

VCDE Workspace Monthly Teleconference

Teleconference Information

Date: August 18, 2011

Time: 1:00 PM – 3:00 PM ET

Moderator: Brian Davis

Executive Summary

The meeting began with the discussion of possible dates, locations and agenda topics at the upcoming joint Architecture-VCDE WS face-to-face meeting as well as the Architecture and VCDE WS participant perspectives on the recent ICR-TBPT face-to-face meeting. John Speakman provided an update on the caBIG program. Larry Wright presented new functionality within the LexEVS browser. Finally, Paul Courtney provided a view from NCI DCCPS of metadata needed for cancer control and population sciences.

Meeting Materials

- [Discussion: Possible dates, locations and agenda topics for Arch-VCDE WS Face-to-Face meeting in October - Sherri deCoronado](#)
- Architecture and VCDE WS participant perspective of the ICR-TBPT Face-to-Face meeting - participants (no slides)
- [caBIG program update - John Speakman](#)
- New functionality in the NCI Thesaurus Terminology Browser - Larry Wright (no slides, see notes below for links)
- [Metadata for Cancer Control and Population Sciences - Paul Courtney](#)

Attendees

First Name	Last Name	Affiliation
David	Triglianos	Apelon Inc.
Pankaj	Agarwal	Duke University
Salvatore	Mungal	Duke University
Sherita	Ala'i	EMMES Corporation
Mary	McAdams	IMS
Jennifer	Stevens	IMS
Troy	Bleeker	Mayo Clinic Cancer Center
Robert	Freimuth	Mayo Clinic Cancer Center
Dale	Suesse	Mayo Clinic Cancer Center
Michael	Riben	MD Anderson
Baris	Suzek	Medical Knowledge Engineers, LLC
Sherri	De Coronado	NCI CBIIT
Gilberto	Fragoso	NCI CBIIT
David	Hau	NCI CBIIT
Diane	Reeves	NCI CBIIT

First Name	Last Name	Affiliation
Robert	Shirley	NCI CBIIT
John	Speakman	NCI CBIIT
Denise	Warzel	NCI CBIIT
Larry	Wright	NCI CBIIT
John	Dzak	Northwestern University
Rakesh	Dhaval	Ohio State University
Albert	Lai	Ohio State University
William	Stephens	Ohio State University
Virginia	Hetrick	Patient Advocate
Carolyn	Petersen	Patient Advocate
Mary	Cooper	SAIC
Tommie	Curtis	SAIC
Brenda	Maeske	SAIC
Musharaf	Rashid	SAIC - Frederick
Paul	Courtney	SAIC-Frederick
Larry	Brem	SAIC-Frederick
Roxanne	Martinez	TerpSys
Claire	Wolfe	TerpSys
Mukesh	Sharma	Washington University
Shelly	Niwah	Westat
Jim	McCusker	Yale University
Linda	Schmandt	University of Pittsburgh
Brian	Davis	3rd Millennium, Inc.
Aleksey	Kahn	Booz Allen Hamilton
Riki	Ohira	Booz Allen Hamilton

MEETING NOTES

[Discussion: Possible dates, locations and agenda topics for Arch-VCDE WS Face-to-Face meeting in October – Sherri deCoronado](#)

- Logistics:
 - Possible Locations: Washing DC (hotel), cancer centers – please let us know if you're interested in hosting, Estes Park, CO
 - Possible Dates: October 18-19 or 19-20. Very busy month, those dates are least objectionable.
- Possible Agenda:
 - caBIG Program Direction and Goals: Q& A, building & sustaining community
 - Interoperability paradigms: wide world view, needs of historic customers
 - Open Source Development: impact on Infrastructure
 - Q: Baris: On standards governance: will those be discussed?

- Sherri: yes, that's important and was also a topic of interest in the ICR-TBPT F2F.
- Baris: another topic of interest - what is VCDE & Architecture role in standards development and adoption?

Architecture and VCDE WS participant perspective of the ICR-TBPT Face-to-Face meeting – participants (no slides)

- Baris Suzek: Views of interoperability is different between Clinical Sciences and Basic Sciences. Could be a topic for our F2F. Clinical is interested in transfer of documents, Basic research is interested
- Virginia Hetrick: Medical centers can refer to same procedures differently. Medical test results sometimes don't indicate the normal range. Semantic issues need to be addressed. Sorting through by someone not familiar with the issue.
- Jim McCusker: Software Development paradigm: heard from Mike Smoot (of Cytoscape) how do we make interoperability a priority in an open software development world: either they will come out of use cases or organizations like NCI end up governing projects.
 - Life Sciences side: what sort of interoperability is needed. Very different on the clinical side. Heard comments like "just give me the data dictionary".
 - Models have to outlive data: data ages like fine wine and systems age like fish. If we want to make our data interoperable it has to outlive the systems.
- Sal Mungal: With development it will be PI driven. PIs need speed in integrating their CDEs into forms. All on board with governance, but need to streamline the process. This was not touched in ICR-TBPT F2F.
- Bob Freimuth: Really enjoyed the open discussion format and ability to raise issues in the presence of a diverse group of people. Mike Smoot from Cytoscape presented their open source development paradigm. Cytoscape core was developed by a tight development team that created plugin architecture so that others ("community") could create plugins. The effort is maintained by the core group. Challenges: support for backwards compatibility; have to be careful about the design architecture; turnover. We need to discuss some of these issues that we are likely to face at caBIG as we move to open source development.
 - Sherri: what are others models, do you think we should have a panel discussing their experience with different paradigms.
 - Bob: Yes, the different examples at ICR-TBPT were very helpful. None of the groups have gone so far as to add metadata and when we do it we're going to come up with new challenges.

