

Statement of Work

For

Columbia University – Herbert Irving Cancer Center

Integrative Cancer Research (ICR) Workspace Developer

caWorkbench Grid Enablement

1. Background

The National Cancer Institute (NCI) will deploy an integrating biomedical informatics infrastructure, the cancer Biomedical Informatics Grid (caBIG), to expedite the cancer research community's access to key bioinformatics platforms. In partnership with the cancer research community, the NCI is creating a common, extensible informatics platform that integrates diverse data types and supports interoperable analytic tools. This platform will allow research groups to tap into the rich collection of emerging cancer research data while supporting their individual investigations.

This Integrative Cancer Research Workspace will provide tools and systems to enable integration and sharing of information among cancer researchers. These tools will facilitate the integration of data not only from different centers, but also data of different types, thereby enabling translational and integrative research. These tools also provide for the integration of clinical and basic research data. The Workspace is tasked to develop a well-documented and validated toolset for use throughout the cancer research community. Workspace activities will include platforms and standards to facilitate the sharing of datasets and repositories, and those appropriate for testing the caBIG infrastructure are being asked to participate. A major goal of this workspace will be a demonstration of how a shared informatics platform can allow a comprehensive, federated grid of information to be made available to the cancer research community.

The Columbia University group is currently developing caWorkbench 3.0 under the ICR workspace and in cooperation with the caArray group. caWorkbench3.0, an open source bioinformatics platform written in Java, makes sophisticated tools for data management, analysis and visualization available to the community in a convenient fashion. It evolved from a project which was originally sponsored by the National Cancer Institute Center for Bioinformatics (NCICB). Some of the most fully developed capabilities of the platform include access to caArray repositories through



use of the MAGEOM API, microarray data analysis, pathway analysis and reverse engineering, sequence analysis, transcription factor binding site analysis, and pattern discovery. The work proposed in this current SOW builds on Columbia's Year 1 caBIG project and aims at further integrating caWorkbench into the caGRID framework by providing a simple user interface to connect interoperable grid-enabled and client-side components into useful workflows.

caWorkbench 3.0 incorporates several algorithmic components, implementing methods for the analysis gene expression and sequence data. At present such components are running either as client side routines or as server components accessible through a SOAP interface. In this project, a subset of these components will be transformed into caGRID analytical services. In addition, caWorkbench will be extended so that it can discover and use them. Further, Columbia will design and implement specific use cases that will demonstrate the transparent utilization of grid resources via the caWorkbench platform, based on available grid services. Finally, Columbia will develop a language (caSCRIPT) that uses the caGRID 1.0 workflow infrastructure to support the creation of scripts for the assembly of complex workflows from a set of individual components. In this effort, Columbia will coordinate closely with the caGRID 1.0 development team to act as adopters of the caGRID 1.0 workflow framework.

2. Objectives

The objective of this project is to complete the caGrid-enablement of the following caWorkbench components: ARACNE, SOM (self-organizing maps) and Hierarchical Cluster Analysis. A new caWorkbench component (caGRIDExplorer) will be developed to offer functionality that allows users to choose which analytical service to execute; use cases will be developed to showcase realistic uses of the grid, involving acquiring data from remote sources and submitting them to grid services. Finally the development of caSCRIPT will permit the codification of a select subset of these use cases into scripts that can be saved and executed multiple times. Using the underlying machinery of the caGRID 1.0 workflow framework, caSCRIPT will control the workflow logic among client-side components and adapters, the latter providing access to either server-side or grid-enabled components. Thus caSCRIPT will allow the transparent creation of grid-enabled workflows without exposing the user to the complexity of the grid. This effort will involve close collaboration with the caGRID 1.0 team.

