

# The Future of the caBIG<sup>®</sup> Clinical Trials Software Development Mission: Engaging the Open Source Community

Robert Annechiarico  
caBIG<sup>®</sup> CTMS Knowledge Center  
Duke Cancer Institute

## Disclaimer

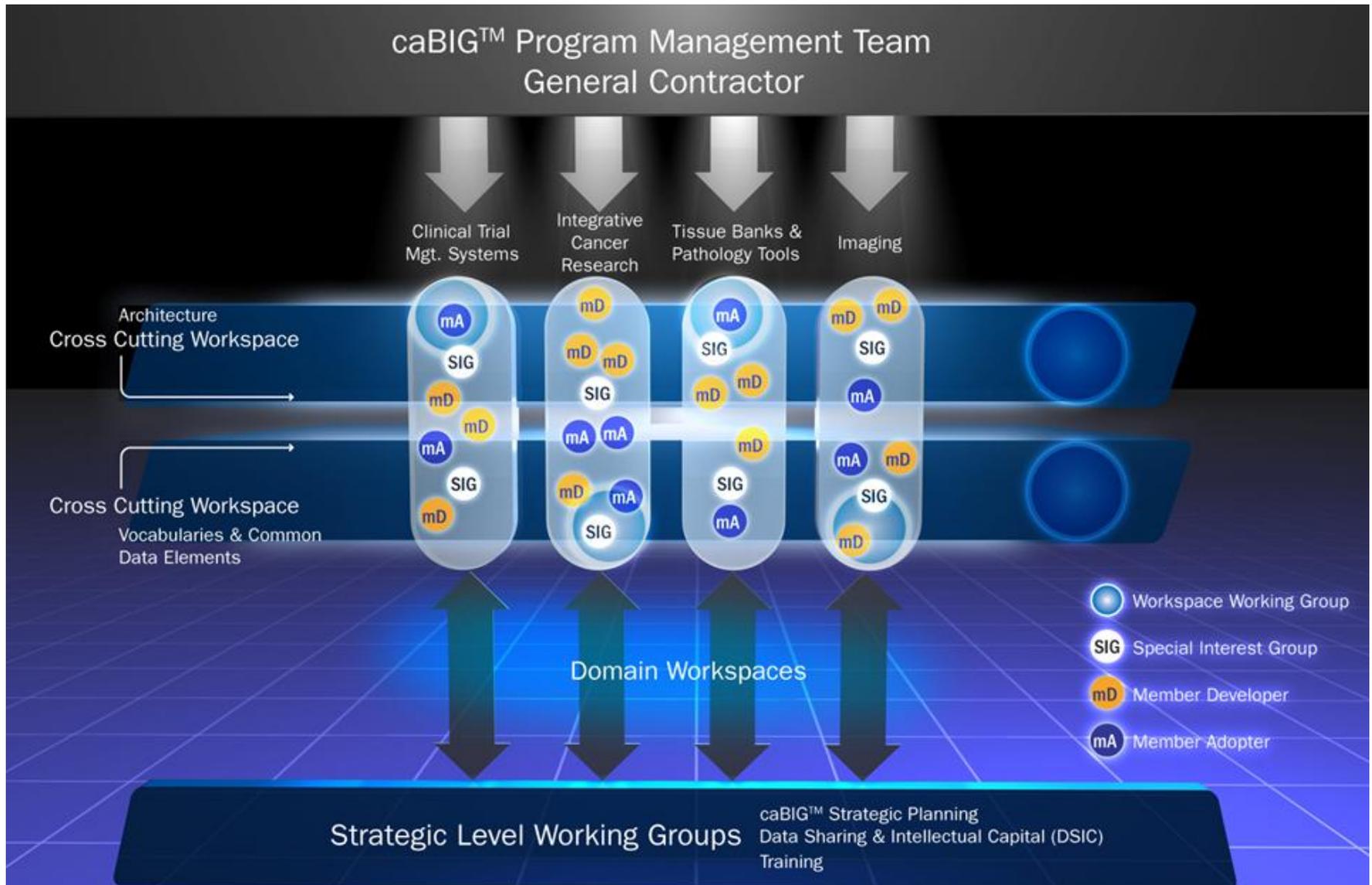
- I am not an employee of the National Cancer Institute
- The caBIG CTMS Knowledge Center operates under a contract from SAIC-Frederick, the prime contractor for NCI
- The opinions/observations expressed in this presentation are of the author(s) and not of NCI, SAIC-F, nor Duke University

