

CAARRAY 2.0

User's Guide



NATIONAL[®]
CANCER
INSTITUTE

Center for Bioinformatics

Revised November 1, 2007

CREDITS AND RESOURCES

caArray Development and Management Teams			
Development	Quality Assurance	Documentation	Project and Product Management
Eric Tavela ²	Tom Boal ⁵	Eric Tavela ²	Mervi Heiskanen ¹
Bill Mason ²	Ron Keene ⁵	Brent Gendleman ²	Anand Basu ^{1c}
Todd Parnell ²	Xiaopeng Bian ¹	Todd Parnell ²	Brent Gendleman ²
Dan Kokotov ²		Paul Duvall ⁴	Juli Klemm ¹
Rashmi Srinivasa ²		Levent Gurses ⁴	
Scott Miller ²		Jill Hadfield ¹	
Systems and Application Support		Training	
Wei Lu ³		Don Swan ³	
Vanessa Caldwell ³			
Michael Gomes ³			
¹ National Cancer Institute Center for Bioinformatics (NCICB)		² 5AM Solutions	³ <u>Terrapin Systems</u>
⁴ Stelligent	⁵ NARTech		

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

TABLE OF CONTENTS

Credits and Resources	i
Using the caArray User's Guide	1
Introduction to the caArray User's Guide	1
Organization of this Guide	1
User's Guide Text Conventions	2
Chapter 1	
About caArray	5
caArray Overview	5
Relationship of caArray to caBIG	6
Chapter 2	
Getting Started in caArray	7
Launching caArray	7
Browsing caArray	8
caArray New User Account and Login	10
Requesting a User Account	11
Using caArray Online Help	13
Navigating the caArray User Interface	13
Elements in the caArray User Interface	14
caArray Welcome Page Navigation Menu	15
User Interface Footer	16
My Experiment Workspace	16
Chapter 3	
Navigating and Searching caArray	19
Browsing caArray	19
Searching the Repository	22
Experiment Search Results	24

Chapter 4

Creating and Managing Experiments27

Overview of an Experiment	27
Creating an Experiment	28
Steps for Creating an Experiment	28
Overview Tab	29
Contacts Tab	32
Annotations Tab	32
Managing Controlled Vocabulary Terms	38
Data Tab	43
Publications Tab	44
Submitting an Experiment	45
Updating An Experiment Proposal	45
Editing an Experiment	45
Adding Data to an Experiment	46
Deleting an Experiment	46
Managing Experiment Visibility	47

Chapter 5

Submitting Data to an Experiment49

Managing Data	49
Uploading Data Files	50
Validating Data Files	51
Importing Data	53
Supplemental Files	55
Downloading Files	55

Chapter 6

Extracting Data from caArray57

Downloading Data from caArray	57
Grid availability	58
Extracting Data Programmatically by API	58

Chapter 7

User Account Management61

Roles in caArray	61
Managing User Accounts	62
Managing Collaboration Groups	62
Creating a Group	63
Viewing Group Details	63
Editing Collaboration Group Details	64

Appendix A	
caArray References	67
Appendix B	
caArray Glossary	69
Index	73

USING THE CAARRAY USER'S GUIDE

This chapter introduces you to the *caArray 2.0 User's Guide* and suggests ways you can maximize its use.

Topics in this chapter include:

- [Introduction to the caArray User's Guide](#) on this page
- [Organization of this Guide](#) on this page
- *User's Guide Text Conventions* on page 2

Introduction to the caArray User's Guide

The *caArray 2.0 User's Guide* is the companion documentation to the caArray software application. The *caArray User's Guide* includes information and instructions for the end user about using caArray.

Organization of this Guide

The *caArray 2.0 User's Guide* contains the following chapters:

Using the caArray User's Guide—This chapter introduces you to the *caArray 2.0 User's Guide* and suggests ways you can maximize its use.

Chapter 1 About caArray—This chapter provides a caArray overview as well as a description of the relationship of caArray to caBIG™.

Chapter 2 Getting Started in caArray—This chapter provides details about launching caArray, navigating the interface, accessing online help and other links in the interface.

Chapter 3 Navigating and Searching in caArray—This chapter describes the process for creating and working with Contacts, namely individual and group contacts in caArray.

Chapter 5 Creating and Managing Experiments—This chapter details instructions for creating and working with Experiments in caArray.

Chapter 6 Submitting Data to an Experiment—This chapter describes the processes for uploading, validating and importing array content and annotation data files into a caArray Experiment.

Chapter 6 Extracting Data from caArray—This chapter describes the processes for downloading data from the caArray repository

Chapter 7 User Account Management—This chapter describes the process for creating user accounts and collaboration group accounts and managing the group accounts of caArray.

Appendix A References—this appendix includes descriptions and links to references closely related to caArray technology and bioscience

Appendix B Glossary—This appendix defines acronyms, objects, tools and other terms referred to in the chapters of this caArray user's guide.

Index—This section of the guide provides a complete index.

User's Guide Text Conventions

Table 2.1 illustrates how text conventions are represented in this guide. The various typefaces differentiate between regular text and menu commands, keyboard keys, toolbar buttons, dialog box options and text that you type.


Convention	Description	Example
Bold & Capitalized Command Capitalized command > Capitalized command	Indicates a Menu command Indicates Sequential Menu commands	Admin > Refresh
TEXT IN SMALL CAPS	Keyboard key that you press	Press ENTER
TEXT IN SMALL CAPS + TEXT IN SMALL CAPS	Keyboard keys that you press simultaneously	Press SHIFT + CTRL and then release both.
Monospace type	Used for filenames, directory names, commands, file listings, and anything that would appear in a Java program, such as methods, variables, and classes.	URL_definition ::= url_string
Icon	A toolbar button that you click	Click the Paste button () to paste the copied text.
Boldface type	Options that you select in dialog boxes or drop-down menus. Buttons or icons that you click.	In the Open dialog box, select the file and click the Open button.
<i>Italics</i>	Used to reference other documents, sections, figures, and tables.	<i>caCORE Software Development Kit 1.0 Programmer's Guide</i>
<i>Italic boldface monospace type</i>	Text that you type	In the New Subset text box, enter <i>Proprietary Proteins.</i>
Note:	Highlights a concept of particular interest	Note: This concept is used throughout the installation manual.

Table 2.1 caArray Guide Text Conventions

Convention	Description	Example
Warning!	Highlights information of which you should be particularly aware.	Warning! Deleting an object will permanently delete it from the database.
{ }	Curly brackets are used for replaceable items.	Replace {root directory} with its proper value, such as c:\cabio

Table 2.1 caArray Guide Text Conventions (Continued)

CHAPTER 1

ABOUT CAARRAY

This [sectionchapter](#) provides a caArray overview as well as a description of the relationship of caArray to caBIG™.

Topics in this [sectionchapter](#) include:

- [caArray Overview on this page](#)
- *Relationship of caArray to caBIG* on page 6

caArray Overview

caArray is an open-source, user-driven, role-based, web and programmatically accessible data management system. caArray guides the annotation and exchange of array data using a federated model of local installations whose results are sharable across the cancer Biomedical Informatics Grid (caBIG™) (<https://cabig.nci.nih.gov/>). Identified primarily as a data service on the Grid, caArray furthers translational cancer research through acquisition, dissemination and aggregation of semantically interoperable array data to support subsequent analysis by tools and services on and off the Grid. As array technology advances and matures, caArray will extend its logical library of assay management.

The following services are provided by caArray: [add links](#)

- Creating and managing array experiments
- Annotating experiments
- Uploading, validating, and importing array data
- Browsing and searching across experiments
- Extracting data from caArray, either directly using the interface or programmatically using a java API
- Managing of experiment visibility

See also [Relationship of caArray to caBIG](#)

Relationship of caArray to caBIG

The National Cancer Institute (NCI) has launched the caBIG™ (cancer Biomedical Informatics Grid™) initiative to accelerate research discoveries and improve patient outcomes by linking researchers, physicians, and patients throughout the cancer community.

The mission of caBIG™ is to provide an infrastructure for creating, communicating and sharing bioinformatics tools, data and research results, using open data standards and shared data models. This supports the development of new types of analysis within and across experiments and allows new forms of collaboration, enabling the sharing of data sets and a range of analytical tools.

The primary goal of caArray, an integral component of the caBIG™ array of tools, is to further translational cancer research through acquisition, dissemination and aggregation of high quality array data to support subsequent analysis. The opportunity for caArray use among the cancer centers and their collaborators through caBIG™ will ultimately benefit the cancer community.

caArray development continues to proceed with an open architecture and supportive documentation to allow for future enhancements, particularly with regard to interfacing with additional analysis tools. The ability to query across platforms, and to exploit web services and/or databases from other components of caBIG™ when they are available is far-reaching. The goal is to create an extensible array system that is non-platform-specific and potentially customizable, enabling development that will continue to expand the vision of caBIG™.

[See also caArray Overview](#)

CHAPTER 2

GETTING STARTED IN CAARRAY

This [sectionchapter](#) introduces you to the caArray interface and its navigation as well as to global operations used in all of the caArray viewing windows.

Topics in this [sectionchapter](#) include:

- [Launching caArray on this page](#)
- *Requesting a User Account* on page 11
- *Using caArray Online Help* on page 13
- *Navigating the caArray User Interface* on page 13

Launching caArray

Note: caArray 2.0 supports the following browsers:

- PC's running Microsoft XP: Firefox 2.0+ and Internet Explorer 6.0+
- Apple Macs running OS X: Firefox 2.0+ and Internet Explorer 5.5

If you have questions about this, contact NCICB Application Support, <http://ncicbsupport.nci.nih.gov/sw/>.

The caArray application can be accessed from NCICB using the following URL: <http://array.nci.nih.gov>.

For instructions about downloading and installing caArray 2.0 at your site, see your local administrator and/or refer to the *caArray Local Installation Guide* that can be downloaded from the caArray download page: <http://ncicb.nci.nih.gov/download/downloadcaarray.jsp>.

[Add info about launching caArray or change title](#)

Note: If you are using a local installation of caArray, contact your Principal Investigator/Laboratory Manager/System Administrator for the correct URL for your use.

Browsing caArray

caArray allows a user to Browse the repository, both before logging and after logging in. The Browse options are the same for Anonymous Users and Logged in users. For more information, see *Browsing Before Login* and *Browsing After Login*.

Browsing Before Login

Once you are on the caArray Portal Welcome login page, without being logged in, you can browse caArray public data or you can perform a search of caArray public objects. *Figure 2.1*).

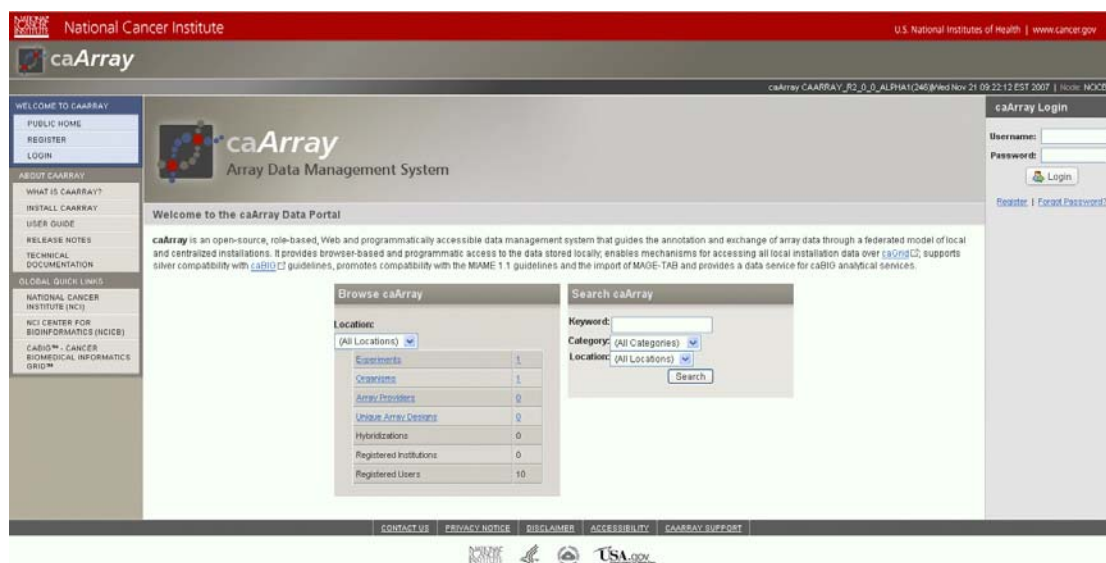


Figure 2.1 Prior to login, you can use these dialogs to browse or search the caArray database

The Browse caArray dialog box on the left center of the page lists the number of objects saved in the database in each of the listed categories. The text categories shown in blue hypertext format can be clicked to open additional pages that display information about those categories.