Measurable objectives:

- 1 Setup and test a caGRID node, to be used for development, deployment and testing of the caGRID-ified analytical services.
- 2 Develop a Functional Requirements and Design Specification document, reviewed by Adopter, and the Architecture and Vocabularies and Common Data Elements (VCDE) Workspaces.
- 3 Create a Risk Management Matrix for the project
- 4 Document a Test Approach that ensures requirements are met



- 5 Write code to achieve the following milestones:
 - Transformation of the following caWorkbench components into caGRID compliant analytical services:
 - ARACNE
 - Self Organizing Maps
 - Hierarchical clustering
 - Development of caGRIDExplorer, a caWorkbench component supporting browsing of caGRID Indexing services and selection of analytical services to execute.
 - Development of caSCRIPT, a scripting language that will utilize the caGRID 1.0 workflow framework to facilitate the assembly of complex workflows from a set of individual components.
- 6 Develop use cases showcasing how caWorkbench can facilitate access to data and analytical services on caGRID.
- 7 Execute on Test Approach
- 8 Write a lessons Learned document

3. Scope of Work

The overall role of the Developer in this project will be to create a Functional Requirements and Design Specification document for the caWorkbench Grid Enablement Project based on use cases, and to implement the requirements as high-quality software

The following activities are for the Developer. See Section 5 for descriptions of acceptance requirements.

Task 1 – Project Management – Cancer Center Project Task Lead will manage Cancer Center level project activities.

- Task 1.1 - Use of General Contractor provided online tools for tracking of project deliverable, i.e., cancer Management Portal (caMP)
- Task 1.2 - The Cancer Center Project Task Lead will plan regular (and when needed ad hoc) communications to share project information. Communications may include face-to-face meetings, teleconferences, videoconferences or use of the caBIG website and forums.
- Task 1.3 -The Cancer Center Project Task Lead will generate monthly status reports and submit to the General Contractor not later than the 5th working day of each quarter.

 **Artifact: Monthly Status Report**



Task 2– Project Activities

- Task 2.1 - Risk Management Matrix: Identify the potential risks in the project and document the plan for managing these risks.

📁 **Artifact: Risk Management Matrix**

- Task 2.2 – Develop Use Cases for the caWorkbench grid enablement project

The use case document will be developed in two phases, draft and final. The draft document does not require Adopter or Cross-Cutting WS sign-off prior to submission to NCI senior leadership for approval. However the draft use case document will be circulated to the appropriate Adopter, Architecture and VCDE Workspace Reviewers to initiate the review process of this document. The review process will be iterative in nature, resulting in the final reviewer-approved version of this document

- Task 2.2.1 Develop Draft Use Cases

📁 **Artifact: Draft Use Case document**

- Task 2.2.1 Develop Final Use Cases

📁 **Artifact: Final Use Case document, approved by Adopter, Architecture and VCDE Workspaces**

- Task 2.3 – Develop Requirements and Specifications Document for the grid enablement of selected caWorkbench analytical services - ARACNE, SOM (self-organizing maps) and Hierarchical Cluster Analysis, and a new caWorkbench component (caGRIDExplorer) to that allows users to choose which analytical service to execute.

The Requirements and Specifications Document will be developed in two phases, draft and final. The draft document does not require Adopter or Cross-Cutting WS sign-off prior to submission to NCI senior leadership for approval. However the draft Requirements and Specifications document will be circulated to the appropriate Adopter, Architecture and VCDE Workspace Reviewers to initiate the review process of this document. The review process will be iterative in nature, resulting in the final reviewer-approved version of this document

- Task 2.3.1 Develop Draft Requirements and Specification Document

📁 **Artifact: Draft Requirements and Specification Document**

- Task 2.3.2 Develop Final Requirements and Specification Document

📁 **Artifact: Final Requirements and Specification Document, approved by Adopter, Architecture and VCDE Workspaces**

- Task 2.4 - Test Approach: Document what approaches will be taken to assure requirements met. Where appropriate, the Test Approach Document should define a validated test data set against which software will be regularly evaluated.



