NIH CDE Workshop - September 30, 2015

Exploring the Role of CDEs in NIH Data Sharing

Workshop Goals:

- Convene NIH community that is interested in Common Data Elements (CDEs) to support NIH-wide understanding of current activities and opportunities.
- Identify current barriers/challenges for the adoption and use of CDEs by NIH-funded researchers, both intramural and extramural.
- Identify possible ways to modify development, implementation, and use of CDEs to increase adoption and value to research
- Identify incentives and opportunities for involvement of relevant communities in CDE development, use, and re-use.
- Develop evaluation plans for CDEs to test their assumed utility.
- Identify opportunities to improve coordination in the development of CDEs for research use and in infrastructure for developing and making them accessible.
- Determine how best to support CDE activities in the context of BD2K

Expected Outcomes from the workshop:

- Generate and disseminate materials about existing CDEs and infrastructures.
- Build buy-in for CDEs across NIH, through improved understanding of what they are, when they might be used, resources that are available, and how to contribute so available CDEs are useful for their purpose/need.
- Strengthen and broaden the NIH-wide CDE community that can develop an NIH-wide governance process for CDEs (as many cross IC boundaries), to manage CDE harmonization.
- Make progress toward an NIH-wide strategy for development of CDEs and related infrastructure/tooling.
- Develop proposals for including CDEs in BD2K
- Identify at least one mechanism for providing evidence of the utility and value of CDEs for advancing science. Define a CDE interoperability challenge.
- Catalog ongoing use cases for data sharing that incorporates CDEs. When and how are CDEs used most effectively?

Planning Committee:

OD: Phil Bourne, Jennie Larkin, Leslie Derr, Angel Horton, Sonynka Ngosso
NLM: Betsy Humphreys, Mike Huerta, Lisa Lang, Jerry Sheehan
NCI: Warren Kibbe, Sherri De Coronado, Dianne Reeves, Denise Warzel
NCATS: Elaine Collier
NIEHS: Cindy Lawler

CDE Workshop Agenda

Location: Lindberg Room, National Library of Medicine (Building 38)

9:00 AM	Welcome and introduction to FAIR principles (Phil Bourne) 10 minutes + 5 discussion
9:15 AM	Introduction to the new structure and role of Scientific Data Council 10 minutes + 5 discussion
9:30 AM	Data Sharing: How CDEs might fit into NIH's Data Management and Sharing Policies (Dina Paltoo and JP Kim) 20 minutes
9:50 AM	Charge to Birds of a Feather (BOAF) Sessions (Betsy Humphreys) Preview of focus of each, and charge to the leads (15 minutes). Groups will sort and move to their designated rooms.
10:20 AM	BOAF Breakouts
12:00	LUNCH
1:00 PM	Concurrent Birds-of-a-Feather (BOAF) sessions continue
2:15 PM	Break
2:30 PM	Report-back from BOAF Groups and Discussion 10 minute presentation /10 minute discussion
4:10 PM	Identify Next Steps (Phil Bourne)
4:30 PM	Adjourn

NIH CDE Workshop Breakout Descriptions

1. **Role of NIH and ICs in CDEs**.

   Hosts: Dina Paltoo (OSP), Rick Moser (DCCPS/NCI)

   Purpose: Define best ways to expand and strengthen the focus and engagement in CDE community across NIH to all ICs and their Programs. Draft possible governance models including multiple levels.

   Questions: What is the optimal role for individual Programs, IC leadership, and trans-NIH collaboration? How can NIH increase understanding of what CDEs are and how they are and could be used? Should there be NIH-wide governance for CDEs since many CDEs cross multiple domains and thus multiple ICs? Should there be criteria for NIH wide endorsed CDEs separate from IC endorsement? Should NIH have a trans-NIH wide policy on use of CDEs? How should NIH approach the interface of NIH CDEs with standards efforts in other venues, e.g., ONC, PCORI, FDA, CDC, VA, and others.

   Target Audience: For those developing, managing, implementing CDEs within their ICs

   Location: Lindberg room, NLM, B38

2. **Operationalizing CDEs: Development, Use, and Harmonization.**

   Hosts: Ellen Werner (NHLBI), Lisa Lang (NLM), Dianne Reeves (CBIIT/NCI)

   Purpose: Describe how and when CDEs are implemented and used most effectively. Share lessons learned by ICs in process of developing, implementing, and harmonizing CDEs. Describe the use cases where CDEs are optimally implemented.

   Questions: What is best way to engage research investigator communities? Program staff? IC leadership? What is best way to start process if relevant domain CDEs do not exist or if they do exist but have gaps? What resources, models and best practices are available for developing, re-using and implementing CDEs within or across domains? Are there tools or methods that can facilitate process? What are different approaches to aligning or harmonizing data elements within or across CDEs? How should we categorize differences and similarities in CDE elements? What works best – requirements or encouragements? What incentives and barriers can help or hinder incorporating CDEs in programs? How can we best provide information on resources/models and best practices for developing, re-using, and implementing CDEs irrespective of domain?