Note: The pages that open from the Browse dialog box list all caArray Experiments that have not been explicitly removed from visibility. Only Public Experiments can be “opened”; high-level metadata also display for non-Public Experiments, but there is no link to open them.

For more information about browsing the caArray database, see *Browsing caArray* on page 19.

The Search caArray dialog on the right center of the page allows you to launch a search of the caArray database for public objects. For more information about executing a caArray search, see *Searching the Repository* on page 22.

Options allowing you to submit, view, modify, and add microarray experiment data to caArray are dependent on your user privileges, once you are logged in. See *Table 2.2* for more information.

Note: You must obtain a user account in order to log in. For more information, see *Requesting a User Account* on page 11.

Navigating the caArray User Interface

Elements in the caArray User Interface

Browsing After Login

Once you have logged into caArray, the Browse feature is available from any page. To start the Browse function after login, click the Browse option on the left sidebar (*Figure 2.2*).



Figure 2.2 Browse option on left sidebar

caArray opens the Welcome to caArray Data Portal page, where you can launch a browse through the system, and as described in the previous section *Browsing Before Login*.

caArray New User Account and Login

From the Welcome to caArray Data Portal page, you can register as a new user or log in if you already have a user account ([Figure 2.2](#)):

U.S. National Institutes of Health | www.cancer.gov

caArray CAARRAY_R2_0_0_ALPHA1(246)(Wed Nov 21 09:27:12 EST 2007 | Node: NCIC)

caArray
Array Data Management System

the caArray Data Portal

Open-source, role-based, Web and programmatically accessible data management system that guides the annotation and exchange of array data through a federated model of local installations. It provides browser-based and programmatic access to the data stored locally, enables mechanisms for accessing all local installation data over [caGrid](#), supports interoperability with [caBIG](#) guidelines, promotes compatibility with the MIAME 1.1 guidelines and the import of MAGE-TAB and provides a data service for caBIG analytical services.

Browse caArray

Location: (All Locations) ▼

Experiments	1
Organisms	1
Array Providers	0
Unique Array Designs	0
Hybridizations	0
Registered Institutions	0
Registered Users	10

Search caArray

Keyword:

Category: (All Categories) ▼

Location: (All Locations) ▼

caArray Login

Username:

Password:

[Register](#) | [Forgot Password?](#)

[CONTACT US](#) | [PRIVACY NOTICE](#) | [DISCLAIMER](#) | [ACCESSIBILITY](#) | [CAARRAY SUPPORT](#)

Figure 2.3 caArray login page

- To login, complete the steps below.
- To register for a new user account, see [Requesting a User Account the following section](#) for more information.

To log in, follow these steps:

1. Navigate to the caArray home page. Use the URL to the NCICB instance <http://array.nci.nih.gov> ([Figure 2.3](#)), or contact your System Administrator for the URL to the local instance of caArray.
2. Enter your user ID and password in the upper right corner of the Welcome to caArray Data Portal page
3. Click **Login**.

After caArray verifies your credentials, the application opens to the caArray workspace, providing access to all features allowed by the permissions granted to you. For more information about finding your way around caArray, see [Navigating the caArray User Interface](#) on page 13.

Note: If you attempt unsuccessfully several times to login, you are locked out of the system. A message is sent to you explaining that the account is locked, and you are given a link to application support team to request that they unlock the account. If you do nothing, caArray will be unlocked automatically after 30 minutes.

Requesting a User Account

To request a caArray user account, you must complete steps 1 through 4.

1. Go to the NCICB caArray login page <http://array.nci.nih.gov> or use the URL provided by your System Administrator for the caArray instance at your institution.
2. Click the **Register** hypertext link, either in the left sidebar or under the caArray Login section in the upper right of the page. This opens the account registration form (*Figure 2.4*).

Figure 2.4 New user account registration form

3. In the Become a caArray User form, enter the appropriate information¹.
 - **Security Information**
 - **Do you have an LDAP account** [a user profile with your institution]?

If **Yes**, enter your username and password for the purposes of verifying that it is correct. If validated, you will be granted limited access to the application immediately as a “caArray User”; this will still require confirmation from the NCICB support teams.

If **No**, caArray indicates that the LDAP credentials do not check out. You are asked to reenter them. If they do not check out a second time, you are denied an account. You can **Cancel** or talk with your System Administrator about the problem.

1. Items with an asterisk or highlighted are required.

- **Username***
- **Password***
- **Requested role(s)*** --Select one or more of the roles. Roles are described in the following table [Table 2.1](#).

Note: Visibility/accessibility for Experiments is set per experiment. For more information see *Managing Experiment Visibility* on page 47.

Role	Description	Permissible 2.0 Actions
System Administrator	Person responsible for the effective operation of caArray	Manages users
Principal Investigator [PI]	Owns experiments and studies and/or projects	Submit data Write Experiment designs Submission of annotation Submission of array data
Lab Administrator	Same as PI in caArray 2.0	Same as PI in caArray 2.0
Lab Scientist	Same as PI in caArray 2.0	Same as PI in caArray 2.0
Biostatistician	Same as PI in caArray 2.0	Same as PI in caArray 2.0

Table 2.1 caArray role descriptions

- Account Details
 - First Name*
 - Middle Initial
 - Last Name*
 - Email [address]*
 - Organization*
 - Address [Lines 1* and 2]
 - City*
 - State/Province*
 - Postal [or Zip] Code*
 - Country* --Select from the drop-down list
 - Phone*
 - Fax

4. Click **Submit Registration Request** to execute the request, or click **Cancel** to halt the registration.

Once the request for a new account is sent to NCICB, it takes 24-72 hours to process. You will receive an email response when the account has been activated.

Note: Once you register, you have immediate accessibility to caArray, although your new account takes a period of time to process.

When your account is registered, the UserID and password you are assigned determines your access rights for the software.




Using caArray Online Help

The caArray online help explains how to use all of the features of the caArray portal.

Note: You can open online help without being logged into caArray.

To access online help in caArray, click the **Help** icon/menu options available in the upper right corner of the user interface and on the left sidebar under the About caArray section.

Once you are in online help, several options help you locate topics of interest.

- Click the **Show in Contents** button () at the top left of the help page to open the complete online help table of contents and locate your current topic in the table.
- The breadcrumb trail at the top of the page shows the relative location of the current help topic relative to neighboring topics. Click a breadcrumb link to display that help topic.
- Click the **Back** or **Forward** links at the top of the page to display help topics you have previously viewed.
- Follow hypertext links or the **Related Topics** buttons in the help topics to open other closely related topics. If the current help page has related topics associated with it, you can also view them by clicking the **Related Topics** button () at the top right of the help page.
- Locate topics using the table of contents that displays in the left pane of the online help project or the **Index** tab that displays at the top of the Table of Contents pane.
- Perform word searches of Help by entering query text in the search text box.
- Print the current topic by clicking the **Print** button () at the top right of the help page.

Note: caArray 2.0 does not have context sensitive help.

Navigating the caArray User Interface

The caArray provides a user-friendly interface to display options relating to login, accessing documentation or performing tasks for proposed or existing experiments.

Elements in the caArray User Interface

caArray Welcome Page Navigation Menu

User Interface Footer

My Experiment Workspace

Elements in the caArray User Interface

The terms in [the following table Table 2.2](#) identify and describe elements of the caArray user interface...

Term	Definition
Left Vertical Navigation Menu	Hypertext links associated with the caArray application, caArray documentation and Global Quick Links.
Online help	caArray documentation accessed from your computer screen providing help and answers for questions about using the software. Help icon/menu options are available in the upper right corner of the user interface and on the left sidebar under the About caArray section. Note: caArray 2.0 does not have context sensitive online help. You can open online help and use the TOC, index or perform a text search.
Browse caArray	The Browse dialog lists database object categories and the number of public objects in each categories. Click each hypertext link to browse details of the listed objects.
Search {caArray database objects}	The Search caArray dialog box, available pre- and post login to all users, allows you to launch a search of public objects. Enter keywords or select a category or location to define the search at a basic level. For more information about caArray searches, see <i>Searching the Repository</i> on page 22.
Task tabs	Located across the top of some of the caArray user interface windows. Task tabs represent the tasks/annotations you create as components of an experiment
<u>My Experiments/ Public Experiments</u>	caArray task tabs in the My Experiments Workspace. These comprise sets of Experiment information and annotations performed on or associated with gene expression or SNP microarrays.

Table 2.2 *Elements of the caArray browser window*

caArray Welcome Page Navigation Menu

The left sidebar of the caArray Welcome page provides links to an array of information.

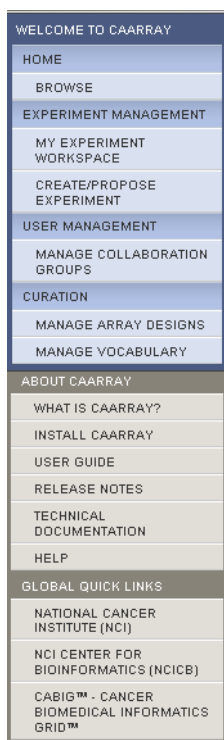


Figure 2.5 The caArray Welcome page navigation or left sidebar menu

The **Welcome to caArray** links direct you to Welcome caArray pages:

- These three options appear only before login:
 - **Public Home**--the Welcome/Login page
 - **Register**--the form for registering a new account
 - **Login**--the Welcome/Login page
- These four categories appear only after login
 - **Home**--Browse and Search dialogs display
 - **Experiment Management**--Displays Experiments to which you have access; offers you the ability to create/propose Experiments
 - **User Management**--Allows you to create and manage collaborationt users groups

Note:Only System Administrators can create users. For more information, see *Managing User Accounts* on page 62.
 - **Curation**--Includes functions related to creating and editing vocabulary terms and protocols.

The **About caArray** links direct you to various components of caArray documentation:

- **What is caArray**--Opens a brief introduction to caArray
- **Install caArray**--Opens the caArray Installation Guide pdf

- **User Guide**--opens the caArray User's Guide pdf
- **Release Notes**--opens Release Notes for caArray
- **Technical Documentation**--opens the caArray Technical Guide
- **Help**--Opens the full online help project

The **Global Quick Links** provide sources for caArray-related bioinformatics information on the Internet. These include links to the following websites:

- **National Cancer Institute (NCI)** (<http://www.cancer.gov/>)
- **NCI Center for Bioinformatics (NCICB)** (<http://ncicb.nci.nih.gov/>)
- **caBIG™ Cancer Biomedical Informatics Grid™** (<https://caarraydb.nci.nih.gov/caarray/>)

User Interface Footer

Options available in the footer are described as follows: [add other links](#)

- **Contact Us**--[Contact information](#) for NCICB
- **Privacy Notice**--NIH Web Privacy Notice
- **Disclaimer**--NIH Disclaimers
- **Accessibility**--NCI Web Accessibility Feedback Form
- **User Support**--Contact information for NCICB Application Support (<http://ncicb.nci.nih.gov/NCICB/support>)

My Experiment Workspace

Note: This section describes in limited detail the elements of the Experiment user interface. For more information about working with Experiments, see *Chapter 3 Navigating and Searching caArray* and *Chapter 4 Creating and Managing Experiments* in this guide.

The three most common uses of caArray are described as follows:

- Submit new data, or modify existing data (with appropriate permissions)
- Search public or private data already in the caArray system
- User and group management (with appropriate permissions)

Upon login, in the user interface, the My Experiment Workspace displays tabs and options relating to existing experiments or other experiments you may propose.

