A Common Terminology Services (CTS) Back-end to Protégé

Harold R Solbrig
Christopher G Chute, M.D., Dr. P.H.
Division of Biomedical Informatics
Mayo Clinic
Outline

• Purpose of the project
• Characteristics of Terminology Links
• The Common Terminology Services (CTS) specification
• Protégé as a CTS Client
• Protégé as a CTS Server
• Issues, discussion and next steps
Outline

• Purpose of the project
• Characteristics of Terminology Links
• The Common Terminology Services (CTS) specification
• Protégé as a CTS Client
• Protégé as a CTS Server
• Issues, discussion and next steps
Purpose of the Project

1) Use Protégé to create an interface between terminology and the information model

2) Use Protégé to extend standard terminologies and (potentially) author new content.
Terminology

- An integrated collection of terms, definitions, annotations, and relationships
- Defines the specialized language of a discipline or subject area
Terminology and Ontology

Terminology - Define the language / tokens used in a given domain

Ontology - Describes domain knowledge in a generic way and provides understanding of a domain*

*Asunción Gómez-Pérez – Tutorial on Ontological Engineering: IJCAI'99
Information Model

- Describes what information is required to accomplish a particular task or process
- Consists of named collections of data elements and their relationships
- Typically includes definitions of each of the elements, how they are used, possible values (at least on a conceptual level)
Ogden’s Semiotic Triangle

Ogden’s Semiotic Triangle

Thought or Reference

Refers to

Stands for

Symbolises

Referent

Symbol

“Rose”, “ClipArt”

The Communication Process

- **Referent**
  - "Rose"
  - "ClipArt"
  - "Red"

- **Symbol**
  - "I see an ClipArt image of a red rose"

- **Concept**
  - Refers To
  - Symbolises
  - Stands For
The Communication Process

Information Model

CONCEPT

Refers To

Symbolises

Stands For

“I see an ClipArt image of a red rose”

Symbol

CONCEPT

Refers To

Symbolises

Stands For

“Rose”, “ClipArt” “Red”

“Rose”, “ClipArt” “Red”

Symbol

Red roses can be used to signify love

(Part of) Ontology

Terminology – tokens and definitions
Terminology and Information Models

• There can be many information models that reference the same terminology

• It is the shared terminological references that allows be searched, compared and aggregated
Outline

• Purpose of the project

• **Characteristics of Terminology Links**

• The Common Terminology Services (CTS) specification

• Protégé as a CTS Client

• Protégé as a CTS Server

• Issues, discussion and next steps
Terminology Links

Classes

Slots

Possible Values
Concept References

Classes

Possible Values

Slots
Concept References

- Concept resources can include:
  - Existing taxonomies
  - Trade and specialty nomenclature
  - Standardized code sets (e.g. country codes)
  - Reporting and classification schemes
  - Ontologies
  - ...
Concept References

• Resource characteristics
  • Format and structure can vary widely
  • Availability can vary from simple ASCII lists of code value pairs to complex ontology services such as the Apelon DTS or DIG
Referencing Concepts

• Creating a manageable solution
  • Create a generic model of terminology content
  • Define an interface layer that allows queries to be posed in terms of the generic model
ConceptReference Example

![ConceptReference Example](image)
Outline

• Purpose of the project
• Characteristics of Terminology Links
• The Common Terminology Services (CTS) specification
• Protégé as a CTS Client
• Protégé as a CTS Server
• Issues, discussion and next steps
Why Terminology Services?

Terminology Resources are Heterogeneous

- **Structure** - varies from code/value pairs to complex formalisms such as NCI Thesaurus or SNOMED-CT
- **Size** – varies from 3 or 4 entries to well over half a million
- **Format** – can be CSV’s, SQL Tables, XML Documents, OWL
Why Terminology Services?

Terminology = content + software

• Sophisticated text search algorithms
• DL style inference
• Distribution, authoring, performance...
The Common Terminology Services (CTS) API

- (Relatively) Simple API
- Read-only
- A specification on how to do things like
  - List supported code systems
  - Get code system information
  - Search coded concepts by text
  - Traverse relationships
- Draft HL7 / ANSI Standard
Determine whether the supplied concept codes are related

Parameters:

- `codeSystem_id` - the code system of the parent and child codes
- `source_code` - the concept code that occurs as the source of the relationship
- `target_code` - the concept code that occurs as the target of the relationship
- `relationship_code` - the concept code that identifies the relationship
- `relationQualifiers` - an optional list of relationship qualifier codes. If the relation qualifiers that match all of the qualifiers in the list will be considered.
- `directRelationsOnly` - TRUE means test direct relationships only, FALSE means if the relationship is not transitive, the result is the same no matter the setting areCodesRelated returns TRUE if one of the following conditions holds:

1. There is a direct relationship of type relationship_code between source_code and target_code.
2. There is a direct relationship of type relationship_code between target_code and source_code.
3. Source_code equals target_code and the relationship is reflexive.
4. `directRelationsOnly` is FALSE, relationship_code is transitive and there is a reflexive closure of relationship_code starting at source_code.
5. `directRelationsOnly` is FALSE, relationship_code is transitive and symmetric and relationship_code equals target_code.
6. `directRelationsOnly` is FALSE, relationship_code is symmetric and relationship_code equals target_code.
CTS Interface Specification

Client

Server

Interface
CTS Interface Specification

- HL7 Messaging
- Decision Logic
- CTS Demo

Other services:
- LDAP Service
- SOAP Service
- LexGrid “Lit”
- ...

Interface
Java Client

LDAP Server
Java Client
SQL Server
No Concept Selected - Showing Connection Summary

You are connected to the LDAP CTS Implementation.

**Code System:** The Stanford Chimaera Project Beer Ontology

**Code System ID:** 1.3.6.1.4.1.2114.108.1.9.102

**Copyright:**

**Runtime Details**

**Service Name:** Mayo CTS Vaoi Runtime

**Relations**

- Graph View
- Zoom
- Rotate
Outline

- Purpose of the project
- Characteristics of Terminology Links
- The Common Terminology Services (CTS) specification
- Protégé as a CTS Client
- Protégé as a CTS Server
- Issues, discussion and next steps
Protégé as a CTS Client

Protégé

LDAP Service

SOAP Service

LexGrid "Lit"

...
Protégé as a CTS Client

Need to create a terminology model

- Code System
- Concept Code
- Designation
- Description
- Annotation
- Relationship
## Terminology Model

**Code System**

<table>
<thead>
<tr>
<th>CodeSystem (from Logical View)</th>
</tr>
</thead>
<tbody>
<tr>
<td>codeSystemDescription : String [0:1]</td>
</tr>
<tr>
<td>codeSystemName : String [1:1]</td>
</tr>
<tr>
<td>codeSystemCopyright : String [0:1]</td>
</tr>
<tr>
<td>codeSystemVersions : String [0:n]</td>
</tr>
<tr>
<td>codeSystemFullName : String [1:1]</td>
</tr>
<tr>
<td>codeSystemId : String [1:1]</td>
</tr>
</tbody>
</table>
## Terminology Model
### Code System

<table>
<thead>
<tr>
<th>Name</th>
<th>Documentation</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>CodeSystem</td>
<td>A collection of concept codes from a common namespace that represent a collection of concepts from a particular perspective or view.</td>
<td></td>
</tr>
</tbody>
</table>

### Template Slots

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Cardinality</th>
<th>Other Facets</th>
</tr>
</thead>
<tbody>
<tr>
<td>supportedProperties</td>
<td>Instance</td>
<td>multiple</td>
<td>classes={ExternallyDefined-SLOT}</td>
</tr>
<tr>
<td>supportedRelationships</td>
<td>Instance</td>
<td>multiple</td>
<td>classes={ExternallyDefined-SLOT}</td>
</tr>
<tr>
<td>codeSystemDescription</td>
<td>String</td>
<td>single</td>
<td></td>
</tr>
<tr>
<td>codeSystemName</td>
<td>String</td>
<td>required single</td>
<td></td>
</tr>
<tr>
<td>codeSystemCopyright</td>
<td>String</td>
<td>single</td>
<td></td>
</tr>
<tr>
<td>codeSystemVersions</td>
<td>String</td>
<td>multiple</td>
<td></td>
</tr>
<tr>
<td>codeSystemFullName</td>
<td>String</td>
<td>required single</td>
<td></td>
</tr>
<tr>
<td>codeSystemId</td>
<td>String</td>
<td>required single</td>
<td></td>
</tr>
<tr>
<td>supportedLanguages</td>
<td>Instance</td>
<td>required multiple</td>
<td>classes={LanguageReference}</td>
</tr>
</tbody>
</table>
Terminology Model
Code System - Example
Terminology Model
Terms & Designations
**Terminology Model**

**Terms**

<table>
<thead>
<tr>
<th>Term</th>
<th>Documentation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Term</td>
<td>A string of text that can be used to represent the meaning or intent of a concept code in a given language and (optional) context or contexts.</td>
</tr>
</tbody>
</table>

**Template Slots**

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Cardinality</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>termText</td>
<td>String</td>
<td>required</td>
<td>single</td>
</tr>
<tr>
<td>termId</td>
<td>String</td>
<td>single</td>
<td></td>
</tr>
<tr>
<td>termLanguage</td>
<td>Instance</td>
<td>required</td>
<td>single</td>
</tr>
<tr>
<td>termSource</td>
<td>Instance</td>
<td>multiple</td>
<td>classes=(.SourceReference)</td>
</tr>
</tbody>
</table>

**Relationships**

- THING
- :SYSTEM-CLASS
- TerminologyEntities
- CodedConcept
- Term
- ConceptDefinition
- ConceptAnnotation
- ConceptDesignation
- ConceptReferences
Terminology Model
Terms - Example
Terminology Model
Designations

Template Slots

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Cardinality</th>
<th>Other Facets</th>
</tr>
</thead>
<tbody>
<tr>
<td>appliesInContexts</td>
<td>Instance</td>
<td>multiple</td>
<td>classes={UsageContextReference}</td>
</tr>
<tr>
<td>preferredForLanguage</td>
<td>Boolean</td>
<td>required</td>
<td>default={false}</td>
</tr>
<tr>
<td>representativeTerm</td>
<td>Instance</td>
<td>single</td>
<td>classes={Term}</td>
</tr>
<tr>
<td>designationCodeSystem</td>
<td>Instance</td>
<td>single</td>
<td>classes={CodeSystem} value={ISO639-1 [216.840.1.113983.6.84]}</td>
</tr>
</tbody>
</table>
Terminology Model
Designations - Example
Terminology Model
Definitions, Annotations, etc.
Terminology Model
Definitions, Annotations, etc.
CTS Client
New File Dialog

Select Format

- Standard Text Files
- Database
- OWL Database
- OWL Files
- CTS Backend
- RDF Schema

OK Cancel
CTS Backend

A class that derives its definition from an external classification system or coding scheme.
CTS Client Preferences

User Preferences:
- ✔ Ignore load errors
- ☐ Use local copy
- ☐ Enable lazy load
- ✔ Locally editable meta classes
- ✔ Locally editable terminology classes

Search result limit: 20000
Timeout: 20000

Save
CTS Client
Coding Scheme Selection

Select coding schemes to import for hL7 (Total 156):
- AcknowledgementCondition
- AcknowledgementDetailType
- AcknowledgementType
- ActClass
- ActCode
- ActInvoiceElementModifier
- ActMood
- ActPriority
- ActReason
- ActRelationshipCheckpoint
- ActRelationshipJoin
- ActRelationshipSplit
- ActRelationshipType
- ActSite
- ActStatus
- ActUncertainty
- AddressPartType
- AdministrativeGender
Outline

• Purpose of the project
• Characteristics of Terminology Links
• The Common Terminology Services (CTS) specification
• Protégé as a CTS Client
• Protégé as a CTS Server
• Issues, discussion and next steps
Protégé as a CTS Service

- HL7 Messaging
- Decision Logic
- CTS Demo

Interface

Protégé
### Concept Details

**Concept Code:** PBB
**Designation:** Yuch Beer (en)
**Code System:** 1.3.6.1.4.1.2114.108.1.9.102
**Code System Version:** 0.1
**Code Status:** true
**Property:** Instance :conceptDesignations (en)
**Property:** String :hasAlcoholicContent (en)

### Relations

- **madeFrom-RELATION**: Yuch Beer
- **awarded-RELATION**: Award (Award)
- **hasAlcoholicStrength-RELATION**: 

### Code System

- Carapils
- Cascade
- Chinook
- Chocolate
- DryStout
- Festival
- Galena
- Grain
- Hillertau
- hasAlcoholicStrength
- Hops
- ImperialStout
- IndiaPaleAle
- Ingredient
- KentGoldsings
- Lager
- locatedIn
- madeFrom
- Malt
- Microbrewery
- Mid
- MountHood
- Munich
- Name
- Pale
- PaleAle
- PIBB
- Perle
- Pilsner
- Roster
Outline

• Purpose of the project
• Characteristics of Terminology Links
• The Common Terminology Services (CTS) specification
• Protégé as a CTS Client
• Protégé as a CTS Server
• Issues, discussion and next steps
Issues, etc.

- **Slow load time**
  - Appears comparable w/ OWL
  - Partially resolved w/ lazy loading
- **Need a SymbolReference**
  - Dynamically resolved Symbol
  - Ties in w/ DirectedBinaryRelation
- **Relations**
  - Need to reconcile w/ DirectedBinaryRelation
- **Need to make CTS Server a Protégé Server**
Protégé as CTS Server

CTS Demo

Protégé (client)

Protégé Client

Protégé Server
Issues, etc. (continued)

- Need to flesh out lazy loading
  - References occur all over
- Need to implement a hierarchy trimmer
  - Only keep references
  - Only keep graph branch points
Credits

• Deepak Sharma – author
• Dan Armbrust - CTS author

This work was supported in part by a grant from the US National Library of Medicine: LM07319.