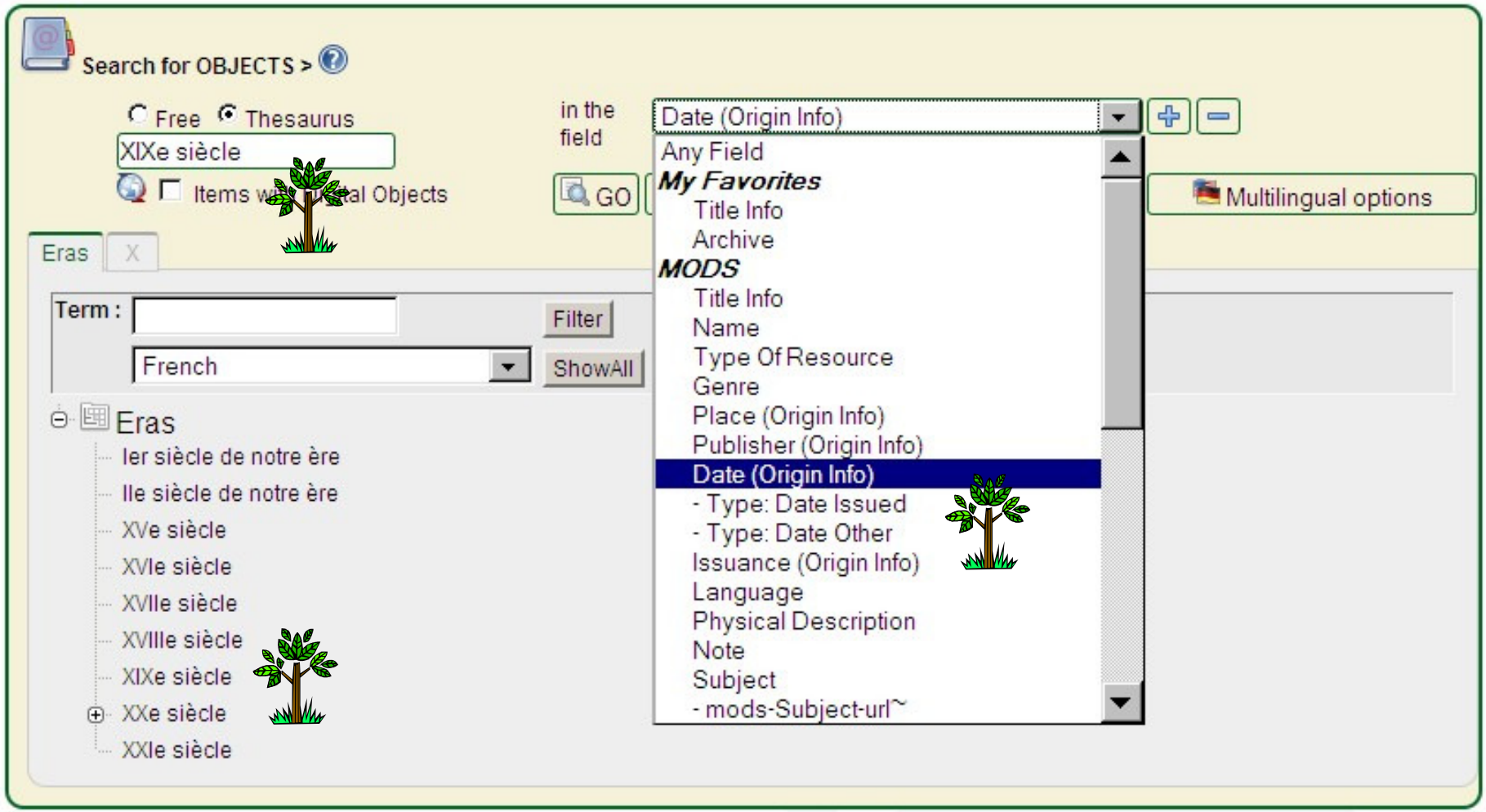
Example: Query expansion

- DISMARC

