caBIG Program Update

MESSAGE FROM NCICB

Every day I'm amazed at the participation and enthusiasm surrounding the caBIG initiative. This truly is a grass roots effort and we know that this vision could not be achieved without your support. Each and every one of you play a critical role in this initiative and, on behalf of the NCI, I thank you for the time and effort you have volunteered thus far to keep the momentum going while we continue to actively address contracts issues.

The contracting process has made significant progress since participating Cancer Centers received the initial contracts basic agreement in late March. In response to questions and needs raised by various Centers; NCI, Govworks, and Booz Allen Hamilton have restructured the current contracts basic agreement with Centers. These changes clearly reflect the various roles that the Government, Booz Allen, and the Cancer Centers play in this initiative and align the legal responsibilities accordingly. Once these agreements are signed, we will quickly move to the next stage of issuing task orders for caBIG project activities based on Statements of Work.

caBIG project activities in the various Workspaces and Working Groups are moving along at a steady pace, and the groups are poised to commence project activities as soon as task orders are issued. Throughout these initial stages of caBIG, there have been requests from the caBIG community for more detailed guidance on caBIG compatibility. To address these requests, a caBIG Compatibility Guidelines document was developed and released to the community on June 18, 2004. These Guidelines are meant to provide specific direction to caBIG participants and other interested members of the cancer research community about the requirements for software that will interact with the emerging Grid. This is an exciting early step in the development of a truly integrated bioinformatics environment.

The remainder of this newsletter will be dedicated to presenting an overview of the caBIG Compatibility Guidelines document. As with other issues of the caBIG Program Update, helpful references and information can be found at the end of this issue.

Once again, I extend my thanks to all of you who continue to help us move forward with caBIG. Your commitment is invaluable.

Sincerely,

Kenneth H. Buetow, Ph.D.

Director

NCI, Center for Bioinformatics

Contact

Mark Adams, PhD

adamsm@mail.nih.gov

What is caBIG Compatibility?

NCICB and the caBIG Project Team have developed preliminary guidelines outlining the requirements for caBIG compatibility. This document is a response to the caBIG community's request for guidance on what it means to be caBIG compatible and it reflects the Team's thinking that went into the founding and development of the caBIG pilot.

These guidelines represent the initial steps towards a community standard for interoperable biomedical informatics software. In addition to providing guidelines for caBIG participants and interested stakeholders, this document provides a set of definitions that caBIG participants can use to measure the maturity level of potential caBIG applications. A sample scorecard is included which can be used to evaluate existing applications and the steps necessary to make an existing vendor or custom application caBIG compatible, and to create new applications at a specified level of maturity.

The caBIG Compatibility Guidelines document is a living document that will continue to evolve as emerging guidelines and definitions are determined by the caBIG community. In particular, the caBIG Cross Cutting Workspaces will play a key role in the ongoing development of these guidelines. All stakeholders (Developers, Adopters, Users and other contributors) are encouraged to familiarize themselves with the requirements for caBIG compatibility. It is hoped that the caBIG Compatibility Guidelines will begin to provide clarity for the development of caBIG-compliant applications, and will provide all interested parties, including commercial software vendors, with what is required to create software that can interoperate with the emerging caBIG infrastructure.

caBIG Principles vs. caBIG Compatibility

How are caBIG's principles different from compatibility? The caBIG Principles are the core values of the program. These include: open source, open development, open access and federated data. caBIG principles apply to ALL caBIG projects funded by NCI. However, it is not a requirement to meet ALL caBIG principles in order to be minimally caBIG compatible. Unfunded, existing custom or vendor solutions can be minimally caBIG compliant without following all of the caBIG Principles.

How are the Compatibility Guidelines structured?

The guidelines define and describe caBIG compatibility in terms of specified maturity levels: Legacy, Bronze, Silver and Gold. The overall goal of caBIG is to have all caBIG compliant applications eventually attain the Gold Level of maturity. It is anticipated that, in the near term, new development activities undertaken as part of caBIG will meet the Silver Level of compliance. Bronze Level maturity will be applicable for meeting the minimum level of maturity for caBIG compliance. This level of maturity is a transitional state along the pathway to the Silver and Gold Levels of compliance. In the immediate term, it may be necessary to retrofit existing legacy systems to achieve the Bronze Level of compliance. The natural progression will be to ensure that those existing applications meet the finalized Gold Level requirements once they have been fully defined by the caBIG community.

caBIG Principles

Open Source

Open Access

Open Development

Federated

The four criteria that define the maturity model for caBIG compliance are the following: Interface Integration, Vocabularies/Terminologies & Ontologies, Data Elements, and Information Models. The main characteristic that defines the maturity of a software application is its level of interoperability; not just in the ability to exchange information, but also its ability to use that information. Interface Integration deals with the ability to exchange the information, whereas the remaining three criteria (Vocabularies/Terminologies & Ontologies, Data Elements, and Information Models) deal with the ability to use the exchanged information.