The test approach document will be developed in two phases, draft and final. The draft test approach document does not require Adopter or Cross-Cutting WS sign-off prior to submission to NCI senior leadership for approval. However the draft test approach document will be circulated to the appropriate Adopter and Architecture Workspace Reviewers to initiate the review process of this document. The review process will be iterative in nature, resulting in the final reviewer-approved version of this document.

- Task 2.4.1 – Draft Test Approach Document

📁 **Artifact: Draft Test Approach document**

- Task 2.4.2 – Final Test Approach Document. Requires approval of Adopter and Architecture Workspace

📁 **Artifact: Final Test Approach document, approved by Adopter and Architecture Workspace**

- o Task 2.5 – Develop UML Models: Define the information models in UML. There will be one model for each analytical service (ARACNE, SOM, Hierarchical Clustering).

The draft UML model does not require Adopter or Cross-Cutting WS sign-off prior to submission to NCI senior leadership for approval. However the draft UML model will be circulated to the appropriate Adopter, Architecture and VCDE Workspace Reviewers to initiate the review process of the model. The review process will be iterative in nature, resulting in the final reviewer-approved version of the UML model. These groups will review the models for the accuracy of the relationships, rather than commenting on the functionality, as these are existing packages whose functionality will not be changed for this project.

- Task 2.5.1 – Create Draft UML Models for selected analytical services

📁 **Artifact: Draft UML models for selected analytical services**

- Task 2.5.2 – Create Final UML Models for selected analytical services. Requires approval of Adopter, Architecture and VCDE Workspaces

📁 **Artifact: Final UML models for selected analytical services, approved by Adopter, Architecture and VCDE Workspaces**

- o Task 2.6-Generation of Common Data Elements (CDEs) from final UML models. All data structures that will be seen on the GRID will have XMI representations, purely internal classes and data structures will not undergo XMI or CDE creation. The final, approved version of the UML models will be used as the substrate for CDE creation. The UML models can be saved as XMI files. Two tools supplied by the caCORE SDK (website) will be used to take the XMIs file derived from the UML models and create CDEs in the caDSR. The Semantic Connector will be used to fully annotate the UML models. The UML



Loader will be used to load the semantically annotated XMI files into the caDSR, thus creating CDEs from the original UML models.

- Task 2.6.1- Generation of semantically annotated XMIs- The semantic connector tool will provide concept codes for descriptions of attributes in the model. Fully annotating the models will be an iterative process as there will be review and curation between the EVS team and the developer.

Artifact: Complete, Semantically annotated XMI files (XMI 1.2 format)

- Task 2.6.2- Generation of CDEs in caDSR- The UML loader tool will take semantically annotated XMI files and convert the objects of the models into CDEs. The CDEs created by loading the UML models into the caDSR will be on the staging server of the caDSR. The CDEs on the staging server will be reviewed by the caDSR team and the developer before they are moved to the production server of the caDSR. Once the UML models are loaded and the CDEs are on the caDSR staging server, the development team can submit an interoperability review package to the Cross-cutting WS

Artifact: Full CDE Use Report

- Task 2.7 – Develop Draft APIs which describe objects and are based on UML models above. APIs will need to document input and output objects. The draft APIs need to be reviewed by the Adopter and by the Architecture and VCDE Workspaces. These groups will review the APIs for the accuracy of the relationships, rather than commenting on the functionality, as these are existing packages whose functionality will not be changed for this project.

- Task 2.7.1 – Develop Draft APIs

Artifact: Code and documentation for draft APIs

- Task 2.7.2 – Develop Final APIs. Requires approval of Adopter, Architecture and VCDE Workspaces

Artifact: Code and documentation for final APIs, approved by Adopter, Architecture and VCDE Workspaces

- Task 2.8 –caBIG™ Silver Compatibility Review Submission Package: Various artifacts packaged together and submitted to the Cross-cutting Workspaces for final review and approval as a caBIG™ Silver Compatible application. Many of the deliverables listed below have been generated in previous tasks.

