# 2017.12 Technical Face-To-Face Meeting Notes

**Document Information**

**Author:** Craig Stancl, Scott Bauer, Cory Endle
**Email:** craig.stancl2@nih.gov, scott.bauer@nih.gov, cory.endle@nih.gov
**Team:** LexEVS
**Contract:** 16X237
**Client:** NCI CBIIT
National Institutes of Heath
US Department of Health and Human Services

**Contents of this Page**


- Tuesday, December 5th, 2017
  - EVS Status and Future Direction
  - EVS Technical Infrastructure, Issues, and Options
  - User Group Discussion - caDSR
  - User Group Discussion - FDA and CDISC
  - User Group Discussion - CTRP / CTS-API
- Wednesday, December 6th, 2017
  - EVS Architecture
  - EVS Architecture - Technical Discussion with Systems team
  - EVS Project Group Discussion (During regular call-in time)
  - NCI Systems Discussions
  - User Group Discussion - caDSR
- Thursday, December 7th, 2017
  - API Services
  - Prioritization and debrief


The **purpose of this document** is to document the 2017.12 technical face to face meeting details between the NCI and the LexEVS Team**.**

# 2017 December Face-to-Face Meeting Notes

## Tuesday, December 5th, 2017

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 9:00 AM - 11:00 AM | 4-W-034 | EVS Status and Future Direction<br><br>**Discuss EVS current state, trends, and future directions**<br>• Larry to give brief overview of EVS infrastructure, resources, and services.<br>• Review overall technical workflow and architecture.<br>• Group discussion of future possible directions and priorities. | Broad cross-EVS participation | **EVS Project Architecture** |

Attendees:

Jason Lucas, Scott Bauer, Larry Wright, Cory Endle, Kim Ong, Tracy Safran, Rob Wynn, Gilberto Fragoso, Margaret Haber, Kumar, Sherri De Coronado, John Campbell, Bron, Luba, Sana Din, Craig Stancl

Discussion Points:

- House keeping Items
  - Reviewed agenda and approved - unless there are changes along the way.
  - WebEx will be live all day.
  - Goal is to record key tasks and wikis
- Goal is to set context for the rest of the meetings this week and to start identifying the issues to be addressed.
- Larry would like to start to complete the complete EVS Project Architecture (including LexEVS)
- General workflow for architecture:
  - Gather terminology content protege and meme loaded into LexEVS terminology service accessed via java api, rest service api, browsers
  - Architecture now needs to include the triple store database and usage (REST API and native REST API). This service is to support clinical trials (CTRP). Ability to make changes into the production service was a driving factor in going to TripleStore architecture. Loading into triplestore can be done nightly if needed. Currently the loads have been nightly.