My Experiment Workspace									
<div> <div>Work Queue (4)</div> <div>Completed Experiments (1)</div> <div>Public Experiments (0)</div> </div> <div>Add a New Experiment</div>									
Displaying 1-4 of 4 Total									
Experiment Title	Array Type	Organism	Disease State	Status	Date Created	# Samples	Properties	Edit	Delete
armet-affy-arabi-316282	Affymetrix	Mouse	Acute Lymphoblastic Leukemia, Adult	Draft	09/23/2007	121			
arnot-affy-chick-445892	Affymetrix	Human	Cancer	Draft	10/01/2007	121			
freij-affy-human-91666	Affymetrix	Mouse	Brain Stem Glioma, Childhood	Draft	09/23/2007	121			
prosc-affy-mouse-450860	Affymetrix	Human	Cancer	Draft	10/01/2007	121			

Figure 2.6 caArray My Experiment Workspace

The My Experiment workspace displays two tabs:

1. **Work Queue**--This page lists Experiments, either in Draft or In Process with which you are associated. This includes non-Public as well as those designated Public.
2. **Public Queue**--Lists all Experiments in caArray that have not been specifically removed from visibility. Only Public Experiments can be opened.

For more information, see *Managing Experiment Visibility* on page 47.

The listed Experiments on each tab are visible in table format, according to the permissions assigned to you. The column headings display fields populated with metadata for each corresponding experiment.

Note: Column headings that are underlined are sortable by clicking on the heading.

- **Experiment ID**--The ID autogenerated by caArray; click the hypertext link to open Experiment details
- **Experiment Title**--name designated for the experiment by the Principle Investigator or his/her designee
- **Assay Type**--the category of array assay type for this Experiment; either **Gene Expression** or **SNP**
- **Samples**--the number of samples used in the Experiment. Click the hypertext link to open the Annotations tab to the Samples details page.
- **Status**--The current status of the Experiment: **Draft** or **In Process**
- **Permissions**--The accessibility given the corresponding Experiment. Click the icon to assign or modify the Experiment Permissions. See *Managing Experiment Visibility on page 47*.
- **Edit**--For Experiments with the appropriate permissions, allows the user to edit the Experiment. For more information, see *Editing an Experiment* on page 45.

Each of these Experiment elements is described in separate [sections](#) [chapters](#) in this [user's guide](#) [online help](#).

Online help and additional resource links remain fixed through all the user interface pages.

CHAPTER 3

NAVIGATING AND SEARCHING CAARRAY

This [sectionchapter](#) describes the processes for browsing the caArray repository and for conducting searches of the repository.

Topics in this [sectionchapter](#) include:

- [Browsing caArray on this page](#)
- [Searching the Repository on page 22](#)

Browsing caArray

In caArray, you can browse the application, a feature that can be launched either before login by any user or after login by a registered user.

Note: A non-logged in user can view only public data. A logged in user can view public data and non-public data to which (s)he has assigned permissions.

You can browse Experiments, Organisms, Providers, Unique Array Designs from the home page of a single installation.

Once you are on the caArray Portal Welcome login page, the Browse caArray dialog box on the left center of the page lists the number of objects saved in the database in each of the listed categories. The text categories shown in blue hypertext format can be clicked to open additional pages that display information about those categories ([Figure 3.1](#)).

Note: At any point in using the Browse or Search features, you can return to the home page by clicking the browser's back button.

From this page, you can also launch a search of the caArray repository. For more information, see *Searching the Repository* on page 22. .

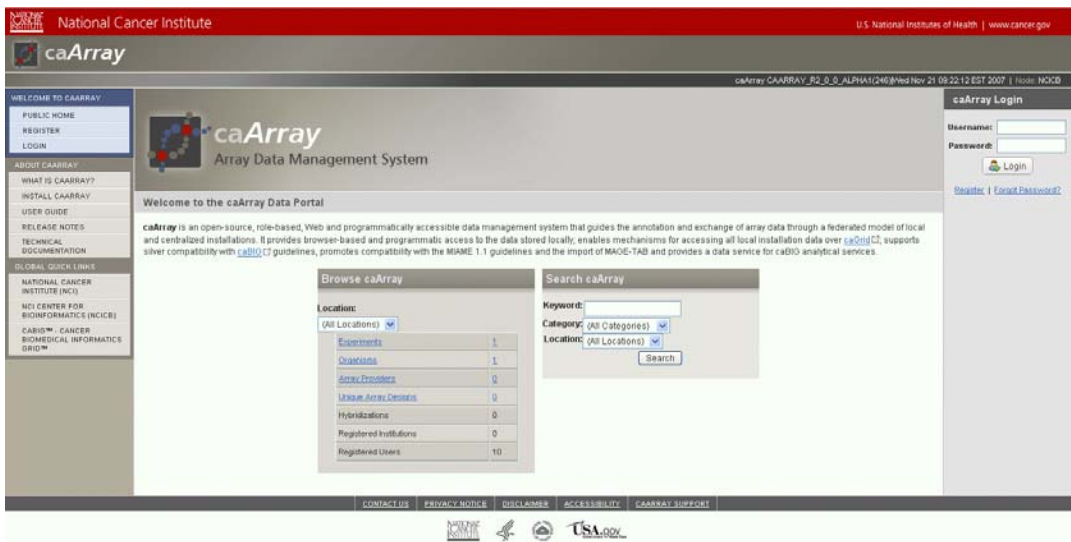


Figure 3.1 Prior to login, you can use these dialogs to browse or search the caArray database

To use the browse feature, follow these steps:

1. Click any of the blue hypertext links in the Experiment properties category list of the Browse dialog box. The following table Table 3.1 describes the categories that may display on the page. The page that opens depends on the category you selected.

Experiments	Both the experiments and corresponding number links open the Browse by Experiments page.
Organisms	Both the Organisms and corresponding number links open the Browse by Organisms page. The page is organized alphabetically by tabs that correspond to different organism names. The count of experiments available within each named group also displays on the tab. Experiments are grouped within the tabs according to the organism(s) identified when the Experiment was created.
Array Providers	Both the Array Providers and corresponding number links open the Browse by Array Providers page. An Array Provider is generally the company or group that manufactured the array used in the Experiment. The page is organized alphabetically by tabs that correspond to different Array manufacturer's names. The count of experiments available within each named group also displays on the tab. Experiments are grouped on the tabs according to the Array Providers identified when the Experiment was created. caArray 2.0 supports Affymetrix, GenePix and Illumina.

Table 3.1 Browse dialog box categories

Unique Array Designs	<p>An Array Design is a file that explains the design of a microarray. This includes such information as the array layout and design, its substrate, surface type, attachment type, the array strand type and the coordinates of each gene on the array.</p> <p>Both the Unique Array Designs and corresponding number links open the Browse by Unique Array Designs page. The count of experiments available within each named array design also displays on the tab. Experiments are grouped on the tabs according to the unique array design identified when the Experiment was created.</p>
Hybridizations	The number of hybridizations in the repository is visible, for information only. You cannot open hybridizations from this page.
Registered Institutions	The number of Registered Institutions in the repository is visible, for information only. You cannot open Registered Institutions from this page.
Registered Users	The number of Registered Users in the repository is visible, for information only. You cannot open Registered Users from this page.

Table 3.1 Browse dialog box categories

Note: Location refers to the caArray instance, either at your institution at NCICB.

- Once the tab or page opens when you click any of these categories ([Table 3.1](#)) described in the previous table, the same metadata displays on all pages for the list of Experiments located for that category ([Table 3.2](#)) (described in the following table). Most of these properties were identified when the Experiment was created or edited.

Note: Only Public Experiments or non-Public Experiments which have not been explicitly removed from visibility can display in the browse results. You can open only Public Experiments and non-Public Experiments with which you are associated.

Experiment ID	The autogenerated identification assigned by caArray. Click the hypertext link to open the corresponding experiment tabs which contain all current Experiment information. Only the Public data can be opened or Private data to which you have been given access.
ID	The Experiment title defined manually, naming and/or briefly describing the Experiment

Table 3.2 Experiment metadata categories

Assay Type	<ul style="list-style-type: none"> • Gene Expression: experiment using microarrays intended to measure levels of transcribed genes • SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA
Primary Contact	<p>The person named as the Point of Contact for the Experiment.</p> <p>Note: The PI and POC can be the same person, but do not have to be so.</p> <p>Click the hypertext link or the envelope icon (✉) to open an email form where you can draft an email to this contact, if named.</p>
Organism	The organism that is the source of the sample biomaterials used in the Experiment
Disease State	The disease state of the source materials used in the Experiment
Samples	The number of samples identified in the Experiment. Click the hypertext link to open the Experiment to the Samples details page.
Updated	The date of the most recent update of the Experiment draft

Table 3.2 Experiment metadata categories

The details can be sorted by clicking column headings. caArray paginates the result sets in groups of 25.

For found experiments, the actions available either to an anonymous user or a logged in user are available for each experiment.

Searching the Repository

In caArray, you can search the repository, a feature that can be launched before you login or by a non-registered “Anonymous User”.

The caArray search feature allows you to locate caArray content based on user-defined search criteria. Once you find the information you seek, you can open the experiment to review or edit details, if you have proper permissions; additionally, you can extract the data, follow hyperlinks to additional data, or you can return to the search feature to refine the query parameters.

To launch a search for a caArray Experiment, follow these steps:

1. From the caArray Portal Welcome page before login, locate the Search dialog box on the right center of the page.

The screenshot shows a search interface titled "Search caArray". It includes a "Keyword:" text field, a "Category:" dropdown menu currently showing "(All Categories)", a "Location:" dropdown menu currently showing "(All Locations)", and a "Search" button at the bottom right.

Figure 3.2 Search dialog box

OR

2. After login, locate the Search area of the page, in the upper right-hand corner.

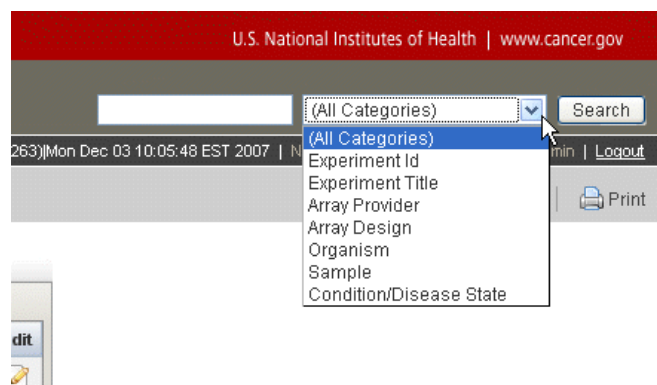


Figure 3.3 Section of the caArray page for launching a search

3. Define the search criteria by using the search options described in the following table Table 3.3.:

Search Option	Description
Category	<p>Select one of the Experiment properties categories listed:</p> <ul style="list-style-type: none"> • All of the below • Experiment ID • Experiment Title • Assay Provider • Array Design • Organism • Sample • Disease State <p>Only experiments in the category selected will be searched. If you do not choose a category, All of the Below (default) remains selected, and caArray will search all Experiments.</p>
Keyword	<p>Enter one or more words, separated by spaces. Example: breast cancer</p> <p>Note: Queries are case insensitive; wild cards implied on both sides of the query string. No logic statements, such as AND or OR or SQL statements are supported in these search features.</p>
Location	<p>Search the instance of caArray at your institution or at NCICB.</p>

Table 3.3 Search criteria options

4. Click **Search** to execute the search.

Note: If you click Search without defining query parameters, the search is unrestricted, and all Experiments in caArray display on the Search Results page.

Experiment Search Results

Experiment search results display on a new page, Search Results. If no results are found, a message informing you of that fact displays on the Search Results page.

Note: Only Public Experiments or non-Public Experiments which have not been explicitly removed from visibility can display in the search results. You can open only Public Experiments and non-Public Experiments with which you are associated.

Search results are listed in table format, with columns displaying properties for each Experiment; fields are described in [Table 3.4](#) the following table. Most of these properties were identified when the Experiment was created or edited.

Note: Column headings that are underlined are sortable by clicking on the heading.