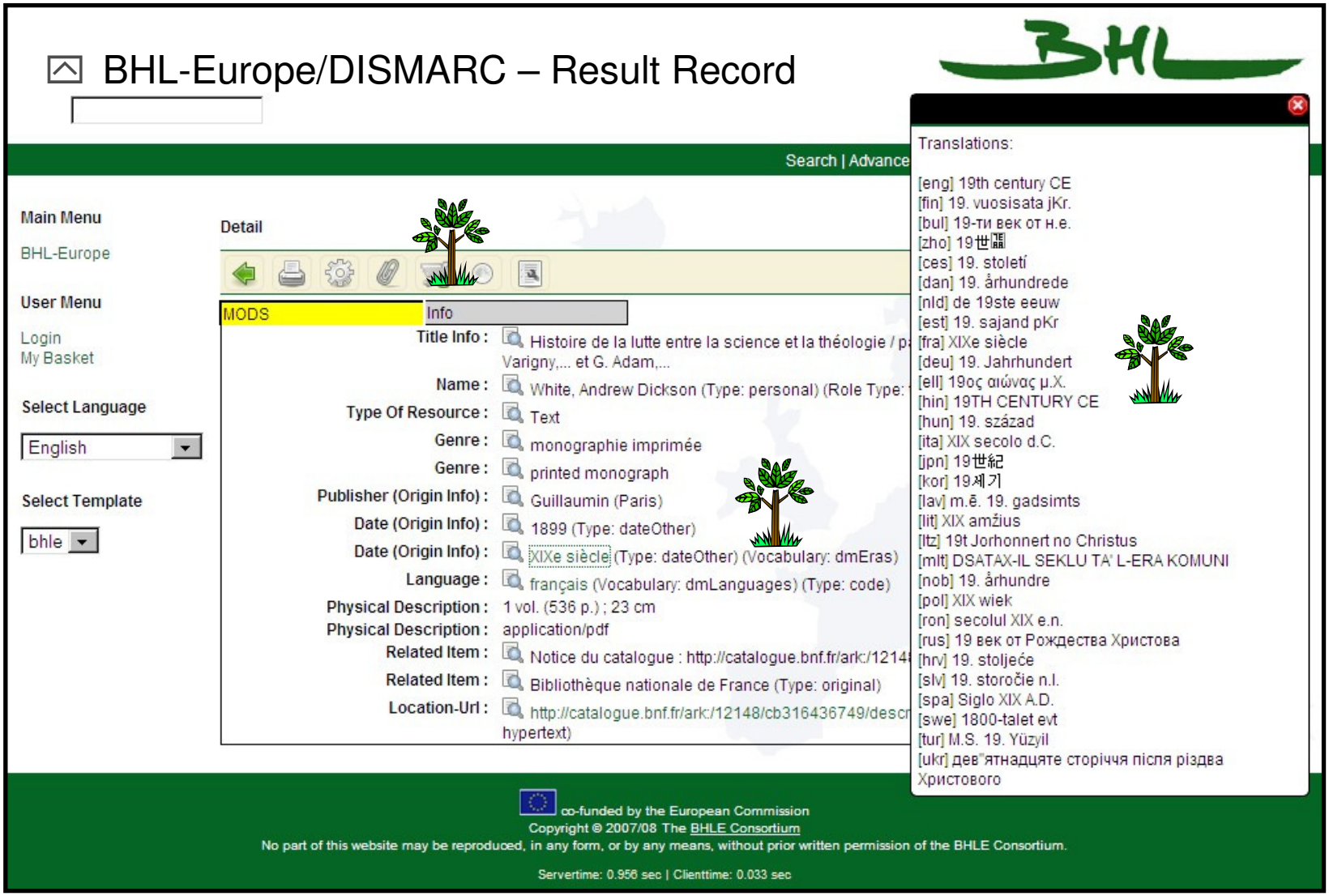


Example: Search Term Selection (1)