# caBIG Refresher Course

- **Began about 8 years ago as a broad initiative to develop software and meta-data standards for the cancer research community**
- **Perform gap analyses and provide thoughtful response and recommendations for an informatics strategy, eventually taking the the form of software, metadata standards, and support**
- **“Buy-in” strategy was to completely engage the cancer clinical and informatics communities through:**
  - Direct supplemental funding to Cancer Centers for:
    - Software Development
    - Subject Matter Expert (SME) resources
    - Support adopt/adapt strategies at the local level
- **It was a financial “shot in the arm” for struggling cancer center informatics budgets**



# The caBIG Development Model



# Where are we today?

- **Tried to be all things to all people**
- **Some of the *wins* in CTMS**
  - cancer Adverse Event Reporting System (caAERS)
    - Comprehensive S/AE reporting tool
    - Logical support for caAERS-adEERS
  - Patient Study Calendar
  - Harmonized eCRF Project (win-pending)
  - caDSR (Common to all workspaces)
- **NCI Board of Scientific Advisors (BSA) Report (3/2011)**
  - Premise of the caBIG program sound, but,
    - NCI should not be in the software development business, per se
    - Instituted governance and review mechanisms (in progress)

*In other words NCI put the brakes on to fall-back and evaluate/recommend*

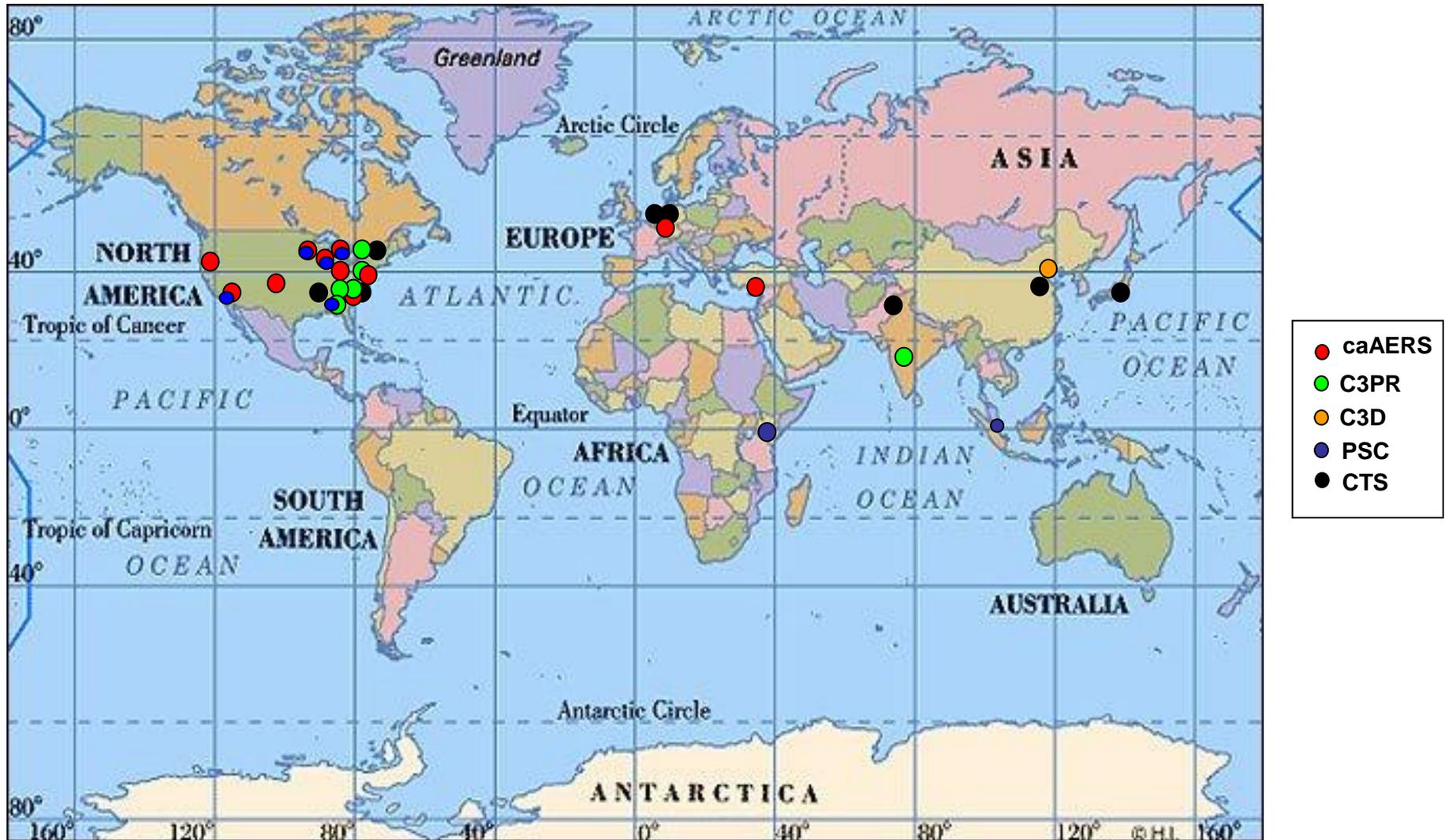


# caBIG compatibility

## Common data element (CDE) details

Legacy	Bronze	Silver	Gold
No Structured metadata is recorded	<p>Data element descriptions have sufficient detail for a subject matter expert to unambiguously interpret</p> <p>Data elements are built using controlled terminology</p> <p>Metadata is stored and publicized in an electronic format that is separate from the resource that is being described</p>	<p>Common Data Elements (CDEs) built from controlled terminologies and according to practices validated by the VCDE workspace are used throughout.</p> <p>CDEs are registered as ISO/IEC 11179 metadata components in the cancer Data Standards Repository (caDSR)</p>	<p><i>All features of Silver, plus:</i></p> <p>Common Data Elements (CDEs) designated as caBIG Standards by the VCDE workspace are used.</p> <p>Metadata is advertised and discoverable via the caBIG grid services registry</p>
<b>Examples</b>			
Free-text pathology reports	GeneOntology from GO website	NCI Thesaurus	GeneOntology registered in EVS NCI Thesaurus

# CTMS Workspace Reach



*That we know of...*

## Transition Planning: Open-source Software Development Initiative (OSDI)

- **caBIG is still part of the plan**
  - Critical elements preserved
- **Aggressively Engage the OS community**
  - Potential to provide “seed” money
- **Pilot program under way**
  - **caLIMS (cancer Lab Info Management System)**

# Why a caBIG<sup>®</sup> Open Source Development Initiative?

- **Informatics tools need to keep pace with the science they support**
  - New platforms for molecular analysis emerge regularly and researchers need to be able to apply state-of-the-art information technology to their research
