

CAGRID PORTAL 2.0.2

User's Guide



This is a U.S. Government work.

July 3, 2008



Model caBIG™ Open Source Software License



Release Date: January 7, 2008

Copyright Notice. Copyright 2008 The Ohio State University Research Foundation (OSURF), Argonne National Labs (ANL), SemanticBits LLC (SemanticBits), and Ekagra Software Technologies Ltd. (Ekagra) ("caBIG™ Participant"). The caGrid 1.2 software was created with NCI funding and is part of the caBIG™ initiative. The software subject to this notice and license includes both human readable source code form and machine readable, binary, object code form (the "caBIG™ Software").

This caBIG™ Software License (the "License") is between caBIG™ Participant and You. "You" (or "Your") shall mean a person or an entity, and all other entities that control, are controlled by, or are under common control with the entity. "Control" for purposes of this definition means (i) the direct or indirect power to cause the direction or management of such entity, whether by contract or otherwise, or (ii) ownership of fifty percent (50%) or more of the outstanding shares, or (iii) beneficial ownership of such entity.

License. Provided that You agree to the conditions described below, caBIG™ Participant grants You a non-exclusive, worldwide, perpetual, fully-paid-up, no-charge, irrevocable, transferable and royalty-free right and license in its rights in the caBIG™ Software, including any copyright or patent rights therein, to (i) use, install, disclose, access, operate, execute, reproduce, copy, modify, translate, market, publicly display, publicly perform, and prepare derivative works of the caBIG™ Software in any manner and for any purpose, and to have or permit others to do so; (ii) make, have made, use, practice, sell, and offer for sale, import, and/or otherwise dispose of caBIG™ Software (or portions thereof); (iii) distribute and have distributed to and by third parties the caBIG™ Software and any modifications and derivative works thereof; and (iv) sublicense the foregoing rights set out in (i), (ii) and (iii) to third parties, including the right to license such rights to further third parties. For sake of clarity, and not by way of limitation, caBIG™ Participant shall have no right of accounting or right of payment from You or Your sublicensees for the rights granted under this License. This License is granted at no charge to You. Your downloading, copying, modifying, displaying, distributing or use of caBIG™ Software constitutes acceptance of all of the terms and conditions of this Agreement. If you do not agree to such terms and conditions, you have no right to download, copy, modify, display, distribute or use the caBIG™ Software.

1. Your redistributions of the source code for the caBIG™ Software must retain the above copyright notice, this list of conditions and the disclaimer and limitation of liability of Article 6 below. Your redistributions in object code form must reproduce the above copyright notice, this list of conditions and the disclaimer of Article 6 in the documentation and/or other materials provided with the distribution, if any.
2. Your end-user documentation included with the redistribution, if any, must include the following acknowledgment: "This product includes software developed by the Ohio State University Research Foundation (OSURF), Argonne National Labs (ANL), SemanticBits LLC (SemanticBits), and Ekagra Software Technologies Ltd. (Ekagra)." If You do not include such end-user documentation, You shall include this acknowledgment in the caBIG™



Software itself, wherever such third-party acknowledgments normally appear.

3. You may not use the names "The Ohio State University Research Foundation", "OSURF", "Argonne National Labs", "ANL", "SemanticBits LLC", "SemanticBits", "Ekagra Software Technologies Ltd.", "Ekagra", "The National Cancer Institute", "NCI", "Cancer Bioinformatics Grid" or "caBIG™" to endorse or promote products derived from this caBIG™ Software. This License does not authorize You to use any trademarks, service marks, trade names, logos or product names of either caBIG™ Participant, NCI or caBIG™, except as required to comply with the terms of this License.
4. For sake of clarity, and not by way of limitation, You may incorporate this caBIG™ Software into Your proprietary programs and into any third party proprietary programs. However, if You incorporate the caBIG™ Software into third party proprietary programs, You agree that You are solely responsible for obtaining any permission from such third parties required to incorporate the caBIG™ Software into such third party proprietary programs and for informing Your sublicensees, including without limitation Your end-users, of their obligation to secure any required permissions from such third parties before incorporating the caBIG™ Software into such third party proprietary software programs. In the event that You fail to obtain such permissions, You agree to indemnify caBIG™ Participant for any claims against caBIG™ Participant by such third parties, except to the extent prohibited by law, resulting from Your failure to obtain such permissions.
5. For sake of clarity, and not by way of limitation, You may add Your own copyright statement to Your modifications and to the derivative works, and You may provide additional or different license terms and conditions in Your sublicenses of modifications of the caBIG™ Software, or any derivative works of the caBIG™ Software as a whole, provided Your use, reproduction, and distribution of the Work otherwise complies with the conditions stated in this License.
6. THIS caBIG™ SOFTWARE IS PROVIDED "AS IS" AND ANY EXPRESSED OR IMPLIED WARRANTIES (INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT AND FITNESS FOR A PARTICULAR PURPOSE) ARE DISCLAIMED. IN NO EVENT SHALL THE OHIO STATE UNIVERSITY RESEARCH FOUNDATION ("OSURF"), ARGONNE NATIONAL LABS ("ANL"), SEMANTICBITS LLC ("SEMANTICBITS"), AND EKAGRA SOFTWARE TECHNOLOGIES LTD. (EKAGRA) OR ITS AFFILIATES BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS caBIG™ SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Revision History

The following is the revision history for this document.

Date	Version	Description	Revised By
June 29, 2008	2.0.2.3	Final 2.02 Revision	Carolyn Kelley Klinger
June 26, 2008	2.0.2.2	Updated Revision History, Acknowledgements, and link to this guide. Copyedited text.	John Eisenschmidt
June 23, 2008	2.0.2.1	Updated for caGrid Portal 2.0.2.	Carolyn Kelley Klinger
January 9, 2008	2.0.1	Updated for caGrid Portal 2.0.	Eddie VanArasdall

Contents

Revision History	iv
Illustrations.....	vii
About This Guide	1
Purpose	1
Audience.....	1
Prerequisites.....	1
Topics Covered	1
Text Conventions Used	2
Credits and Resources	3
Chapter 1 About the caGrid Portal.....	4
Purpose of the caGrid Portal	4
caGrid Portal Features	4
Discovery	4
Service Metadata Exploration	5
Data Service Query.....	5
Browsing caBIG Tools.....	6
Query Sharing	6
caGrid Account Registration	6
Grid Status Overview	8
News and Events	8
Visual Overview of the caGrid Portal Home Page.....	9
Chapter 2 Registering and Logging In	11
About Registration	11
Registering for a Login Account	12
Logging in with Your caGrid Portal Account.....	15
Logging in with Your NIH Account.....	16
About the Welcome Menu	16
Chapter 3 Using the Map Portlet.....	17
About the Map Portlet.....	17
Selecting a Category View	18
Selecting a Service, Participant, or POC	19
Selecting a Data Service.....	19
Selecting an Analytical Service.....	21
Selecting a Participant Institute.....	22
Selecting a Hosting Research Center.....	24
Selecting a POC.....	25
Chapter 4 Using the Discovery Portlet.....	27
About the Discovery Portlet	27
Using the Directory Tab	28
Using the List View Subtab	28
Using the Map View Subtab.....	30
Using the Search Tab	30
Searching for a Service.....	31
Searching for a Participant.....	35
Searching for a POC	36
Using the Details Tab	38
Viewing Details for Services.....	38

Viewing Details for a Participant	41
Viewing Details for a POC	42
Chapter 5 Using the Data Service Query Portlet	43
About the Data Service Query Portlet	43
Query-Building Scenarios.....	44
Scenario 1: Starting without a URL or Query.....	44
Scenario 2: Starting with a URL but No CQL Query.....	45
Scenario 3: Starting with a URL and a CQL Query.....	45
Building a Query Using the Query Builder Tool.....	46
Building CQL Queries	46
Building DCQL Queries.....	50
About the History Tab.....	60
Working with Shared Queries.....	62
Sharing a Query.....	62
Searching for a Shared Query	63
Chapter 6 Using Other Portal Features.....	65
Using the News Portlet.....	65
About the caGrid Links Portlet.....	66
Using the caGrid Status Portlet	68
Viewing Details for a Recently Registered Service or Participant.....	68
Viewing a List of Participants or Services	69
About the Calendar Portlet.....	69
Chapter 7 Using the caBIG Tools Portlet	71
About the caBIG Tools Portlet.....	71
Sorting caBIG Tools	72
Viewing caBIG Tool Details	75
Index	77

Illustrations

Figure 1-1. caGrid Portal home page	9
Figure 2-1. Register links.....	12
Figure 2-2. Registration page	13
Figure 2-3. Successful login message.....	15
Figure 2-4. Welcome message with Logout link.....	15
Figure 2-5. Welcome menu	16
Figure 3-1. Map portlet	17
Figure 3-2. Categories list.....	18
Figure 3-3. Callout balloon for a data service center	20
Figure 3-4. Discovery Portlet showing details for a selected data service.....	20
Figure 3-5. Callout balloon for a selected participant	22
Figure 3-6. Discovery portlet showing details for a participant institute.....	23
Figure 3-7. Callout balloon for an analytical service center	24
Figure 3-8. List of hosting research centers with More Details links.....	25
Figure 4-1. Directory tab > List View subtab – results for All Services category	28
Figure 4-2. Map View subtab.....	30
Figure 4-3. Services subtab with the Simple search mode selected	31
Figure 4-4. Services subtab with Advanced search mode selected	32
Figure 4-5. Discovered services	32
Figure 4-6. Services subtab with the EVS search mode selected	33
Figure 4-7. Results of an EVS search	34
Figure 4-8. Discovered services	34
Figure 4-9. Participants subtab.....	35
Figure 4-10. POCs subtab	36
Figure 4-11. Details for a selected data service	38
Figure 4-12. Selecting a UML class for a query.....	39
Figure 4-13. Details for a selected analytical service	40
Figure 4-14. Details for a selected participant	41
Figure 4-15. Details for a selected POC	42
Figure 5-1. Starting a query from a selected result.....	44
Figure 5-2. Working query	45
Figure 5-3. Query Builder subtab with selected UML class <code>ArrayDesign</code>	46
Figure 5-4. Attributes and associations for selected class.....	48
Figure 5-5. Specifying a predicate and value for a selected attribute	49
Figure 5-6. Query Builder subtab with added criterion	49
Figure 5-7. Query XML tab	50
Figure 5-8. History tab with query information	50
Figure 5-9. Query Builder specifying a service and URL.....	51
Figure 5-10. Services with the same UML class to aggregate from	52
Figure 5-11. Summary of data services selected for aggregation	52
Figure 5-12. History tab with query information	52
Figure 5-13. Query Builder specifying a service and URL.....	54
Figure 5-14 Query Builder tool with selected <code>NucleicAcidSequence</code> class.....	54
Figure 5-15 Tree view of attributes and associations for the <code>NucleicAcidSequence</code> class	55
Figure 5-16 Query criterion tree view for caBIO <code>NucleicAcidSequence</code> class	55
Figure 5-17 Service information for semantically equivalent UML class	56
Figure 5-18 Selecting the join condition.....	56

Figure 5-19 Entering criterion value for Organism scientificName attribute	57
Figure 5-20 Enter criterion value for Gene name attribute	57
Figure 5-21 Complete federated query for caBIO NucleicAcidSequence class	58
Figure 5-22 XML federated query for caBIO NucleicAcidSequence class	58
Figure 5-23 History Tab with the running DCQL query.....	59
Figure 5-24 Results of federated query for caBIO NucleicAcidSequence class	59
Figure 5-25. Results tab showing query results.....	60
Figure 5-26. Form for adding a query name and description.....	62
Figure 5-27. Saved query	63
Figure 5-28. Shared Queries tab > Find Shared Queries subtab – Search form.....	64
Figure 6-1. News links	65
Figure 6-2. News page	66
Figure 6-3. Status portlet	68
Figure 6-4. Calendar portlet.....	69
Figure 7-1.Tools portlet.....	72
Figure 7-2. caBIG tools sorted by Area of Focus.....	73
Figure 7-3. caBIG tool details	75

About This Guide

This section introduces you to the *caGrid Portal 2.0.2 User's Guide* and provides an orientation to using the guide. It also includes support contact information.

Topics in this section:

- *Purpose* on this page
- *Audience* on page 1
- *Prerequisites* on page 1
- *Topics Covered* on page 1
- *Text Conventions Used* on page 2
- *Credits and Resources* on page 3

Purpose

This guide introduces you to the caGrid Portal. It explains how to use the portal features to do the following:

- Discover caGrid services, caBIG participants, and caGrid points of contact
- Explore the service metadata of a discovered service
- Build queries using the CQL and DCQL query languages
- Share queries and search for shared queries
- Register for an account so that you can interact with secure caGrid services
- View the status of caGrid services
- View caGrid-related news using the portal or using an RSS 2.0-compliant news reader
- View a calendar of caGrid-related events.

Audience

This guide is intended for the following users:

- Non-developer members of the caBIG community
- caGrid service developers
- Anyone interested in caBIG and caGrid

Prerequisites

To get the most out of this guide, you should have the following skills:

- Ability to browse the Internet using a Web browser
- Understanding of caGrid metadata concepts (for building queries)

Topics Covered

This brief overview explains what you will find in each chapter and appendix.

- Chapter 1, *About the caGrid Portal* (p. 4), provides an overview of the caGrid Portal and its feature set.
- Chapter 2, *Registering and Logging In* (p. 11) covers the registration process and the procedures for logging in as a new portal user or an NIH user.
- Chapter 3, *Using the Map Portlet* (p. 1717), explains how to use the Map portlet to select a service or institution and link to more detailed information.
- Chapter 4, *Using the Discovery Portlet* (p. 2727), explains how to use the Discovery portlet to browse, search, and view details for caGrid services, caBIG participants, and caGrid POCs.
- Chapter 5, *Using the Data Service Query Portlet* (p. 4343), explains how to use the Data Service Query portlet to build and share queries.
- Chapter 6, *Using Other Portal Features* (p. 65), covers the News, caGrid Links, Grid Status, and Calendar portlets.
- Chapter 7, *Using the caBIG Tools Portlet* (p. 71), explains how to use the caBIG Tools portlet to sort and browse caBIG tools.

Text Conventions Used

This section explains conventions used in this guide. The various typefaces represent interface components, keyboard shortcuts, toolbar buttons, dialog box options, and text that you type.

Convention	Description	Example
Bold	Highlights names of option buttons, check boxes, drop-down menus, menu commands, command buttons, or icons.	Click Search .
<u>URL</u>	Indicates a Web address.	http://domain.com
text in SMALL CAPS	Indicates a keyboard shortcut.	Press ENTER.
text in SMALL CAPS + text in SMALL CAPS	Indicates keys that you press simultaneously.	Press SHIFT + CTRL.
<i>Italics</i>	Highlights references to other documents, sections, figures, and tables.	See <i>Figure 4.5</i> .
<i>Italic boldface monospace</i> type	Represents text that you type.	In the New Subset text box, enter <i>Proprietary Proteins</i> .
Note:	Highlights information of particular importance.	Note: This concept is used throughout this document.
{ }	Surrounds replaceable items.	Replace {last name, first name} with the Principal Investigator's name.

Credits and Resources

<i>caGrid Portal Development and Management Teams</i>		
<i>Development</i>	<i>Quality Assurance</i>	<i>Program and Project Management</i>
Joshua Phillips ²	Aynur Abdurazik ³	Krishnakant Shanbhag ¹
Manav Kher ²		Peter Yan ³
		John Eisenschmidt ⁴
<i>Documentation</i>		
	Carolyn Kelley Klinger ⁵	
	Eddie VanArdall ⁵	
^{1.} National Cancer Institute Center for Bioinformatics (NCICB)	^{2.} SemanticBits, LLC	^{3.} Science Applications International Corporation (SAIC)
^{4.} 5AM Solutions	^{5.} Lockheed Martin Management System Designers	

<i>Contacts and Support</i>		
NCICB Application Support	http://ncicbsupport.nci.nih.gov/sw/ Telephone: 301-451-4384 Toll free: 888-478-4423	

Chapter 1 About the caGrid Portal

This chapter introduces you to the caGrid Portal and describes its feature set.

Topics in this chapter:

- *Purpose of the caGrid Portal* on this page
- *caGrid Portal Features* on this page
- *Visual Overview of the caGrid Portal Home Page* on page 9

Purpose of the caGrid Portal

The caGrid Portal is a Web-based application that enables you to discover and interact with the services that are available on the caGrid infrastructure. The portal serves as the primary visualization tool for the caGrid middleware and provides a standards-based platform for hosting caBIG-related tools. It also serves as a caBIG information source. Using the caGrid portal, you have instant access to information about caBIG participants, caGrid points of contact (POCs), and caGrid-related news and events.

caGrid Portal Features

The following sections discuss the various caGrid Portal features:

- *Discovery* on this page
- *Service Metadata Exploration* on page 5
- *Data Service Query* on page 5
- *Browsing caBIG Tools* on page 6
- *Query Sharing* on page 6
- *caGrid Account Registration* on page 6
- *Grid Status Overview* on page 8
- *News and Events* on page 8

Discovery

One of the goals of the caGrid infrastructure is to enable you to easily discover and access services. To achieve this goal, caGrid services register with a central index service. This service allows other applications to locate caGrid services and retrieve the necessary information for interacting with those services.

The caGrid Portal uses caGrid component APIs to provide an interface that you can use to query information gathered from one or more index services. Regardless of whether you are a service developer or non-developer, you can use this interface to determine how to use and benefit from caGrid.

In addition to service information, the portal also provides information about caBIG participant institutions and caGrid points of contact (POCs). The goal is to enable you to quickly determine who the point of contact is for specific activities and easily obtain contact information for that person.

The process of locating services, participants, and POCs is known as *discovery*. The portal provides the following modes of discovery: *map view*, *directories*, *keyword search*, and *EVS search*. The following subsections describe these modes.