<i>Search Results Properties</i>	<i>Search Results Fields Descriptions</i>
<u>Experiment ID</u>	The autogenerated identification assigned by caArray. Click the hypertext link to open the corresponding experiment tabs which contain all current Experiment information.
ID	The Experiment title defined manually, naming and/or briefly describing the Experiment
<u>Assay Type</u>	Gene Expression: experiment using microarrays intended to measure levels of transcribed genes SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA
Primary Contact	The person named as the Point of Contact for the Experiment. Click the hypertext link or the envelope icon (✉) to open an email form where you can draft an email to this contact, if named.
<u>Organism</u>	The organism that is the source of the sample biomaterials used in the Experiment
Disease State	The disease state of the source materials used in the Experiment
Samples	The number of samples identified in the Experiment. Click the hypertext link to open the Experiment to the Samples details page.
<u>Updated</u>	The date of the most recent update of the Experiment draft

Table 3.4 Experiment metadata categories

The details for categories underlined in [Table 3.4](#) can be sorted by clicking column headings. caArray paginates the result sets in groups of 25.

You can open any Experiment to which your assigned permissions grant you access. For Private Experiments to which you have not been assigned permission, only the ability to contact POC is available.

To open the Experiment details, click any Experiment ID or click the Samples number to open the Experiment to the Samples details page. You can review the Experiment or, contact the POC for the Experiment, or with appropriate permissions, edit it or extract the Experiment.

- For information about editing an Experiment, see *Editing an Experiment* on page 45.
- For information about contacting the Experiment PI, see *Primary Contact* in [Table 3.4](#).
- For information about extracting data from an Experiment, see *Downloading Data from caArray* on page 57.

Note: At any point in using the Browse or Search features, you can return to the home page by clicking the browser's back button.

CHAPTER 4

CREATING AND MANAGING EXPERIMENTS

This [sectionchapter](#) describes the processes for proposing/creating caArray Experiments and conducting searches of the repository.

Topics in this [sectionchapter](#) include the following:

[Overview of an Experiment on this page](#)

[Creating an Experiment on page 28](#)

[Updating An Experiment Proposal on page 45](#)

[Managing Experiment Visibility on page 47](#)

Overview of an Experiment

In a caArray Experiment, you can capture all relevant information as you create the project. This includes general information about the experiment, such as the experimental design and experimental factors, associated publications, biological samples, quality control and data processing steps, and so forth. You then upload all of the research data files related to the hybridizations performed in the Experiment, and link the sample information you submitted to the hybridizations.

The basic elements of a caArray Experiment are described in [the following table Table 4.1](#):

Term	Definition
Experiment Overview	Basic information about an experiment such as IDs, service and assay types, provider of array and array designs, source of biomaterials, and disease state

Table 4.1 [Elements of a caArray Experiment](#)

Term	Definition
Contacts	Principal Investigator and/or Point of Contact for the Experiment
Annotations	Experimental factors and design, biomaterials used, hybridizations
Data	Experimental data files uploaded and imported into caArray; imported data available for extraction from caArray; Supplemental data
Publications	Publications associated with the Experiment, primarily, journal articles

Table 4.1 *Elements of a caArray Experiment*

Adding an Experiment in caArray involves two main segments:]

1. Creating an Experiment with appropriate characteristics and annotations
2. Uploading the experimental research data files into caArray and associating them with the appropriate Experiment

With the appropriate permissions, you can create (“propose”) an experiment, save the draft, edit it, and finally submit an Experiment with its corresponding annotations. A minimum set of information must be entered for an Experiment before you can submit it, but almost as soon as a project is begun, you can save it as a draft to be retrieved and completed at a later time. Alternatively, you can have another designee with appropriate permissions complete the draft and submit the experiment.

Overview Tab

Contacts Tab

Annotations Tab

Data Tab

Publications Tab

Creating an Experiment

When you create an Experiment in caArray, you begin entering information on the Overview tab. Once you have saved the information on the Overview tab as an Experiment draft, additional tabs for entering Experiment information then display.

Overview Tab

Contacts Tab

Annotations Tab

Data Tab

Publications Tab

Steps for Creating an Experiment

To create an Experiment in caArray, follow these steps:

1. If you plan to use the NCICB instance of caArray, go to the NCICB caArray login page <http://array.nci.nih.gov>, and log in. If you plan you use the local installation of caArray at your center, see your local System Administrator for the URL.

Once you are on the caArray Portal Welcome login page, the browser displays the Experiment workspace.

2. On the left sidebar, click **Create/Propose Experiment**. This opens the Overview tab for entering overall characteristics for the Experiment.
3. Proceed to the [Overview Tab](#).



Figure 4.1 Create/Propose Experiment on left sidebar

Overview Tab

When you create a new experiment in caArray, the Overview tab initially displays alone in the user interface. Once you enter the appropriate information on this tab and save it, other tabs, also necessary for adding Experiment information display as well.

To complete the Overview tab, follow these steps:

1. On the Overview tab, enter the appropriate information for Overall Experiment Characteristics¹ as described in the [the following table Table 4.2](#).

Experiment: New Experiment

Overview

Overall Experiment Characteristics

The Overall Experiment Characteristics represent the minimum set of attributes required to submit an experiment for review. Required fields are marked with 'asterisks'.

Experiment Title*:

Status: Draft

Experiment Identifier:

Service Type*:

Assay Type*:

Provider*:

Array Designs:

Organism*:

Tissue Sites*:

- Leg
- Brain
- Cell Line
- Feed

Selected Tissue Sites:

Tissue Types*:

- whole_organism
- cell lysate
- organism_part
- DNA
- total_RNA

Selected Tissue Types:

Cell Types:

- Chromaphine Cells
- Brain Tissue

Selected Cell Types:

Conditions:

- Adrenocortical Carcinoma

Selected Conditions:

Figure 4.2 : Overview tab for an Experiment

Overview Tab Fields	Description
Experiment Title	Title designate by the PI or other user creating the Experiment
Status	Draft or Completed?
Experiment Identifier	This project Identifier is autogenerated by caArray upon the initial save of the Experiment. The Experiment Identifier is not editable. The ID is generated using the PIs last name followed by a 5 character number. Example: klemm-90765. The Experiment ID is geneally a hypertext link that opens the Experiment.

Table 4.2 Fields for Overall Experiment Characteristics

1. Fields with a red asterisk * are required.

Overview Tab Fields	Description
Service Type	<p>Service Type is description of relationship between PI and lab. Select from the drop-down menu the appropriate service type. Options are:</p> <ul style="list-style-type: none"> • Full--typical relationship between PI and lab; with this selection, the Experiment is created in caArray and “submitted to a lab” using the standard workflow. <p>Note: Full Service Experiments are placed in Public workspaces</p> <ul style="list-style-type: none"> • Publish--This option exists for getting a significant amount of pre-existing experiment data into caArray, thereby providing a vehicle whereby the data can be annotated and making the data available to the caArray community. When you select this option, the data validation feature is turned off. <p>Note: Publish Service Experiments remain in the PI's Experiment queue for making public when desired.</p> <p><u>Partial and Analysis options not available in 2.0.</u></p>
Assay Type	<p>Select from the drop-down menu the appropriate assay type. Options are the following:</p> <ul style="list-style-type: none"> • Gene Expression--experiment using microarrays intended to measure levels of transcribed genes • SNP--experiment using microarrays intended to detect nucleotide changes in chromosomal DNA
Provider	<p>Select from the drop-down menu the provider of the array.</p> <p>Note: Only Affymetrix , Illumina and GenePix are supported in caArray 2.0.</p> <p>Once selected, caArray automatically loads a corresponding list of Array Designs (next field)</p>
Array Designs	<p>Select one or more array design from the automatically-generated list of array designs corresponding to the Provider you selected.</p>
Organism	<p>Select from the drop-down menu the organism that is the source of the sample biomaterial</p>

Table 4.2 Fields for Overall Experiment Characteristics

2. After entering the information, click the **Save** button at the bottom of the page. Upon saving, caArray validates required fields and saves the Experiment as a draft. A confirmation messages displays, verifying that the proposal is saved. If the validation fails, caArray display a message indicating which field(s) need correction.

When you save the draft successfully, other tabs used for adding additional information for the Experiment display.

3. Proceed to the **Contacts** tab *Contacts Tab*.

Contacts Tab

A caArray Contact is a person associated with an Experiment, either as a Principal Investigator (PI) or the Point of Contact (POC). The Contact does not have to be a registered user of caArray.

To enter contact information for the Experiment, follow these steps:

1. On the Contacts tab, enter information for the fields described in [the following table Table 4.3](#).

Contact Fields	Description
P.I First Name	First and last names of the Principle Investigator. <u>These fields automatically populate with the name of the person creating the Experiment.</u>
P.I Last Name	
Email Address	Email address of the P.I.
Phone	Phone number of the P.I.
Is the P.I. the P.O.C. (point of contact for the Experiment)?	Select Yes or No . If answer is No , caArray opens additional text boxes where you can add contact information for the project Point of Contact (POC). Note: The POC does not have to be registered user of caArray.

Table 4.3 Contact fields

2. Click **Save**.
3. Proceed to the [Annotations Tab](#)Annotations tab.

Annotations Tab

The Annotations tab opens with seven subtabs for entering annotation data for the Experiment you are creating. Enter the appropriate information as described [below in the following topics](#) for each of the seven subtabs.

[Experimental Design](#)

[Experimental Factors](#)

[Sources Tab](#)

[Samples Tab](#)

[Extracts Tab](#)

[Labeled Extracts Tab](#)

[Hybridizations Tab](#)

Experimental Design

The Experimental Design describes the intent of the research and a description that is common to all hybridizations performed in the experiment.

1. Enter on the Experimental Design subtab (under the Annotations Tab) the appropriate information as described in the following table Table 4.4¹.

Experimental Design Fields	Description
Experimental Design Type*	Select the Experimental Design type among the options in the drop-down menu.
Experimental Design Description	Enter a description for the Experimental Design used for the Experiment.
Quality Control Types	Select the QC type in the displayed list.
Quality Control Description	Enter a description for the Quality Control used for the Experiment.
Replicate Types	Select one or more Replicate Types 2 kinds from the displayed list. Replicates can be either technical (arrays) or biological (laboratory animals or samples, etc.)
Replicate Description	If there are replicates used in the experiment, describe the number of replications and how the replicates were generated.

Table 4.4 Experimental Design fields

2. Click **Save**.
3. Proceed to the **Experimental Factors** subtab.

Experimental Factors

Experimental Factor are the intended sources of variation in the experiment. The Experimental Factors subtab (under the Annotations Tab) displays any previous Experimental Factors that may have been added previously to the Experiment you are creating.

1. On the Experimental Factors subtab, click the **Add a New Experimental Factor** button.
2. In the form that opens, enter the information as described in the following table Table 4.5².

Experimental Factors Fields	Description
Factor Name*	Enter a name for the Experimental Factor.
Description	Enter a description for the Experimental Factor.
Category	Select the appropriate category for the Experimental Factor in the displayed list.

Table 4.5 Experimental Factor fields

-
1. Fields with a red asterisk * are required.
 2. Fields with a red asterisk * are required.

3. Click **Save (or Cancel to halt the action)**. The name of the factor now displays on the Experimental Factors subtab. Click **Cancel** to return to the subtab without addin the factor.
4. Repeat steps 1 - 3 as often as needed to enter all the Experimental Factors for this Experiment.
5. Click **Save**.
6. Proceed to the [Sources TabSources](#) subtab. See also [Biological Source Material on this page](#).

Biological Source Material

Materials of biological origin are used in microarray experiments, and the state and characteristics of those biomaterials can be objectively documented and described. Biological materials can be treated, extracted, labeled and hybridized on microarrays to study relative expression levels of the genes represent on the arrays. All categories of biological materials are created in caArray to be included as essential components in a caArray Experiment.

Many biological materials' characteristics are defined in caArray by terms found in the caArray Controlled Vocabulary Terms. For more information, see [Manage Vocabularies \(Submit Exper UC\)](#).

In caArray, biological materials are divided into four different categories based on the treatment status of the material. The biological materials consist of **Sources**, **Samples**, **Extracts** and **Labeled Extracts**, defined below and illustrated in [the figure below Figure 4.3](#) in their hierarchical relationship.