- Brief Description of the data system and its design (Powerpoint format, will be used as a presentation to the Cross-cutting Workspaces)
 - UML Model for the data system in Enterprise Architect Format
 - Semantically annotated XMI file describing the data system.



- Completed Semantic Connector report
- UML Loader checklist
- Value Domain Report
- Vocabulary Report
- Standards Report
- Full CDE Use Report
- API documentation (such as Javadocs)

Artifact: caBIG™ Compatibility Review Submission Package

- Task 2.9 – Create XML Schemas for each exposed data type and publish them to the caGrid GME service framework

Artifact: XML Schemas for exposed data types

- Task 2.10 – Use caGrid analytical toolkit to bring packages to the grid

Artifact: Code generated using caGrid analytical toolkit

- Task 2.11 – Work with the caGrid team to create a test grid node to demonstrate functionality. Install virtualized (grid enabled) analytical services at Columbia.

- Task 2.11.1 – Install current version of caGrid
- Task 2.11.2 – Plug APIs into the caGrid infrastructure

Artifact: Documentation of installation and testing processes and outcomes for the Columbia grid node

Artifact: Documentation of installation and testing processes.

- Task 2.12 – Develop a caWorkbench component (caGRIDExplorer) which acts as a browser and a search engine (using metadata keywords) of caGRID indexing services and which enables the discovery of analytical services. Using caGRIDExplorer users will be able to designate which instance of a supported grid service (ARACNE, SOM, Hierarchical Clustering) they would like to execute.

Artifact: Code for caGRIDExplorer.

- Task 2.13 – Develop Requirements and Specifications Document a scripting language (caSCRIPT) that utilizes the caGRID 1.0 workflow framework and expresses data access and analysis workflows that utilize caWorkbench components. The developer will collaborate with the caGRID 1.0 team in the preparation of this document.

The Requirements and Specifications Document will be developed in two phases, draft and final. The draft document does not require Adopter or Cross-Cutting WS sign-off prior to submission to NCI senior leadership for approval. However the draft Requirements and Specifications document will be circulated to the appropriate Adopter, Architecture and VCDE Workspace Reviewers to initiate the review process of this document. The review process will be iterative in nature, resulting in the final reviewer-approved version of this document



- Task 2.13.1 Develop Draft Requirements and Specification Document for caSCRIPT
 - 📁 **Artifact: Draft Requirements and Specification Document for caSCRIPT**
 - Task 2.13.2 Develop Final Requirements and Specification Document
 - 📁 **Artifact: Final Requirements and Specification Document for caSCRIPT, approved by Adopter, Architecture and VCDE Workspaces**
- Task 2.14 – Implement caSCRIPT.
 - Task 2.14.1 – Develop a caSCRIPT editor (as a caWorkbench component) to enable the composition, storage, retrieval and execution of caSCRIPT scripts caSCRIPT will utilize the underlying workflow framework of caGRID 1.0.
 - 📁 **Artifact: Code for caSCRIPT editor.**
 - Task 2.14.2 – Develop 3 scripts to showcase the execution of non-trivial, scientifically-relevant analysis workflows. The workflows will be based on the use cases developed in Task 2.15. In the planning and execution of this task, the Developer will closely coordinate with the caGRID 1.0 team.
 - 📁 **Artifact: caSCRIPT scripts for selected workflows.**
- Task 2.15 – In collaboration with the Adopter, plan and carry out 3 case studies, showcasing how caWorkbench can be used to carry out non trivial analysis tasks, involving accessing data from caGRID data sources and analyzing such data through caGRID-enabled analytical services. Present outcomes to ICR Workspace.
 - 📁 **Artifact: Powerpoint presentation of implemented case studies**
- Task 2.16 – Technical/Architecture Guide – Describes architecture, systems requirements, APIs, and other tools that integrate with the software being developed as well as their implementation (https://cabig.nci.nih.gov/working_groups/Training_SLWG/Documents/working_groups/Training_SLWG/Documents/caBIG_programmers_manual_template_091405_jbh.doc); subject to caBIG™ Documentation and Training Review Process. The Technical/Architecture Guide will be developed in two phases, draft and final. The draft Technical/Architecture Guide should be submitted to the Training Working Group for review and feedback, but does not require Training Working Group sign-off prior to submission to NCI for approval. However, the Final Technical/Architecture Guide does require review and approval by the Training Working Group prior to its submission to NCI leadership.
 - Task 2.16.1 – Draft Technical/Architecture Guide
 - 📁 **Artifact: Draft Technical/Architecture Guide**