Map View

The portal home page displays a map of the U.S. that shows the geographic location of services and participant institutions. You can click the icon for a service or institution and link to more detailed information.

Categories

Services, participants, and POCs are grouped into categories. You can browse these categories and click hyperlinks that lead to more detailed information.

Keyword Search

The portal provides an intuitive, keyword search interface that enables you to enter one or more keywords and match them against one or more fields.

EVS Search

This feature allows the user to enter free text that the Portal matches against concepts found in the NCI Thesaurus. As you type characters into the search field, the Portal presents candidate concepts. After you select a concept, the Portal starts the search. Services whose metadata refer to the concept or any sub-concept are presented in the search results.

Service Metadata Exploration

To support the caBIG goals of syntactic and semantic interoperability, caGrid services expose a rich, standard set of metadata. This metadata not only describes the operations that the services support; it also links the data types that those operations use to concepts defined in the NCICB Enterprise Vocabulary Services (EVS), common data elements defined in the Cancer Data Standards Repository (caDSR), and XML schemas defined in the Global Metadata Exchange (GME) service.

caGrid services that support the caGrid Query Language (CQL) also expose an XML representation of the UML domain model that the service supports. This model is also linked to the EVS, caDSR, and GME. While caGrid services support standard Web service APIs for retrieving metadata, the caGrid project provides client API components that facilitate the process.

The portal uses caGrid client APIs to build an interface that enables you to explore caGrid service metadata. You can view the semantic descriptions that come from the EVS, the data type descriptions from the caDSR, and XML schemas that are registered with the GME. The interface makes it easier for you to determine whether a service meets your needs and what is necessary to interact with the service.

Data Service Query

caGrid services that support CQL are known as *data services*. They expose a UML domain model, support a query operation that processes CQL queries against that domain model, and return XML documents that contain serialized instances of the

UML classes defined in the domain model. These serialized instances conform to the XML schemas that are registered with the GME.

The portal provides an interface that enables you to select a service and target UML class, build a query, execute the query, view the results in tabular form, and export to a Microsoft Excel or XML document. While using the discovery tool or while exploring service metadata, you can select the service and UML class to be queried. You can then use a graphical query-building interface to construct CQL or DCQL, or you can manually type the query text.

If you are a registered caGrid Portal user (discussed under *caGrid Account Registration* on this page), the portal maintains a complete history of your queries so that you can always retrieve and work with previously executed queries. The results of all queries that you execute during a single session (while your browser is open) are available for viewing.

Browsing caBIG Tools

To increase the cancer research community's access to key bioinformatics tools, platforms, and data, the caGrid Portal exposes the same inventory of applications, software tools, database technologies, and Web-based applications that appear on the [caBIG Community Website](#). As on the caBIG website, caGrid users can search for these tools based on name, compatibility level, keyword, area of focus, workspace, and technical specifications.

Query Sharing

Although data service domain models are linked to semantic metadata and data type specifications, you still need to do a fair amount of exploring to construct useful queries. If you have query-building knowledge, the portal enables you to share that knowledge with the rest of the community. Any portal user can publish CQL queries and provide documentation to accompany the query. Others can search for CQL queries based on the target service, the target UML class, the creator of the query, or the associated name and description of the query.

For more information on queries, see *Using the Data Service Query Portlet* on page 4343.

caGrid Account Registration

The caGrid security infrastructure, known as GAARDS, uses an authentication approach that is based on Public Key Infrastructure (PKI). PKI requires that users have an X.509 certificate and private key, also known as *grid credentials*.

The caGrid Dorian service is used to provide users with grid credentials. To obtain grid credentials, you must register either with the Dorian service or another Identity Provider (IdP) that has established a trust relationship with Dorian.

Secure caGrid services apply an authorization policy to incoming requests based on the identity of the client (*i.e.*, the user who is making the request). The identity is determined through the authentication process by inspecting the client's grid credentials. So, in order to interact with secure services, caGrid users need to obtain grid credentials.

While nearly all the functionality that the portal provides is available to unauthenticated users, only authenticated users can share queries and maintain a

query history across browser sessions. The portal uses GAARDS to authenticate users. So, to use the query sharing and persistent query history features, you need to obtain grid credentials.

The portal provides an interface that enables you to register for an account with the caGrid Dorian service that is hosted by the NCICB. This service provides a Level of Assurance 1 (LOA1), which means anyone can request an account and the request will automatically be approved. Once your request is approved, you can provide log in to the portal using your username and password. The portal will then use the grid credentials obtained from Dorian when interacting with secure caGrid services.

Grid Status Overview

The caGrid Portal provides a quick, visual description of the scale and activity of caGrid. The map on the front page of the portal not only indicates where services are being hosted, but whether those services are currently running. Icons of various colors indicate the service status.

In addition to the map, a status portlet presents up-to-date counts of all participants, all services, data services, and analytical services. It also provides links to more information about the five most recently registered services.

News and Events

A news portlet on the front page of the portal displays a running list of the latest caGrid-related news items. You can click each item to link to the full news article. You can also click a link to subscribe to the news feed using your browser or any other RSS 2.0-compliant news feed reader.

The portal provides a news feed editor application that is available to the portal administrator. This application can be used to maintain one or more RSS news channels.

The portal home page also includes a full-featured calendar application. This application enables you to browse through caGrid-related events by day, week, month, and year. A portal administrator can import and export calendar events using the standard iCal format, which is described at <http://tools.ietf.org/html/rfc2445>.

Visual Overview of the caGrid Portal Home Page

The caGrid Portal is located at <http://cagrid-portal.nci.nih.gov/web/guest/home>. This address launches the portal home page, shown in Figure 1-1.

Each numbered callout represents one of the features discussed in *caGrid Portal Features* on page 4.

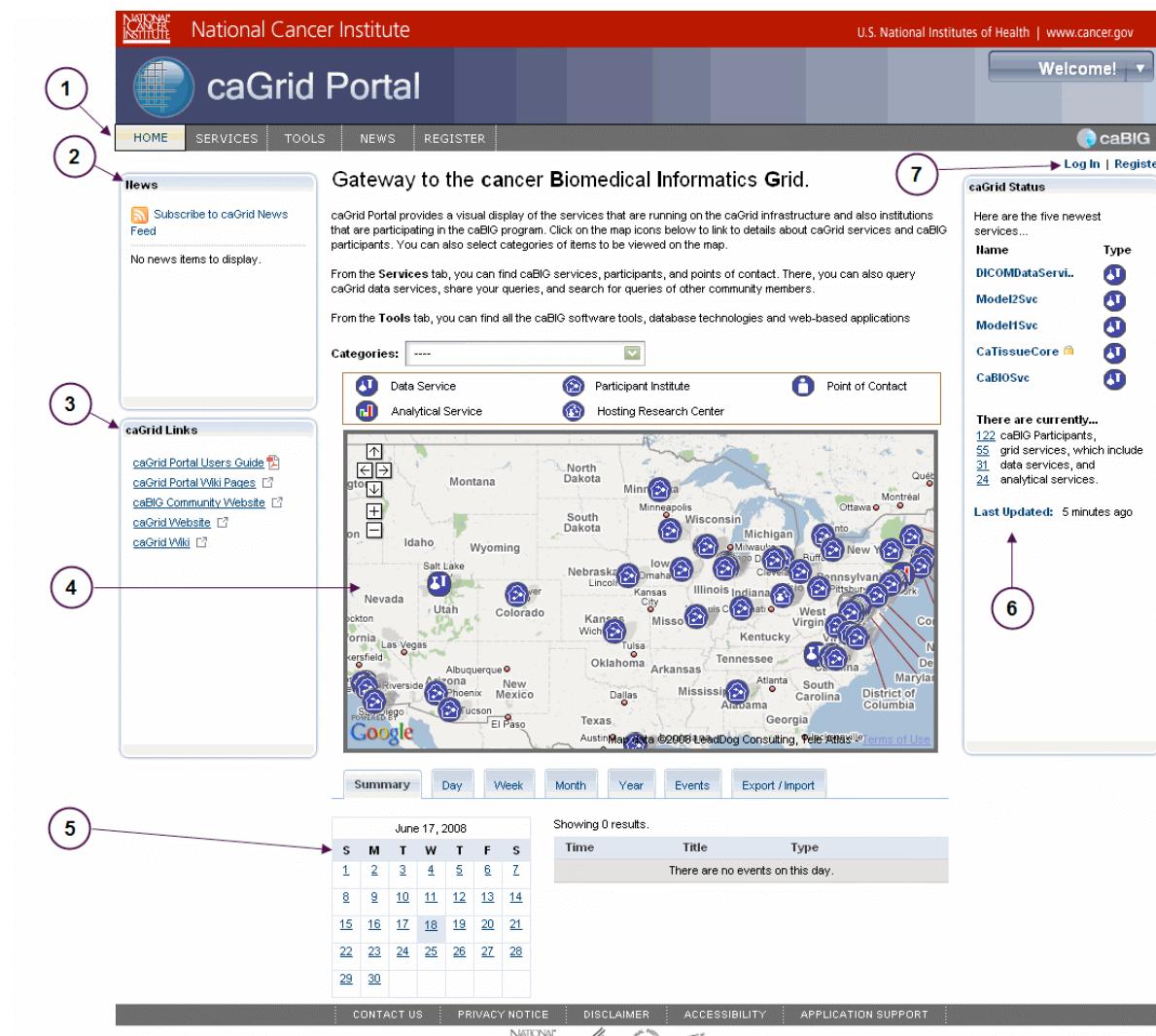


Figure 1-1. caGrid Portal home page

For a description of each numbered callout in Figure 1-1, see Table 1-1 on page 10.

Callout Number	Feature/Description
1	<p>Main Navigation Bar with four links:</p> <ul style="list-style-type: none"> • Home: Returns to the home page from any other portal page. • Services: Opens the Tools page, where you can use the Discovery and Data Service Query tools to search for services, participant institutions, and POCs. You can also build and share queries from this page. • Tools: A directory of caBIG tools is provided here. • News: Displays a News page with announcements and links to caGrid-related news. • Register page: Enables you to register as a caGrid Portal user.
2	<p>News portlet: Enables you to link to caGrid-related news stories and subscribe to the caGrid News Feed.</p>
3	<p>caGrid Links portlet: Lists links to external caGrid-related sites. Each link destination opens a new window or tab.</p>
4	<p>Map portlet with legend: Shows the geographic location of services and participant institutions. You can click the icon for a service or institution and link to more detailed information.</p>
5	<p>Calendar: Enables you to browse through caGrid-related events by day, week, month, and year.</p>
6	<p>caGrid Status portlet: Provides a quick, visual description of the scale and activity of caGrid.</p>
7	<p>Log In and Register links: Links for logging in with your existing account, or registering as a caGrid Portal user.</p>

Table 1-1. Description of feature callouts in Figure 1-1

Chapter 2 Registering and Logging In

Although you can freely browse and use the caGrid Portal without logging in to an established account, consider becoming a *registered* user. This chapter explains the benefits of registration, the procedure for registering, and the procedures for logging in as a new portal user or an NIH user.

Topics in this chapter:

- *About Registration* on this page
- *Registering for a Login Account* on page 12
- *Logging in with Your caGrid Portal Account* on page 15
- *Logging in with Your NIH Account* on page 16
- *About the Welcome Menu* on page 16

About Registration

As discussed in *caGrid Account Registration* on page 6, becoming a registered caGrid Portal user gives you Grid credentials so that you can interact with caGrid services. If you are a registered user, the portal maintains your query history from previous sessions and enables you to share queries with other users. In addition, future releases of the portal will enable you to further customize your user experience.

Note: If you already have a login account through the National Institutes of Health (NIH), you can log in to the portal using your existing credentials. For more information, see *Logging in with Your NIH Account* on page 16.

Registering for a Login Account

To register for a caGrid Portal login account, follow these steps:

1. Click either of the following **Register** links:
 - ° The link appearing on the main navigation toolbar, or
 - ° The link appearing in the upper right of the home page.

Figure 2-1 shows pointers to both links.

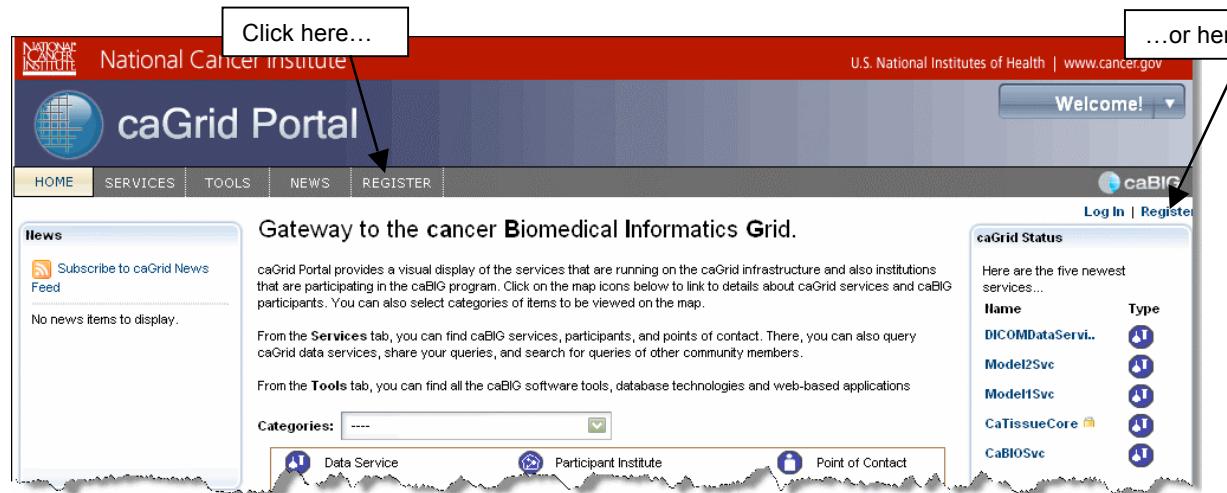


Figure 2-1. Register links

The Registration page appears, as shown in Figure 2-2.

National Cancer Institute
caGrid Portal

Registration

Why Register?
If you have an NIH username and password, then you do not need to register. You can just select the NCICB AuthenticationService IdP on the [Log In screen](#). If you don't have an NIH account then you will need to register in order to take advantage of some of the portal's features, e.g. sharing queries and interacting with secure caGrid services. Registering for an account here will enable you to log in to the portal using the NCICB Dorian Identity Provider. You cannot use this account to log in to the portal using the NCICB AuthenticationService Identity Provider.

Username Constraints:
Must be between 4 and 15 characters in length.

Password Constraints:
Must be between 10 and 20 characters and CANNOT contain a dictionary word and MUST contain at least one upper case letter, at least one lower case letter, at least one number, and at least one symbol (~!@#\$%^&*()_+={}[]|:;<>,.)

First Name:	*
Last Name:	*
Username:	*
Password:	*
Email:	*
Phone:	*
Organization:	*
Street Address 1:	*
Street Address 2:	*
City:	*
State:	AK
Postal Code:	*
Country:	AD *

Submit

* Indicates that a field is required.

CONTACT US | PRIVACY NOTICE | DISCLAIMER | ACCESSIBILITY | APPLICATION SUPPORT |

Figure 2-2. Registration page

2. Complete the registration fields.

Note: Pay special attention to the password constraints. A password cannot be a dictionary word and must be constructed as follows:

Length: 10 to 20 characters (10-character minimum)

Characters: at least one uppercase letter, one lowercase letter, one number, and one of the following symbols:
 ~!@#\$%^&*()_{}[]|:;<>,.

3. When you are finished, click **Submit**.

- ° If your username or password did not conform to the required constraints, an invalid username or password message appears in red.
- ° If your submission was successful, the following message appears in green: *Your account was approved, your current account status is Active.*

Note: Although you are still on the Registration page, do not click **Submit** again if you received a successful confirmation message.

4. Do either of the following:

- Click the **Home** link on the navigation bar to return to the home page, or
 - Click the **Log In** link to log in using the account that you just created.
5. Log in by following the instructions in *Logging in with Your caGrid Portal Account* on page 15.

Logging in with Your caGrid Portal Account

To log in with your caGrid Portal (non-NIH) account, follow these steps:

1. If you are not currently viewing the home page, click the **Home** link on the navigation bar.
2. Click the **Log In** link in the upper right corner of the page.
3. Enter your username and password.
4. Ensure that the **Identity Provider** field shows **NCICB Dorian**.
5. Click **Log In**.
 - ° If your login is not recognized, the following message appears:
Invalid username and/or password.
 - ° If your login was successful, the following message appears:
You are logged in as: {your name}

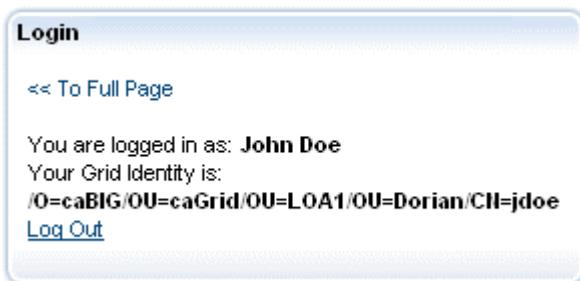


Figure 2-3. Successful login message

6. Click the **<< To Full Page** link to return to the home page.

In the upper right corner of the home page, a welcome message confirms that you are logged in.

Note: At the end of your session, be sure to log out by (1) clicking the **Log Out** link below the welcome message, or (2) clicking the drop-down arrow to the right of the Welcome message and selecting **Sign Out**.



Figure 2-4. Welcome message with Logout link

Logging in with Your NIH Account

To log in with your NIH account, follow these steps:

1. If you are not currently viewing the home page, click the **Home** link on the navigation bar.
2. Click the **Log In** link in the upper right corner of the page.
3. Enter your username and password.
4. Ensure that the Identity Provider field shows **NCICB Authentication Service**.
5. Click **Log In**.
 - If your login is not recognized, the following message appears:
Invalid username and/or password.
 - If your login was successful, the following message appears:
You are logged in as: {your name}
6. Click the **<< To Full Page** link to return to the home page.