        - OBO is currently being looked as as a third delivery channel.
        - Expected to have expanded use of services and downloads
            - Adverse events were the most downloads - and then used and built in other systems (CDISC and FDA, etc).
        - Report Writer extracts value sets from LexEVS
            - Current work happening to create a SPARQL based report writer. Planned for early February.
                - This would be only on QA team.
                - External use would require a security layer (doesn't exist today)
                - Gilberto noted that report writer cannot currently take a search and return result set (with preferred names). Noted this is "simple search" and could also be part of the term browser.
                - Existing templates should be able to be run without authentication.
- The TripleStore still cannot provide all the terminology data needed for EVS and is stored in LexEVS.
- Mappings - need to determine how to capture and allow access to mappings. The LexEVS model and triple store model do not provide the needed flexibility today.
- Synchronization of data sources and coordination of distribution of data is an open issue. Consideration is needed to provide an umbrella API (Federated) that serves both LexEVS and TripleStore content.
- Gilberto noted that CTS2 services should be revisited and he'd like to review missing functionality and complexity (noted by CTRP).
- LexEVS historically has been based on standards since the early inception of the tooling.
    - As a terminology service, all the content has been loaded into Lexgrid data model.
    - Focus from standards has shifted to providing usable services to end users.
    - There are possibilities for enhancing the service today that still provides interoperability.
    - Thesaurus based use cases should be considered when determining goals.
    - CTRP usage of TripleStore was speed in loading content. LexEVS loading of transitive table is long process. Tracy noted that LexEVS could start to use TripleStore technology to remove that bottleneck when loading content.
    - Noted that existing applications may not be ready to transition to new serivces.
    - Primary goal of EVS is to proviede terminolgoy content to NCI customers and users to support the sciences.
        - Noted that no interest in re-desigining
    - LexEVS provides a consistent model.
- Mappings
    - Currently default to the LexGrid XML format. The other is RRF loader mappings.
    - No plans to load mappings into SPARQL.
    - ICDO3 Map from Meta - Kim looked at performace - not the best results.
    - Mappings in LexEVS in coding schemes.
    - FTP will be main distribution for Maps. (Tab delimited to text)
    - Review of CTS2 mapping support will help decide if additional functionality can be added.
    - No support for contextual mapping currently exists.
- Systems Priorities
    - LexEVS has spent time this last year and utilizing Docker. This will provide efficient deployments.
        - Current deployments are completed by using a documented deployment document.
        - Containers will reduce the middle man needed for deployments.
        - Systems team has been trying to create an "approved" NCI container so the LexEVS team can use.
        - Will still need to continue running tests.
        - Need to talk with systems about environment.
        - Docker is operational at NCI.
        - Docker usage for data
            - Gilberto suggested promotion of data using docker.
                - Provide the "database" container to the systems team.
                - This would help as it is deployed up the tiers.
        - Docker distributions to be used for end users.
            - ASU was working on this type of container.
            - Current scripts do 90% of what is needed. Need to change configuration so it doesn't remove the services.
- QA
    - Currently there are test scripts created by Kim and Sana
        - Covers the majority of the usage.
        - Test scripts still being developed.
        - Tin is still available for reference.
        - Tests should be avaiable in Jenkins. (LexEVS team is currently doing this)
        - Docker shouldn't cause concern for QA.
    - Tech stack support
        - Confirm that NODEJS and other technologies are supported.
- Security Scans
    - 508 needs to be addressed.
        - Heroku and SWAGGER pages may need to be reviewed.
- Lucene, SOLR and elastic search
    - Jason noted that there might be need for discussion during the API focused discussion on Thursday.

Decision Points:

- Action Items
    - Mapping to be discussed further with the editors.
    - Capture the Architecture to describe the workflow.
    - LexEVS to look at utilizing the triplestore to speed up load (remove the need to load transitive table)
    - LexEVS REST (CTS2) services should be revisited and reviewed for missing functionality and complexity
    - Investigate the use of Docker to deploy data (Database container).
    - Investigate the use of Docker containers to end users.
    - Docker environment updates from systems team.
    - Consider 508 compliance going forward.
    - Confirm tech stack status for use of NODEJS and other related technologies.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 11:00 AM - 12:00 PM | 4-W-034 | EVS Technical Infrastructure, Issues, and Options<br><br>**Flesh out architecture and workflow diagrams, identify key areas of discussion**<br><br>• Review overall architecture and expand/update<br>• Identify areas where current infrastructure is changing or problematic | Primarily EVS technical team members<br><br>(several have conflicting clinical trials meeting) | **EVS Project Architecture** |

Attendees: Jason Lucas, Scott Bauer, Kim Ong, Rob Wynn, Gilberto Fragoso, Cory Endle, Craig Stancl, Shamine, Sherri de Coronado

Discussion Points:

- Mapping
  - LexEVS doesn't allow for ability to map into Meta
    - Only saves code to code mapping and qualifiers
    - Synonym information is not saved. Currently you need to call into the coding scheme to get additional information.
    - There is no coding scheme loaded for ICDO3 so cannot get the additional information.
      - Sherri to determine if ICDO3 could be loaded.
      - Gilberto suggestedto have ICDO3 as an active coding scheme but not displayed in browser. Kim would need to do this.
    - Scott suggested to investigate the possibility of mapping for independent term to a source in meta; or another loaded terminology; or 2 sources in meta.
    - Rob suggested MetaMorphosis usage, but would need further investigation.
    - Execel spreadsheet (provided by Steph) provides ICDO code to Meta mappings.
    - SwisPROT mapping
      - Not a coding scheme.
      - Just a tab delimited file from website.
      - When loading into LexEVS, there is no target coding scheme
      - Scott suggested loading target entities that allows users to look elsewhere for resolution.
      - Traci suggested the use of URL Resolver - but doesn't need to be resolvable.
      - UniPROT resources
        - http://www.uniprot.org/uniprot/P62258
        - http://sparql.uniprot.org/
- Architecture and Workflow
  - **EVS Project Architecture**