- Task 2.16.2 – Final Technical/Architecture Guide; subject to the caBIG™ Documentation and Training Review Process.
 📁 **Artifact: Final Technical/Architecture Guide**
- Task 2.17 – Installation Guide – Outlines the supported configurations and technical installation instructions for a software application; subject to caBIG™ Documentation and Training Review Process. The Installation Guide will be developed in two phases, draft and final. The draft Installation Guide should be submitted to the Training Working Group for review and feedback, but does not require Training Working Group sign-off prior to submission to NCI for approval. However, the Final Installation Guide does require review and approval by the Training Working Group prior to its submission to NCI leadership.
 - Task 2.17.1 Draft Installation Guide
 📁 **Artifact: Draft Installation Guide**
 - Task 2.17.2 Final Installation Guide
 📁 **Artifact: Final Installation Guide**
- Task 2.18 – Administration Guide – Describes process for updating and maintaining the application, importing and deleting data, creating authorization for users and user groups (https://cabig.nci.nih.gov/working_groups/Training_SLWG/Documents/working_groups/Training_SLWG/Documents/T5aTemplateforcaBIGApplicationSystemInstallation-AdministrationReferenceManualv050.doc); subject to caBIG™ Documentation and Training Review Process. The Administration Guide will be developed in two phases, draft and final. The draft Administration Guide should be submitted to the Training Working Group for review and feedback, but does not require Training Working Group sign-off prior to submission to NCI for approval. However, the Final Administration Guide does require review and approval by the Training Working Group prior to its submission to NCI leadership.
 - Task 2.18.1 – Draft Administration Guide
 📁 **Artifact: Draft Administration Guide**
 - Task 2.18.2 – Final Administration Guide
 📁 **Artifact: Final Administration Guide**
- Task 2.19 – Release Notes – Lists new features and functionalities, known bugs, and their status, etc.
 - Task 2.19.1 – Release Notes
 📁 **Artifact: Release Notes**
- Task 2.20 – Document developed functionality by (1) creating new sections in the caWorkbench online help system, and (2) updating the caWorkbench user manual. The updated End-User Guide will be developed in two phases, draft and final. The draft document should be submitted to the Training Working Group for review and feedback, but does not require Training Working Group sign-off prior to



submission to NCI for approval. However, the Final End-User Guide does require review and approval by the Training Working Group prior to its submission to NCI leadership.

- Task 2.20.1 –Draft updates to the caWorkbench End-User Guide.

📁 **Artifact: Draft caWorkbench End-User Guide**

- Task 2.20.2 – Create the Final updates to the caWorkbench End-User Guide; subject to the caBIG™ Documentation and Training Review Process.

📁 **Artifact: Final caWorkbench End-User Guide**

- Task 2.21 – Update the caWorkbench training materials covering the new developed functionality. These materials will be used for training NCICB personnel and will also be made available to the general public from the caWorkbench web page. The Training Module will be developed in two phases, draft and final. The draft module(s) should be submitted to the Training Working Group for review and feedback, but does not require Training Working Group sign-off prior to submission to NCI for approval. However, the Final Training Module(s) does require review and approval by the Training Working Group prior to its submission to NCI leadership.

- Task 2.21.1 – Create a draft Training Module for caWorkbench.

📁 **Artifact: Draft Training Module(s)**

- Task 2.21.2 – Finalize the Training Module for caWorkbench, subject to the caBIG™ Training and Documentation Review Process.