- **Intense community interest in sharing in the development and direction of the tools and infrastructure on which they are dependent**
  - Within and beyond cancer research
    - Ex. Children's Health Study adoption of PSC
  - Within and beyond US academic research institutions

***Meeting the challenge of rapidly creating robust informatics tools to support these evolving needs will benefit from an open, collaborative approach to software development.***

# Aren't caBIG<sup>®</sup> tools already open source?

- **Yes, they are. But...**
  - there is not a consistent mechanism for accessing the current development version of the code or for contributing custom code back to the main distribution of the tool
  - customizations that may be freely offered and broadly useful are not readily available to other groups.
  - organizations that have a locally customized version of the code who want to upgrade to the latest version, cannot readily and therefore will be out of synch with the regular releases

***Providing a means to easily connect these community contributors to the codebase and a means for them to participate in the governance about what ends up in the supported code trunk will provide a more transparent and consistent means to keep all of these in alignment.***

# The Golden Age of Open Source?

<http://blogs.the451group.com/opensource/2010/08/09/the-golden-age-of-open-source/> (Matthew Aslett)

- **Open Source 1.0** - software developed by communities of individuals and academia, with key projects including Apache, GNU, the Linux kernel, PostgreSQL and BSD Unix.
- **Open Source 2.0** - vendors beginning to engage with those existing developer communities
- **Open Source 3.0** - vendor-dominated open source development/distribution projects: MySQL, JBoss, et al
- ***Open Source 4.0*** - *emphasis is on collaboration and community rather than on control*

# Open Source Development Initiative *Essential Elements*

- **Community** - A community of interested participants must be built and *maintained*. A key component of this is identifying and implementing the appropriate incentives for participation.
- **Governance** - Perhaps the most critical consideration. A successful open development project cannot be a free for all. There is a need for a decision making structure, whether formal or informal, that makes strategic decisions depending on changing user requirements and other factors.
- **Licensing** - Software contributed to the project should be provided under a common open source license.
- **Software Infrastructure** - To support a geographically dispersed community of developers, projects require a version control system, bug and task management, and testing tools that meet the needs of the project.
- **Communication Tools** - Includes tools for community developers to communicate with one another (listservs, forums, wikis) as well as mechanisms to communicate with the user community
- **Testing and Certification** - An automated build and testing environment will contribute to ensuring the quality of contributed code; contributors should also provide tests with their contributions.

# But what does it Cost, REALLY?



Free as in puppy...



Not as in Beer!