In the upper right corner of the home page, a welcome message confirms that you are logged in. (See Figure 2-4 on page 15.)

Note: At the end of your session, be sure to log out by clicking the **Log Out** link below the welcome message.

About the Welcome Menu

The Welcome message in the upper right corner of the caGrid Portal includes a drop-down menu with four commands:

- **Home:** Returns to the home page
- **Sign Out:** Logs you out of the caGrid Portal
- **My Places:** Links to public pages.



Figure 2-5. Welcome menu

Chapter 3 Using the Map Portlet

As discussed in Chapter 1, the caGrid Portal home page includes a Map portlet that shows the geographic location of services and participant institutions in the United States. This chapter explains how to use the Map portlet to select a service or institution and link to more detailed information.

Topics in this chapter:

- *About the Map Portlet* on this page
- *Selecting a Category View* on page 1818
- *Selecting a Service, Participant, or POC* on page 1919

About the Map Portlet

The Map portlet provides a visual perspective of caGrid services, participants, and POCs. It serves as an access point for getting started with discovery.

As shown in Figure 3-1, the map is preceded by a **Categories** drop-down list and a legend explaining the map icons. The next two sections explain the list and the icons.

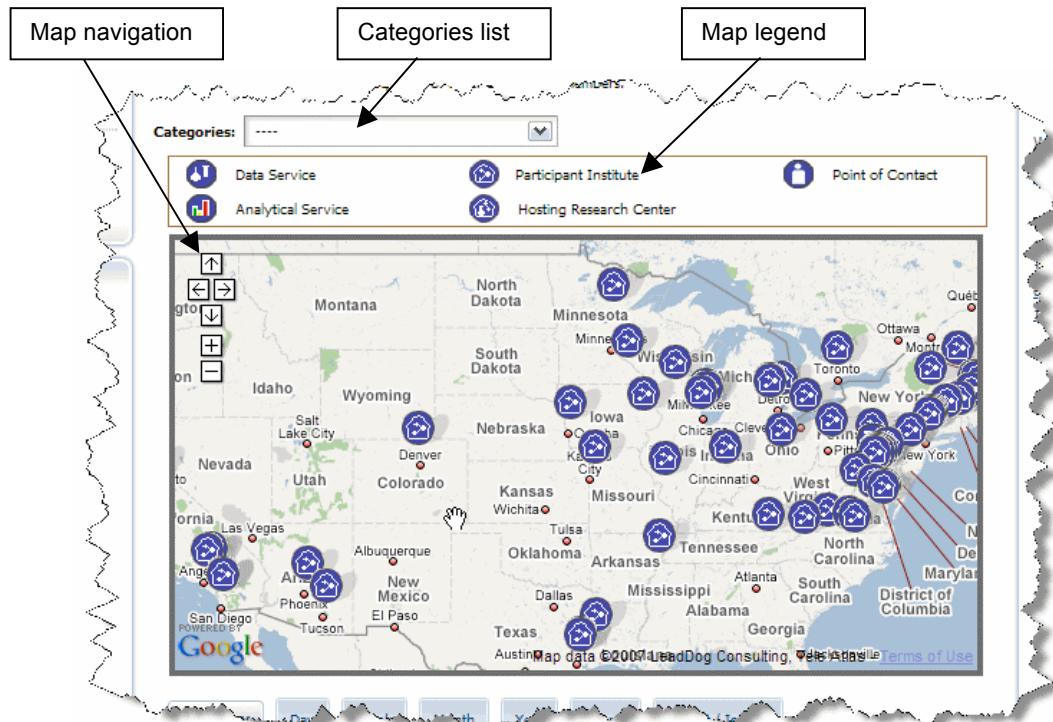


Figure 3-1. Map portlet

The Map portlet uses the Google Maps API and has standard Google Maps navigation features. These include navigation arrows, zoom in (+) and zoom out (-) buttons, and the “grabbing hand” mouse pointer for moving around the map.

Selecting a Category View

The **Categories** list enables you to filter the map view to show specific categories. The default view, shown as a series of dashes (----), displays all participant institutes.

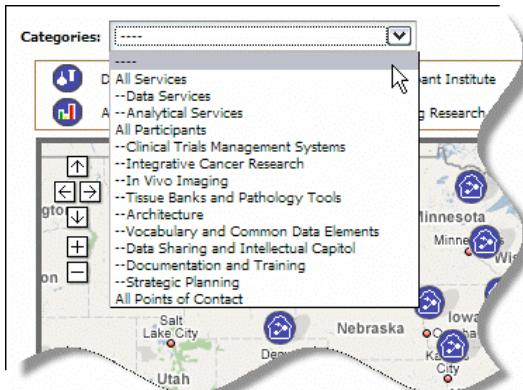


Figure 3-2. Categories list

The list also provides the following views:

- All Services
 - Data Services
 - Analytical Services
- All Participants
 - Clinical Trials Management Systems
 - Integrative Cancer Research
 - In Vivo Imaging
 - Tissue Banks and Pathology Tools
 - Architecture
 - Vocabulary and Common Data Elements
 - Data Sharing and Intellectual Capitol
 - Documentation and Training
 - Strategic Planning
- All Points of Contact (not currently available)

To select a category view, follow these steps:

1. Click the drop-down arrow to the right of the **Categories** list.
2. Select a category from the list.

Selecting a Service, Participant, or POC

Map portlet icons are clickable links. Table 3-1 lists each icon type and includes a cross-reference to the relevant section for working with the selected type.

Note: Icons for active entities appear in dark blue. Identical icons appearing in gray indicate inactive entities. Those appearing in red indicate partially active entities. Table 3-1 shows only icons for active entities.

Icon	Description
	Data Service See <i>Selecting a Data Service</i> on this page.
	Analytical Service See <i>Selecting an Analytical Service</i> on page 2121.
	Participant Institute: See <i>Selecting a Participant Institute</i> on page 2222.
	Hosting Research Center See <i>Selecting a Hosting Research Center</i> on page 2424. Note: A hosting research center represents one or more services (data or analytical) that are hosted at the same location.
	Point of Contact (POC) The Map portlet does not currently display POC information. For more information on finding and viewing details for a POC, see the following sections: <ul style="list-style-type: none">• <i>Searching for a POC</i> on page 3636• <i>Viewing Details for a POC</i> on page 4242

Table 3-1. Map icon descriptions for active entities

Selecting a Data Service.

To locate and view details for a data service, follow these steps:

3. Select the **Data Services** category from the drop-down list.

The map refreshes to show only data services.

4. Click a **data service** icon to select the represented service.

A callout balloon appears above the icon. The balloon displays the following information:

- a link to more details
- the name of the hosting center
- the current status of the hosting center (active, inactive, or partially active).

Figure 3-3 shows the results of clicking the NCICB hosting center *GridPIR*.

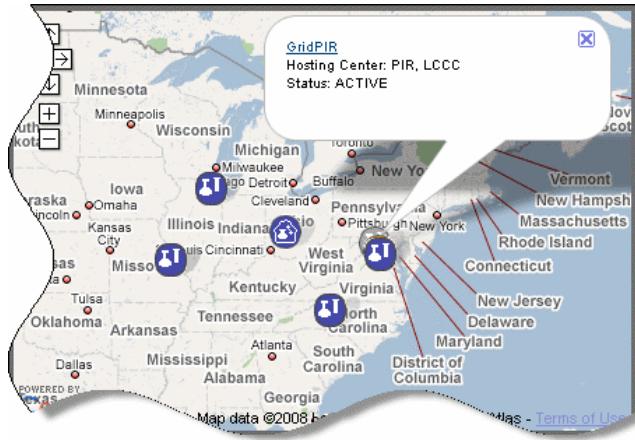


Figure 3-3. Callout balloon for a data service center

5. Click the underlined link.

The Services page appears with the **Details** tab and **Service** subtab of the Discovery portlet selected, as shown in Figure 3-4.

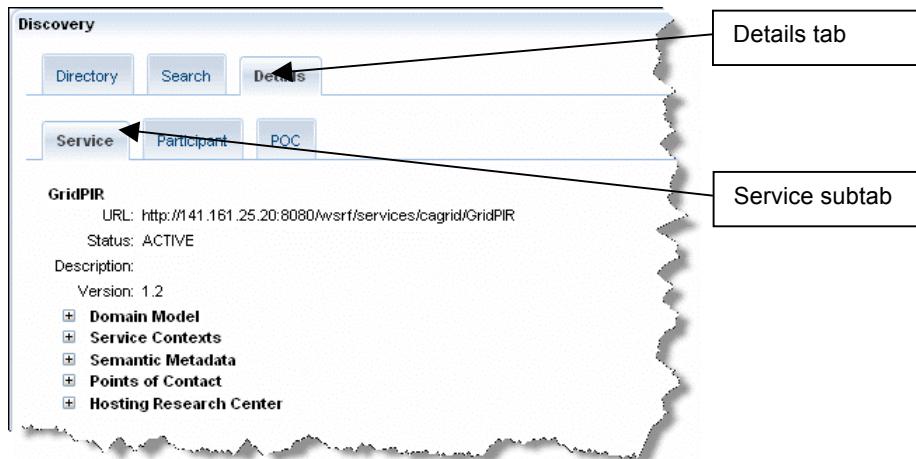


Figure 3-4. Discovery Portlet showing details for a selected data service

For more information about the Discovery portlet and the Service subtab, see *Viewing Details for Services* on page 3838.

Selecting an Analytical Service.

1. Select the **Analytical Services** category from the drop-down list.

The map refreshes to show only analytical services.

Note: The portal represents some analytical services using the **analytical service** icon  . Others that are part of a group of services in the same geographic location show a **hosting research center** icon  . You can find analytical services by clicking either icon, but the results will differ slightly, as explained in the next step.

2. Click the appropriate icon:

- ° If you clicked the **analytical service** icon  , the callout balloon displays the same information used for a data service center (shown in Figure 3-3 on page 2020):

- a link to more details
 - the name of the hosting center
 - the current status of the hosting center (active, inactive, or partially active).

Click the details link to view more information on the **Service** subtab of the Discovery portlet, as shown in Figure 3-4 on page 2020.

- ° If you clicked the **hosting research center** icon  , follow the steps explained in *Selecting a Hosting Research Center* on page 2424.

Selecting a Participant Institute.

1. Select the **All Participants** category from the drop-down list, or select a specific sub-category (for example, **Clinical Trials Management Systems**).

The map refreshes to show only the selected participant category.

2. Click a **Participant Institute** icon  to select the represented service.

A callout balloon appears above the icon. The balloon displays the following information:

- a participant name link
- a URL to the participant's home page.

Figure 3-5 shows an example of a callout balloon for Lombardi Comprehensive Cancer Center.

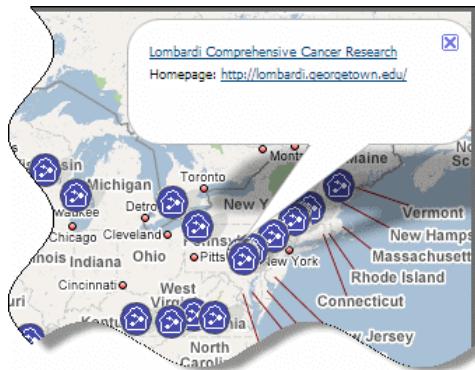


Figure 3-5. Callout balloon for a selected participant

3. Click the participant name link (top link).

The Services page appears with the **Details** tab and **Participant** subtab selected, as shown in Figure 3-6.

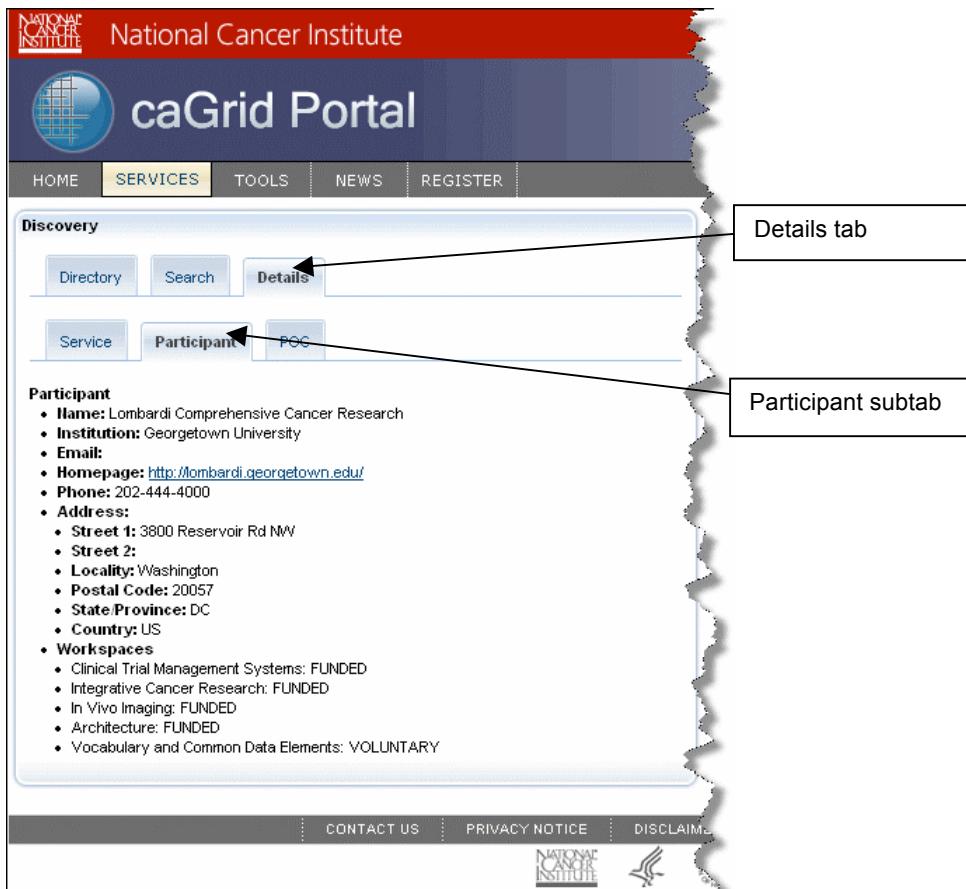


Figure 3-6. Discovery portlet showing details for a participant institute

For more information about the information available on this page, see *Viewing Details for a Participant* on page 4141.

Selecting a Hosting Research Center.

1. Find a **Hosting Research Center** icon , then click the icon.

A callout balloon appears above the icon. The balloon displays the following information:

- ° the number of services available at the selected location
- ° a **View** link.

Figure 3-7 shows an example of a callout balloon for a hosting research center.



Figure 3-7. Callout balloon for an analytical service center

2. Click the **View** link.

The Services page appears with the **Directory** tab and **List View** subtab selected. The **List View** subtab displays a list of services.

3. In the list at the bottom of the page, do the following:

- a. Locate the service that you want to view.
- b. Select the **More Details...** link on the right, as shown in Figure 3-8.

The screenshot shows the National Cancer Institute caGrid Portal's Discovery interface. At the top, there are tabs for HOME, SERVICES (which is selected), TOOLS, NEWS, and REGISTER. Below the tabs, there are two sub-tabs: List View (selected) and Map View. The main content area is titled "Discovery" and contains a table with three rows. Each row represents a hosting research center. The columns are Name, Institution, and Homepage. The "Homepage" column for each row contains a URL (e.g., <http://cont.hsc.usc.edu>) followed by a "More Details..." link. The entire screenshot is framed by a decorative border.

Name	Institution	Homepage
University of Southern California?Norris Comprehensive Cancer Center	University of Southern California	http://cont.hsc.usc.edu More Details...
USC/Norris Comprehensive Cancer Center - University of Southern California	University of Southern California	http://cont.hsc.usc.edu More Details...

Figure 3-8. List of hosting research centers with More Details links

For more information about the information available on the List View subtab, see *Viewing Details for an Analytical Service* on page 3939.

4. Click the **Home** link in the upper left to return to the home page.

Selecting a POC.

As explained in *Selecting a Category View* on page 1818, the Map portlet does not currently display POC information. For more information on finding and viewing details for a POC, see the following sections:

- *Searching for a POC* on page 3636
- *Viewing Details for a POC* on page 4242

Chapter 4 Using the Discovery Portlet

In Chapter 1, the *Discovery* section on page 4 introduced you to the concept of *discovery*. This chapter introduces the Discovery portlet, which enables you to discover caGrid services, caBIG participants, and caGrid POCs.

Topics in this chapter:

- *About the Discovery Portlet* on this page
- *Using the Directory Tab* on page 2828
- *Using the Search Tab* on page 3030
- *Using the Details Tab* on page 3838

About the Discovery Portlet

The Discovery portlet offers a versatile interface for browsing, searching, and viewing metadata. The interface includes drop-down lists, keyword entry fields, and search field selection boxes.

The various Discovery portlet features are available in a series of tabs with subtabs. Table 6-1 describes each tab and its subtabs.

Feature	Description
Directory tab	Enables you to browse for services, participants, or POCs by selecting directories from a drop-down list. You can use a List view or Map view (discussed below).
List View subtab	Displays browsing results in a paginated list. Each result includes a link for viewing more details (metadata).
Map View subtab	Displays selected items on a map of the U.S. and enables you to link to further information. This tab works the same way as the Map portlet discussed in Chapter 3 (p. 1717).
Search tab	Enables you to search for services, participants, and POCs using keywords and search fields.
Services subtab	Used to search for a service by entering one or more keywords and selecting fields to constrain the search. This tab includes both a simple and advanced field set.
Participants subtab	Used to search for a participant by entering one or more keywords and selecting fields to constrain the search.
POCs subtab	Used to search for a POC by entering one or more keywords and selecting fields to constrain the search.

