LexEVS and CTS2 Meeting Minutes - 2016.09.30

Attendees

Name	Role	Present
Wright, Larry	NIH/NCI	x
Fragoso, Gilberto	NIH/NCI	x
De Coronado, Sherri	NIH/NCI	x
Safran, Tracy	NIH/NCI [C]	x
Ong, Kim L	IS	x
Lucas, Jason R	IS	x
Bauer, Scott	Mayo	x
Stancl, Craig	Mayo	x
Endle, Cory	Mayo	x
Solbrig, Harold	Mayo	x
Jiang, Guoqian	Mayo	x
Wynne, Robert	NIH/NCI [C]	
Tran, Tin	NIH/NCI [C]	
Carlsen, Brian	NIH/NCI [C]	
Wong, Joanne	NIH/NCI [C]	
Kuntipuram, Kumar	NIH/NCI [C]	
Haber, Margaret	NIH/NCI	x
Campbell, John	NIH/NCI	x

Goals

We need to broaden the scope of and participation in this series, and at the same time sharpen the focus.

1. LexEVS and CTS2 are planned to continue as key components of NCI terminology services, but NCI also plans to extend beyond them, notably with RDF triple store SPARQL and REST components. This raises important design and coordination issues.
2. REST services need substantial extensions, within and beyond CTS2. NCI plans to offer REST APIs that can support the requirements of its own and others' applications, and expects a significant component to be REST services from the RDF triple store. An obvious test case is ability to support EVS browser functionality (even while these browsers will likely use the Java API for some time to come).
3. We face some immediate needs from other users, notably CTRP and GDC, so need to clarify design approaches and produce new services based on them in the relatively near term.

Agenda

Review and discuss (Kim's /Jason's) concerns
 	
• Each of these are to support the browser (via java API)
	1. Handling of metadata
	   a. Metadata to be stored as RDF
	2. Supported search algorithms (currently, only exactMatch, startsWith, and contains are available)
	   a. CTS2 service - would like to have CTS2 support the full LexEVS Search Extension
	   b. Need to confirm that the Lucene algorithm is available.
	3. CTS2 Handling of namespace in read service. Could namespace be included as a parameter in a REST call?
	   a. entity read service is able to include namespace.
	   b. Need to confirm the responses to determine both namespaces are supported for single concept code
	   c. Need to provide ability to return concept with different namespaces.
	4. Search against multiple coding schemes (count, ranking, iterator, pagination)
	   a. Required for browser
	   b. Currently allows search against all, but not a designated set of coding schemes.
	5. Search against multiple mappings and value sets (count, ranking, iterator, pagination)
	   a. Required for browser
	   b. Currently allows search against all, but not a designated set of coding schemes, mapping and value set API may require additional public methods
	6. Mapping search (count, ranking, iterator, pagination) - currently iterator and pagination are broken in remote LexEVSAPI)
	   a. CTS2 will no longer be running against remote (only local)
	   b. Broken iterator no longer issue if running local.
	7. Support for property and relationship search?
	   a. Need further notes on this. Graph implementation should cover relationship search. Properties are currently restricted to presentations, it may mean an extension to supply better coverage to generic property types.
	8. Constrained/restricted search (e.g., restrictToCode on CNS; search only on a particular branch of tree, for instance, Gene, Gene Product -- CTRP like use cases).
	   a. Not readily available in CTS2 (as in LexEVS)
	   b. Some searching on properties (designations)
	9. Tree construction (expandable status, distance-2 child node count)
	
	   a. Need to have ability to know if a node is expandable or not.
	   b. Can be provided by leaf from graph api
	10. Resolving of a graph (transitive closure).
	   a. Graph API can provide this - but not implemented today.
	11. Current use of CRC (cyclic redundancy code) in the response, would like to understand better the use?
	   a. This is provided in CTS2 response.
	   b. This an identifier applied to a definition (IE)
	
	      i. A name for the definition.
	      ii. In this case, it can be anything to identify the version.
	      iii. The request is to resolve this identifier on the server rather than passing it back and forth between client and server. This would require an API change.
	12. Issues we face for Triplestore extension: Various OWL formats (e.g., in NCIt the unique idenfier is code, but in GO, it's oboInOWL:id)
	   a. CTS2 allows for many identifiers
	   b. Gilberto to document and share the concerns for consideration
	   c. Similar parallels to URI resolver?
	13. Handling of NCIm and other non-OWL or RDF coding schemes.

CTRP and GDC Issues (Larry, others)	• Nothing immediate to address
	• Ability to collect a "structured neighborhood of concepts related to a user query"
	   ○ For example, Investigator of a clinical trials protocol will search for a gene, protein, etc. but not specific - and then present a neighborhood in a structured way for the investigator to review.
	   ○ Provide ability to efficiently traverse a set of relationships
	• Both groups will looking at using REST APIs to get terminology content
	• Nothing additional today, but possible future requirements.

Additional Considerations	• Modeling of cancer concepts ○ Need to identify the notes/requirements. Possibly notes from previous F2F. • Partonomy

Next Steps	• Review notes from previous F2F.
	• Determine priority for browser support
	• determine requirements that fit into CTS2
	• Determine an architecture to support requirements that do not fit into CTS2
	• Determine integration of LexEVS, CTS2, Triple Store
	   ○ separate rest services?
	   ○ discussion necessary