- **Installation and maintenance**
  - **Hardware**
  - **Systems Personnel**
  - **Training**

# Resources for the caBIG® OSDI - Wiki



National Cancer Institute

U.S. National Institutes of Health | [www.cancer.gov](http://www.cancer.gov)



NCI Wiki



New Account



Helpful Tips

Dashboard > [caBIG® Community Code Contribution and Integration WG](#)

Browse ▾

[Log In](#)

> [caBIG® Open Development Initiatives](#)



## caBIG® Open Development Initiatives

3 Added by Ann Wiley, last edited by Juli Klemm on Jun 24, 2011 (view change)

Tools ▾

<https://wiki.nci.nih.gov/x/VASmAQ>

### Overview

Welcome to the caBIG® Open Development Initiatives Wiki. This is the home for information about ongoing activities in the program to promote and extend community development efforts across the caBIG community.

- If you are a first time visitor to this site, we recommend you visit [The Cancer Community and the Open Source Community](#) page
- For information about current open development processes for caBIG projects and how you can contribute TODAY, please visit our [Current Open Development Opportunities in caBIG®](#) page
- The [caBIG Community Code Working Group](#) has researched models of open source governance and incentives for open development:
  - [Core development team governance](#) ("Google-like model")
  - [Community governance](#) ("Apache/Mozilla-like model")
  - [Incentives for open development](#)
- There are any number of open development initiatives and information related to their best practices, workflow, governance and tools. We are compiling a [list of resources](#) that have been found to be most useful to the current CBIIT effort.
- Want to provide feedback or thoughts on open development in caBIG®?
  - We welcome your Comments on any page or post a Blog - see [Recent Blog Posts](#).
  - Start a discussion thread on the [caBIG® open development forum](#)

Powered by Atlassian Confluence 3.4.6, the Enterprise Wiki | [Report a bug](#) | [Atlassian News](#) | [Contact Administrators](#)

cancer Biomedical Informatics Grid®



# Resources for the caBIG® OSDI – Discussion Forum

NATIONAL CANCER INSTITUTE National Cancer Institute U.S. National Institutes of Health | www.cancer.gov

caBIG Knowledge Center A part of the Enterprise Support Network Login create account account help

Home Knowledge Centers Discussion Forums Bugs/Feature Requests Development Code Repository

caBIG-KC.nci.nih.gov - General Forum NCI - caBIG General Forum Search... Search Advanced search

Board index < Open Development Discussion

Open Development Discussion <http://tinyurl.com/65bu8kp>

NEWTOPIC\* Search this forum... Search 3 topics • Page 1 of 1

Topics	Replies	Views	Last post
 "What's in it for me?" considerations by klemmj » Mon Jul 11, 2011 7:20 am	0	1	by klemmj » Mon Jul 11, 2011 7:20 am
 Software Testing in an Open Development Community by klemmj » Wed Jun 15, 2011 3:44 pm	2	5	by klemmj » Wed Jun 29, 2011 11:23 am
 Inviting feedback on the caBIG® Open Development Wiki by klemmj » Thu May 26, 2011 4:24 pm	0	4	by klemmj » Thu May 26, 2011 4:24 pm

# Be influential – Get Involved TODAY!



National Cancer Institute

U.S. N



NCI Wiki

Dashboard › caBIG® Community Code Contribution and Integration WG › caBIG® Open Development Initiatives

Browse ▾

› Current Open Development Opportunities in caBIG®



## Current Open Development Opportunities in caBIG®

Added by [Juli Klemm](#), last edited by [Beth DiGiulian](#) on Jul 15, 2011 ([view change](#))

### Contributing to caBIG® Projects

As the next generation of caBIG open development evolves, it is important to emphasize that a number of tools have processes in place today for accepting contributions from users. We enthusiastically welcome and encourage you to work with these teams.

- [Contributing to caBIG® Projects](#)
  - [caArray](#)
  - [caIntegrator](#)
  - [caGrid](#)
  - [caTissue](#)
  - [geWorkbench](#)
  - [GenePattern](#)
- [Contributing to Other Projects Sponsored by the NCI Center for Biomedical Informatics and Information Technology](#)
  - [LexEVS](#)
  - [Cancer Data Standards Repository \(caDSR\)](#)

# Acknowledgements

**Mohammad Farid, Duke University  
CTMS KC Operations Manager**

**Kimberly Johnson, Duke University  
CTMS KC Cooperative Group Liason/SME**

**Madeline Carroll, RN, Duke University  
CTMS KC SME**

**Juli Klemm  
Director of the NCI/CBIIT OSDI**