

To: Vocabularies and Common Data Elements Workspace
From: Workspace Leads
Date: April 19, 2004
RE: **CDE development and processing model**

There is an urgent need in the cancer center community in general and, more specifically, in the caBIG Domain Workspaces, for engagement with the Vocabulary and Common Data Elements Workspace Working Group (V/CDE workspace) with regard to the development and administration of CDEs. Therefore, one of our first objectives as a working group is to come to consensus on definition of key working terms and the most optimal and relevant governance model for the creation, collection and curation of CDEs, within the caBIG context.

In order to facilitate the consensus process, we would like to offer potential scenarios for your consideration and to serve as the basis for discussion during our workspace teleconference on Wednesday (April 21 at 1-3 pm).

As a starting point, we need to define the key roles and responsibilities in this endeavor. There are at least four principal roles or groups of players in the process. They are the following.

1. **CDE Users:** These are the scientists, clinical study groups, data managers, cancer centers, application development teams (individuals, entities that require, and provide terms and concepts for construction of new CDEs for implementation in their cancer studies).

CDEs. *Common Data Elements* are data descriptors (metadata) used in NCI-sponsored research.

2. **CDE Administrators:** The administrators role/group is to evaluate new content and to provide editing and designation of CDEs, within and between CDE Contexts.

‘Context’ has a distinct technical meaning for our purposes. The most sensible proposal may be to define a ‘caBIG Context’ within the caDSR. Borrowed directly from the caDSR online introduction, the following definitional comments apply to this model.

Contexts. *Context* is one of the fields in the ISO/IEC 11179 CDE model. This field is used to distinguish different CDE development efforts that are being managed by different authorities. Each Context has a curatorial authority that manages the creation, editing, and designation of CDEs for the Context. The Context with the largest body of CDEs that have been approved for use in NCI clinical trials is managed by the NCI Cancer Therapy Evaluation Program (CTEP). "Designation" of a CDE means that a CDE from one Context is permitted for use in another; this is an important mechanism for minimizing the redundancy and duplication of CDEs across Contexts.

3. **CDE Harmonization Team:** In mid-2003, the NCI began a formalized *harmonization* program for CDEs in the caDSR. CDE review, acceptance and harmonization takes place across all domains. The CDE Harmonization activities are directed by the NCI Team.

- 4. caDSR Software Team:** caDSR software developers/analysts who manage system activities and make system improvements, e.g. bulk data loads, large-scale corrections. The center of this activity is the NCICB.

The activities of the caDSR Software Team are primarily the responsibility of the NCICB. However, the V/CDE workspace must propose the most suitable actors/workflow model for the CDE User and Administrator activities. Three scenarios are currently proposed for consideration. Please refer to the attached graphic depictions.

Figure 1: Scenario 1: Domain WS-Central

Figure 2: Scenario 2: Domain WS-Distributed

Figure 3: Scenario 3: V/CDE WS-Central

Scenario 1: Domain WS-Central

Scenarios 1 and 2 are similar. In these scenarios, the CDE Users are cancer center members residing and working in the 'domain' workspaces. Any of these members can propose and initiate CDE creation within the caBIG context. The CDE Context development and administration responsibilities will also reside in the domain workspaces, except that this work will be centralized, and carried out by a selected group of domain workspace designates. Finally, the acceptance, review and harmonization activities will be carried out by the V/CDE workspace in conjunction with the NCI.

Scenario 2: Domain WS-Distributed

Scenario 2 is a variant of Scenario 1. In this case, the cancer centers/data managers (CDE Users) will not only propose new CDEs for development, but also serve in the capacity of CDE Context Administrators. The V/CDE workspace along with NCI will assume the CDE review, acceptance and harmonization responsibilities. This scenario takes into account the potential for a V/CDE workspace bottleneck in the processing of new CDE material.

Scenario 3: V/CDE WS-Central

In Scenario 3, the cancer centers (the CDE Users) will be responsible for CDE proposal and initiation. The CDE Context Administration will be carried out by a designated group from within the V/CDE workspace. Finally, the review and harmonization activities will be handled by the V/CDE workspace in cooperation with the NCI.

It is important that we receive your feedback and any alternative proposals on the appropriate scenarios for the CDE development governance model, and attempt to reach a clear consensus as soon as possible, so that we can proceed with the caBIG-fundamental activities of this workspace.

In the coming weeks, we will consider alternatives for the review and integration of vocabularies into the caBIG context, in the same way we are currently addressing CDE development.

We look forward to hearing from everyone at our next meeting opportunity on Wednesday. In the interim, please feel free to send your comments to the V/CDE workspace Primary Contact- Christine Richardson (crichardson@kevr.com).

Christine Richardson
Mike Keller
Greg Eley
Peter Covitz