- **Source** is any biological site from which the tissue for the array is derived before any preparation of the tissue for the array takes place. *Example:* human brain tumor tissue that has been treated with an anti-cancer drug.
- **Samples** are the original source biomaterials after initial treatment events. *Example:* A tissue or biopsy material treated to create a cell lysate.
- **Extracts** are Samples after a treatment event in which DNA or RNA is extracted. *Example:* A Sample tissue treated with RNA extraction method yields the Extract, RNA.
- **Labeled Extracts** are Extracts that have been labeled for detection of the nucleic acids. *Example:* The Extract RNA is labeled with a fluourescent dye, yielding the labeled extract RNA.

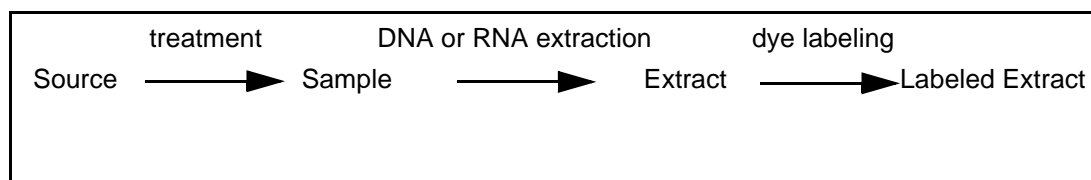


Figure 4.3 Biomaterials components and their relationship in caArray

The nature of each type of biomaterial is defined in varying dialog boxes in caArray. The relevant fields for each is described in [the associated topics Table 4.6, Table 4.7, and Table 4.8](#).

Sources

Samples

Extracts

Labeled Extracts

Sources

The term Source refers specifically to original source material before it has been subjected to any treatment. The nature of Sources in caArray is described in the following table [Table 4.6¹](#).

Source Fields		
Field	Description	Example
Source Name	Name given to the Source material Note: It is a good idea to clearly name Samples	
Description	Description of the Source material.	
*Organism	The species from which the Source was obtained.	<i>Homo sapiens</i>
Related Samples	Sample(s) that have been identified in caArray as being derived from this Source	

Table 4.6 Fields for documenting Sources

Samples

Extracts

Labeled Extracts

Samples

The term Sample refers specifically to original source material after it has been initially treated. The nature of Samples in caArray is described in fields in the following table [Table 4.7²](#):

Sample Fields		
Field	Description	Example
Sample Name	Name given to the Sample Note: It is a good idea to clearly name Samples	
Description	Description of the Sample.	

Table 4.7 Fields for documenting Samples

-
1. Items with an asterisk are required fields.
 2. Items with an asterisk are required fields when adding a new Sample. See *Sources Tab* on page 37.

Sample Fields		
Field	Description	Example
*Organism	The species from which the Sample was obtained	<i>Homo sapiens</i>
Tissue Site	The name of the tissue site from which the Sample was obtained	NIH Clinical Center
Source(s)	The name of the Source from which the Sample was derived	Lung Biopsy tissue--PatientX
Extract(s)	Extract(s) that have been identified in caArray as being derived from this Sample	Private

Table 4.7 *Fields for documenting Samples (Continued)*

Sources

Extracts

Labeled Extracts

Extracts

An Extract represents the nucleic acid extracted from a Sample for use in microarray experiments. The nature of Extracts in caArray is described in fields in [the following table Table 4.8](#):

Extract Fields		
Field	Description	Example
Extract Name	Name given to the Extract	
Description	Description of the Extract	cDNA, total RNA
Nucleic Acid Type	The type of nucleic acid extract: DNA or RNA <u>cDNA??</u>	
Related Sample(s)	The name of the Sample from which the Extract was derived	<i>Homo sapiens</i>
Labeled Extracts	Labeled Extract(s) that have been identified in caArray as being derived from this Extract	

Table 4.8 *Fields for documenting Samples*

Sources

Samples

Labeled Extracts

Labeled Extracts

A Labeled Extract is an Extract that has been labeled with signal emitters such as fluorescent materials for nucleic acid detection. Labeled Extracts are detected by

automated equipment. The nature of Labeled Extracts in caArray is described in the fields in [the following table Table 4.9](#):

Labeled Extract Fields		
Field	Description	Example
Label Extract Name	The name given to the Labeled Extract	
Description	Description of the Labeled Extract	
Labels Used	The signal emitter used to create this Labeled Extract from the Extract	
Amount	<u>The volume of the label or of the LE??</u>	
Protocol	The name of the protocol used to label the LE	
Related Extract	Extract from which this Labeled Extract was derived	
Hybridizations	The name of hybridizations performed with this labeled extract??	

Table 4.9 *Fields for documenting Labeled Extracts*

The Sources subtab (under the Annotations Tab) displays any Sources that may have been added previously to the Experiment you are creating.

- [Sources](#)
- [Samples](#)
- [Extracts](#)

Sources Tab

To add a Source, follow these steps:




1. On the Sources tab, click the **Add a new Source** button in the upper right-hand corner of the tab.
2. In the Sources form that opens, enter the information as described in [the following table Table 4.10](#).

Source Fields	Description
Source Name	Name assigned to the source
Description	Description of the source
Tissue Site	There are three ways you can enter terms for annotating this attribute. See the following section Managing Controlled Vocabulary Terms for more information about using this feature.

Table 4.10 *Fields for documenting a Source*

Source Fields	Description
Tissue Type	There are three ways you can enter terms for annotating this attribute. See the following section Managing Controlled Vocabulary Terms for more information about using this feature.
Cell Type	There are three ways you can enter terms for annotating this attribute. See the following section Managing Controlled Vocabulary Terms for more information about using this feature.
Disease State	There are three ways you can enter terms for annotating this attribute. See the following section Managing Controlled Vocabulary Terms for more information about using this feature.

Table 4.10 *Fields for documenting a Source*


3. Click **Save** to save the Source to the Experiment.
 - Once the draft is saved, the Organism you specified when you began creating the Experiment displays in the **Organism** field.
 - If you have already specified Samples for this Experiment, you specified the Source for the Samples. They now display in the **Related Samples** column.
4. Repeat steps 1-3 as often as necessary to add all appropriate Sources to the Experiment.
5. To copy a Source, click the corresponding **Copy** icon (). caArray copies the Source attributes, renames it using the existing Source name and adding an incremental number. The copied Source now displays under the original.
6. To edit the new Source, click the corresponding **Edit** button () and edit the data. Click **Save** to save the edits.
7. To delete a Source, click the Delete icon () in the corresponding row.
8. Proceed to the [Samples Tab](#) *Samples Tab*.

Managing Controlled Vocabulary Terms


In caArray, when you are creating or editing Experiments, many Experiment attributes are available for entering descriptive terms or annotations. These attributes are: [more?](#)

- **Tissue Site**
- **Cell Type**
- **Disease State**
- **Material Type**

You can enter terms for any of these attributes in three ways:

1. If the value for the attribute or condition displays in the site list, click the adjoining icon () to move it into the **Selected {attribute}** panel.
2. If the appropriate value is not displayed, to find a term of interest that might already be in the caArray dictionary, begin typing a term in the **Filter** text box.

The available terms that display in the panel below are limited according to the text you enter. A message displays if a corresponding term cannot be found.

- To add a new term, click the **Add** button (). This takes you to the Manage {Attribute or Condition} page where you can add a new term. The following table describes fields for defining the vocabulary term.

Vocabulary Term Category	Description of Fields
Term	
Value*	Enter the new term. <i>Example:</i> DNA
Description	Enter the description of the term, as appropriate. <i>Example:</i> deoxyribonucleic acid
Source	
Create a New Source [for the Term you are adding]	Select Yes or No <ul style="list-style-type: none"> If No, select from the drop-down list in the next field, the Source for the term. In many cases, the source will be an existing controlled vocabulary such as the NCI Thesarus, or the MGED Ontology (MO). If Yes, the dialog box expands with new fields where you can add the name, URL and version for the new source.
Source*	Select from the drop-down menu the source for the new term you are adding. This field disappears if you select Yes in the previous field.
Accession	
Accession URL	Enter the exact URL for accessing the new term. <i>Example:</i> http://mged.sourceforge.net/ontologies/MGEDontology.php#DNA
Accession Value	Enter the value given the term in the source vocabulary. <i>Example:</i> MO_945

Table 4.11 Fields for entering a new vocabulary term

- Once you have entered the appropriate information, click **Save**.
This returns you to the original tab, where you can continue defining the Experiment attribute.

Samples Tab

A Sample is the biological material to be applied to the array, prior to the extraction of the nucleic acid. *Example:* Tissue sample or biopsy material. The Samples subtab (under the Annotations Tab) displays any previous Samples that may have been added previously to the Experiment you are creating. [Fields earlier in chapter??](#)

- Click the **Add a New Sample** button to add a new Sample.

- In the Sample form that opens, enter the information described in [the following table Table 4.12](#).




Samples Fields	Description
Sample Name	Name assigned to the sample
Description	Description of the sample
External ID	An identification value given to the Sample outside of caArray, for example, in the lab that did the work
Source(s)	There are three ways you can enter terms for annotating this attribute. See the previous section Managing Controlled Vocabulary Terms for more information about using this feature.
Material Type	There are three ways you can enter terms for annotating this attribute. See the previous section Managing Controlled Vocabulary Terms for more information about using this feature.

Table 4.12 Fields for documenting Samples

- Click **Save**.

Note: When other data about the new Sample has already been entered in this Experiment, the fields that display on the Samples tab will be populated - Organism and Extracts derived from the Samples. If those fields are not currently populated, they will become such when the appropriate information is entered.

The name of the Sample(s) now displays on the Samples subtab.

- Repeat steps 1 - 3 as often as needed to enter all the Samples used in this Experiment.
- To copy a Sample, click the corresponding **Copy** icon (). caArray copies the Sample attributes, renames it using the existing Sample name and adding an incremental number. The copied Sample now displays under the original.
- To edit the new Sample, click the corresponding **Edit** button () and edit the data. Click **Save** to save the edits.
- To delete a Sample, click the Delete icon () in the corresponding row.
- Proceed to the [Extracts tab](#) [Extracts Tab](#).

Extracts Tab

An Extract is a Sample after a treatment event in which DNA or RNA is extracted. for the array.




The Extracts subtab (under the Annotations Tab) displays any Extracts that may have been added previously to the Experiment you are creating.

- Click the **Add a New Extract** button to add a new Extract.

- In the Extract form that opens, enter the information described in the following table [Table 4.13](#).

Extracts Fields	Description
Extract Name*	Name assigned to the Extract
Description	Description of the extract
Sample*	Sample from which the Extract was derived. Select a listed Sample or if the appropriate value is not displayed, begin typing a term in the Filter text box. The available terms that display in the panel below are limited according to the text you enter. A message displays if a corresponding term cannot be found.
Material Type	There are three ways you can enter terms for annotating this attribute. See <i>Managing Controlled Vocabulary Terms</i> on page 38 for more information about using this feature.

Table 4.13 Fields for documenting an Extract

- Click **Save**. The name of the Extract now display on the Extracts subtab.
Note: When other data about the new Extract has already been entered in this Experiment, the fields that display on the Extracts tab will be populated - **Nucleic Acid Type, and Labeled Extracts derived from the Extracts.** If those fields are not currently populated, they will become such when the appropriate information is entered.
- Repeat steps 1 - 3 as often as needed to enter all the Extracts used in this Experiment.
- To copy an Extract, click the corresponding **Copy** icon (). caArray copies the Extract attributes, renames it using the existing Extract name and adding an incremental number. The copied Extract now displays under the original.
- To edit the new Extract, click the corresponding **Edit** button () and edit the data. Click **Save** to save the edits.
- To delete an Extract, click the **Delete** icon () in the corresponding row.
- Proceed to the [Labeled Extracts Tab](#)[Labeled Extracts tab](#).

Labeled Extracts Tab

A Labeled Extract is an Extract that has been labeled for detection of the nucleic acids for the array.




The Labeled Extracts subtab (under the Annotations Tab) displays any Labeled Extracts that may have been added previously to the Experiment you are creating.

- Click the **Add a New Labeled Extract** button to add a new Labeled Extract.

- In the Labeled Extract form that opens, enter the information described in [the following table Table 4.14](#).