📁 **Artifact: Final Training Module(s)**

- Task 2.22– Purchase hardware and software needed for the development effort:

- Linux server to house the caGRID software.
 - 2 JBuilder Licenses.
 - 2 Enterprise Architect Licenses.

📁 **Artifact: Hardware and software purchase reports**

- Task 2.23 – MAGE-API Query support: A graphical user interface will be implemented to support query-based microarray retrieval from MAGE-API data-sources. This will become an option in the data-loader if a MAGE-API source is selected. In particular, it will be possible to select experiments by variety of criteria, including:

- Tissue type
 - Chip/platform type.

- Inclusion of a marker set. This option will sub-select only experiments containing markers that belong to activated marker panels
- Library ID
- Investigator name
- Organism
- Name
- PubMed ID

📁 **Artifact: Code for graphical user interface to support query-based microarray retrieval from MAGE-API data sources**

- Task 2.24 – Lessons Learned Document

📁 **Artifact: Project Lessons Learned**

The General Contractor will provide necessary standard operating procedures and templates to support project activities.

4. Deliverables

	Deliverable	DUE DATE	TEMPLATE AVAILABLE?	PAYMENT PER DELIVERABLE (INCLUDES INDIRECTS)
1.	Monthly Status Report (Task 1.3)	5 th business day of each month	Yes	
2.	Risk Management Matrix (Task 2.1)	1 month from contract award	Yes	
3.	Draft Use Case document (Task 2.2.2)	2 months from contract award	Yes	
4.	Final Use Case document, approved by Adopter, Architecture and VCDE Workspaces (Task 2.2.2)	4 months from contract award	Yes	
5.	Draft Requirements and Specification Document (Task 2.3.1)	3 months from contract award	Yes	
6.	Final Requirements and Specification Document, approved by Adopter, Architecture and VCDE Workspaces (Task 2.3.2)	4 months from contract award	Yes	
7.	Draft Test Approach	6 months from	Yes	



	Deliverable	DUE DATE	TEMPLATE AVAILABLE?	PAYMENT PER DELIVERABLE (INCLUDES INDIRECTS)
	Document (Task 2.4.1)	contract award		
8.	Final Test Approach Document, approved by Adopter and Architecture Workspace (Task 2.4.2)	8 months from contract award	Yes	
9.	Draft UML models for selected analytical services (Task 2.5.1)	7 months from contract award	No	
10.	Final UML models for selected analytical services, approved by Adopter, Architecture and VCDE Workspaces (Task 2.5.2)	8 months from contract award	No	
11.	Complete, Semantically annotated XMI files (XMI 1.2 format) (Task 2.6.1)	9 months from contract award	No	
12.	Full CDE Use Report (Task 2.6.2)	10 months from contract award	Yes \	
13.	Code and documentation for draft APIs (Task 2.7.1)	9 months from contract award	No	
14.	Code and documentation for final APIs, approved by Adopter, Architecture and VCDE Workspaces (Task 2.7.2)	10 months from contract award	No	
15.	caBIG™ Compatibility Review Submission Package (Task 2.8)	10 months from contract award	Yes	
16.	XML Schemas for exposed data types (Task 2.9)	9 months from contract award	No	
17.	Code generated using caGrid analytical toolkit (Task 2.10)	10 months from contract award	No	
18.	Documentation of installation and testing processes and outcomes for the Columbia grid node (Task 2.11)	11 months from contract award	No	
19.	Code for caGRIDExplorer (Task 2.12)	11 months from contract award	No	