Decision Points:

- Action Items:
  - Sherri to determine if ICDO could be loaded into LexEVS
  - Investigate the possibility of mapping for independent term to a source in meta; or another loaded terminology; or 2 sources in meta.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 1:00 PM - 2:00 PM | 5-W-032 | User Group Discussion - caDSR<br><br>**User Teams to share how they are using EVS and discuss requirements /priorities for the future.**<br><br>• APIs: Java, REST (CTS2 or 3-store), SPARQL, FTP<br>• Backwards compatibility of server/client/data releases<br>• Incl: Java/jar file issues and future<br>• Incl: New terminology server API/content/other needs. | caDSR contact - Denise, Philippa, developers | |

Attendees: Vikram, Natalia, Luba, Larry Wright, Scott Bauer, Jason Lucas, Cory Endle, Craig Stancl, Tracy Safran, Phillipa Barnes, Margaret Haber, John Campbell, Rob Wynn, Bron Kessler, Kim Ong, Sherri de Coronado, Denise Warzel, Sana Din, Liz

Discussion Points:

- Currently uses the API using the Jar file and dependencies.
  - Curresnt uses LexEVS Java API
  - No use of REST API except for limited use.
- Denise noted issues when data model or data have changed
  - Curation tool and SIW - no plan to change/update.
    - No reason to change the API unless it was going to be depricated.
- Denise noted that no release somplete to access to resolved value sets
  - LexEVS can provide, but caDSR needs to update.
- There are proof of concept services being evaluated.
  - CDE Recommender Service
  - CDE Validator Service

    - Looking at using SPARQL
- Traci asked what of the Java API is currently being used by caDSR so that when building the REST API, the team could focus on those services.
    - Search for Concept Code
        - Return name, definition, def source
        - Return super concepts or sub concepts
    - Search for Top level concept for value set/domain
        - Return resolved codes
    - Search for CDISC SDTM Variable Terminology
    - List all value sets
- Release Roadmap
    - Early 2018 - plan to make recommendation and decision
- Java 7 / Java 8 Jar
    - would require caDSR to do a maintenace release in 2018 (Q1)
    - Until this is complete, the EVS team needs to maintain 6.4 and 6.5 sets of data.

Decision Points:

- Action Items
    - caDSR team to provide a list of what is used from the Java API to determine what would need to be exposed in a REST API. Phillipa could meet with the team Wednesday at 3PM.
    - caDSR to update to Java 8 jar in 2018Q1.
    - Send retirement notice regarding the 6.4 retirement and caCORE REST.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 2:00 PM - 3:00 PM | 5-W-032 | **User Group Discussion - FDA and CDISC**<br><br>**User Teams to share how they are using EVS and discuss requirements/priorities for the future.**<br><br>• APIs: Java, REST (CTS2 or 3-store), SPARQL, FTP<br>• Backwards compatibility of server/client/data releases<br>• Incl: Java/jar file issues and future<br>• Incl: New terminology server API/content/other needs. | Editors<br><br>Liz, Erin, Brenda | |

Attendees:

Bron, Lub, Larry Wright, Scott Bauer, Rob Wynn, Jason Lucas, Liz, Gilberto Fragoso, Tracy Safran, Erin Mulbrandt, Lori Whiteman, Margaret Haber, Terry Quinn, Sherri de Coronado, Sana Din

Discussion Points:

- FDA
    - Report Writer
        - Terry generates 25 files (FDA and others) every month. Would like to be able to batch command and provide the dates needed for report writer.
        - Files are posted to FTP site.
            - These files contain subsets
        - Changes are indentifed by doing an exact compare of both sets of data (addition, changes, deletions)
        - Rob and Tracy are working on new report writer that should help make this process tolerable.
- CDISC
    - On the cancer.gov page for CDISC Terminology - as there are alot to scroll through, request for a table of contents to make it more usable.
    - OWL/RDF updates to metadata model for CDISC (Rob, via TopBraid)
    - Request to update the CDISC new term suggestion request form.
        - Update Request type dropdown
        - Update Code List dropdown
        - Possible type-ahead
    - CDISC Publication Column Headers
        - header naming switch planned (significant change) - 2018Q2
            - i.e. "CDISC Submission Value"
        - Rob noted there are changes needed for the reports

Decision Points:

- Action Items
    - Request for a table of contents on the CDISC Terminology page. (Cancer.gov)
    - Request to update the CDISC term suggestion request form.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|

| | | | | |
|---|---|---|---|---|
| 3:00 PM - 4:00 PM | 5-W-032 | **User Group Discussion - CTRP / CTS-API**<br><br>**User Teams to share how they are using EVS and discuss requirements/priorities for the future.**<br><br>• APIs: Java, REST (CTS2 or 3-store), SPARQL, FTP<br>• Backwards compatibility of server/client/data releases<br>• Incl: Java/jar file issues and future<br>• Incl: New terminology server API/content/other needs. | CTRP / CTS-API - managers, developers, Tiger team (Gisele, Samantha, David, Brian, Peter, Tracy, Jason, others) | |

Attendees:

Bron, Lub, Larry Wright, Scott Bauer, Rob Wynn, Jason Lucas, Liz, Gilberto Fragoso, Tracy Safran, Margaret Haber, Sherri de Coronado, Sana Din, Gisele, Samantha, David, Kim Ong

Discussion Points:

- Moving to hierarchial structure.
  - Search NCIT natively (no longer to use caDSR)
- Data needs to be avaiable for precise matches
- https://www.cancer.gov/about-cancer/treatment/clinical-trials/advanced-search
  - Larry noted that when finding clinical trials (search) the same stage could be listed several times in the dropdown.
  - Drugs and Drug family is problematic when determing what should come to the top of the list. Need to look at agent/therapy categories.
- Accrual coding
  - Need to understand how to capture the mapping data. (Meeting on Dec 12)
- Partial matching on terms
  - i.e. search for partial term and provide weighted results (relevancy ranking). More exact on top and then less weighted results.
- David to follow up on use of REST services (CTS and LexEVS REST).

Decision Points:

- Action Items:
  - Follow up on the use of REST Services (CTS, LexEVS REST)
  - Investigate the issue - Drugs and Drug family is problematic when determing what should come to the top of the list.
  - Determine mapping for accrual coding.

---

# Wednesday, December 6th, 2017

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 9:00 AM - 10:00 AM | 3-W-030 | **EVS Architecture**<br><br>**Discuss Potential of using a variety of architectures**<br><br>Proposed topics for discussion:<br><br>• **Micro services**<br>  ◦ Considerations:<br>  ▪ Determine how to synchronize data on the back-end. LexEVS DB and Triple Store need to be in sync when NCIt information (such as value sets) changes.<br>  ▪ Determine the potential of a loader that relies on SPARQL queries (after SPARQL query load, kick off LexEVS loader)<br><br>• **LexEVS integration with EVS Triple Store**<br>  ◦ Considerations:<br>  ▪ Determine use of triple store calls in parallel with LexEVS DB<br>  ▪ Determine performance improvements over LexEVS DB<br>  ▪ Determine what calls could be made to the triple store instead of LexEVS.<br>  ▪ Determine use of Stardog built in graph database.<br>  ▪ Determine performance considerations for hierarchy traversal for graph resolution.<br>• **Future implementation considerations** | Gilberto Fragoso<br><br>Kim Ong<br><br>Tracy Safran<br><br>Rob Wynne<br><br>Larry Wright<br><br>Margaret Haber<br><br>Sherri De Coronado<br><br>Bron Kisler<br><br>Systems team<br><br>John Cambell /Ruth Monterio users of the SQARL | MicroServiceProp.pptx<br><br>TripleStore.pptx |

Attendees:

Jason Lucas, Kim Ong, John Campbell, Larry Wright, Bron Kisler, Rob Wynn, Craig Stancl, Cory Endle, Scott Bauer, Kumar, Luba, Sherri De Coronado, Margaret Haber, Gilberto Fragoso, Liz, Denise, Tracy Safran

Discussion Points:

- Overview of EVS Architecture
  - Need to had value sets and mappings.
  - Would like to include all the ways that reports are created. (content channels - as separate slide)
  - Need to incude additional sources being loaded into SPARQL.
  - Consider adding channels from triplestore to LexEVS.