Labeled Extracts Fields	Description
Labeled Extract Name*	Name assigned to the Extract
Description	Description of the extract
Extracts*	Extract(s) from which the Labeled Extract was derived. Select a listed Extract or if the appropriate value is not displayed, begin typing a term in the Filter text box. The available terms that display in the panel below are limited according to the text you enter. A message displays if a corresponding term cannot be found.
Material Type	There are three ways you can enter terms for annotating this attribute. See <i>Managing Controlled Vocabulary Terms</i> on page 38 for more information about using this feature.

Table 4.14 Fields for documenting a Labeled Extract

- Click **Save**. The name of the Labeled Extract now display on the Extracts subtab.
- Note:** When other data about the new Labeled Extract have already been entered in this Experiment, the fields that display on the Extracts tab will be populated - Related Extracts, Hybridizations. **NEXT BUILD** If those fields are not currently populated, this will happen when the appropriate information is entered.
- Repeat steps 1 - 3 as often as needed to enter all the Extracts used in this Experiment.
 - To copy the Extract, click the **Copy** icon () in the Copy column. caArray copies the Extract attributes, renames it using the existing Labeled Extract name and adding an incremental number. The copied Labeled Extract now displays under the original.
 - To edit the new Labeled Extract, click the corresponding **Edit** button () and edit the data. Click **Save** to save the edits.
 - To delete a Labeled Extract, click the **Delete** icon () in the corresponding row.
 - Proceed to the [Hybridizations Tab](#)[Hybridizations tab](#).

Hybridizations Tab

The Hybridizations subtab (under the Annotations Tab) displays any Hybridization information that may have been added previously to the Experiment you are creating.

Where does one enter Protocol information in 2.0? This is obviously crucial information.
NEXT BUILD

- Click the **Add a New Hybridization** button to add a new Hybridization.

2. In the Hybridizations form that opens, enter the information described in the following table Table 4.15.



Hybridizations Fields	Description
Hybridization Name*	Name assigned to the Extract
Description	Description of the hybridization
Labeled Extract*	Labeled Extract used in the hybridization protocol. Select a listed Labeled Extract or if the appropriate value is not displayed, begin typing a term in the Filter text box. The available terms that display in the panel below are limited according to the text you enter. A message displays if a corresponding term cannot be found.

Table 4.15 Fields for documenting a hybridization

3. Click **Save**. The new Hybridization object displays on the Hybridization tab.

Note: On the saved draft, the uncompressed size of the hybridization file is defined by caArray when it becomes available.

This would only display after the hybridization file(s) have been imported, wouldn't it?

4. Repeat steps 1 - 3 as often as needed to enter all the hybridization data used in this Experiment.
5. To edit the new Hybridization entity, click the corresponding **Edit** button () and edit the data. Click **Save** to save the edits.
6. To delete a Labeled Extract, click the **Delete** icon () in the corresponding row.
7. Click **Save**. The name of the Hybridization now displays on the Hybridizations subtab.
8. Proceed to the Data tab.

Data Tab

The Data tab is the vehicle for uploading, validating, importing and downloading data relating to caArray Experiments. When you click on the **Data** tab, four subtabs where you initiate data-related tasks display. They are described in the following table Table 4.16.

Data Tabs	Description
Manage Data	From this tab, you can perform data-related tasks such as uploading, validating and importing data into caArray. Additional tasks such as changing data file types and designating supplemental files also takes place here.
Imported Data	This subtab list all files that have been imported into caArray.
Supplemental Files	This tab lists files and documents that have been uploaded to caArray.

Table 4.16 Tabs for performing data-related tasks

Data Tabs	Description
Download Data	From this tab, you can download data that has been imported into caArray.

Table 4.16 Tabs for performing data-related tasks

Note: To import data, the Experiment must be in Edit mode.

All of these data-related tasks are described in detail in [Chapter 5 Submitting Data to an Experiment](#) and [Chapter 6 Extracting Data from caArray](#).

Publications Tab

The Publications tab allows you to create, edit or delete associations of publications with the Experiment you are creating or editing. This feature allows you to add journal articles, reviews, and books.

To add publication associations, follow these steps:

1. Locate and open the Experiment for which you want to associate publications.
2. Click the **Publications** tab.
3. Click the **Add a New Publication** button.
4. On the form that opens, enter the appropriate information for the article or publication in the fields provided (and described in [the following table Table 4.17](#)).¹

Publications Fields	Description
Title*	Title of the publication
Authors	Author(s) of the publication
URL	URL for locating the publication
Editor	Editor of the publication, where appropriate
Pages	The page number(s) of the article your are referencing, where appropriate,
Publisher	Publisher of the publication
PubMedID	ID for locating the publication in PubMed
Volume	Volume where article is found
Year	Year of publication
Publication	Name of the publication where article is found
Type	Select in the drop-down menu the publication type.
Status	Select in the drop-down menu the publish status: Published, In Preparation, Submitted

Table 4.17 Fields for documenting Publications

1. Fields with a red asterisk * are required.

5. After entering the appropriate information, click **Save**. A message displays verifying that the publication was successfully saved with the Experiment.

The System saves the Experiment with the associated publications and returns you to your Experiment workspace, which is now updated with the state of the project.

Submitting an Experiment

During the time you are creating an Experiment and saving drafts, the Experiment is not visible to others. As long as it is Public, it is listed through browsing, but others cannot view the Experiment detail pages.

At any point after filling out the overview tab, you can decide to submit the Experiment to make it visible to others. To do so, click the **Submit Experiment** button in the upper right of any of the Experiment pages.

For information about setting the visibility of the Experiment, see *Managing Experiment Visibility* on page 47.

Updating An Experiment Proposal

At any point, after you have saved an Experiment draft, you, as its creator, can open the draft and edit it, add or delete data, or delete the Experiment.

Editing an Experiment

Adding Data to an Experiment


Deleting an Experiment

Editing an Experiment

At any point after an Experiment draft has been saved or even after an Experiment has been submitted to caArray, with appropriate permissions, you can open the draft and edit it. If you open an Experiment, and for various reasons, it is not editable, the Edit button will not be available.

Note: A good rule of thumb to follow is if you can see an **Edit** button or **Delete** option, then you can modify or delete an object. If you cannot see those options, then you cannot edit or delete the object.

To edit an Experiment, follow these steps:

1. Perform either a browse or search to locate the Experiment you want to edit. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
2. On the row corresponding to the Experiment you want to edit, click the **Edit** button () and edit the data. All information is editable except the automatically generated Experiment ID and the Status.
3. Click **Save** to save the edits to the draft.

Note: To import data, the Experiment must be in Edit mode.

Adding Data to an Experiment

Deleting an Experiment

Adding Data to an Experiment

Once an Experiment is in Draft status, as the Experiment creator you can add data such as annotations, data and publications to the Experiment. To do so, follow these steps:

1. Perform either a browse or search to locate the Experiment you want to modify. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
2. Open the Experiment by clicking on the ID hypertext link or the Samples link.
3. Navigate to the tab for the attribute or object you want to add. The tabs where you can add data have an **Add {attribute}** button in the upper right corner. To mention a few, you could add Samples, Labeled Extracts, upload data or associate publications with your Experiment.
4. Click the **Add {attribute}** button in the upper right hand corner of the tab.
5. Enter the appropriate information you wish to add. For more specific information, see *Creating an Experiment* on page 28.
6. Click **Save** to complete adding the data.

Editing an Experiment

Deleting an Experiment

Deleting an Experiment


To delete data in an Experiment, or to delete the Experiment itself, follow these steps:

1. Perform either a browse or search to locate the Experiment you want to edit. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.

To delete the Experiment:

2. On the row corresponding to the Experiment you want to delete, click the **Delete** icon ().

To delete Experiment properties:

3. Open the Experiment by clicking on the ID hypertext link or the Samples link.
4. On the row corresponding to the data you want to delete (i.e. Source, Sample, Extract, Labeled Extract, Hybridizations, etc) , click the **Delete** icon ().
5. Click **Save** to save the changes.


Editing an Experiment

Adding Data to an Experiment

Managing Experiment Visibility

Once you create an Experiment draft, it is listed on the **Work Queue** tab of the My Experiment Workspace. At that point, you can review and or modify the access of other users to the project.

To assign or modify Experiment visibility, follow these steps:

1. Perform either a browse or search to locate the Experiment. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
1. Click the Permissions icon () in the row listing the Experiment.
2. The Experiment Permissions page that opens displays two visibility categories:
 - a. **Browsability**--The default setting is **Not Browseable**. To change the setting to **Browseable**, click the **Toggle** button.
 - b. **Policies**--The only option here is to click the **Use TCGA Policy** check box. The Cancer Genome Atlas (TCGA) has established a policy in which only 4 fields of annotation can be provided to the open public:
 - Clinical Diagnosis
 - Histologic Diagnosis
 - Tissue Anatomic Site
 - Pathologic Status
 If you check the box, the same access standard will apply to your Experiment.
3. At this point, visibility cannot be assigned to Collaboration Groups.

CHAPTER 5

SUBMITTING DATA TO AN EXPERIMENT

This [section](#)[chapter](#) describes the processes for submitting data such as annotation and array content into caArray Experiments.

The following topics are part of this chapter:

- [Managing Data on this page](#)
- [Uploading Data Files](#) on page 50
- [Validating Data Files](#) on page 51
- [Importing Data](#) on page 53
- [Supplemental Files](#) on page 55
- [Downloading Files](#) on page 55
- [Importing MAGE-TAB](#) is the only mechanism for entering annotations that are not displayed as generically available and editable fields in the annotation tabs. The unique data will be visible but uneditable. [OWN SECTION??](#)
- [Importing array designs?--OWN SECTION either on data tab or in manage section--bugs importing in exp; if imported manage arrage designs, still be listed on data tab.](#)

Managing Data

In caArray, you can submit data to an Experiment by performing the following tasks on the Experiment Data tab:

1. **Uploading data:** moving data into caArray from a local or networked location
2. **Validating data:** determining by caArray if the data uploaded is valid according to pre-determined rules
3. **Importing data:** making validated data available through the user interface and or an API for download from caArray.

The Data tab is the vehicle for uploading, validating, importing and downloading data relating to caArray Experiments. When you click on the **Data** tab, four subtabs where you initiate data-related tasks display. They are described in [the following table Table 5.1](#).

Data Tabs	Description
Manage Data	From this tab, you can perform data-related tasks such as uploading, validating and importing data into caArray. Additional tasks such as changing data file types and designating supplemental files also takes place here.
Imported Data	This subtab list all files that have been imported into caArray.
Supplemental Files	This tab lists files and documents that have been uploaded to caArray.
Download Data	From this tab, you can download data that has been imported into caArray.

Table 5.1 Tabs for performing data-related tasks

All kinds of data can be uploaded and otherwise managed in caArray, although the majority of data will likely be annotation and array content files.


Must be in Edit mode

Note: It is not possible to import source or sample annotations directly into caArray 2.0. You can, however, import MAGE-TAB files that contain source and sample information. See *Importing Data* on page 53.

Uploading Data Files

Through the process of uploading annotation and array data, the content becomes available for validation and import into caArray. Imported files can be shared or even deleted.

To upload data into caArray, with appropriate permissions for the Experiment, follow these steps:

1. Log into caArray, either your local instance or that at NCICB.
2. Perform either a browse or search to locate the Experiment you want to edit. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (). This opens the Experiment to the Overview tab.
4. Select the **Data** tab, and **Manage Data** subtab.
5. Click the **Upload New Files** button.
6. In the form that opens, browse for the file you want to upload. To select several files for upload at once, click the **Add More Files** button. This opens more File/Browse options where you can locate several files.
7. Click the **Upload** button. Click **Cancel** if you decide to halt the task.

Note: Files of any size or type can be uploaded, including .zip files. The upload can run in the background while you continue your work in caArray. A background monitor on the screen indicates the process continues.

caArray launches the upload process. When the upload has finished, the list of files displays on the page, as well as their status and file type. As you continue to work with the data, their status updates (**Uploading**, **Uploaded**, **Validating**, **Validated**, **Importing** and **Imported**) (*Figure 5.1*).