The "Legacy" Maturity Level is not considered caBIG compliant because there is no interoperability present between the Legacy system and another system. A given Cancer Center using Legacy Level software would not be able to transmit or exchange their information to other Cancer Centers because programming interfaces to the system do not exist. Additionally, there are no external standard vocabularies, common data elements, or information models that would allow for standardization and use of the information.

The "Bronze" Maturity Level meets the minimal set of requirements for an application to be caBIG compliant. By using standardized controlled vocabularies that are publicly accessible, as well as local terminologies, common data elements (CDEs) from a repository shared across caBIG applications, provides capabilities to exchange key information electronically with other Cancer Centers. The Bronze Level does not require meeting of key caBIG principles, namely open source, open access, open development and federated data. It does, however, allow for exchange of selected, key information in common formats with common meaning.

The "Silver" Maturity Level, in addition to meeting all the Bronze Level requirements, does need to fully meet the caBIG principles of open source, open access, and open development. A key differentiator from the Bronze Level is that "Silver" requires well-defined Application Programming Interfaces (APIs) and/or standard messaging systems based on common information models (like HL7 v3) that allow for an increased ability to exchange information among Cancer Centers. The application will use standard vocabularies approved by public standards bodies and/or the caBIG Vocabulary & Common Data Elements Workspace; and a common standard electronic representation for CDEs such as ISO 11179. The information models will be defined in a standard modeling language such as UML. These three facets (Vocabularies, CDEs, and Models) will provide increased ability to use information exchanged among Cancer Centers. It is anticipated that new development activities will meet the Silver Level of compliance until Gold standards are more clearly defined.

The "Gold" Maturity Level will be the maximum set of requirements for an application to be considered caBIG compliant, and will be fully defined as guidelines and definitions evolve through the efforts of the caBIG community. In particular, unlike "Bronze" or "Silver", the "Gold" standard fully meets the caBIG principle and goal of federated data. The application will not only comply with the Silver requirements, but it will also be interoperable with the data grid architecture to be defined by the caBIG community.

caBIG Maturity Levels

Legacy

Bronze

Silver

Gold

Next Steps for the Compatibility Guidelines

The caBIG Compatibility Document is a living document that will evolve through the efforts of the caBIG community. It provides guidelines that can be used to determine if a given application is caBIG compatible and to inform the development of new applications. Coordinated through the Cross Cutting Workspaces (Architecture Workspace and Vocabularies & Common Data Elements Workspace), future versions of the guidelines will be a collective effort by the caBIG community, and will reflect the needs and aspirations for the program. The Cross Cutting Workspaces, NCICB and the caBIG community, will work together to further define the "Gold" Level of standards.

Where does the Compatibility Guidelines document live?

The Compatibilty Guidelines are posted on the caBIG Website at:

http://caBIG.nci.nih.gov/caBIG/cabiglfs/caBIG_Compatibility.pdf

How can I provide feedback on the Compatibility Guidelines?

There are a number of ways you can provide feedback. You may post comments to the caBIG Compatibility forum at http://ncicbforums.nci.nih.gov/cabigforum or by contacting your Workspace or Working Group Coordinator. You may also email comments to Arumani Manisundaram at manisundaram_arumani@bah.com.

caBIG Information and Resources

There are many ways to access additional information and resources on the caBIG initiative and activities. Following is a list current resources:

- caBIG Website:

  http://caBIG.nci.nih.gov

- caBIG Workspaces:

  http://caBIG.nci.nih.gov/workspaces

- caBIG Strategic Level Working Groups:

  http://caBIG.nci.nih.gov/caBIG/working_groups

- Inventory of caBIG infrastructure, applications and datasets:

  http://caBIG.nci.nih.gov/inventory

- caBIG Interactive Overview (multi-media presentation on caBIG):

  http://www.nci.nih.gov/directorscorner/caBIG

- National Cancer Institute Center for Bioinformatics (NCICB) homepage:

  http://ncicb.nci.nih.gov

In addition to these more general resources, a specific caBIG Forum has been established to provide a resource through which all of the caBIG participants can communicate and coordinate with other caBIG members. The forums are readable by anyone, but posting is limited to registered participants.

The caBIG forums can be reached at the following URL:

  http://ncicbforums.nci.nih.gov/cabigforum

If you do not yet have a login, and you are a registered participant in the caBIG project, you can get one by sending email to Leslie Derr at derrl@mail.nih.gov.

Raising Questions or Concerns

If you are a caBIG participant and you have any questions about caBIG Workspace and Working Group activities, please consider the following avenues to raise your questions:

- The caBIG Forum at

  http://ncicbforums.nci.nih.gov/cabigforum

- Workspace or Working Group Meetings

For specific, directed questions that you feel would be better addressed individually, please feel free to reach out to your respective Workspace/Working Group Coordinator from the caBIG Project Team, or Mark Adams at adams_mark@bah.com.