        - Add detail for Browser and assiciated dependencies
        - Change from SPARQL to TripleStore.
- Overview of LexEVS Stack
    - Scott noted that the Distributed LexEVS should be considered to be deprecated.
        - Serialization is the primary concern.
    - Tracy noted that the REST service is used more than the Java API.
    - Most NCI customers/users need simplified API.
    - Users of Distributed API
        - Matching Program for Editors
        - caDSR
    - Lucene recommendations
        - Move all Lucene code into the DAO layer (this is not complete today)

Decision Points:

- Action Items:
    - Update Architecture to include value sets and mappings.
    - Update Architecture to include all the ways that reports are created. (content channels - as separate slide)
    - Update Architecture to incude additional sources being loaded into SPARQL.
    - Update Architecture by adding channels from triplestore to LexEVS.
    - Update Architecture to add detail for Browser and assiciated dependencies
    - Update Architecture to change from SPARQL to TripleStore.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 10:00 AM - 12:00 PM | 10:00-10:30<br><br>3-W-030<br><br>11:00-12:30<br><br>TE-420<br><br>**(Can't fill gap 10:30-11:00)** | **EVS Architecture - Technical Discussion with Systems team**<br><br>**Discuss technical aspects of potentially using a variety of architectures**<br><br>Proposed topics for discussion:<br><br>• **Micro services**<br>  ◦ Considerations:<br>  ▪ Embedded Tomcat implementations<br>  ▪ Alternative web service platforms<br>  ▪ Container/Port clashes<br><br>• **LexEVS integration with EVS Triple Store**<br>  ◦ Considerations:<br>  ▪ SPARQL clients<br>  ▪ Docker options<br>• **Future implementation considerations**<br>  ◦ Java<br>  ◦ Python<br>  ◦ Node.js/javascript<br>  ◦ Others? | Systems team | |

**10:00 - 10:30**

Attendees:

Jason Lucas, Kim Ong, John Campbell, Rob Wynn, Craig Stancl, Cory Endle, Scott Bauer, Gilberto Fragoso, Tracy Safran

Discussion Points:

- EVS REST Service Overview
    - John provided overview of the REST service.
    - There exists a UI for loading and report writer.
- Anthill Pro project migration to Jenkins
    - Teams will need to work with systems team to ensure migration is successful to Jenkins.

**11:00 - 12:30**

Attendees:

Jason Lucas, Kim Ong, John Campbell, Rob Wynn, Craig Stancl, Cory Endle, Scott Bauer, Gilberto Fragoso, Tracy Safran, Margaret Haber, Larry Wright, Luba, Bron

Discussion Points:

- Overview of NCIt Browser Architecture
    - Need to include History Table
- LexEVS VS Resolution

    - Noticable performance enhancement for resolving value sets directly against the NCIt
    - Still need to demonstrate text search.
    - Value Sets would no longer need to be loaded.  The definition is implied in the source.
- Complete Architecture Overview
    - Architecture overview slides should be completed and shared on the wiki.
- MicroServices
    - Decomposing LexEVS services could start with MetaData service.
    - Coordination of many microservices are problematic.
    - Larry noted that if there is nothing immediately gained with microservices, then it is best to focus on other client/stakeholder needs.
    - Picklists were discussed (Picklist reference:  https://wiki.nci.nih.gov/x/vI9-AQ)
    - Microservice opportunities:
        - Browser Meta Data
        - Hierarchy
        - Node Graph (via Triple Store)
        - Node Set (via Triple Store)
        - Value Sets (via Triple Store)
    - Node Graph should continue to be investigated.

Decision Points:

- Action Items:
    - Update NCIt Browser Architecture to include History Table
    - Further evaluate LexEVS VS resolution by implementing the text search.
    - Complete Architecture overview slides and publish on the wiki.
    - Investigate Node Graph (microservice or other).

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 1:00 PM - 2:00 PM | 3-W-030 | **EVS Project Group Discussion (During regular call-in time)**<br><br>**Proposed topics for discussion:**<br><br>• **(High Level Overview) Discuss direct calls to NCIt for value sets**<br>  ○ Performance<br>  ○ Workflow<br>  ○ API Implications<br><br>• **Discuss Mappings and cross-walking coding schemes**<br><br>• **SwissProt, ICD-0-3, and MED-RT as the successor of NDRFT**<br>  ○ Associations from/to<br>  ○ Cross walking coding schemes<br>  ○ Loader considerations for Mesh, RxNorm | Kim Ong<br><br>Tracy Safran<br><br>Rob Wynne<br><br>Editor's Representative /Margaret Haber<br><br>Larry Wright<br><br>Sherri De Coronado<br><br>Gilberto Fragoso | Proposed Biomarker Terminology Sets_2017-12-05.pptx |