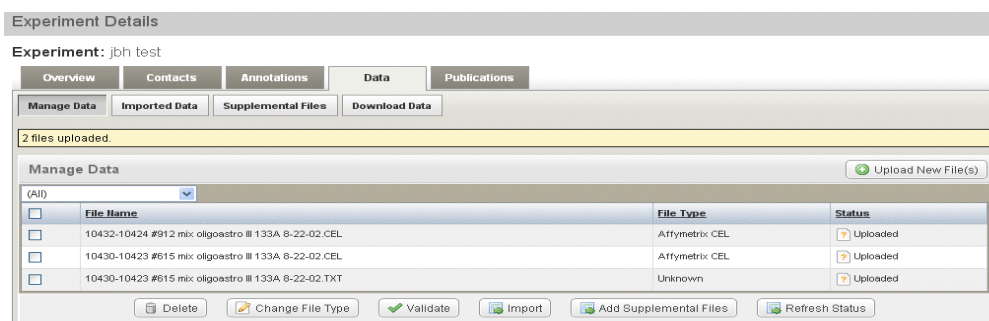


Figure 5.1 Manage Data subtab displays files in caArray and their status

Note: If a file is compressed, such as with a .zip file, caArray automatically extracts the files, discards the original and displays each of the files in the .zip to the user, indicating that each has been uploaded.

Check boxes corresponding to each file allow you to select one or more at a time individually for further management.

Once files are uploaded, the files workflow should be continued by Validating the File(s) and Importing the File(s). From the Manage Data tab, you can also change file types, designate files as supplementary, and delete files.

Deleting a File

On the Data tab, Manage Files subtab, check the box corresponding to one or more files you wish to discard and click the **Delete** button. Confirm the deletion. The file(s) are physically deleted from caArray.

Validating Data Files

Once data have been uploaded into caArray, anyone associated with an Experiment can validate and import annotation and array content files into the project. File validation verifies that data content adheres to a certain format; it does not evaluate the accuracy of the data from scientific viewpoint.

Note: The Validate feature for annotation data files whose Service Type is “Publish” is turned OFF. These files can still be imported, however. For more information, see *Service Type* in *Table 4.2* the table in *Overview Tab*.

The following file types support information sharing and can be uploaded, but because they are not array data files, the validation fails. These files should be identified as Supplemental Files, as described in *Supplemental Files* on page 55.

- Word documents

- Excel spreadsheets
- PowerPoint files
- PDFs

These files are not associated with any samples, but are associated at an Experiment level.

Following array content file types can be validated:


- **MGED**: IDF, SDRF, ADF
- **Affymetrix**: .dat, .cel, .chp

Note: Affymetrix .dat and .exp files cannot be validated, but they can be imported into caArray. See *Importing Data* on page 53.

- **Illumina**: Gene expression and SNP

Note: Any user associated with the PI's research lab can perform the validation.

To validate uploaded data files in caArray, with appropriate permissions for the Experiment, follow these steps:

1. Log into caArray, either your local instance or that at NCICB.
2. Perform either a browse or search to locate the Experiment for which you want to validate files. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (). This opens the Experiment to the Overview tab; you are in Edit mode.
4. Select the **Data** tab, and the **Manage Data** subtab.
5. Check the box corresponding to the file(s) you want to validate, and click the **Validate** button.

Note: If you choose to import a file before validating it, validation is launched automatically prior to import.

- For files where the type cannot be inferred, you are prompted to select the file type before validation can proceed.
- If the File Type is unknown, the file should be moved to the Supplemental Data tab. For more information, see *Supplemental Files* on page 55

caArray performs structural and then content validation against each file you have selected, updating the status of each file periodically (10 seconds) until all files display the validation status in the **Status** column: **Validated** or **Validation Failed**.

Validation Errors

If validation fails, the file cannot be imported and a **Validation Failed** message displays in the row corresponding to the file.

A validation error can be structural or content-based. Validation can fail for the following reasons:

- Format unknown (based on file extension and array type)
- Reference file not found
- file incomplete
- Array design missing--for array data, the associated design isn't available in the system library or project file space.
- Vocabulary failure--annotation terms not found in supported ontology

To view a validation error description, click the hypertext **Failed Validation** link in the **Status** column

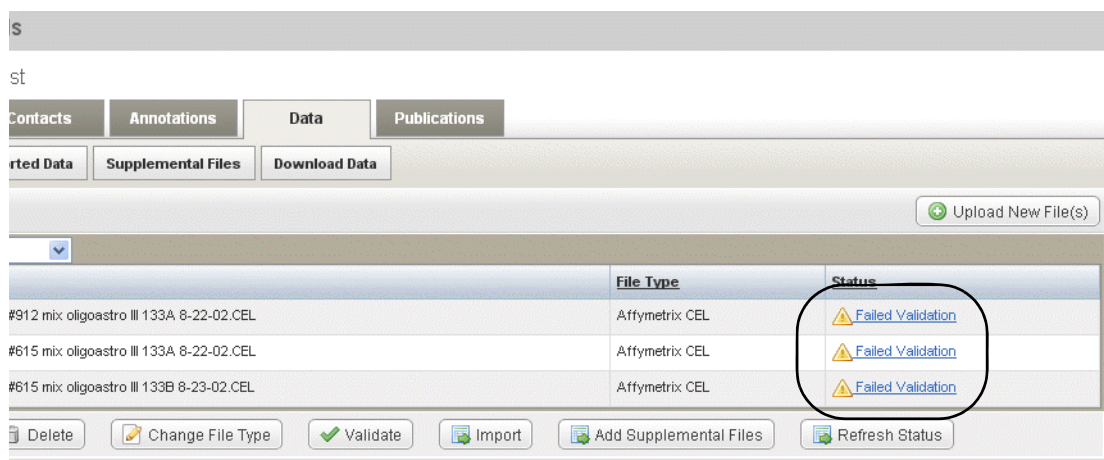


Figure 5.2 Validation failures display on the Manage Data table.

Importing Data

Once files are uploaded into caArray, anyone associated with an Experiment can validate and import annotation and array content files into the project. The import feature makes validated data available for download through the user interface and APIs.

Note: Annotation files uploaded to Experiments whose Service Type is "Publish" cannot be validated. Array content files can be validated.

Files Types that can be Imported into caArray


The file types listed in [Table 5.2](#) the following table can be imported into caArray:

Raw/processed data files	<ul style="list-style-type: none"> • Affymetrix .CEL • Affymetrix .CHP • Affymetrix.DAT (can be imported even though they cannot be validated in caArray) • Affymetrix.EXP (can be imported even though they cannot be validated in caArray) • GenePix .GPR • Illumina data files
Array design files	<ul style="list-style-type: none"> • .CDF • MAGE-TAB ADF (Array Design Format) • Illumina .CSV and .txt • GenePix GAL
MAGE-TAB files	<ul style="list-style-type: none"> • MAGE-TAB with single SDRF (Sample and Data Relationship Format) • MAGE-TAB with multiple SDRFs • IDF (Investigation Description Format) only, no referenced SDRFs • SDRF with only source > Hybridizations • ADF (Array Design Format) • MAGE-TAB with existing samples in system. Warn that data already exists for samples. Allow update, no update, or file association only

Table 5.2 File types that can be imported into caArray

Steps for Importing Data

To import data, with appropriate permissions for the Experiment, follow these steps:

1. Log into caArray, either your local instance or that at NCICB.
2. Perform either a browse or search to locate the Experiment for which you want to import files. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (). This opens the Experiment to the Overview tab; you are in Edit mode.
4. Select the **Data** tab, and the **Manage Data** subtab. Select the **Data** tab, and **Manage Data** subtab.
5. Check the box corresponding to the file(s) you want to import, and click the **Import** button.

Note: If the file has not been previously validated, caArray performs the validation.

- If validation fails, import does not proceed. Validation errors can be viewed as described in *Validating Data Files* on page 51.
 - If the validation completes successfully, caArray continues with the import and automatically auto-refreshes the status of the file set until the Import is complete and the Import Status of the file displays.
6. After a successful import, the files automatically move to the Imported Data subtab and the **Status** of file set is **Imported**.
 7. After a successful import, you can delete the file set. See *Deleting a File* on page 51.

Supplemental Files

Many file types can be uploaded into caArray, but only validated array content files can be imported into the application. Other files types can be designated on the Manage Data page as “Supplemental Files”.

To identify uploaded files as Supplemental files, follow these steps:

1. Select the **Data** tab, and **Manage Data** subtab.
 2. Check one or more boxes for the appropriate file(s).
 3. Complete one of these following methods:
 - a. Click the **Change File Type** button .
 - b. On the drop-down list, scroll down and select **Supplemental File**.
 - c. Click **Save**.
- OR
- a. On the Manage Data subtab, click the **Add Supplemental Files** button.

In both cases, the selected files are moved to the Supplemental Files tab.

Downloading Files

With appropriate permissions, you can extract files that have been imported into caArray. For more information, see [Chapter 6 Extracting Data from caArray](#).

CHAPTER 6

EXTRACTING DATA FROM CAARRAY

This [sectionchapter](#) describes the processes for extracting data from the caArray repository.

Topics in this [sectionchapter](#) include:


- [Downloading Data from caArray on this page](#)
- [Extracting Data Programmatically by API on page 58](#)

Downloading Data from caArray


Any user can download public data from any publicly available instance of caArray. No permission restrictions are required to acquire the data, either at an experiment level or sample level within or across projects.

Any registered user, after login, can download public data (only) associated with private data for which (s)he has the appropriate permissions.

To download data of interest from caArray, follow these steps:

1. Log into caArray, either your local instance or that at NCICB.
2. Perform either a browse or search to locate the Experiment for which you want to import files. For more information, see *Browsing caArray* on page 19 or *Searching the Repository* on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (). This opens the Experiment to the Overview tab; you are in Edit mode.

Note: Select the **Data** tab, and the **Manage Data** subtab. Select the **Data** tab, and **Manage Data** subtab. You can sort the columns of the list by clicking on the column headers.

4. In the **File Name** column, click the blue hypertext name of the file. That places the data set, including total file size, in the download queue, visible in the right panel on the tab.
 - You can select several files to be placed in the download queue. Once initiated, the download proceeds until all designated files are downloaded.
 - To remove selected files from the queue, click the **Remove** icon(s) () corresponding to the data file or click the **Cancel** button.

5. How are the following indicated?

- a. MAGE-TAB files (if any),
- b. Hybridization Data
- c. Supplemental Data

With the additional ability to

- d. Filter by file extension; user can select one or more extensions, or select All. Don't see any way to filter??
 - e. Select all; Don't see All or None options.
 - f. None; removes any files already selected
6. To initiate the download, click the **Launch Download Job** button.
 - The length of time for the download is dependent upon the file size.
 - You can continue to work in caArray during the download process. True??
 7. In the dialog box that opens, indicate whether you want to open or save the file to be downloaded. To save, navigate to the destination where the file will be saved.

When the download is complete, caArray displays an on-screen message telling you that the download is finished or that it failed.

Grid availability

For those Institutions that make their instance of caArray available to caGRID, the data is also accessible from all instances that have followed suit and available to the integrated tools that use the caGRID service.

Note: Search features across the Grid are not available in caArray 2.0.

Extracting Data Programmatically by API

Using External Systems, data in the caArray database can be extracted programmatically using an API. The external systems from which the data extraction can be launched are:

- geWorkBench
- caIntegrator
- WebGenome

- Grid client through a caGrid service

For more information about using the API to extract data, see the *caArray 2.0 Technical Guide* which can be downloaded from this site [URL: .In the Deployment chapter for now.](#)

CHAPTER 7

USER ACCOUNT MANAGEMENT

This [sectionchapter](#) describes the process for creating and managing accounts for users and groups of collaborators in caArray.

Note: The options for performing user management tasks are visible in caArray on the left sidebar of the browser only if you have these Admin privileges. [true?](#)

Topics in this [sectionchapter](#) include:

- [Roles in caArray on this page](#)
- [Managing User Accounts on page 62](#)
- [Managing Collaboration Groups on page 62](#)

Roles in caArray

When a new user account is created within caArray, the user can be assigned one or more roles that then determines that user's access to specified private data.