Feature	Description
Details tab	Displays metadata for a selected service, participant, or POC.
Service subtab	Displays metadata for a selected service.
Participant subtab	Displays metadata for a selected participant.
POC subtab	Displays metadata for a selected POC.

Table 4-1. Discovery portlet tab descriptions

Using the Directory Tab

Use the **Directory** tab when you want to select a category and view its results in a list or on a map. This tab has two subtabs: **List View** and **Map View**. The next two sections explain how to use these subtabs.

Using the List View Subtab

The **List View** subtab enables you to select a category from the **Directories** list and view the results in a paginated results list. You can also select a saved result set from your current browser session.

Figure 4-1 shows the List View subtab with results displayed for the **All Services** category. The result list displays ten results per page. Note that each result includes a **More Details** link. Data services also include a **Query** link that enables you to build a query for the selected service.

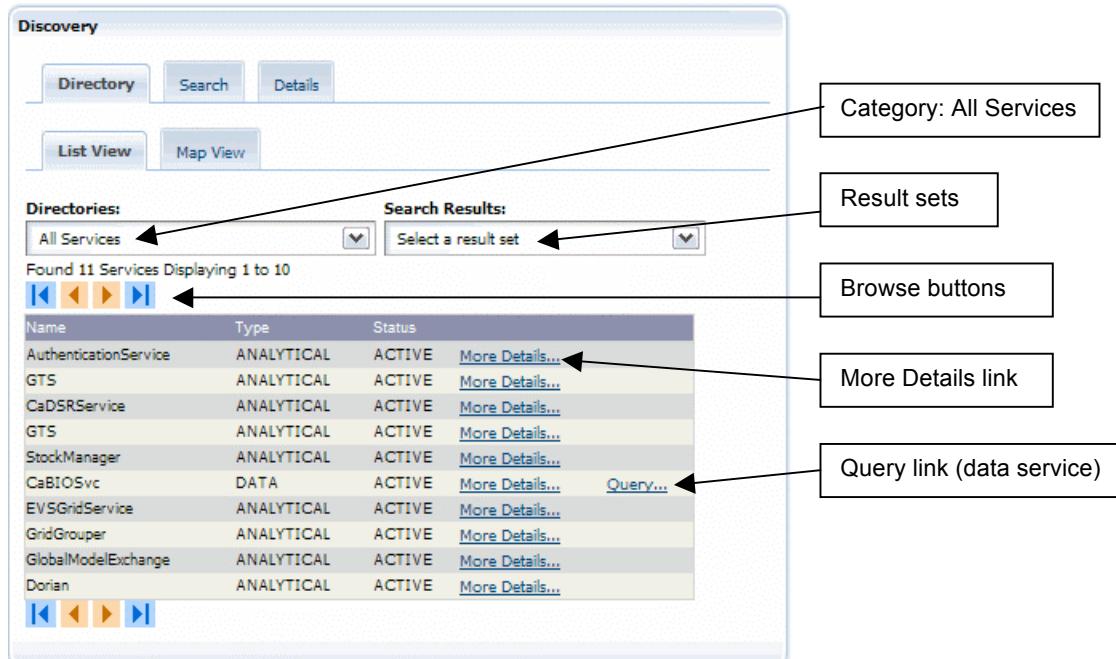


Figure 4-1. Directory tab > List View subtab – results for All Services category

To view results for a selected category, follow these steps:

1. Select the **Directory** tab.
2. Select the **List View** subtab.
3. Do either of the following:
 - ° Select a category from the **Directories** list; or
 - ° Select a result set from the **Search Results** list.The results appear in the lower area of the subtab.
4. Use the browse buttons to page through the list:
 - ° Click the **Forward** button  to page forward.
 - ° Click the **Back** button  to page backward.
 - ° Click the **Return**  to return to the beginning of the list.
 - ° Click the **Skip to End**  to skip to the end of the list.
5. To view details for a result, click the **More Details** link.

For more information about viewing details for a result, see *Using the Details Tab* on page 3838.
6. If the results include data services and you want to build a query for a selected service, click the **Query** link.

For more information about building queries, see *Building a Query Using the Query Builder Tool* on page 4646.

Using the Map View Subtab

After you select a category and view the results, you can switch to the **Map View** subtab for a graphical view of the results. Or, instead of starting with the List View subtab, you can start with the Map View subtab and select a category or result set using its drop-down lists.

To view search results using the Map View subtab, follow these steps:

1. Use the **List View** subtab to select a category and display the results.

For complete instructions, see *Using the List View Subtab* on page 2828.

2. Click the **Map View** subtab.

The map displays the results as shown in Figure 4-2.

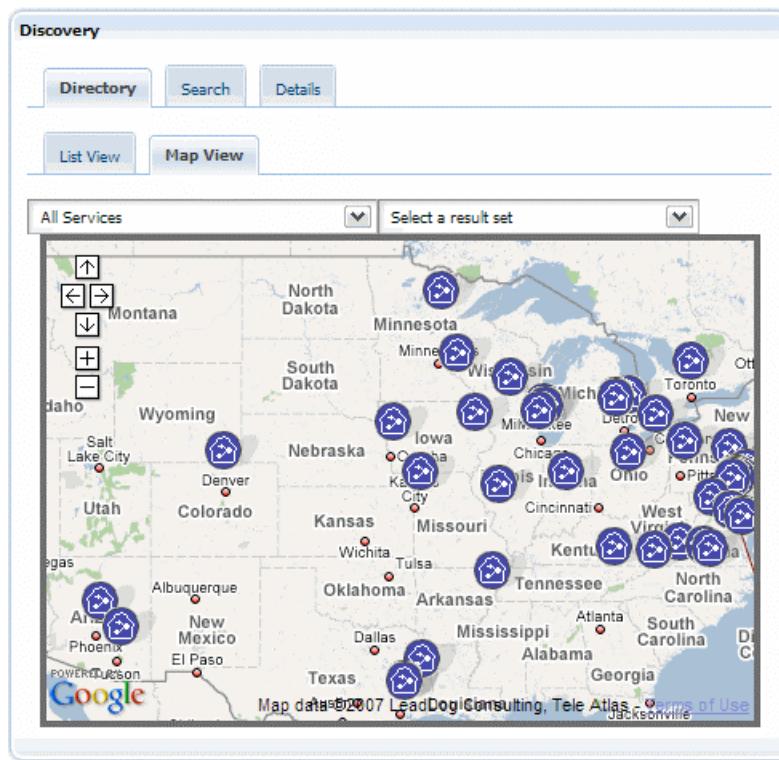


Figure 4-2. Map View subtab

The Map View subtab functions exactly like the Map portlet. For more information, see the following sections in Chapter 3, *Using the Map Portlet*, on page 1717.

- *About the Map Portlet* on page 1717
- *Selecting a Category View* on page 1818
- *Selecting a Service, Participant, or POC* on page 1919

Using the Search Tab

The **Search** tab provides a comprehensive solution for finding caGrid services, caBIG participants, or and caGrid POCs. Each entity has its own search subtab.

The following topics explain how to use each subtab to execute a search:

- *Searching for a Service* on this page 31
- *Searching for a Participant* on page 3535
- *Searching for a POC* on page 3636

Searching for a Service

Use the **Services** subtab to search for a service. You have three options for how to perform the search: Simple, Advanced, and EVS.

The Simple and Advanced search modes involve entering one or more keywords and then optionally selecting fields to narrow your results. The Simple search mode includes standard metadata fields while the Advanced search mode expands the available fields by including metadata from the domain model that services expose.

Using the EVS search mode involves entering concept names from the NCI Thesaurus. You can then find services that are related to these concepts. As you type each character in a concept name, the concept-based discovery Portlet suggests concepts that you can select. Results include all services involving the selected concept.

All searches use sub-string, case-insensitive matching.

For more information, see *Performing a Simple or Advanced Search* on page 3131 and *Performing an EVS Search* on page 3333.

Performing a Simple or Advanced Search

To perform a Simple or an Advanced search, follow these steps:

1. Click the **Search** tab.
2. Click the **Services** subtab.
3. Select the **Simple** or the **Advanced** search option. The Advanced search option includes more fields by which you can narrow your search results.

Figure 4-3 shows the layout of the Services subtab when the Simple search mode is selected.

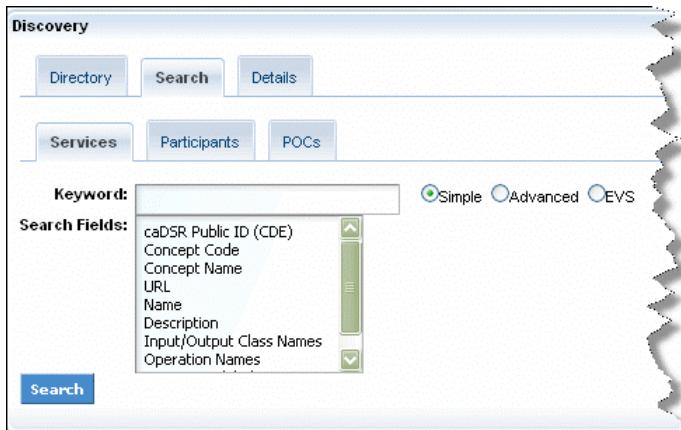


Figure 4-3. Services subtab with the Simple search mode selected

Figure 4-4 shows the layout of the Services subtab when the Advanced search mode is selected.

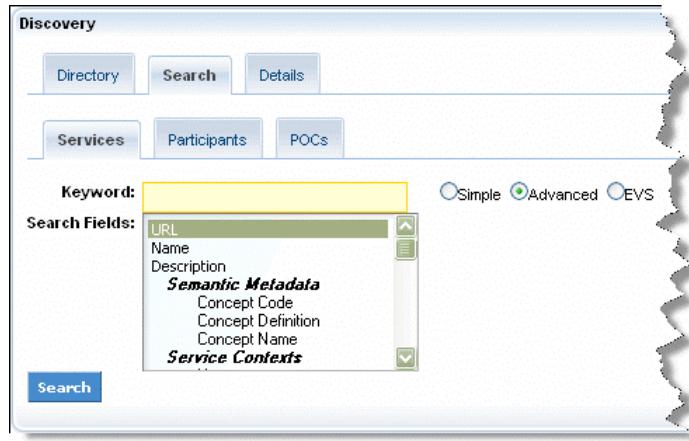


Figure 4-4. Services subtab with Advanced search mode selected

4. Enter one or more keywords in the **Keyword** field.
5. Optionally, restrict your search by selecting one or more fields in the **Search Fields** selection box.

Note: If you do not select any fields in the Simple or Advanced search modes, caGrid Portal includes all fields in its search.

To select consecutive fields in the **Search Fields** selection box, select a field, press and hold the SHIFT key, and then select the last field. To select fields that are not consecutive, select a field, press and hold the CTRL key, and then select additional fields.

6. Click **Search**. The results appear in the lower area of the subtab.

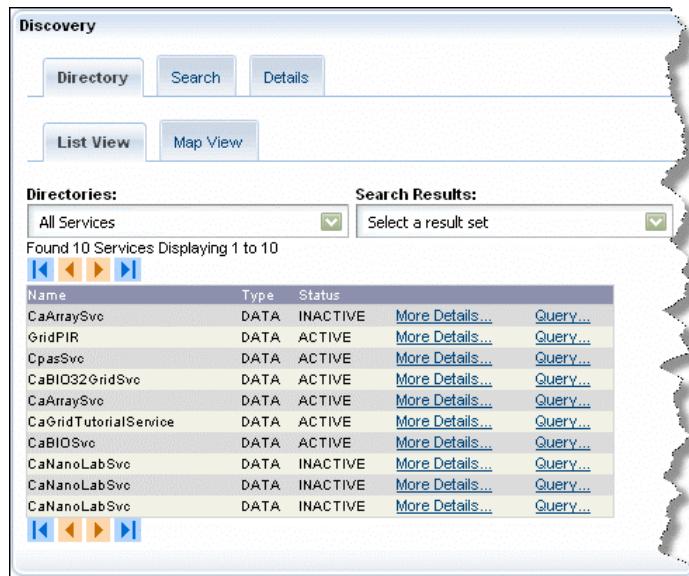


Figure 4-5. Discovered services

7. Use the browse buttons to page through the list:

- Click the **Forward** button  to page forward.
 - Click the **Back** button  to page backward.
 - Click the **Return**  to return to the beginning of the list.
 - Click the **Skip to End**  to skip to the end of the list.
8. To view details for a result, click the **More Details** link. For more information about viewing details for a result, see *Using the Details Tab* on page 3838.
 9. If the results include data services and you want to build a query for a selected service, click the **Query** link. For more information about building queries, see *Building a Query Using the Query Builder Tool* on page 4646.

Performing an EVS Search

To perform an EVS search, follow these steps:

1. Click the **Search** tab.
2. Click the **Services** subtab.
3. Select the **EVS** search option.

Figure 4-3 shows the layout of the Services subtab when the EVS search mode is selected.

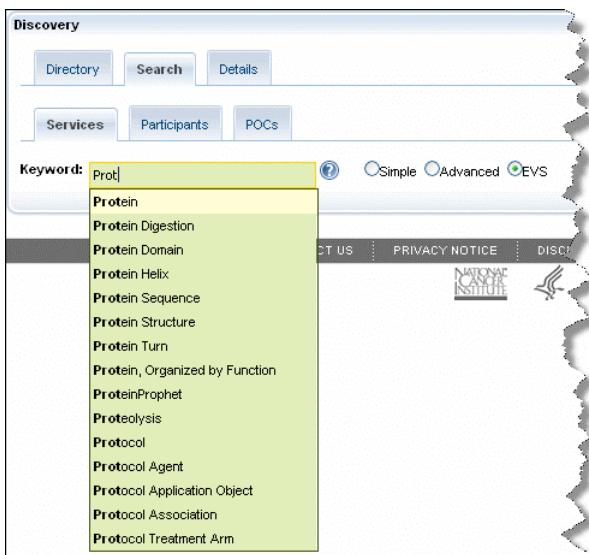


Figure 4-6. Services subtab with the EVS search mode selected

4. Discover services by entering concept names into the **Keyword** field. As you type each letter, the auto completer searches concepts relevant to services in the Portal. If it doesn't find a match, it looks up concepts in the NCI Thesaurus.
5. Select the concept matching the services you want to discover and then press ENTER. The concept code, concept definition, and relevant services for the selected concept appear in the tab.

Figure 4-7 shows the results of an EVS search.

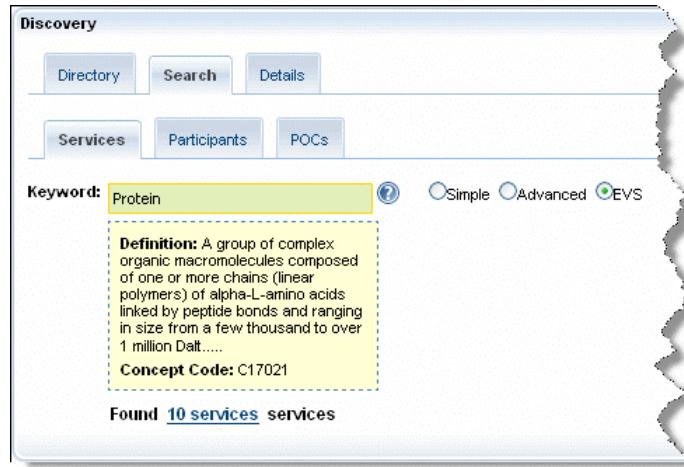


Figure 4-7. Results of an EVS search

6. Click the link noting how many services match the concept you selected. The results appear in the lower area of the subtab.

The screenshot shows the 'Discovery' interface with the 'List View' tab selected. It displays a table of services under the heading 'Directories: All Services'. The table has columns for Name, Type, and Status. Each service row includes a 'More Details...' link and a 'Query...' link. Navigation buttons at the top of the list allow for browsing through the results.

Name	Type	Status	
CaArraySvc	DATA	INACTIVE	More Details... Query...
GridPIR	DATA	ACTIVE	More Details... Query...
CpasSvc	DATA	ACTIVE	More Details... Query...
CaBIO32GridSvc	DATA	ACTIVE	More Details... Query...
CaArraySvc	DATA	ACTIVE	More Details... Query...
CaGridTutorialService	DATA	ACTIVE	More Details... Query...
CaBIOSvc	DATA	ACTIVE	More Details... Query...
CaNanoLabSvc	DATA	INACTIVE	More Details... Query...
CaNanoLabSvc	DATA	INACTIVE	More Details... Query...
CaNanoLabSvc	DATA	INACTIVE	More Details... Query...

Figure 4-8. Discovered services

7. Use the browse buttons to page through the list:
 - Click the **Forward** button to page forward.
 - Click the **Back** button to page backward.
 - Click the **Return** to return to the beginning of the list.
 - Click the **Skip to End** to skip to the end of the list.
8. To view details for a result, click the **More Details** link. For more information about viewing details for a result, see *Using the Details Tab* on page 3838.

9. If the results include data services and you want to build a query for a selected service, click the **Query** link. For more information about building queries, see *Building a Query Using the Query Builder Tool* on page 4646.

Searching for a Participant

Use the **Participants** subtab to search for a participant by entering one or more keywords and selecting fields to constrain the search. The search form includes a **Keyword** field and a **Search Fields** selection box, as shown in Figure 4-9.

Figure 4-9. Participants subtab

To search for a participant, follow these steps:

1. Click the **Search** tab.
2. Click the **Participants** subtab.
3. Enter one or more keywords in the **Keyword** field.
4. Select fields from the **Search Fields** selection box.

Note: If you do not select search fields, all of the fields are used in the search.

5. Click **Search**.

The results appear in the lower area of the subtab.

6. Use the browse buttons to page through the list:
 - ° Click the **Forward** button to page forward.
 - ° Click the **Back** button to page backward.
 - ° Click the **Return** to return to the beginning of the list.
 - ° Click the **Skip to End** to skip to the end of the list.
7. To view details for a result, click the **More Details** link.