caBIG program update – John Speakman

- Things have changed: caBIG no longer the only game in town, scope has broadened, numerous groups wanting to contribute inside and outside of NCI-

- supported community, weight of existing tools constraining from pursuing new items, limited government appetite for programs.
- March 2011 report of NCI BSA working group:
 - NCI response was presented on 6/20.
 - Scientific Advisory Group has been established: currently ad hoc committee of the BSA, will provide a bridge to scientific community, will prioritize projects and align to scientific goals, advise BSA and NCI Director about steps for improvement, first meeting held end of July.
 - Guiding Principles: retain stewardship of the development of community – driven technical standards and interoperability frameworks; certain funded projects through grants.
 - Technical Support: continue to provide support through KCs, continue to certify Support Service Providers, infrastructure
 - Changes to anticipate: shift to open source development: caBIG roles as stakeholder but not as controlling entity; focus on community-developed informatics; renew focus on relation with other NCI divisions
 - Not changing: building and maintaining the community, protection of investment to-date, overarching focus on interoperability and data sharing.
 - Q: Troy Bleeker: so an example – put out a grant to develop a feature within existing software?
 - John: Wouldn't put out a grant for a particular feature or existing software which should go to the community. The grant would be more for a novel idea.
 - Troy: Could you speak to motivation to contribute to a particular tool.
 - John: For example, someone is paying them to do it. Also, looking at other open source development community, you see people doing for the glory of it.

New functionality in the NCI Thesaurus Terminology Browser – Larry Wright (no slides, see links below)

Larry Wright gave a Demonstration of the new browser functionality for NCI Thesaurus and NCI Metathesaurus.

NCI Term Browser: <http://nciterms60-stage.nci.nih.gov>

NCIm Browser: <http://ncim60-stage.nci.nih.gov>

the VKC NCI Term Browser draft wiki page:

[https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/NCI Term Browser](https://cabig-kc.nci.nih.gov/Vocab/KC/index.php/NCI_Term_Browser)

and the EVS Use and Collaborations wiki page:

<https://wiki.nci.nih.gov/display/EVS/EVS+++Overview+of+Use+and+Collaborations>

with highlighted link to new detailed use profile:

[Supplemental Information](#)

- A Help page describes everything that's new.
- Search page selects the last selected dictionaries by default.
- Need to accept the license agreement
- The result shows all the dictionaries where the term was found

- “Shopping cart” allows to save items and export them later to XML or CSV
- Value Sets: what used to be on FTP servers are now available through the NCI Term Browser.
- Mappings tab was added. Map of term that was added. Link to meta-thesaurus.
- Q: Mike Riben: MD Anderson has a locally deployed browser, will this new functionality be available for locally deployed browsers?
 - Larry: Yes, as soon as it is ready we’ll be discussing.
- Q: Jim McCusker: OB and GEM ontology are both listed and they’re modules in their own right – is this being addressed (properties just as important as classes).
 - Larry: we do represent all the relationships through the loaders of ontologies. This requires a good deal of tweaking. The intent is that all the features be represented. If you think that things are not right or missing, please let us know.
- Jim and Mike Riben: Is there any way of getting this as a SPARQL end-point?
 - Larry: Please submit your requirements and we’re going to look at ways of getting them implemented.
- Tommie Curtis: What was the time period for the previous versions?
 - Larry: I will look into this and send out a response.
- Action Item: Everyone please look at the changes and provide feedback/additional requirements.
- Action Item: Jim McCusker and Mike Riben and Larry Wright follow up on requirement for SPARQL endpoint.

[Metadata for Cancer Control and Population Sciences – Paul Courtney](#)

- Metadata for Cancer Control and Population Sciences projects
 - GEM – grid enabled measures
 - NCCOR – looking at childhood obesity
 - PhenX – looking at phenotype and genotype associations
- Q: Sherri: What is your timeline for discussions for putting it together?
 - Paul: we started our conversation with PhenX just a month and a half ago and we’re in the beginning stages, but we do have meetings set up. We’re in the beginning stages of understanding how each system handles the information.
 - Sherri: Do you think you will have any terminology needs?
 - Paul: Yes, quite sure we will and we will reach out to you.
- Q: Dave Hau: Besides mapping between the three, are you also mapping to the CDC terminology and outside of cancer specific reporting.
 - Paul: That is going to be part of what we’re doing. It will be important to tap into existing and widely used standards.
 - Dave: We’re starting to work with CDC as well. It will really help for people using data for comparison effectiveness across wider scope.
- Q: Baris: How does this relate CFR (cancer family registry)?
 - Paul: Is managed out of the cancer epidemiology program and will be part of it.

- Baris: Georgetown assumed responsibility for CFR and has data dictionaries though not quite in format that presented.
- Action item: Paul and Baris to touch base after the call.

Next Meeting: Friday, September 9, 2011; 11:00 AM – 1:00 PM ET