Attendees:

Tracy Safran, Rob Wynn, Liz, Larry Wright, Kim Ong, Jason Lucas, Gilberto Fragoso, Luba, Craig Stancl, Cory Endle, Scott Bauer, Bron Kiesler, Terry Quinn, Stephanie, Sherri De Corondao, Sana Din, Nick, Margaret Haber, Lori Whiteman, Jordan Li, george Chang, Erin Muhlbradt, Brenda Petty

Discussion Points:

- Value Sets Concerns from Editors
    - Value set Hierarchy
        - LexEVS represents the value set correct, but a way to display hierarchy needs to be provided.
        - Those not defined though concept/subset associations do not display correctly in the hierarchy.
            - These are loaded manually.
        - subsetOf relationship may be a way to accomplish hierarchy.
        - Further investigation of Hierarchy needed.
- NDF-RT/MED-RT
    - Margaret will need to follow up with FDA to further understand what is need.
- ICD-O-3.1
    - Maps are not well aligned with what is currently published.
    - ICD-O-3 source is not stand-alone.
    - Consider to first publish as a spreadsheet.

Decision Points:

- Action Items:
    - Further investigation of VS Hierarchy Display needed.
    - Further investigation of what FDA needs (NDF-RT/MED-RT)
    - Small group to meet to discuss ICD-O-3.1

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 2:00 PM - 3:00 PM | 3-W-030 | **NCI Systems Discussions**<br><br>**Proposed topics for discussion:**<br><br>• **Discuss CI and Docker Status/Roadmap**<br>  ◦ Discuss the current status of the Docker scripts used to build/test LexEVS components.<br>  ▪ Discuss the current NCI Docker images used in LexEVS tests.<br>  ◦ Discuss NCI's current status and future plans to use Docker.<br><br>• **Discuss Tech Stack Upgrades**<br>  ◦ Discuss DB upgrade: MySQL 5.6 vs. MariaDB<br>  ◦ Discuss migrating from Anthill Pro to Jenkins | Jacob and Systems team<br><br>Gilberto Fragoso<br><br>Rob Wynne<br><br>Tracy Safran<br><br>Kim Ong<br><br>Larry Wright<br><br>Margaret Haber<br><br>Sherri De Coronado<br><br>Q/A (Sana) | |

Attendees:

Tracy Safran, Rob Wynn, Liz, Larry Wright, Kim Ong, Jason Lucas, Gilberto Fragoso, Luba, Craig Stancl, Cory Endle, Scott Bauer, Bron Kiesler, Sherri De Corondao, Sana Din, Margaret Haber, Jacob, Kumar,

Discussion Points:

- Docker/CI
  - CentOS update nearly complete for Docker containers.
    - Cory to complete testing to ensure users are working correctly.
  - For deployments, Jacob noted that they would clone/copy the containers from dev and move up the tiers.
  - Tier specific properties would need to be passed to the container as environment variables.
  - There are no security model issues on the tiers (docker deployment).
  - Deployment using docker containers (built on CBIIT image of Tomcat and Java) to external sites is a concern due to the CBIIT dependencies.
    - Containers would need to be built with public dependencies - Tomcat and Java)
  - Dev system would need to be made available for Docker.
  - Future docker deployment would be one container up the tiers.
  - Systems team has started looking into Docker swarm.
  - Performance on VMs has historically been a concern.  This will need to be evaluated as we move forward with Docker.
- MariaDB
  - Is supported.
- Elastic Search/SOLR
  - Elastic Search and SOLR are supported by the systems team.
  - There are existing projects that use Elastic Search and SOLR
- Docker Boot
  - Would need to be evaluated.
- Node.js
  - Is supported.