For more information about viewing details for a result, see *Using the Details Tab* on page 3838.

Searching for a POC

Use the **POCs** subtab to search for a POC by entering one or more keywords and selecting fields to constrain the search. The search form includes a **Keyword** field and a **Search Fields** selection box, as shown in Figure 4-10.

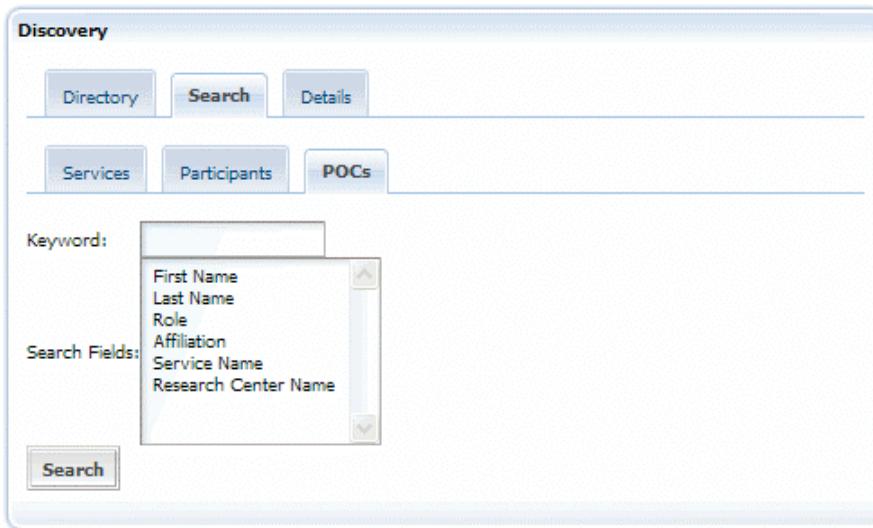


Figure 4-10. POCs subtab

To search for a POC, follow these steps:

1. Click the **Search** tab.
2. Click the **POCs** subtab.
3. Enter one or more keywords in the **Keyword** field.
4. Select fields from the **Search Fields** selection box.

Note: If you do not select search fields, all of the fields are used in the search.

5. Click **Search**.

The results appear in the lower area of the subtab.

6. Use the browse buttons to page through the list:
 - Click the **Forward** button to page forward.
 - Click the **Back** button to page backward.
 - Click the **Return** to return to the beginning of the list.
 - Click the **Skip to End** to skip to the end of the list.
7. To view details for a result, click the **More Details** link.

For more information about viewing details for a result, see *Using the Details Tab* on page 3838.

Using the Details Tab

The **Details** tab displays metadata for services, participants, and POCs. The following topics discuss the level of detail for each entity:

- *Viewing Details for Services* on this page
- *Viewing Details for a Participant* on page 4141
- *Viewing Details for a POC* on page 4242

Viewing Details for Services

caGrid Services provide a standard set of metadata. This metadata includes descriptions of the service's operations, information about the hosting research center and, if the service is a data service, the complete domain model that is supported by the service.

In addition, data types that the service operations use are defined by XML schemas and are registered with the Global Metadata Exchange (GME) service.

Viewing Details for a Data Service

To view details for a data service, follow these steps:

1. Use the **Directory** or **Search** tabs to find the desired data service.

For more information, see the following topics:

- *Using the List View Subtab* on page 2828
- *Searching for a Service* on page 3131

2. Click the **Details** tab.

Details for the selected data service appear on the **Service** subtab, as shown in Figure 4-11.

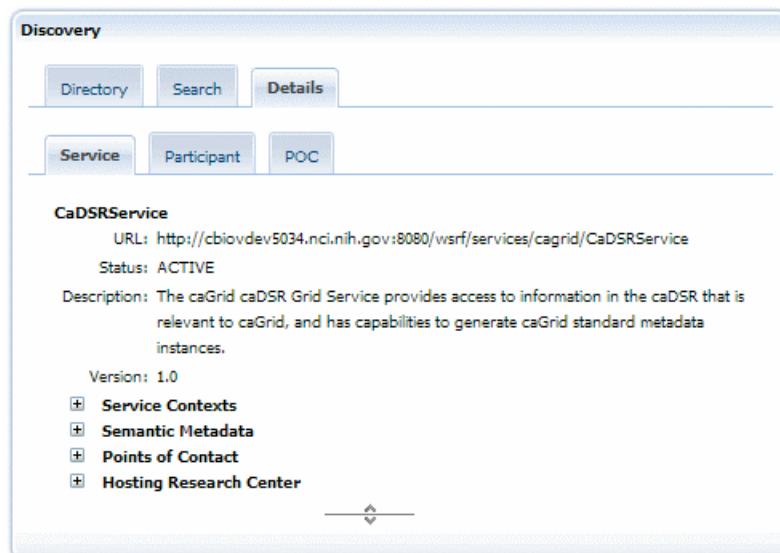


Figure 4-11. Details for a selected data service

3. Click the plus sign preceding each top-level of metadata to drill down to the sub-levels.

For example, when viewing details for a data service, you can drill down through the complete domain model that the service supports. You can then select a class for a query by clicking the **Select For Query** button, as shown in Figure 4-12.

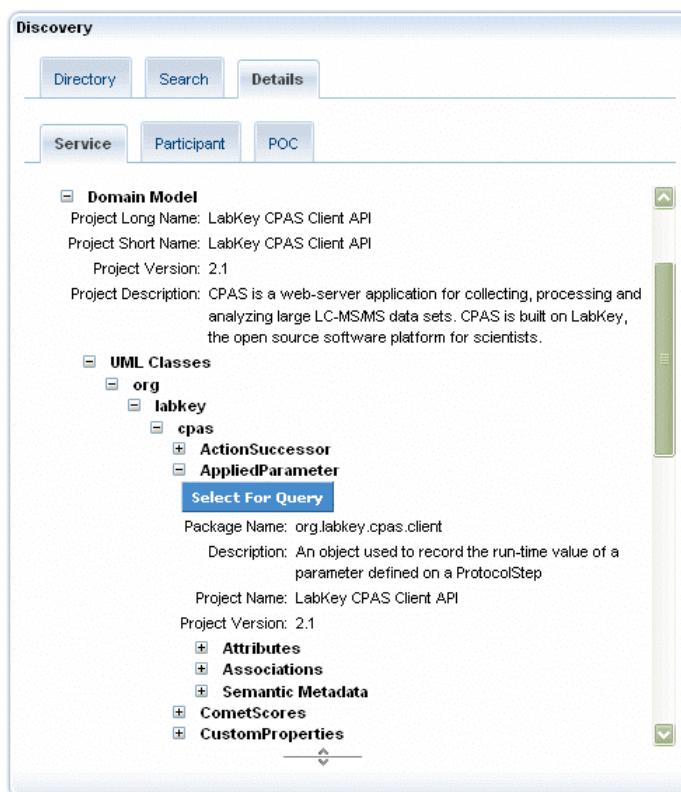


Figure 4-12. Selecting a UML class for a query

Viewing Details for an Analytical Service

To view details for an analytical service, follow these steps:

1. Use the **Directory** or **Search** tabs to find the desired analytical service.

For more information, see the following topics:

- ° *Using the List View Subtab* on page 2828
- ° *Searching for a Service* on page 3131

2. Click the **Details** tab.

Details for the selected analytical service appear on the **Service** subtab, as shown in Figure 4-13.

The screenshot shows the 'Discovery' interface of the caGrid Portal. At the top, there are tabs for 'Discovery', 'Directory', 'Search', and 'Details'. Under 'Details', there are sub-tabs for 'Service', 'Participant', and 'POC'. The main content area is titled 'CaDSRService' and displays the following information:

- URL: <http://cbiovdev5034.nci.nih.gov:8080/wsrf/services/cagrid/CaDSRService>
- Status: ACTIVE
- Description: The caGrid caDSR Grid Service provides access to information in the caDSR that is relevant to caGrid, and has capabilities to generate caGrid standard metadata instances.
- Version: 1.0
- Metadata sections (each with a plus sign):
 - Service Contexts
 - Semantic Metadata
 - Points of Contact
 - Hosting Research Center

Figure 4-13. Details for a selected analytical service

3. Click the plus sign preceding each top-level of metadata to drill down to the sub-levels.

Note: Analytical services do not include a domain model.

Viewing Details for a Participant

To view details for a participant, follow these steps:

1. Use the **Directory** or **Search** tabs to find the desired participant.

For more information, see the following topics:

- *Using the List View Subtab* on page 2828
- *Searching for a Participant* on page 3535

2. Click the **Details** tab.

Details for the selected participant appear on the **Participant** subtab, as shown in Figure 4-14.

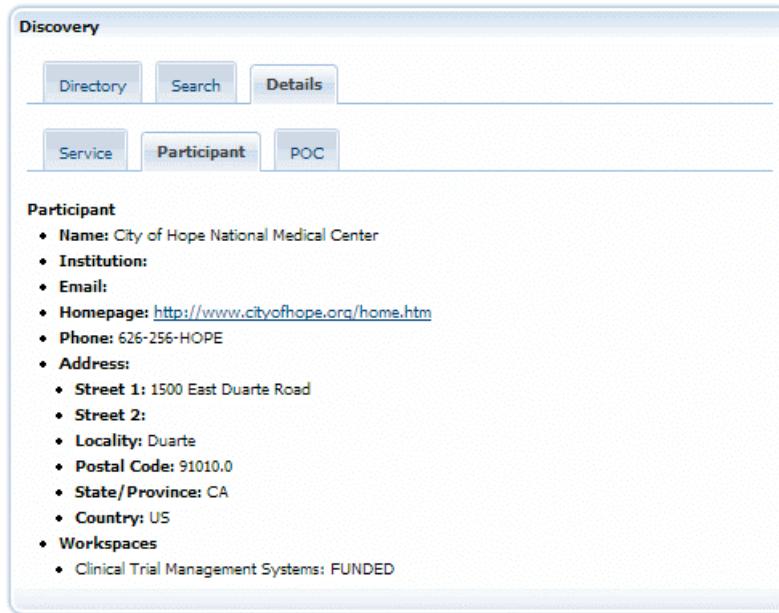


Figure 4-14. Details for a selected participant

Viewing Details for a POC

To view details for a POC, follow these steps:

1. Use the **Directory** or **Search** tabs to find the desired POC.

For more information, see the following topics:

- *Using the List View Subtab* on page 2828
- *Searching for a POC* on page 3636

2. Click the **Details** tab.

Details for the selected POC appear on the **POC** subtab, as shown in Figure 4-15.

The screenshot shows a 'Discovery' interface with a navigation bar at the top. The 'Details' tab is active. Below the tabs, there are three sub-tabs: 'Service', 'Participant', and 'POC'. The 'POC' sub-tab is also active. The main content area is titled 'Point of Contact' and lists the following details for Scott Oster:

- Name: Scott Oster
- Email: oster@bmi.osu.edu
- Phone:
- Roles
 - Role: Maintainer
 - Affiliation: OSU
 - Research Center: The Ohio State University
 - Role: Developer
 - Affiliation: OSU
 - Service: <http://cbiovdev5034.nci.nih.gov:8080/wsrf/services/cagrid/CaDSRService>

Figure 4-15. Details for a selected POC

Chapter 5 Using the Data Service Query Portlet

The Data Service Query portlet enables you to query services that support the CQL query language. This chapter explains how to build and share queries.

Topics in this chapter:

- *About the Data Service Query Portlet* on this page
- *Query-Building Scenarios* on page 4444
- *About the History Tab* on page 60
- *Working with Shared Queries* on page 62

About the Data Service Query Portlet

Using the Data Service Query portlet, you can build and execute a query (CQL or DCQL), view the results in tabular form, and export to Microsoft Excel or XML form. As long as you are a registered portal user, the portal maintains a complete history of your queries so that you can always retrieve and work with your previously executed queries.

The Data Service Query portlet interface is organized in a series of tabs. Table 5-1 describes each tab.

Tab	Description
Select UML Class	Enables you to enter a URL for a selected data service and view a list of UML classes.
Query	Provides two subtabs: <ul style="list-style-type: none">• Query Builder, which provides a graphical interface for building CQL and DCQL queries• Query XML, which enables you to enter a data service URL and query directly and then execute or share the query
History	Shows all of your stored queries.
Results	Displays query results in tabular or XML form.
Shared Queries	Provides a Find Shared Queries subtab that includes both a Keyword field and a Search Fields selection box. You can use these tools to search for shared queries.

Table 5-1. Data Service Query portlet tab descriptions

Tip: For more information about CQL and DCQL concepts, see http://www.cagrid.org/mwiki/index.php?title=Data_Services:CQL and http://www.cagrid.org/wiki/Federated_Query:DCQL.

Query-Building Scenarios

To execute a query, you need the following:

- The URL of the caGrid data service to which the query will be sent
- The query itself

The portal provides several approaches for specifying these requirements.

The following scenarios explain each approach.

Scenario 1: Starting without a URL or Query

If you do not have the service URL or the query, follow these steps:

1. Locate a service using either the **Directory** or **Search** tab of the Discovery portlet.
For more information, see *Searching for a Service* on page 3131.
2. If you want to view metadata for a UML class before selecting it for a query, follow substeps 2a through 2d. Otherwise, skip to step 3:
 - a. Click the **More Details** link for the desired service.
 - b. Expand the **Domain Model**.
 - c. Drill down to the metadata for a selected class.
 - d. Click **Select for Query**.
3. To select the desired service from the results list, follow these steps:
 - a. Click the **Query** link for the desired service.

On the Data Service Query portlet > **Select UML Class** tab, the **Service URL** field displays the URL for the selected service. A list of UML classes appears below the field, as shown in Figure 5-1.

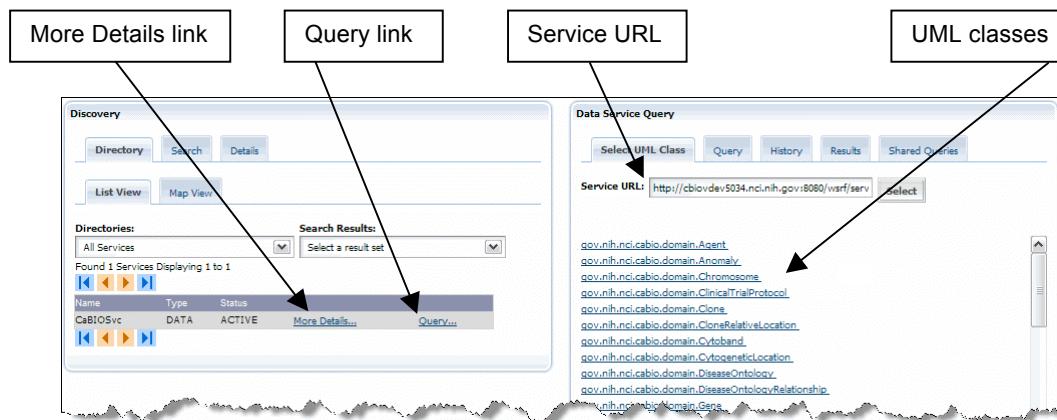


Figure 5-1. Starting a query from a selected result

- b. Click the desired UML class.

On the Data Service Query portlet, the **Query** tab > **Query Builder** subtab displays a working query, as shown in Figure 5-2.



Figure 5-2. Working query

Scenario 2: Starting with a URL but No CQL Query

If you have the service URL but no CQL query, follow these steps:

1. On the Data Service Query portlet, click the **Select UML Class** tab.
2. Enter the URL in the **Service URL** field.
3. Click **Select**.
4. Click the desired UML class from the list.

The **Query** tab > **Query Builder** subtab displays a working query.

Scenario 3: Starting with a URL and a CQL Query

If you have both the URL and the CQL query, follow these steps:

1. On the Data Service Query portlet, click the **Query** tab.
2. Click the **Query XML** subtab.
3. Enter the URL in the **URL** field.
4. Enter the CQL query in the **Query** text field.
5. Click **Submit Query**.

Building a Query Using the Query Builder Tool

caGrid data services all use a uniform query language, Common Query Language (CQL), to express queries against a data source using an object oriented language.

To express queries over a combination of caGrid data services, you must use Distributed caBIG Query Language (DCQL) instead. DCQL queries include two types, aggregate and distributed.

You can use the Query Builder tool to create both CQL and DCQL queries.

This section includes the following topics:

- *Building CQL Queries*, below
- *Building DCQL Queries* on page 5050

Building CQL Queries

The **Query** tab > **Query Builder** subtab provides a graphical interface for building queries. Although you can use this tool without knowing CQL, you can more easily use the tool if you understand CQL concepts. You can learn more at http://www.cagrid.org/mwiki/index.php?title=Data_Services:CQL.

To use the Query Builder subtab, you must first select the service and URL as explained in *Scenario 1: Starting without a URL or Query* on page 4444. After you have selected a UML class for the query, the Query Builder tab resembles Figure 5-3.



Figure 5-3. Query Builder subtab with selected UML class *ArrayDesign*

To modify and submit the query, follow these steps:

1. Click **Edit Query Modifiers** to select one of the options described in Table 5-2.

Option	Description
Object	Returns full results from the data service.
Count Only (the default)	Returns only the count of objects that match specified criteria.
Distinct Attribute	Returns the unique value of one of the following attributes of the UML class being queried: <ul style="list-style-type: none"> • clusterID • fullName • id • symbol You can select only one of these attributes.
Selected Attributes	Displays the same four attributes offered by the Distinct Attribute option, but enables you to select multiple attributes. If you select this modifier, the results will include only these attribute values for each object in the result set.

Table 5-2. Options for editing query modifiers

2. When you have finished editing modifiers, click **Update**.

The Result Type updates to show the selected attributes.

3. To add a criterion to the query, click **Add Criterion**.

The Query Builder subtab displays a list of attributes and associations for the selected class, as shown in Figure 5-4.

The screenshot shows the 'Data Service Query' interface. At the top, there are tabs: 'Select UML Class', 'Query' (which is selected), 'History', 'Results', and 'Shared Queries'. Below these are two sub-tabs: 'Query Builder' (selected) and 'Query XML'. A link '=< Back' is also present.

Select an attribute of ArrayDesign, or of an associated class.

ArrayDesign

- [assayType](#)
- [biqid](#)
- [id](#)
- [IsidAuthority](#)
- [IsidNamespace](#)
- [IsidObjectid](#)
- [name](#)
- [version](#)
- + [substrateType](#)
- + [annotationFile](#)
- + [designDetails](#)
- + [organism](#)
- + [polymerType](#)
- + [microarray](#)
- + [surfaceType](#)
- + [provider](#)
- + [designFile](#)
- + [technologyType](#)
- + [printing](#)
- + [AbstractArrayData \(CaArraySvc\)](#)
- + [AbstractDesignElement \(CaArraySvc\)](#)
- + [Array \(CaArraySvc\)](#)
- + [ArrayDataType \(CaArraySvc\)](#)
- + [ArrayDesign \(CaArraySvc\)](#)
- + [ArrayDesignDetails \(CaArraySvc\)](#)
- + [ArrayGroup \(CaArraySvc\)](#)

Figure 5-4. Attributes and associations for selected class

4. Select an attribute.
- A form appears with the name of the selected attribute.
5. Set the predicate by selecting it from the drop-down list.
 6. Enter a value.

The form should now resemble Figure 5-5.

The screenshot shows the 'Data Service Query' portlet interface. At the top, there are tabs: 'Select UML Class', 'Query', 'History', 'Results', and 'Shared Queries'. Below these are two sub-tabs: 'QueryBuilder' (which is selected) and 'Query XML'. The main area is titled 'Attribute Name' and contains a table with one row. The row has three columns: 'Attribute Name' (containing 'assayType'), 'Predicate' (containing 'EQUAL_TO'), and 'Value(s)' (containing 'geneExpression'). Below the table are three buttons: 'Update' (highlighted in blue), 'Delete', and 'Cancel'.

Figure 5-5. Specifying a predicate and value for a selected attribute

7. Click **Update**.

The Query Builder subtab should now resemble Figure 5-6.

The screenshot shows the 'Data Service Query' portlet interface with the 'QueryBuilder' sub-tab selected. Below the sub-tab are several buttons: 'Edit Query Modifiers', 'Aggregate Query', 'Add Criterion', 'Export To XML', and 'Submit Query' (which is highlighted in blue). Below these buttons, the query details are listed: 'Querying: gov.nih.nci.caarray.domain.array.ArrayDesign', 'Result Type: Object', and 'Criteria: assayType EQUAL_TO geneExpression'.

Figure 5-6. Query Builder subtab with added criterion

Although you can add multiple criteria to a query, the Query Builder does not currently support logical OR groupings other than for multiple values specified for the same criterion. Currently, the tool groups all criteria using a logical AND. If you specify multiple values for a specific criterion, the tool creates multiple criteria and groups them by OR. To further edit your query, you can export it to XML format.

Note: Once you export a CQL query to XML, you cannot edit the query and then switch back to using the Query Builder tool.

8. (Optional) To export your query to XML format, follow these steps:
 - Click **Export to XML**.

The **Query XML** tab appears in front. The **URL** field displays the service URL, and the **Query** field displays the XML structure for the query.

 - Edit the XML structure as needed.

Data Service Query

Select UML Class **Query** History Results Shared Queries

Query Builder **Query XML**

URL: `http://array.nci.nih.gov:80/wsrf/services/cagrid/CaArraySvc`

Query:

```
<ns1:CQLQuery xmlns:ns1="http://CQL.caBIG/1/gov.nih.nci.cagrid.CQLQuery">
<ns1:Target name="gov.nih.nci.caarray.domain.array.ArrayDesign">
<ns1:Group logicRelation="AND">
<ns1:Attribute name="assayType" predicate="EQUAL_TO" value="geneExpression"/>
</ns1:Group>
</ns1:Target>
</ns1:CQLQuery>
```

▼

Submit Query

Figure 5-7. Query XML tab

- When you have finished building the query, click **Submit Query**.

The **History** tab appears in front. This tab displays the query State, Start Time, Finish Time, and Service, as shown in Figure 5-8.

Data Service Query

Select UML Class **Query** **History** Results Shared Queries

Operations	State	Start Time	Finish Time	Service
---	COMPLETE	Jun 22, 2008 1:55:21 PM	Jun 22, 2008 1:55:23 PM	Service Name CaArraySvc
---	COMPLETE	Jun 22, 2008 1:51:09 PM	Jun 22, 2008 1:51:15 PM	Service Name CaArraySvc

Figure 5-8. History tab with query information

For more information about the History tab, see *About the History Tab* on page 60.

Building DCQL Queries

caGrid supports federated queries through a infrastructure service called the Federated Query Processor (FQP) and a language known as Distributed caGrid Query Language (DCQL). FQP service supports the execution of DCQL queries.

After you build DCQL queries using the Query Builder tool, the Portal then asynchronously executes them against an available FQP service.

There are two types of DCQL queries that you can construct in the caGrid Portal: aggregate and distributed. Aggregate queries aggregate similar data from multiple services. Distributed queries have distributed joins across multiple services on the grid.

The following sections explain how to create aggregate (below) and distributed join queries using DCQL (on page 53).

Building Aggregate DCQL Queries

Aggregate queries run against two or more data services and return an aggregated result set.

General steps for building an aggregate DCQL query in the caGrid Portal are as follows:

1. Discover caGrid Data Services providing a UML Class name.
2. Constrain the query using some query criteria.
3. Determine possible aggregation scenarios.
4. Execute the aggregate query.

For example, an aggregate query might involve retrieving all `ArrayDesigns` from the caArray data service hosted at Georgetown University **and** the caArray data service hosted at NCICB.

To build this aggregate DCQL query, follow these steps:

1. Select the service and URL for the query as explained in *Scenario 1: Starting without a URL or Query* on page 4444.



Figure 5-9. Query Builder specifying a service and URL

2. Click **Aggregate Query**. All services with the chosen UML class, in this case `ArrayDesign`, appear.

Data Service Query

[<< Back](#)

Following Services have the chosen UML Class. Select services to aggregate data from.

http://array.nci.nih.gov:80/wsrf/services/cagrid/CaArraySvc
 http://129.22.245.168:18080/wsrf/services/cagrid/CaArraySvc

Add Targets

http://array.nci.nih.gov:80/wsrf/services/cagrid/CaArraySvc
http://129.22.245.168:18080/wsrf/services/cagrid/CaArraySvc

Figure 5-10. Services with the same UML class to aggregate from

3. Select the checkboxes for all of the services from which to aggregate the chosen UML class.
4. Click **Add Targets**. A summary of the data services you selected for the aggregate query appears.

Data Service Query

Select UML Class **Query** History Results Shared Queries

Query Builder **Query XML**

Aggregate Query Add Criterion Export To XML **Submit Query**

Querying: gov.nih.nci.caarray.domain.array.ArrayDesign
Query Type: DCQL Query
Aggregating from:
http://array.nci.nih.gov:80/wsrf/services/cagrid/CaArraySvc
http://129.22.245.168:18080/wsrf/services/cagrid/CaArraySvc

Figure 5-11. Summary of data services selected for aggregation

5. Click **Submit Query**. The **History** tab appears in front. As shown in Figure 5-12, this tab displays the query state, start time, finish time, and service. The service is in this case is the Federated Query Processor.

Data Service Query

Select UML Class **Query** **History** Results Shared Queries

Operations	State	Start Time	Finish Time	Service
---	RUNNING	Jun 22, 2008 2:27:19 PM		Service Name FQP

Figure 5-12. History tab with query information

For more information about the History tab, see *About the History Tab* on page 60.

Building Distributed DCQL Queries

Distributed queries link semantically equivalent domain objects across multiple data services. You can construct DCQL queries manually and execute them through a command line. The caGrid Portal helps in visualizing the possible joins and then providing the tool to build DCQL queries that contain cross-service joins.

The following text describes how the caGrid portal can be used to query a grid service in a federated fashion. See [here](#) for a video that covers most of the same information.

The steps involved are:

1. Select a caGrid data service.
2. Select a UML class from the selected data service to query on.
3. Specify query criteria on the selected UML class by specifying join conditions based on a UML class from a second caGrid data service.
4. Execute the distributed join query and view the results.

We can correlate data between services by doing a distributed join. For example, we can retrieve all `NucleicAcidSequence` objects from the `caBIO32` grid service hosted at NCI/CBIIT where they are related to the proteins that we retrieve from the `GridPIR` grid service hosted at Georgetown University.

When building a query against the model of one service, the portal automatically displays UML classes in other models if they have at least one concept code in common. The service name next to the UML class name indicates the source of the UML class. Classes that share a concept code are most likely to also contain attributes that can be used as a join condition.

To build this aggregate DCQL query, follow these steps:

1. Select the service and URL for the query as explained in *Scenario 1: Starting without a URL or Query* on page 4444.

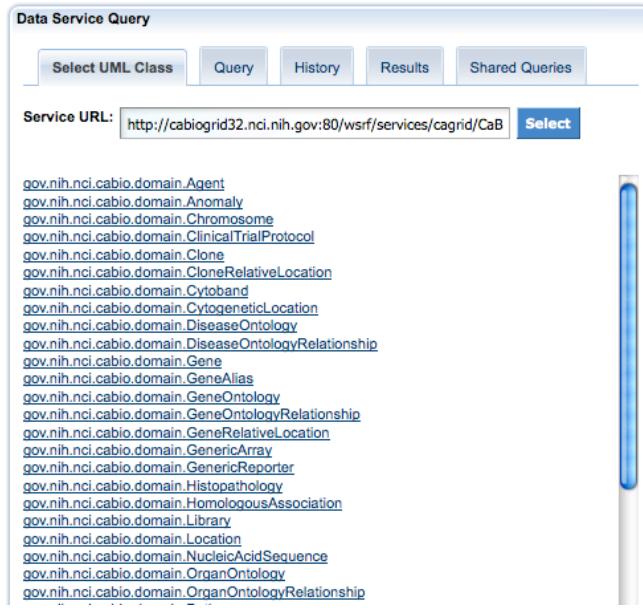


Figure 5-13. Query Builder specifying a service and URL

2. Select the `gov.nih.nci.cabio.domain.NucleicAcidSequence` class.

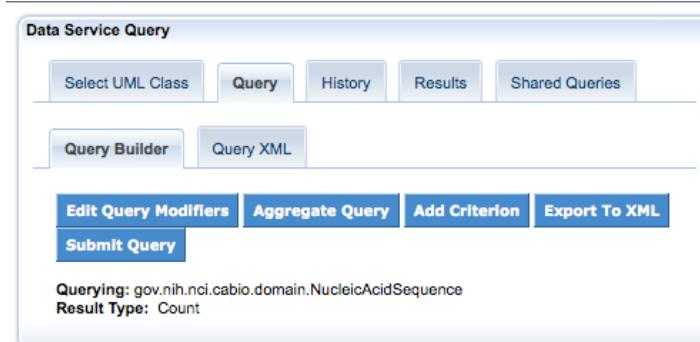


Figure 5-14 Query Builder tool with selected `NucleicAcidSequence` class

3. Click the **Add Criterion** button. All of the possible criteria (class attributes and associations) appear as shown in Figure 5-16.

The screenshot shows the 'Data Service Query' interface. At the top, there are tabs: 'Select UML Class' (selected), 'Query', 'History', 'Results', and 'Shared Queries'. Below these are two buttons: 'Query Builder' and 'Query XML'. A link '=> Back' is present. The main content area displays a tree view of the 'NucleicAcidSequence' class. The tree includes attributes like 'accessionNumber', 'accessionNumberVersion', 'biqid', 'id', 'length', 'type', and 'value'. It also lists associations: 'cloneRelativeLocation', 'locationCollection', 'databaseCrossReferenceCollection', 'geneCollection' (which is selected and highlighted in grey), 'clone', 'NucleicAcidSequence (CaBIOsvc)', and 'NucleicAcidSequence (CaGridTutorialService)'.

Figure 5-15 Tree view of attributes and associations for the NucleicAcidSequence class

4. Expand the **NucleicAcidSequence** class tree node by selecting the **geneCollection** association. Further expand the tree by selecting the **proteinCollection** association as shown in Figure 5-16.

This screenshot shows the same 'Data Service Query' interface as Figure 5-15, but with a different tree view. The 'geneCollection' association under 'NucleicAcidSequence' is expanded, showing its sub-elements: 'Gene', 'Protein', and 'ProteinFeature'. The 'Gene' node has attributes like 'biqid', 'clusterId', 'fullName', 'id', and 'symbol'. The 'Protein' node has attributes like 'biqid', 'copyrightStatement', 'id', 'keywords', 'name', 'primaryAccession', 'secondaryAccession', and 'unProtCode'. The 'ProteinFeature' node is also listed. Other associations like 'chromosome', 'nucleicAcidSequenceCollection', 'geneRelativeLocationCollection', 'geneAliasCollection', 'databaseCrossReferenceCollection', and 'histopathologyCollection' are also visible.

Figure 5-16 Query criterion tree view for caBIO NucleicAcidSequence class

Hovering the mouse over any service name that is next to a UML class provides further information about the service.

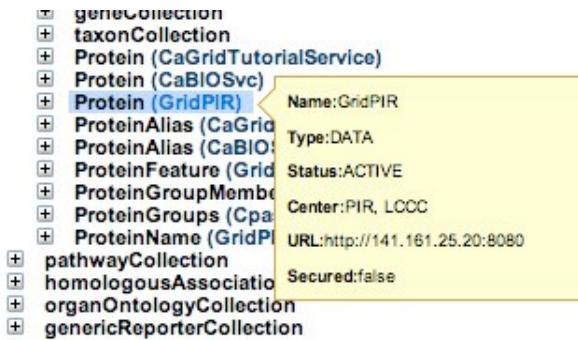


Figure 5-17 Service information for semantically equivalent UML class

6. Select the **Protein** class from the **GridPIR** service. The **Protein** class belongs to an external service (**GridPIR**) and is considered a foreign association in DCQL.
7. Expand the “foreign” UML class node. The Join Condition dialog box appears.
8. Select the Local Attribute as `uniProtCode` and the Foreign Attribute as `uniprotkbEntryName` as shown in Figure 5-18.

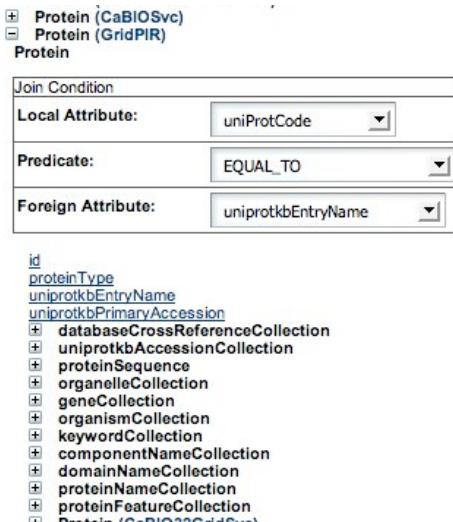


Figure 5-18 Selecting the join condition

9. Expand the tree by selecting the **organismCollection** association tree node on the **Protein** class. Click the **scientificName** attribute of the Organism class. This lets you specify a value on the **scientificName** criterion for the **Protein** class as shown in Figure 5-19.

The screenshot shows the 'Data Service Query' interface. At the top, there are tabs: 'Select UML Class', 'Query' (which is selected), 'History', 'Results', and 'Shared Queries'. Below these are two buttons: 'Query Builder' and 'Query XML'. The main area has a table with columns: 'Attribute Name', 'Predicate', and 'Value(s)'. In the first row, 'scientificName' is the attribute name, 'EQUAL_TO' is the predicate, and 'homo sapiens' is the value. At the bottom are three buttons: 'Update', 'Delete', and 'Cancel'.

Attribute Name	Predicate	Value(s)
scientificName	EQUAL_TO	homo sapiens

Figure 5-19 Entering criterion value for Organism scientificName attribute

10. Repeat steps 3-5. Select the **geneCollection** association tree node on the **Protein** class from **GridPIR**. Click the **name** attribute and enter **brca1** for the criterion as shown in Figure 5-20.

The screenshot shows the 'Data Service Query' interface. At the top, there are tabs: 'Select UML Class', 'Query' (selected), 'History', 'Results', and 'Shared Queries'. Below these are two buttons: 'Query Builder' and 'Query XML'. The main area has a table with columns: 'Attribute Name', 'Predicate', and 'Value(s)'. In the first row, 'name' is the attribute name, 'EQUAL_TO' is the predicate, and 'brca1' is the value. At the bottom are three buttons: 'Update', 'Delete', and 'Cancel'.

Attribute Name	Predicate	Value(s)
name	EQUAL_TO	brca1

Figure 5-20 Enter criterion value for Gene name attribute

The complete query is shown in Figure 5-21.

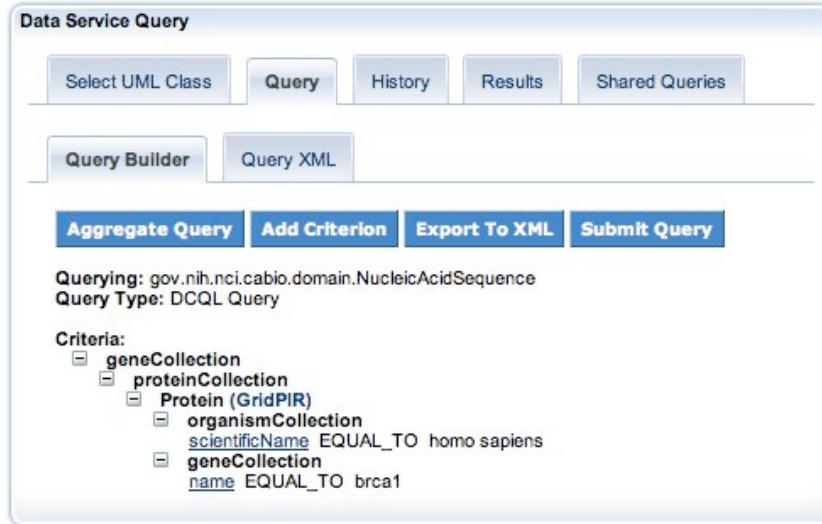


Figure 5-21 Complete federated query for caBIO NucleicAcidSequence class

11. Click the **Export To XML** button. The DCQL query appears in XML format, as shown in Figure 5-22.

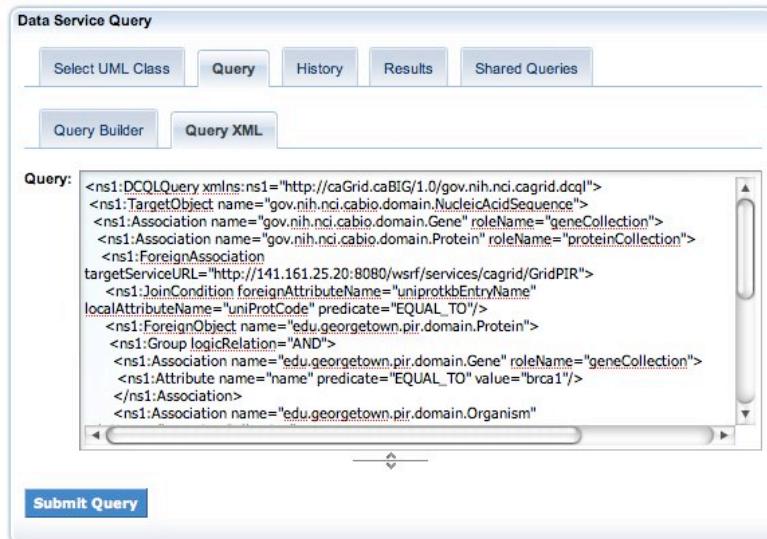


Figure 5-22 XML federated query for caBIO NucleicAcidSequence class

12. Click **Submit Query**. The query starts executing asynchronously. The **History** tab shows the running query as shown below.

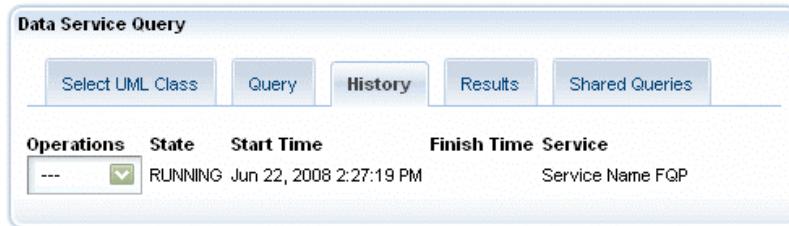


Figure 5-23 History Tab with the running DCQL query

After the query completes, the State text field automatically updates from *Running* to *Complete*.

13. In the Operations list, select **View Results**. The **Results** tab appears where you see the results of the distributed join query as shown in Figure 5-24.

The screenshot shows the 'Data Service Query' interface with the 'Results' tab selected. At the top, there are tabs for 'Select UML Class', 'Query', 'History', 'Results' (which is highlighted in blue), and 'Shared Queries'. Below the tabs, it says 'Displaying 1 to 10 of 210 results.' with links to 'Export to Excel' and 'Export to XML'. There are navigation arrows for navigating through the results. The main area is a table with columns: 'type', 'length', 'accessionNumberVersion', and 'value'. The 'value' column contains DNA sequence data. Some sequences are truncated with 'see full results' links. The table has 10 rows visible, with many more rows indicated by a scroll bar.

type	length	accessionNumberVersion	value
2	665	1	TTGGCTCAGCGGTTCCGTGGCAACAGTAAAGCGTGGATTACAGA
2	234	1	CGGCCGCTCA tcc cc GTTCACATCAAAGTGACTTTGG
2	654	1	t GTGAGCAATAGTTTTAATCACCTGGAGTCACCAAAAGC
2	909	1	AGTTGCCAGAAAAACACCACTCACITTAACTAATCTAATTACTGAAGA
			see full results]
2	554	1	gagacggaggcccttgtcgccaggctggaaatgcagtgccatgtctggcatgtatcgccatgcac
2	666	1	GCACCGAGGATTCTCTGAAGATAACCGTTAATAAGGCAACTTATTGCGAG
			see full results]
2	496	1	t GATAACTGTAACATTCGCTATAAACACTTGACTCTATTGCTAA
1	3642	2	CTTAGCGGTAGCCCCCTGGTTCCGTGGCAACGGAAAAGCGCGGAA
			to see full results]
2	846	1	TGCTCCCTGCCACGGCTCCCTTAcccgccgtgaaggatccctgcgttagccaccc
2	498	1	TGTGGAACCCCCCTTCTCCGCTGCCGCCAACGGTTCGGTCCCTCCG/

Figure 5-24 Results of federated query for caBIO NucleicAcidSequence class

About the History Tab

As noted in the last section, when you run a query from the Query Builder or Query XML subtabs, the **History** tab becomes the active tab. This tab displays information about all of the queries you have executed.

If you are a registered caGrid Portal user, the History list displays all queries that you have executed since you registered. If you are not registered, then you will see only queries that you have executed during the current browser session. The most recent query appears first.

The History tab also includes an **Operations** list, which includes the following operations:

- **View Results** – Displays the query results in the **Results** tab in tabular format as shown in Figure 5-25. The result list shows ten results at a time and includes buttons for browsing by page, skipping to the end of the list, or skipping to the beginning of the list. The **Export to Excel** link allows you to open or save the query in an Excel (*.xls) workbook and the **Export to XML** link allows you to open or save the query in XML format.

clusterId	fullName	symbol
121028	Asp (abnormal spindle) homolog, microcephaly associated (Drosophila)	ASPM
123659	Sphingomyelin phosphodiesterase, acid-like 3B	SMPDL3B
147111	ATPase, H+/K+ transporting, nongastric, alpha polypeptide	ATP12A
149910	Achaete-scute complex homolog 4 (Drosophila)	ASCL4
160786	Argininosuccinate synthetase 1	ASS1
168222	Acrosomal vesicle protein 1	ACRV1
207776	Aspartylglucosaminidase	AGA
100878	Asteroid homolog 1 (Drosophila)	ASTE1
109368	ATPase, class V, type 10B	ATP10B
464876	Additional sex combs like 3 (Drosophila)	ASXL3

Figure 5-25. Results tab showing query results

- **Reload** – Reloads the query into the **Query XML** subtab, where you can edit the query or run it again.
- **Delete** – Removes the query from the query history.

Table 5-3 describes the remaining information displayed on the History tab.

Option	Description
State	The current state of the query. The next five rows list each status and explain what operations you can perform while each status is in effect.
RUNNING	Indicates that the query is still running. You can cancel or delete the query.

Option	Description
COMPLETE	Indicates that the query completed successfully. You can view results, reload, or delete the query.
ERROR	Indicates that the query failed. You can view the error, reload the query, or delete the query.
CANCELED	Indicates that you canceled the query. You can reload or delete the query
TIMED OUT	Indicates that the system canceled the query. You can reload or delete the query
Start Time	The time that the query started running, displayed in hours, minutes, and seconds.
Finish Time	The time that the query finished running, displayed in hours, minutes, and seconds.
Service	The name of the target service.

Table 5-3. History tab information

Working with Shared Queries

If you are currently logged in, you can share new or previously executed queries with other caGrid Portal users. This section explains how to share a query and how to search for shared queries.

Note: Only CQL queries can be shared. DCQL queries cannot be shared at this time.

Sharing a Query

To share a query, follow these steps:

1. If you are not currently using the Data Service Query portlet, click the **Services** tab on the main navigation bar.
2. To share a *new* query, follow these steps:
 - a. Build the query as explained in *Building a Query Using the Query Builder Tool* on page 4646.
 - b. Skip to step 4.
3. To share a *previously executed* query, follow these steps:
 - a. Click the **History** tab.
 - b. Select the **Re-load** command from the **Operations** drop-down list.
 - c. Continue to the next step.
4. On the **Query** tab > **Query Builder** or **Query XML** subtab, click **Share Query**.

The **Shared Queries** tab > **My Queries** subtab becomes the active tab.

5. Enter a name and description for the query using the **My Queries** form shown in Figure 5-26.

Note: The name and description are required.

The screenshot shows the 'Data Service Query' portlet interface. At the top, there is a horizontal menu bar with tabs: 'Select UML Class', 'Query', 'History', 'Results', and 'Shared Queries'. The 'Shared Queries' tab is highlighted. Below the menu, there are two buttons: 'My Queries' (which is currently selected) and 'Find Shared Queries'. Underneath these buttons is a link '=> To shared queries list'. The main content area contains two input fields: 'Name:' and 'Description:'. The 'Name:' field is empty, and the 'Description:' field is also empty. At the bottom of the form are two buttons: 'Save' and 'Cancel'.

Figure 5-26. Form for adding a query name and description

6. Click **Save**.

A message confirms that the shared query has been successfully saved.

7. Click the << To shared queries list link.

The **Shared Queries** tab > **My Queries** subtab becomes the active tab. Your saved, shared query is summarized at the bottom of the form.

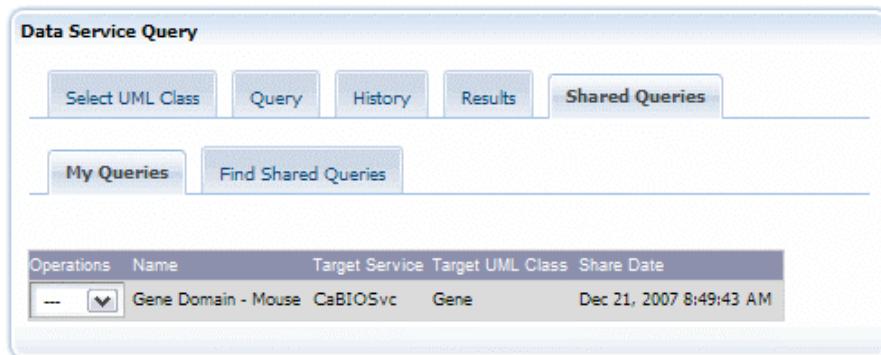


Figure 5-27. Saved query

The query summary shows the following information:

- A slightly different version of **Operations** drop-down list (see step 8)
 - Name
 - Target Service
 - Target UML Class
 - Shared Date
8. (Optional) Select any of the following options from the **Operations** drop-down list:
- **Edit** – Returns to the **Shared Queries** tab > **My Queries** subtab, where you can edit the query name and description.
 - **Load** – Runs the query and displays the **Query** tab > **Query XML** subtab.
 - **View** – Activates the **Shared Queries** tab > **My Queries** subtab and shows the query in XML format.

Searching for a Shared Query

You can search for shared queries using a keyword search and selected constraints, including the target service, target UML class, and query creator.

To search for a shared query, follow these steps:

1. If you are not currently using the Data Service Query portlet, click the **Services** tab on the main navigation bar.
2. Click the **Shared Queries** tab.
3. Click the **Find Shared Queries** subtab.

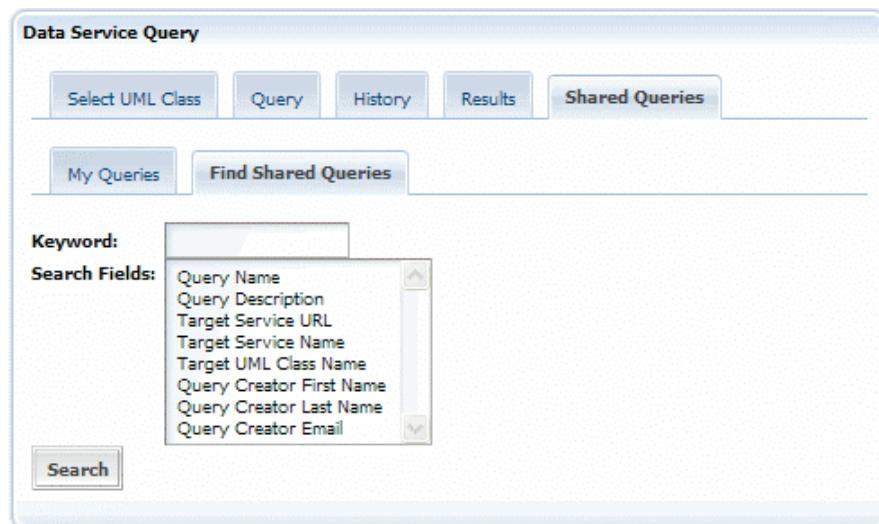


Figure 5-28. Shared Queries tab > Find Shared Queries subtab – Search form

4. Enter one or more keywords in the **Keyword** field.
5. Select desired search constraints in the **Search Fields** selection box.
Note: If you do not select search fields, all of the fields are included in the search.
6. Click **Search**.
The query results appear in a paginated list.
7. (Optional) To edit, load, or view a found query, select an option from the **Operations** list.

Chapter 6 Using Other Portal Features

The previous chapters covered the main caGrid Portal features used for discovering and querying caGrid services, caBIG participants, and caGrid POCs. The portal includes other useful portlets: *News*, *caGrid Links*, *caGrid Status*, and *Calendar*. This chapter provides an overview of the portlets.

Topics in this chapter:

- *Using the News Portlet* on this page
- *Using the caGrid Status Portlet* on page 68
- *About the Calendar Portlet* on page 69

Using the News Portlet

The caGrid Portal gives you the latest caGrid-related news through an RSS news feed. You can view the latest news on the portal site using two links:

- The **News** portlet on the home page
- The **News** link on the main navigation bar

The portlet displays news headlines as links. The posting date and time, as well as a portion of the text follow each headline from the article. Clicking a link opens the News page, where you can read the full article.

This link takes you directly to the News page, shown in Figure 6-2 on page 66.

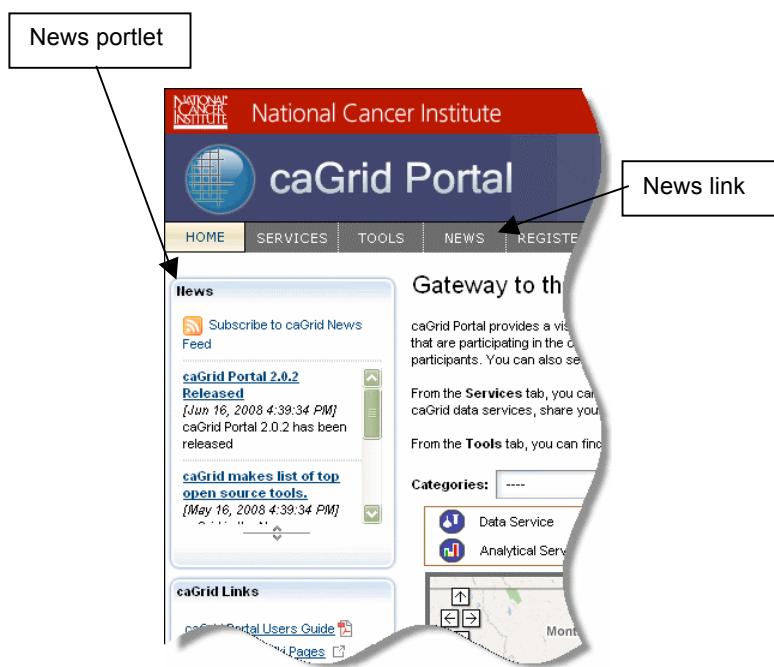


Figure 6-1. News links

On the News page, the main article appears in the middle area of the page. Other article links are listed in the **Other News** section at the bottom of the page.

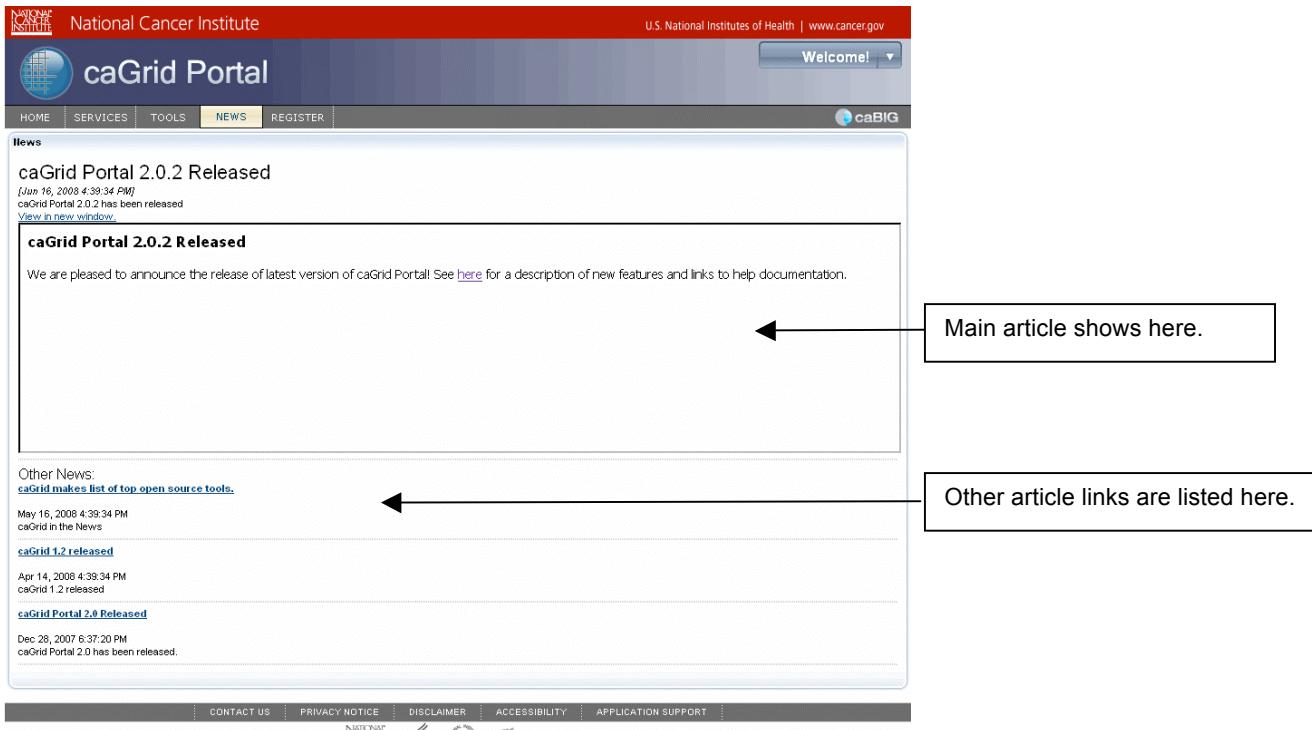


Figure 6-2. News page

Note: You can subscribe to the portal news feed directly from your browser or using any RSS 2.0-compliant news reader.

About the caGrid Links Portlet

The caGrid Links portlet includes links to the following sites:

- The caGrid Portal User's Guide at https://gforge.nci.nih.gov/plugins/scmcvs/cvsweb.php/~checkout~/cagrid-1-0/Documentation/docs/portal/portal2/guides/caGrid-Portal_2.0.2_UsersGuide.pdf?rev=HEAD;content-type=text%2Fplain;cvsroot=cagrid-1-0 is the latest version of this guide.
- caGrid Portal Wiki Pages at <http://www.cagrid.org/mwiki/index.php?title=Portal:Portal>. caGrid developers maintain these pages and use them to document all of the features of the caGrid Portal.
- The caBIG Community Website at <https://cabig.nci.nih.gov/>. This site provides a comprehensive overview of caBIG.
- The caGrid Website at <https://cabig.nci.nih.gov/workspaces/Architecture/caGrid>. This site provides an overview of the caGrid.

- The caGrid Wiki at <http://www.cagrid.org/mwiki/index.php?title=CaGrid>. This site provides a collaborative environment for caGrid users.

Using the caGrid Status Portlet

The caGrid Status portlet provides a quick, visual description of the scale and activity of caGrid.

As shown in Figure 6-3, the Status portlet is located on the right side of the home page, just below the **Login** and **Register** links.

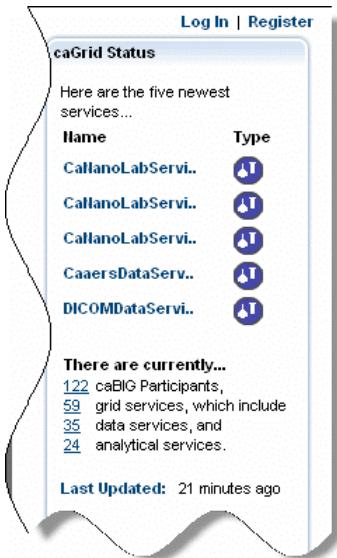


Figure 6-3. Status portlet

The following table explains the information that the portlet provides.

Status Item	Description
Service links	Click the links to the five most recently registered services. Links open the Services page where you can use the Discovery portlet to view details for a specific service or select from a list view.
Service info popup window	Hover your mouse over any service to quickly see the service's name, type, status, hosting center, hosting center URL, and secured status.
Type icons	Icons that denote the service type. See <i>Selecting a Service, Participant, or POC</i> on page 1919 for a table explaining each icon.
Counts	Up-to-date counts of all participants, services, data services, and analytical services.
Timestamp	Indicates the last time at which the portal aggregated data from the index service.

Viewing Details for a Recently Registered Service or Participant

The upper half of the Status portlet lists the five most recently registered services or participants with an identifying icon to the right of each listed item.

To view details for an item, click the name of the item. The Services page appears with the **Details** tab and **Service** subtab of the Discovery portlet selected. From

there you can drill down to more detail as discussed in Chapter 4, *Using the Discovery Portlet*, on page 2727.

Viewing a List of Participants or Services

The lower half of the Status portlet shows current counts for caGrid participants and services. Here you can view totals for all participants, all services, all data services, and all analytical services.

To view a list of services or participants, click the underlined number preceding each item. The Services page appears with the **Directory** tab and **List View** subtab of the Discovery portlet selected. The list display varies according to whether you selected a participant or service. You can select a list item and drill down to more detail as discussed in *Using the List View Subtab* on page 2828.

About the Calendar Portlet

Appearing at the bottom of the caGrid Portal home page, the Calendar portlet provides dates of important caGrid-related events. Figure 6-4 shows this portlet as it appears on the home page.

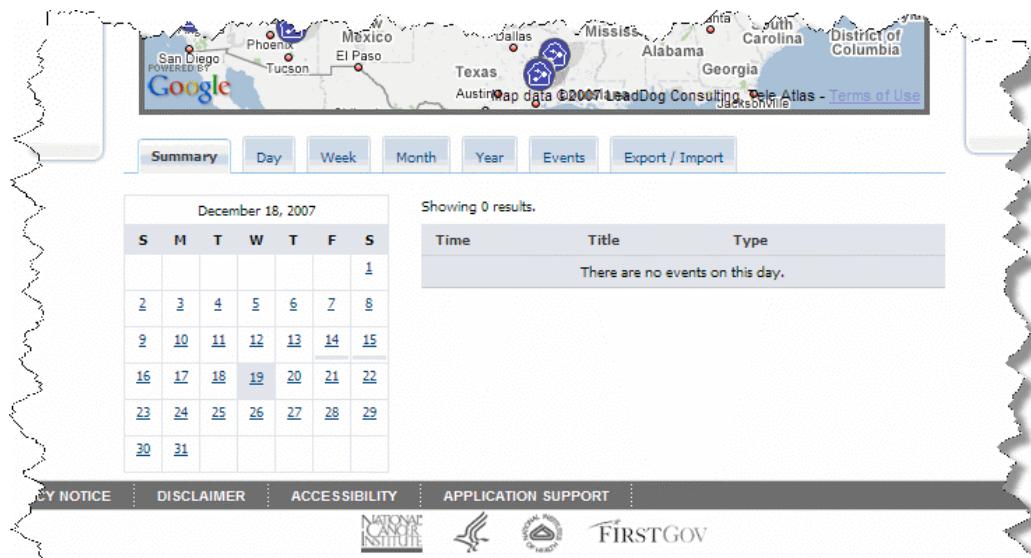


Figure 6-4. Calendar portlet

You can change the view of the calendar by clicking the various tabs described in Table 6-1.

Tab	Description
Summary	Shows a calendar on the left and a list of events on the right.
Day	Shows events for the selected day and enables you to browse forward or backward in one-day increments.
Week	Shows a seven-day view starting with the selected day and enables you to browse forward or backward in seven-day increments.
Month	Shows the current month as a full-sized calendar.

Year	Shows a calendar for each day of the current year.
Events	Shows a listing of events, including the date, time, title, and type. Click the event title to view details for the event.
Export/Import	Enables you to import a file and add an event to your calendar.

Table 6-1. Description of Calendar portlet tabs

Chapter 7 Using the caBIG Tools Portlet

The caBIG Tools portlet allows you to access the same inventory of tools that are available on the [caBIG Community Website](#). This chapter explains how to browse those tools.

Topics in this chapter:

- *About the caBIG Tools Portlet* on this page
- *Sorting caBIG Tools* on page 7244
- *Viewing caBIG Tool Details* on page 60

About the caBIG Tools Portlet

To expedite the cancer research community's access to key bioinformatics tools, platforms, and data, the caBIG Community Website (<https://cabig.nci.nih.gov/inventory>) lists applications (tools) and their documentation, as well as infrastructure and datasets used to support the caBIG™ initiative. The inventory is updated as new items are released.

The caGrid Portal provides another entry point into the tools component of that inventory. When you click the caGrid **Tools** tab, the most current list of caBIG tools appears.

The caBIG™ tools inventory includes over 40 software tools, database technologies, and Web-based applications in the areas of clinical trials management, biospecimens, imaging, genome annotation, proteomics, microarrays, pathways, data analysis and statistical tools, data sharing, infrastructure, vocabularies, and translational research. They are based on open-source software and free. Download packages include installation elements and documentation. In many cases training is also offered.

As on the caBIG website, caGrid users can search for these tools based on name, compatibility level, keyword, area of focus, workspace, and technical specifications.

The features of the caGrid Tools portlet are illustrated in Figure 7-1.

Tools

Sort

Num current inve...

Search Results for caBIG™ Tools

Sort by: **Title** | Area of Focus | Primary Workspace | Maturity Level | Compatibility Level | caGrid Enabled | Installation Level

Definitions for headers

Bioconductor

Primary Workspace: ICR
Maturity Level: Development Phase
Installation Level: Intermediate - technical assistance may be required, download may require supporting infrastructure or software
Product Description:
 Bioconductor is an established open-source collection of software packages for high throughput genome analysis. Packages adapted for caBIG allow preprocessing of microarray data, DNA copy number assessment from gene expression data, and SELDI-TOF mass spec peak finding.

caAdapter

Primary Workspace: CTMS
Maturity Level: Stable Release (Adoption in Progress)
Installation Level: Basic - Wizard or web browser application; minimal technical assistance required
Area of Focus: Data Sharing
caGRID Enabled: No
Compatibility Level: Not yet determined

CONTACT US | PRIVACY NOTICE | DISCLAIMER | ACCESSIBILITY | APPLICATION SUPPORT

NATIONAL CANCER INSTITUTE FIRSTGOV

Figure 7-1. Tools portlet

Sorting caBIG Tools

While the caBIG tools are sorted alphabetically by title by default, you can change the sort order to more easily find and understand each tool.

To sort caBIG tools, follow these steps:

1. If you are not currently using the Tools portlet, click the **Tools** tab on the main navigation bar. The Tools page appears.
2. Click one of the sort headers at the top of the page. See Table 7-1 for a description of each header.

The tools in the list are sorted by the selected header. The sort header currently in effect is highlighted in red.

The screenshot shows a web-based application interface for the caBIG Tools Portlet. At the top, there is a navigation bar with links for 'Search', 'Definitions for headers', and 'Logout'. Below the navigation bar, a header reads 'Search Results for caBIG™ Tools' and includes a 'Sort by:' dropdown menu with options: Title, Area of Focus, Primary Workspace, Maturity Level, Compatibility Level, caGrid Enabled, and Installation Level. A link 'Definitions for headers' is also present. To the right of the sort dropdown, it says '45 results returned'. The main content area displays a list of search results. The first result is 'cancer Text Information Extraction System (caTIES)'. Below the result title, there is a summary box containing the following information:

- Primary Workspace:** TBPT
- Maturity Level:** Stable Release (Adoption in Progress)
- Installation Level:** Intermediate - technical assistance may be required, download may require supporting infrastructure or software
- Product Description:** The cancer Text Information Extraction System (caTIES): A locator to tissue resources via the extraction of coded information from free text surgical pathology reports (SPRs), is using controlled terminologies to populate caBIG-compliant data structures. Provides researchers with the ability to query, browse, and acquire annotated tissue data and physical material across a network.
- Area of Focus:** Biospecimens
- caGrid Enabled:** Yes
- Compatibility Level:** Silver

Figure 7-2. caBIG tools sorted by Area of Focus

The following table explains each of the sort headers available in the caBIG Tools portlet.

Sort Header	Description
Title	This is the name of the tool as chosen by the development team and the caBIG Workspace who fostered the development. Tool titles are sorted alphabetically.
Area of Focus	The Area of Focus is the primary domain-specific audience or technical area that the tool serves. Only one area is specified from the following choices: <ul style="list-style-type: none"> • Biospecimens • Data Analysis and Statistical Tools • Clinical Trials Management • Genome Annotation • Data Sharing • Infrastructure • Imaging • Pathways • Microarrays • Translational Research • Proteomics • Other • Vocabularies

Sort Header	Description
Primary Workspace	<p>The caBIG Workspace where the tool was originally developed and maintained. Only one area is specified from the following choices:</p> <ul style="list-style-type: none"> • ARCH (Architecture Workspace) • CTMS (Clinical Trials Management Systems Workspace) • DSIC (Data Sharing and Intellectual Capital Workspace) • DT (Documentation and Training Workspace) • ICR (Integrative Cancer Research Workspace) • IMAG (Imaging Workspace) • SP (Strategic Planning Workspace) • TBPT (Tissue Banks and Pathology Tools Workspace) • VCDE (Vocabulary and Common Data Elements Workspace)
Maturity Level	<p>Developers of the tool are required to assign a maturity level assessment to the version of the tool. The tool maturity level must be classified as one of the following:</p> <ul style="list-style-type: none"> • Development Phase • Beta-Test or Pilot Release Available (Pre-Adoption) • Stable Release (Adoption in Progress) • Mature Product (Successfully Adopted)
Compatibility Level	<p>To aid in the creation of software that will be able to interoperate within the caBIG program, a set of compatibility guidelines was developed that spells out requirements for interoperability in areas of Interface Integration, Vocabularies/Terminologies and Ontologies, Information Models and Data Elements. Systems that meet the requirements are said to be caBIG Compatible</p> <p>The levels of compatibility are:</p> <ul style="list-style-type: none"> • Gold Level Compatibility • Silver Level Compatibility • Bronze Level Certification • Legacy <p>Learn more about compatibility at https://cabig.nci.nih.gov/guidelines_documentation.</p>
caGrid Enabled	<p>caGrid is the distributed infrastructure built upon established standards and technologies that provides the concrete implementation for this interoperability to happen seamlessly. When systems are built following the caGrid specifications, these systems are able to discover, access and use other similarly built services programmatically. These software systems are then called "caBIG Grid Enabled".</p> <p>In terms of compatibility guidelines, "caBIG Grid Enabled" applications would become the "Gold" compatible systems once the caBIG Cross cutting work spaces start performing "Gold" compatibility reviews.</p>

Sort Header	Description
Installation Level	<ul style="list-style-type: none"> Basic - Wizard or web browser application; minimal technical assistance required Intermediate - technical assistance may be required, download may require supporting infrastructure or software Advanced - technical team required for installation

Table 7-1. Sort header options

Viewing caBIG Tool Details

Click any tool in the list to view details about that tool. Details about each tool are structured similarly and include the following information next to the tool logo:

- Product Summary
- Demo Files & Exercises
- Documentation & Training
- Tool Download Files
- End User Support Resources
- Related caBIG Tools.
- Developer Institution(s) & Adopter Institution(s)
- Related Articles

Figure 7-3 shows the home page for one caBIG tool.

The screenshot displays the National Cancer Institute's caGrid Portal. At the top, there is a red banner with the National Cancer Institute logo and the text "U.S. National Institutes of Health | www.cancer.gov". Below the banner, the portal has a blue header with the "caGrid Portal" logo and navigation links for "HOME", "SERVICES", "TOOLS" (which is highlighted), "NEWS", and "REGISTER". On the right side of the header, there is a "Welcome!" button and a "caBIG" logo. The main content area is titled "Tools" and features a large image of the Bioconductor logo, which includes a green DNA helix and the word "BIOCONDUCTOR". To the right of the logo is a sidebar titled "Quick Links to Page Contents" containing links to various sections: "Product Summary", "Demo Files & Exercises", "Documentation & Training", "Tool Download Files", "End User Support Resources", "Related caBIG™ Tools", "Developer Institution(s) & Adopter Institution(s)", and "Related Articles".

Figure 7-3. caBIG tool details

Index

A

analytical service
selecting (Map portlet), 23
viewing details for, 42

C

caBIG tools
sorting, 78
viewing details, 81
caBIG Tools portlet, 77
Calendar portlet, about, 8, 75
Categories list (Map portlet), 20

D

data service
query, about, 6
selecting (Map portlet), 21
viewing details for, 40
Data Service Query Portlet, about, 45
DCQL queries
aggregate, 54
building, 54
distributed, 56
Details tab
Participant subtab, 43
POC subtab, 44
Service subtab, 40
details, viewing
for caBIG tools, 81
for data service, 40
for new participant, 75
for participant, 43
for POC, 44
Directory tab
List View subtab, 30
Map View subtab, 32
Discovery portlet, about, 29
discovery, about, 4–5

E

events. See Calendar portlet

H

History tab, about, 64
home page, about, 9
hosting research center, selecting (Map portlet), 26

I

icons for Map portlet, 21

L

Links portlet, about, 73
List View subtab, using, 30
logging in
as NIH user, 16
as registered user, 15

M

Map portlet
about, 19
analytical service, selecting, 23
Categories list, 20
data service, selecting, 21
hosting research center, selecting, 26
icons for, 21
participant institute, selecting, 24
POC, selecting, 27
Map View subtab, using, 32
metadata
for analytical service, 42
for data service, 40
for participant, 43
for POC, 44
for services, 5

N

News portlet, about, 8, 71

P

participant
searching for, 37
viewing details for, 43
participant institute, selecting (Map portlet), 24
POC
searching for, 38
selecting (Map portlet), 27
viewing details for, 44
portlet
Calendar, 75
Data Service Query, 45
Discovery, 29
Links, 73
Map, 19
News, 71
Status, 74
Tools, 77

Q

query
building, 49
building a CQL, 49
building a distributed DCQL, 56

building an aggregate DCQL, 54
for data service, 6
portlet overview, 45
scenarios for building, 47
shared, searching for, 68
sharing
 about, 6
 procedure for, 66

R

registration
 about, 11
 procedure for, 12
 technical details for, 6

POC subtab, 38
Services subtab, 33
searching
 for participant, 37
 for POC, 38
 for service, 33
 for shared query, 68
service
 analytical, viewing details for, 42
 data, viewing details for, 40
 searching for, 33
 viewing details for, 40
Status portlet
 about, 8
 using, 74

S

Search tab
 Participant subtab, 37

W

Welcome menu, about, 16