	Deliverable	DUE DATE	TEMPLATE AVAILABLE?	PAYMENT PER DELIVERABLE (INCLUDES INDIRECTS)
20.	Draft Requirements and Specification Document for caSCRIPT (Task 2.13.1)	7 months from contract award	Yes	
21.	Final Requirements and Specification Document for caSCRIPT, approved by Adopter, Architecture and VCDE Workspaces (Task 2.13.2)	8 months from contract award	Yes	
22.	Code for caSCRIPT editor (Task 2.14.1)	10 months from contract award	No	
23.	caSCRIPT scripts for selected workflows (Task 2.14.2)	11 months from contract award	No	
24.	Powerpoint presentation of implemented case studies (Task 2.15)	12 months from contract award	No	
25.	Draft Technical/Architecture Guide (Task 2.16.1)	11 months from contract award	Yes	
26.	Final Technical/Architecture Guide (Task 2.16.2)	12 months from contract award	Yes	
27.	Draft Installation Guide (Task 2.17.1)	11 months from contract award	Yes	
28.	Final Installation Guide (Task 2.17.2)	12 months from contract award	Yes	
29.	Draft Administration Guide (Task 2.18.1)	11 months from contract award	Yes	
30.	Final Administration Guide (Task 2.18.2)	12 months from contract award	Yes	
31.	Release Notes (Task 2.19)	12 months from contract award	Yes	
32.	Draft caWorkbench End-User Guide (Task 2.20.1)	11 months from contract award	Yes	
33.	Final caWorkbench End-User Guide (Task 2.20.2)	12 months from contract award	Yes	
34.	Draft Training Module(s) (Task 2.21.1)	11 months from contract award	Yes	



	Deliverable	DUE DATE	TEMPLATE AVAILABLE?	PAYMENT PER DELIVERABLE (INCLUDES INDIRECTS)
35.	Final Training Module(s) (Task 2.21.2)	12 months from contract award	Yes	
36.	Hardware and software purchase reports (Task 2.22.1)	1 month from contract award	No	
37.	Code for graphical user interface to support query-based microarray retrieval from MAGE-API data sources (Task 2.23)	TBD	No	
38.	Project Lessons Learned (Task 2.24)	12 months from contract award	Yes	

5. Inspection and Acceptance Requirements

All reports and deliverables are subject to the final approval of the NCI. General quality measures identified below will be applied to each work product received from the Workspace under this statement of work and of requirements documents completed as part of this statement of work.

- **Accuracy** - Work products shall be accurate in presentation, technical content, and adherence to accepted elements of style.
- **Clarity** - Work products shall be clear and concise. Any and all diagrams shall be easy to understand and be relevant to the supporting narrative.
- **Consistency to Requirements** - All work products must satisfy the requirements of this statement of work.
- **File Editing** - All text and diagrammatic files shall be editable by the Workspace Manager.
- **Format** - Work products shall be submitted in electronic media in CVS.
- **Timeliness** - Work products shall be submitted on or before the due date specified in this statement of work or submitted in accordance with a later scheduled date determined by agreement between the Workspace and General Contractor.
- **User acceptance** - User acceptance for ICR software products will be documented by stakeholders signature(s) on software design documents and test plan, as well as completed software user testing documentation. The intent of involving stakeholders at all stages of design and testing is to ensure that the software both performs as per their specific needs, but that it also meets the



documented design requirements. These signatures will represent that the relevant stakeholders have participated in the requirements generation, test design and user testing, ensured the software performs according to the parameters specified in the design documents, and properly executed the approved test plan carried out at their site.

- **Developer User Acceptance** – User acceptance for Workspace/Working Group software and solution development products will be documented by Developers' signature(s) on a sign-off page for source code. These signatures will represent that the Developers:
 - Conducted software or solution development according to defined processes
 - Conducted quality assurance and quality control activities (e.g., white-box and unit testing) during the development process
 - Conducted software defect reporting and responded to software defect reports from Adopters.

5.1 Review and Acceptance Processes

This section describes the review and acceptance processes to be implemented for all caBIG™ artifacts/deliverables. With the exception of status reports, all other deliverables are subject to draft and final review processes, and final approval/sign-off by the NCI.

5.1.1 General Review and Acceptance Process

5.1.1.1 Draft Review – The general draft artifacts review process is an iterative process that involves the maximum number of reviewers from the Workspaces, Working Groups, NCI, Special Interest Groups (SIGs), etc. This process is managed by the Workspace/Working Group Leads with the goal of vetting the deliverable, thereby, reducing the level of effort needed for the Final Review. Reviewers will indicate their acceptance of the deliverable prior to its submission to NCI for a final review and approval.

5.1.1.2 Final Review – The final deliverable review process is an activity for the NCI Leadership to approve the production-ready product. This activity will be an efficient, relatively low level effort following an effective draft review process. The NCI Leadership review will result in a sign-off indicating approval of the product.

5.1.2 Cross-Cutting Review and Acceptance

5.1.2.1 Draft Interoperability Review – This specialized draft deliverable review process is an iterative process that involves the Architecture and Vocabularies & Common Data Elements (VCDE) Cross-Cutting Workspaces interoperability review processes. The goal is to be sure the development project achieves semantic and

syntactic interoperability as outlined in the caBIG compatibility guidelines. The interoperability review process involves the maximum number of reviewers from the Workspaces, Working Groups, NCI, SIGs, etc. Only deliverables that require cross-cutting reviews are subject to this process. The Architecture and V&CDE interoperability review processes and documents are available in the caBIG™ Management Portal (caMP) at [URL TBD]. This process is managed by the Cross-Cutting Workspace Leads.

5.1.2.2 Final Interoperability Review – The final deliverable review process is an activity for the NCI Leadership to approve the production-ready product. This activity will be an efficient, relatively low level effort following an effective draft review process. The NCI Leadership review will result in a sign-off indicating approval of the product.

5.1.3 Status Reports

No formal review/approval process is required for Status Reports.

6. Travel and Other Direct Costs

- 1 All travel estimates will conform to the current Federal Travel Regulations (FTRs). All travel expenses should be included as part of the RFP bid.
- 2 Food, travel and lodging payments shall be based on the official government per diem rates found at <http://www.policyworks.gov/org/main/mt/homepage/mtt/perdiem/perd04d.html>.
- 3 Any anticipated Other Direct Costs (e.g. reproduction, long-distance telephone, courier, and teleconference charges) may be included in the cost proposal.

7. Place of Performance

The majority of the work will be performed at the Cancer Center.

8. Period of Performance

The period of performance is 12 months.

9. Data Use, Disclosure of Information and Handling of Sensitive Information

The caBIG grid must be built to share sensitive information from disparate data sources, meet the requirements of the Cancer Centers that participate in the pilot, as well as to include many additional Cancer Centers and other research stakeholders in future years. Therefore, the potential sensitivity of the information collected, information security issues, local Institutional Review Board requirements and the



Health Insurance Portability and Accountability Act of 1996 (HIPAA) requirements will need to be addressed.

Final regulations issued by the Department of Health and Human Services provide privacy and security standards that must be observed in the handling of patient data resulting from biomedical research. HIPAA privacy standards will be used to establish safeguards and restrictions for the use and disclosure of research records. HIPAA security standards will be used to help Cancer Centers implement administrative, physical, and technical safeguards to protect electronic health information. Improper use or disclosure of sensitive information under the rules may be subject to criminal or civil sanctions prescribed in HIPAA.

Performance of caBIG tasks should be guided by the technical and operational principles of the caBIG program. Technical principles are defined in the caBIG Compatibility Guide, to be provided by the General Contractor. Overarching principles include:

- 1 Open source
- 2 Open access
- 3 Open development
- 4 Federation

10. Primary Point of Contact

All questions and concerns regarding this statement of work should be directed to the caBIG Project Manager:

Mark Adams, PhD
Booz Allen Hamilton
6010 Executive Boulevard
Suite 500
Rockville, Maryland 20852
Email – adams_mark@bah.com

The ICR Workspace Project Lead is:

Juli Klemm, PhD
3rd Millennium
391 Totten Pond Road, Suite 104
Waltham, Massachusetts 02451
Email – jklemm@3rdmill.com