Decision Points:

- Action Items:
  - Cory to complete testing of containers using CentOS.
  - Systems team to configure Dev system to be available for Docker and non-docker applications.
  - Investigate usage of MariaDB for JSON and graphing capabilities.
  - Jacob to provide versions of elastic search and SOLR.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 3:00 PM - 4:00 PM<br><br>(added meeting) | 3-W-030 | **User Group Discussion - caDSR**<br><br>**Continued discussion of current API**<br><br>• APIs: Java, REST (CTS2 or 3-store), SPARQL | caDSR - Philippa,<br><br>Vikram, Natalia, Rui | EVS CTRP REST API<br><br>• https://evsrestapi.nci.nih.gov/evsrestapi/swagger-ui.html#/evs-controller<br><br>CTS2 REST API<br><br>• https://lexevscts2.herokuapp.com/cts2<br><br>caDSR LexEVS API calls<br><br>• caDSR-EVS-Calls-NA.xlsx<br>• caDSR-EVS-Calls-VS.xlsx |

Attendees:

Jason Lucas, Kim Ong, Rob Wynn, Natalia, Tracy Safran, Craig Stancl, Cory Endle, Scott Bauer, Larry Wright, Luba, Sana, Philippa, Vikram, Natalia

Discussion Points:

- Team has reviewed class usage in semantic integration workbench (SIW)
  - ie. LexBigService / CodedNodeSets
- Upgrade considerations
  - Client Jars documented on Wiki. (need to share with the caDSR team)
  - Migrating to LexEVS 6.5.0
- Usage of CTS2/REST services has not been discussed by the caDSR team.
- Need to determine sources that are accessed by caDSR.

Decision Points:

- Action Items:
  - Natalia to send spreadsheet of class usage.
  - LexEVS team to send Migration link to caDSR team.
  - Determine gaps between LexEVS REST services and caDSR LexEVS API usage.
  - Provide assistance to caDSR team to parse results coming from CTS2.

---

# Thursday, December 7th, 2017

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 9:00 AM - 12:00 PM | 1-E-030<br><br>**(9:00-12:00)** | **API Services**<br><br>**Proposed topics for discussion:**<br><br>• **Overview of EVS REST API and future direction**<br>  ◦ Discuss what exists today.<br>  ◦ Discuss future plans.<br><br>• **Overview of LexEVS provided API services and future direction**<br>  ◦ LexEVS API<br>  ◦ LexEVS Remote API<br>  ◦ LexEVS REST API (CTS2-based)<br><br>• **Discuss a new REST service**<br>  ◦ Discuss user requirements for a new REST API<br>  ▪ Document the gaps from CTS2 REST API and user requirements.<br>  ▪ Explore how simple searches and searches with an identifier and how this could replace most of the remote API.<br>  ◦ Consider a new REST API to supplement CTS2 REST Services<br>  ▪ Propose examination of current CTS2 REST API<br>  ▪ Capture what doesn't work for the end users.<br>  ▪ Review gap between Remote API and REST services.<br>  ◦ Discuss surfacing terminology history in a REST service (NCIt and ULMS)<br>  ▪ Determine what history information is required by the user to be returned.<br>  • Discuss surfacing history of value sets in a REST service<br>  ▪ Discuss if this fits into LexEVS or is part of a new REST API | Kim Ong<br><br>Tracy Safran<br><br>Rob Wynne<br><br>End users such as CTRP, others<br><br>Larry Wright<br><br>Margaret Haber<br><br>Sherri De Coronado<br><br>Gilberto Fragoso | DiffCTS2_LexEVS.pptx |

Attendees:

Scott Bauer, Craig Stancl, Cory Endle, Gilberto Fragoso, Larry Wright, Lyubov, Tracy Safran, Rob Wynne, Kim Ong, Jason Lucas, Liz, Margaret, Bron Kiesler, Sherri De Coronado, Sana Din

Discussion Points:

- EVS REST API
  - CTRP and GDC are currently using the API.
  - https://evsrestapi.nci.nih.gov/evsrestapi/swagger-ui.html#/evs-controller
  - Working with other groups for API usage.
  - CTRP in the URL identifies the data being searched. Kim noted that in future it would be coding system and version.
  - Modeled specific for the CTRP usecases. Data in the response may have specific CTRP.
  - Larry would perfer to have a common API that could then be customized for specific users.
  - Suggested to perform a gap analysis to look at what caDSR current uses and what is available in the EVS CTRP API.
    - Will need to get parameter instances so we can know exactly what is used. (restrict to code/properties, etc)
    - This should be reviewed by NG, Tracy, Scott, etc and the caDSR team (Natalia, Vikram, etc.)
  - Agreement to have a joint working group to help build a combined REST service to get content from both LexEVS and TripleStore environments
  - Minimal testing has been completed (10 concurrent users).
- REST Service Future Direction
  - The REST service should not have application specific content.
  - There will be a need to consider input from users - caDSR, GDC, CDISC, CTRP
  - Sherri suggested that it may be worthwhile to generalize the existing CTRP REST service and then get user feedback.
  - Considerations for REST services:

        - Keep the CTRP API separate and build a more common API
        - Create a common API and provide specific convience APIs on top of the common API for specific customers.
        - Fill in the gaps (to be identified) with CTS2 based REST service. Create local extensions and look at opportunities to update the specification.
    - Customer needs will come first.
- LexEVS REST (CTS2) and LexEVS Remote Service gap analysis
    - Service Discovery - could be created
    - Paging and Iteration - no gap
    - Entity Search - gaps in search type (LexEVS 6.4 Search Algorithm Implementation Details)
    - Result Sorting - gaps exist
    - Custom Result Filtering - gap exists
    - CTS2 URI Read of Entiy - could be created
    - Code System Operations - gap exists
    - Entity Count - gap exists
    - Relational Operations - gap exists
    - Query Parameters - gap exists
    - Result Parameters - gap exists
    - Operation Functions - gap exists
    - Value Set Operations - gap exists
    - Value Set Search - gap exists

Decision Points:

- Action Items:
    - Perform a gap analysis to look at what caDSR current uses and what is available in the EVS CTRP API.
    - Assemble a joint working group to investigate a combined/coordinated REST service to get content from both LexEVS and TripleStore environments.
    - Identify additions to the existing EVS API to make it more useful to users.

| Time | Location | Topics | Participants | Resources |
|---|---|---|---|---|
| 1:00 PM - 3:00 PM | 5-E-030 | **Prioritization and debrief**<br>**Determine strategic direction and priorities** | Kim Ong<br>Tracy Safran<br>Rob Wynne<br>Larry Wright<br>Margaret Haber<br>Sherri De Coronado<br>Gilberto Fragoso | |

Attendees:

Tracy Safran, Jason Lucas, Rob Wynne, Larry Wright, Gilberto Fragoso, Kumar, Cory Endle, Scott Bauer, Craig Stancl, Lyubov, Margaret Haber, Bron, Sherri de Coronado, Sana Din

Discussion Points:

- Overall Impessions and themes
    - Mappings
        - Extract mapping from Meta efficiently
            - This would save effort of creating custom maps.
        - Extend the model of mappings (is it supported in CTS2) to support different types of maps.
            - ie ICD-O3, Meta and the logic (OR, AND..)
    - Diversity of paths through the architecture
        - Addition of triple store
        - Multiple APIs to address user communities.
        - flow described - input of data through the tooling and delivery
        - 2 views for documentation were identified:
            - Focus on what users needed
            - Focus on overall architecture (technical)
    - Remote API Roadmap
        - Determine the replacement for what of the API is needed. (based on gap analysis)
        - Determine current users and identify what is required for those users.
    - REST API
        - Federation using SPARQL or other tooling
        - Big Data will require that performance be addressed (caching, etc.)
            - Will support the annotation pipelines
        - LexEVS will need to provide REST services for content not available in TripleStore
        - Provide documentation to better help users
    - Report Writer
        - Support for other terminologies. LexEVS REST services/EVS REST Services
    - SWAGGER documentation

        - Differentiate from the general API and CTRP specific API
    - Microservices
        - integration of triple store to support/enhance LexEVS functionality
        - Hierarchy/Transitive Table support.
    - User needs to create a unified service
        - Discuss with stakeholders to gather requirements
        - Determine how to move forward based on the requirements (best practice)
            - Separate APIs
            - Combined APIs
        - Ensure the service simpilfies what the user needs to know about the technical implementation.
        - This could be several months of effort (across teams)
    - User Education - Enable users to use the services
        - Provide better documentation for end users.
        - Provide mapping of source into LexEVS or REST models so users can understand how to query the service in LexEVS
        - Review and update Wiki Organization
        - Provide documentation to aid in building applications that will utilize the services (REST, Java API, TripleStore/SPARQL/ftp)
        - Architecture diagrams to describe the 1) flow of data and 2) technical specifics.
        - Provide timeline for enhancement (REST Services), dates for deprecation, system deployments
        - LexEVS REST Code Migration Guide
    - Build and Deploy (Docker)
        - Continued development of Docker containers with the systems team.
        - Use of Node.js to be discussed with systems team.
        - Investigate use of Docker for data deployments.
        - Migrate CTS2 API from Heroku to NCI.

Decision Points:

- Action Items: