LexEVS Meeting Minutes - Value Set Architecture Planning Session - 2017.01.06

Attendees

Name	Role	Present
Wright, Larry	NIH/NCI	x
Fragoso, Gilberto	NIH/NCI	
Safran, Tracy	NIH/NCI [C]	x
Ong, Kim L	IS	x
Wynne, Robert	NIH/NCI [C]	x
Lucas, Jason	NIH/NCI [C]	x
Bauer, Scott	Мауо	x
Stancl, Craig	Мауо	x
Endle, Cory	Мауо	x

Agenda

Goals:

- Better understand the struggles and issues associated to loading value sets. Once those are identified, understand the actual requirements.
 Once requirements are agreed upon, propose solutions to satisfy them.

Additional requirements	Discussion Points:
Value set resolutionSearch and manipulation	Decision Points:

Review Value Set workflow process (Tracy and Rob)	Discussion Points:	
 The goal is to identify the areas of the workflow that cause the greatest issues. (for example, long load time for 700 value sets) 	 What are the requirements for resolving value sets? Performance parameters? Rob mentioned there are 760 resolved value sets - loaded over the weekend. Last one took about 2.5 days. This is run once a month. Is there a simpler/quicker way to resolve these value sets? Better logging would help to determine if there were any failures for resolving the value sets. Rob suggested that if the script ran over night, this would be acceptable. Scott mentioned we could make the value sets transitive. This would make loading resolved values sets perform better. Rob questioned if we can load multiple values in parallel. Kim asked if we need to resolve all of these value sets every month. Produce a large OWL file that contains entities with the value sets. This file needs value set membership information. Can search against this large file. This is all in the LexEVS environment. Kim would need a convenient API method to generate this output file. There are a few value sets that are not from NCI Thesaurus. We will need to verify that the value sets (about 700 of them) are concept in subset. The hierarchy of the targets are about 2 levels deep. Most of the value sets can be built off from the information in the NCI Thesaurus. Decision Points: Requirements Ability to audit and simplify the workflow. Better logging to help determine if there were any failures for resolving the 760 value sets. Rob suggested that if the script ran over night, this would be acceptable. 	
Discuss and capture issues for further review.	Discussion Points: Decision Points:	

REFERENCES

Value Set Issues Identified during 2016.12 Technical Face-To-Face

• 2016.12 Technical Face-To-Face Prioritization List

Priority	Issue
1	Investigate ability to be able to determine if Resolved VS coding scheme has changed.
1	Investigate ability to update as needed (not have to load all 700 at the same time). Decrease time it takes to publish/deploy.
2	Investigate ability to be able to determine if Value Set Definition has changed.
2	Investigate ability to capture "Hierarchal Value Sets" as coding schemes with hierarchy. (inside value set)