- BHL-Europe/DISMARC

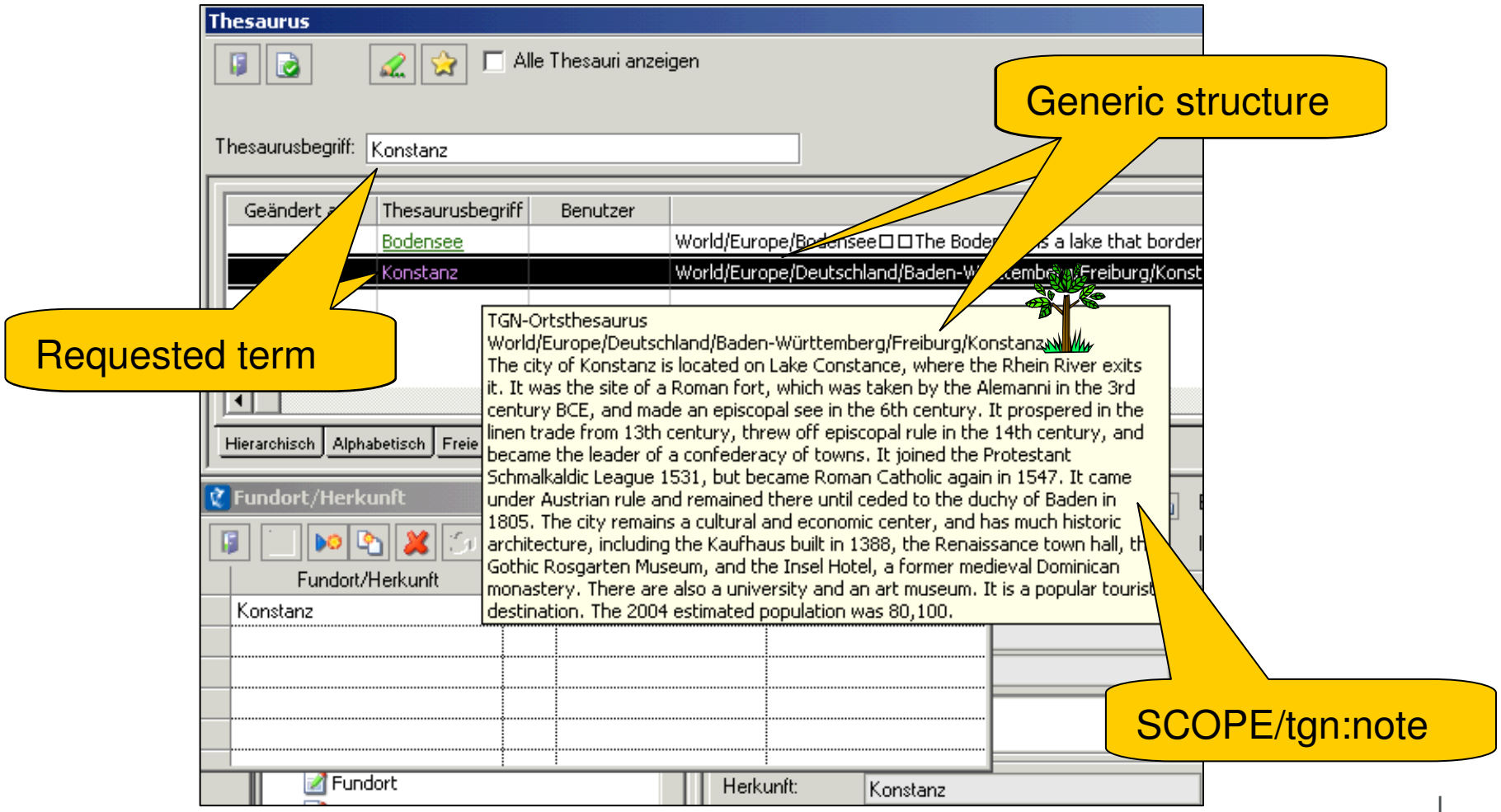


Example: Search Term Selection (2)



Example: Collection Management Integration (1)

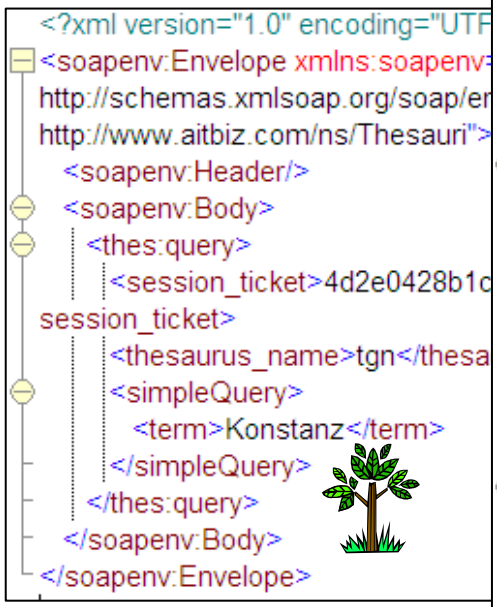
- TGN – Lookup in IMDAS (Screenshot provided by J.Sieglerschmidt)

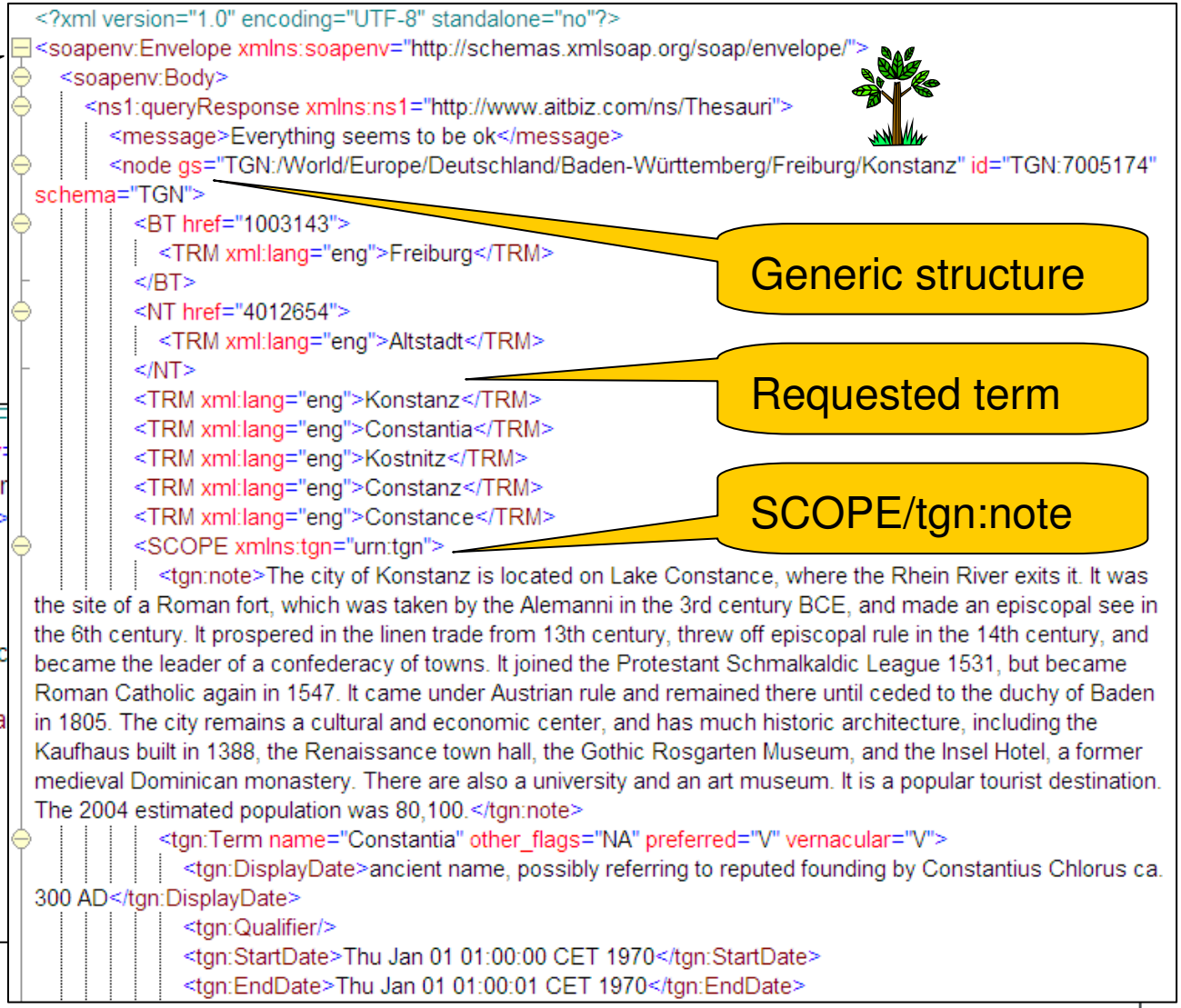


Beispiel: Collection Management Integration (2)

SOAP-Result

SOAP-Request



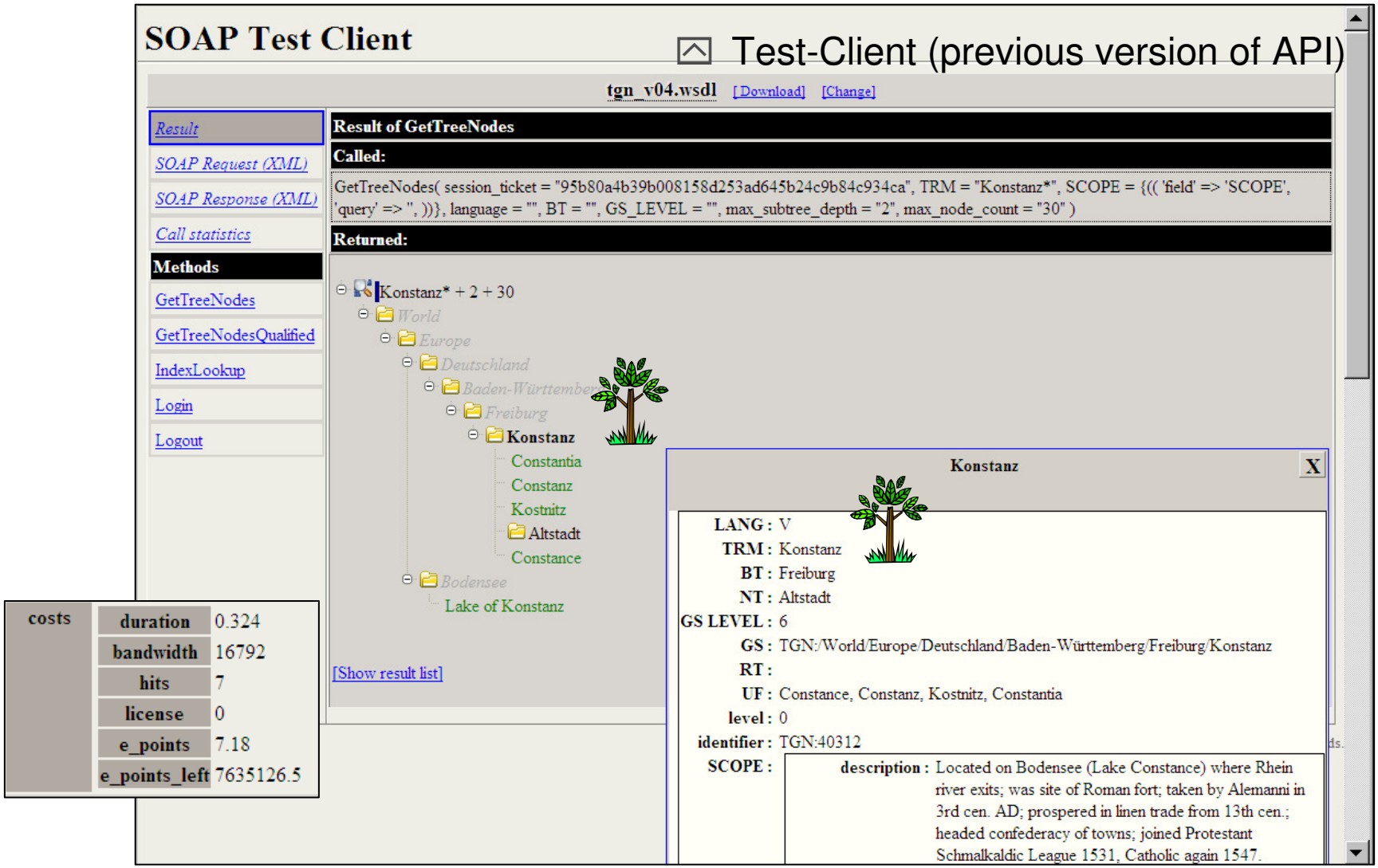


Generic structure

Requested term

SCOPE/tgn:note

Example: Collection Management Integration (3)



Thank You

Steinbeis Innovation Transfercentre

Information Management and
Cultural Heritage Informatics (IMCHI)

Leiter: Univ.Prof.Dr.**Walter Koch** (TU-Graz)

A-8010 Graz, Klosterwiesgasse 32/1

Fon: +43.316.8182228

Fax: +43.316.8182228-30

E-Mail: walter.koch@stw.de (walter.koch@tugraz.at)

Internet: http://www.stw.de/