   Target Audience: For those experienced in developing, managing, and implementing CDEs and those just getting started and have questions.

   Location: MORNING LHC 7th floor conference room (Building 38A, Lister Hill Center), AFTERNOON PSD Conf Room-1W20

3. **CDE Impact and Value Evidence Challenge.**

   Host: Joanne Odenkirchen (NINDS)

   Purpose: Identify controlled experiment(s) to provide evidence (positive or negative) of the utility/value/impact of CDEs on advancement of science and can be completed in 6 months to a year. Or define how NIH could challenge the research community to provide empiric evidence of value/impact/utility of NIH CDEs.

   Questions: What criteria will NIH accept as showing value and impact on science? What data are needed? Who has that data? What mechanisms can NIH use to incentivize the collection of such evidence? What is best mechanism for obtaining

evidence of value of CDEs to science? Can burden or facilitation of conducting research using CDEs be quantified? What will be most persuasive case of value of use of CDEs to researchers, NIH leadership, program staff?

Target Audience: For those involved and those new to CDEs at NIH and those in policy or leadership positions at NIH.

Location: NLM Conf Room B (on same floor as Lindberg Room), B38

4. **Interoperability CDE Challenge.**

Hosts: Kerry Goetz (NEI), Sherri de Coronado (NCI), Yaffa Rubinstein (NCATS)

Purpose: Identify a controlled experiment that can demonstrate the interoperability of more than three current NIH CDEs and can be done in 6 months to year. Outcome should include the criteria for demonstrating "interoperability" across and between CDEs. Identify a mechanism for CDE owners to demonstrate interoperability of their CDEs with other CDEs.

Questions: How do we obtain evidence of interoperability of existing CDEs? What is difference in interoperability and harmonization? What are the characteristics of a CDE that allow it to interoperate with other CDEs? Does interoperability encourage use of CDEs? What is value of interoperability of CDEs to investigators? To funders? How can this be captured?

Target Audience: For those already using CDEs.

Location: TSD Conf Room-B1W17A

5. Infrastructure and Tools for CDEs.

Hosts: Denise Warzel (CBIIT/NCI), Duc Nguyen (NLM-NCI), Matt McAuliffe (CIT)

Purpose: Share and review details of current infrastructure and tools that NIH CDE owners, managers, or aggregators are using. Define differences and similarities related to features; user and manager experience; and expansion, cost, and maintenance issues. Categorize the optimal set of features, usability, and functionality for each group of key users of resources and tools.

Questions: What are best/optimal approaches for NIH support for infrastructure and tools for the development, management, and dissemination of CDEs? How can NIH promote best practices and sharing of tools and information among multiple infrastructures? Are multiple infrastructure needed and how should they work together? How do resources and tools interface with other CDE infrastructure outside of NIH? How should NIH provide the resources, tools, and infrastructure to research communities that are or will use CDEs? Should this be a centralized resource or federated resource? How does NIH CDE infrastructure incorporate other standards that are required, such as from ONC, CMS, FDA, or domain specific technical standards (e.g. radiology)?

Target Audience: Owners and managers of infrastructure or tools used to develop, manage, or disseminate CDEs

Location: LO Conf Room-B2E-11

Attendee Roster

ADDS	Phil Bourne
ADDS	Jennie Larkin
CIT	Matt McAuliffe
NCATS	Elaine Collier
NCATS	Yaffa Rubinstein
NCI	Sherri de Coronado
NCI	Ashley Wilder Smith
NCI	Brenda Maeske
NCI/CBIIT	Denise Warzel
NCI/CBIIT	Dianne Reeves
NCI/CCCT	Gisele Sarosy
NCI/CCPS	Rick Moser
NCI/CCR	Liz Ness
NCI/CTEP	Mike Montell
NCI/DCEG	Melissa Rotunno
NCI/DCP	Lori Minasian
NCI/OD	Jean Zenklusen
NEI	Kerry Goetz
NHGRI	Ken Wiley
NHGRI	Erin Ramos
NHLBI	Mihailo Kaplarevic
NHLBI	Ellen Werner
NHLBI	Allison Wise
NIAID	Ashley Xia
NICHD	Steven Hirschfeld
NIDA	Udi Ghitza
NIDCR	John Prue
NIGMS	Jon Lorsch
NIGMS	Susan Gregarick
NIGMS	Ravi Ravichandran
NIH Library	Lisa Federer
NIH/CC	Lyubov Remennik
NIH/CC	Stephen Klagholz
NIH/OD/CF	Leslie Derr
NIH/OER	JP Kim
NIH/ORWH	Ching-Yi Shieh
NIH/OSP	Dina Paltoo
NIH/OSP	Carrie Wolinetz
NIMH	Greg Farber
NINDS	Joanne Odenkirchen
NLM	Betsy Humphreys
NLM	Clem McDonald
NLM	Mike Huerta
NLM	Jerry Sheehan
NLM	Duc Nguyen
NLM	Christophe Ludet
NLM	Liz Amos
NLM	Lisa Lang
NLM/CC	Vojtech Huser