Terminology Services

Ceri Binding & Douglas Tudhope
Hypermedia Research Unit,
University of Glamorgan

http://hypermedia.research.glam.ac.uk/
Use of Terminology Services at Glamorgan

- **STAR** *(Semantic Technologies for Archaeological Resources)*
  - 3 year AHRC funded project
  - English Heritage & RSLIS, Denmark

- **PERTAINS** *(Personalised Tagging Information Services)*
  - 5 month JISC funded project
  - MIMAS (COPAC, Intute)
Users of Glamorgan Terminology Services

- ArchaeoTools – Archaeology Data Service (ADS), York
- DELOS Digital Library Management System (DLMS)
- Glamorgan PhD, BSc projects
STAR – Web Services and Client Applications

STAR Web Services

- Archaeological Datasets (CRM)
- English Heritage thesauri (SKOS)

STAR Client Applications

- Archaeological Sciences Thesaurus
- Evidence Thesaurus
- Building Materials Thesaurus
- Monument Type Thesaurus
- MDA Object Type Thesaurus

Windows Applications
- Browser Components
- Full text search
- Browse concept space
- Navigate via expansion
- Cross search archaeological datasets

http://hypermedia.research.glam.ac.uk
Technical Implementation Overview

- Services implemented in C#, using “SemWeb” RDF Library
  (http://razor.occams.info/code/semweb/)

- MySQL database backend holds RDF triples representing multiple SKOS thesauri and glossaries – facilitates cross searching of thesauri

- Full-text index added for partial matching on literal strings

  - GetTopmostConcepts(..) : Concept[]
  - GetConceptScheme(..) : ConceptScheme
  - GetConceptSchemes() : ConceptScheme[]
  - GetConcept(..) : Concept
  - GetAllConceptRelatives(..) : ConceptRelative[]
  - GetKeywordMatch(..) : ConceptRelative[]
  - ExpandConcept(..) : ConceptRelative[]

STAR Web Service Calls
STAR Services – Windows Client Applications
STAR Services – Windows Client Applications
Terminology Services

STAR Services – Windows Client Applications

http://hypermedia.research.glam.ac.uk
STAR Services – Windows Client Applications
**STAR Services – Web Client Components**

- **Concept search**
  - Search across multiple thesauri

- **Concept**
  - BROOCH
  - http://reswn1.isd.glam.ac.uk/star/rdf/concepts#EH128
  - Ornament often with a hinged pin and catch, worn fastened to clothing.

- **Concept relatives**
  - BT: JEWELLERY
  - NT: ANNUAL BROC
  - NT: BOW BROOCH
  - NT: DRAGONESQUE BROC
  - NT: LONG BROOC
  - NT: PERANNUAL BROOC

- **Concept schemes**
  - EVIDENCE
  - MONUMENT TYPE
  - TIMELINE THESAURUS (text only)
  - ARCHAEOLOGICAL SCIENCES THESAURUS
  - MDA OBJECT TYPE
  - MAIN BUILDING MATERIALS

- **Top concepts**
  - ANIMAL EQUIPMENT
  - AGRICULTURE AND SUBSISTENCE
  - ARCHITECTURE
  - FURNISHINGS AND FURNITURE
  - TRANSPORT
  - FOOD PREPARATION AND CONSUMPTION

- **Navigate via semantic expansion**

- **Browse available thesauri**

---

[http://hypermedia.research.glam.ac.uk](http://hypermedia.research.glam.ac.uk)
A Case For More Specialised Services?

• Vocabulary types not specified within SKOS
  • Thesauri
  • Classifications
  • Taxonomies
• API based on SKOS could also consider:
  • pre co-ordination
  • facet grammar
  • synthesis rules
Classification schemes – usage observations

• More general, covering wider subject area
• Longer, more descriptive captions
• Notation encompasses specific ordering principle
• Usage – browsing subject area
• Classification, not indexing
• Pre-coordinated captions. KWIC / KWOC?

Laurence-Moon Syndrome — ultrastructure
Laurence-Moon Syndrome — therapy
Laurence-Moon Syndrome — surgery
Laurence-Moon Syndrome — rehabilitation
Laurence-Moon Syndrome — radiotherapy
Dedicated Service API or Query Language?

- **Query Language** (e.g. SPARQL)
  - Flexibility – end user dictates requirements
  - Standardisation of query language
  - Ease of deployment – native implementations

- **Service API**
  - Less knowledge of underlying data schema required
  - Aggregation functionality, pre-processing, algorithms
  - Performance - query optimisation server side
**TARGETS – Tag Suggestion Services**

**Document Metadata**
- Title
- Classification
- Uncontrolled keywords
- Controlled keywords
- Description

**Aim** - improve quality of tag metadata
**Basis** - existing document metadata / DDC (SKOS)
**Algorithm** - matching, weighting, filtering, ranking
**Implementation** - URL based web service interface

**DDC (SKOS)**

**PERTAINS Web Services**

[http://hypermedia.research.glam.ac.uk](http://hypermedia.research.glam.ac.uk)

University of Glamorgan
you live, you learn
PERTAINS – Client Applications

Dewey Decimal Classification (DDC) Search

Note: All copyright rights in the Dewey Decimal Classification system are owned by OCLC. Dewey, Dewey Decimal Classification, DDC and WebDewey are registered trademarks of OCLC.

Select area(s) of interest:
- 000: Computer science, information & general works
- 100: Philosophy & psychology
- 200: Religion
- 300: Social sciences
- 400: Language
- 500: Science
- 600: Technology
- 700: Arts & recreation
- 800: Literature
- 900: History & geography

Suggest terms from selected area(s) of interest

Clear selection

Suggest terms

0 results

http://hypermedia.research.glam.ac.uk
Conclusions

• Significant potential for reuse
  • Third parties can integrate your data
• Pre-built functionality – eases end user adoption
  • Platform neutral access
• Maintain version control
  • Dissemination of updates – instant
• Maintain access control
Terminology Services

Ceri Binding & Douglas Tudhope
Hypermedia Research Unit,
University of Glamorgan

http://hypermedia.research.glam.ac.uk/