<i>Role</i>	<i>Description</i>	<i>Permissible Actions</i>
Anonymous User Note: Not listed under Roles	User without a caArray account or a non-logged in user	Browse and Search tasks
System Administrator	Person responsible for the effective operation of caArray	Manages users Manages user groups <u>Provides for overall management and quality of caArray ??</u>

Role	Description	Permissible Actions
Principal Investigator [PI]	Owns experiments and studies and/or projects	Submit data Review of experiment designs Submission of quality control metadata Upload of normalized data Annotation of parameters used
Lab Administrator	Same as PI in caArray 2.0	Same as PI in caArray 2.0
Lab Scientist	Same as PI in caArray 2.0	Same as PI in caArray 2.0
Biostatistician	Same as PI in caArray 2.0	Same as PI in caArray 2.0
External System Note: Not listed under Roles	Systems other than caArray from which caArray data can be extracted programmatically using an API.	For more informations, see <i>Extracting Data Programmatically by API</i> on page 58.

Please modify these descriptions as appropriate for the context of this manual. I based the descriptions on what you said in email about roles.

When the account is registered and roles are defined, the UserID and password assigned to the user determines access rights for the user.

Managing User Accounts

In caArray 2.0, all tasks related to creating and managing user accounts can be performed only by a System Administrator. The System Administrator must use the NCICB User Provisioning Tool (UPT) v. 3.2 for performing these tasks. All instructions for managing user accounts are described in the UPT 3.2 User's Guide which can be accessed from this website link to NCICB GForge: http://gforge.nci.nih.gov/frs/download.php/2634/UPT_User_Guide.pdf

Note: If you are a person who wants to be caArray user, you can launch a request for a user account from the Welcome/Login page. For more information, see *Requesting a User Account* on page 11

Managing Collaboration Groups

A registered user any?? or just Admin?? in caArray, can create, edit and delete collaboration groups and the users within them. This set of users ("collaborators") can then be given access by a Data Owner (i.e. the user who created the object, project, samples, etc.) to a project, set of projects or particular samples and their underlying array data. To be assigned to a collaboration group, a user must have a caArray user account.

Note: The options for performing collaboration group management tasks are visible in caArray only if you have these Admin privileges. Is this still true??

Viewing Group Details

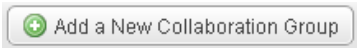

*Creating a Group**Viewing Group Details**Editing Collaboration Group Details*

Creating a Group

A group can be created as a response to a request submitted to the NCICB Application Support (ncicb@pop.nci.nih.gov) or by an [Admin](#) User.

If you are an Admin User, before creating a new group in caArray, you should perform a search to ascertain whether the group already exists. For more information, see *Viewing Group Details* on page 63. Still true in 2.0??

To create a group, as an **Admin User** follow these steps:

1. Click the **Manage Collaboration Groups** option on the left sidebar.
2. On the Manage Collaboration Groups page, click the **Add a New Collaboration Group** button () in the upper right corner of the page.
3. On the Add Group Member page, enter the **Group Name** in the appropriate text box.
4. Scroll through the list of all valid caArray users below the text boxes.
or
5. Perform a search for users using the filter option. Enter the user's last name, and/or select the role category the user may assigned, or the user's institution. You can also select the status of the user, with an enabled or a disabled account.
6. For each member to be added to the group, click the **Add** icon () in the far right column of the screen.
7. Click the **Save** button to save the group information. or click **Cancel** to abandon creation of the group.
8. Confirm the organization of the collaboration group and click What?? to confirm.


caArray saves the group and lists all of the registered users by name, organization and email address. Upon a successful save, does a Group Details page display? does this reopen the Group Search screen by default?

*Viewing Group Details**Viewing Group Details**Editing Collaboration Group Details*

Viewing Group Details

From a list of groups on the Collaboration Groups page, as a logged in Admin User, you can view the group details. To do so, click the hypertext link corresponding to the group. The group details lists all collaborators in the group, along with their corresponding Institution and email address.

The group details page displays group information in read-only format. You can however, launch an email to a collaborator by clicking the email address.

Click?? (use back arrow??) to close the page. Click the **Edit** button () to open the page with edit features.

Creating a Group


Viewing Group Details


Viewing Group Details


Editing Collaboration Group Details

Editing Collaboration Group Details

To edit collaboration group details, follow these steps:

1. From the Manage Collaboration Groups page, click the hypertext link corresponding to the group you select.
2. On the details page for the group, click the **Edit** button ().
3. The page that opens lists group members, their institution and email address. On this page, as an Admin User, you can perform the following edits:
 - Edit the Group Name. (A Group Name should be unique within the system.)
 - View users within the group.
 - Add or delete group members.
 - View user details using the hypertext link corresponding to a user in the group.
 - Delete the group.
4. To perform these edits, do the following: I may change this format later. Please check for accuracy though.

Edit the group name	Enter new name in the Group Name text box. Note: The new member must already have a valid caArray user account.
Add a new group member	Note: The new member must already have a valid caArray user account. Click the Add a New Group Member button at the top right of the page. In the new section of the page that opens, you can search for the group member using one or more criteria. <u>true??</u> Enter the last name, select the Role category, the Institution, and the Status to be searched. Click the Filter button.
Remove a group member	On the Collaboration Group page, in the Remove column, click the icon () that corresponds to the group member.

Edit the group name	Enter new name in the Group Name text box. Note: The new member must already have a valid caArray user account.
Review group member details	Click the name of the group member. The page that opens displays security and user account information about the group member, but this is not editable.
Delete the group	To delete the entire Collaboration Group, click the Delete button ( Delete) at the bottom center of the page.

Creating a Group

Viewing Group Details

Viewing Group Details

APPENDIX

A

CAARRAY REFERENCES

- **NCICB Production Site:** <http://caarraydb.nci.nih.gov/caarray/> – the public instance of the caArray Portal with links to caAMEL, the MAGE-OM API, and grid services.
- **Product Summary Site:** <https://cabig.nci.nih.gov/tools/caArray> - the summary of caArray capabilities and direction
- **Public Information Site:** <http://caarray.nci.nih.gov/> – a public web site that allows anyone to download the latest version, access documentation, launch the portal and visit sites that provide analysis of the data contained in caArray.
- **caArray Work Group Site:** <https://cabig.nci.nih.gov/workspaces/ICR/caArray-wg/> – this public web site provides access to the schedule, monthly meeting notes and links to the listserv for the stakeholder community
- **Microarray Gene Expression Data Society** - <http://mged.org/> The providers and curators of microarray standards, software and models.
- **Welcome Trust Sanger Institute** – http://www.sanger.ac.uk/Software/formats/GFF/GFF_Spec.shtml - for the General Feature Format Specifications

APPENDIX B

CAARRAY GLOSSARY

Acronyms, objects, tools and other terms referred to in the [chapters or appendixes sections](#) of this caArray [online help User's Guide](#) are described in this glossary.

Term	Definition
annotation	the experiment metadata that describe the “what” and “how” a experiment was executed
Anonymous User	A public user who visits caArray without logging in; may not have a user account
array	the raw data that is produced as a result of an experiment
caAMEL	caArray MAGE-ML Loader
caBIG	cancer Biomedical Informatics Grid
caBIO	Cancer Bioinformatics Infrastructure Objects
caCORE	cancer Common Ontologic Representation Environment
caDSR	Cancer Data Standards Repository
caMOD	Cancer Models Database
CGH	Comparative Genomic Hybridization
Collaborator	User designated by a caArray PI to preview a project (before it is made public), to work on a project subse or to provide samples and/or analysis for an experiment
Curator	Data curator who curates and ensures the quality of data in caArray through editing, demoting and promoting content
EBI	European Bioinformatics Institute

Term	Definition
EVS	Enterprise Vocabulary Services
LDAP	Lightweight Directory Access Protocol; software protocol for enabling anyone to locate organization, individuals and other resources such as files and devices in a network, such as a public internet or a corporate intranet. <u>UC summary p. 4 for more</u>
LDAP Account	User profile within your organization
Extract	
Labeled Extract	
MAGE 1.1	MAGE 1.1 is a widely-used microarray data standard or guideline
experiment	<u>a set of annotation and array data that describe the experimental design, the samples used and the array data produced</u>
MAGE-ML software format	XML-based standard for representation of microarray data
MIAME 1.1	MIAME1.1. is a standard or guideline for the minimum amount of information required to make a microarray record useful to others.
MGED Ontology	MGED Ontology is a controlled vocabulary standard that concisely defines terms as they relate to Microarrays and caArray as a whole
MGED	Microarray Gene Expression Data Society
MMHCC	Mouse Models of Human Cancers Consortium
NCI	National Cancer Institute
NCICB	National Cancer Institute Center for Bioinformatics
Project submitter	Registered user who submits projects, including experimental design, samples and qc governing the samples.
Registered User	caArray user who has been assigned a user account; can search database, submit and maintain array data and annotations, and extract data
Sample	<u>the material used for extraction onto an array</u>
Source	<u>the material – typically tissue or cell lines - from which a sample is taken</u>
URI	Uniform Resource Identifier
URL	Uniform Resource Locators
XML	Extensible Markup Language (http://www.w3.org/TR/REC-xml/) - XML is a subset of Standard Generalized Markup Language (SGML). Its goal is to enable generic SGML to be served, received, and processed on the Web in the way that is now possible with HTML. XML has been designed for ease of implementation and for interoperability with both SGML and HTML

Table B.1 caArray Glossary

INDEX

A

- account, requesting new user 11
- adding
 - Experiment data 46
 - Group 63
 - Publication 44
- Annotations tab 32
- Application Support i, 16

B

- biological source materials
 - description 34
 - relationship between 34
- browsing
 - after login 9
 - before login 8
 - caArray 8, 19

C

- caArray
 - About caArray sidebar links 15
 - browsing 8, 19
 - common uses 16
 - compatible browsers 7
 - glossary 69
 - interface description 14
 - launching 7
 - login 8
 - My Experiment Workspace 16
 - objectives 5
 - overview 5
 - references 67
 - relationship to caBIG 6
 - requesting user account 11
 - searching the repository 22
 - search results 24
 - task tabs 16
 - user interface footer 16
 - Welcome page 15
 - Welcome to caArray sidebar links 15
- caArray User's Guide

- introduction 1
- organization 1
- text conventions 2

- caBIG, relationship to caArray 6

- collaboration group

- creating 63
 - editing 64
 - viewing details 63

- collaboration groups, managing 62

- Contact tab 32

- creating

- collaboration group 63
 - Experiment 29
 - group 63
 - user account 62

D

- Data tab

- deleting a file 51
 - description 43
 - downloading data files 57
 - importing files 53
 - Manage Data subtab 49
 - Supplemental Files 55
 - uploading data 50
 - validating data 51

- deleting

- data file 51
 - Experiment 46

- downloading data files 57

E

- editing

- collaboration group 64
 - Experiment 45

- Experiment

- adding data to 46
 - Annotation tab 32
 - Contact tab 32
 - creating 29
 - deleting 46

- editing 45
 - managing data submission 49
 - managing visibility 47
 - overview tab 29
 - submitting 45
 - updating 45
 - uploading data to 50
- Experimental Factors tab 33
- Experiments, overview 27
- Extract
 - adding new 40
 - description 34
 - fields 36
 - tab 34
- extracting caArray files 57
- Extracts tab 40

F

- file types, importable 54

G

- Global Quick Links 16
- Glossary 69
- group
 - creating 63
 - viewing details page 63

H

- help in caArray 13
- Hybridization, adding new 42
- Hybridizations tab 42

I

- importable data file types 54
- importing data files 53

L

- Labeled Extract
 - adding new 41
 - description 34
 - fields 36
 - tab 34
- Labeled Extracts tab 41
- launching caArray 7

M

- managing
 - controlled vocabulary terms 38
 - data 49
 - user accounts 62
 - user groups 62

- My Experiment Workspace 16

N

- NCICB Application Support i, 16

O

- online help
 - icons 13
 - using 13
- overview, chapters in guide 1
- Overview tab 29

P

- Publication, adding 44
- Publications tab 44

R

- references 67
- registering new user 11

S

- Sample
 - adding new 39
 - description 34
 - fields 35
 - tab 34
- Samples tab 39
- searching
 - caArray repository 22
 - from caArray login screen 8
- search results 24
- Source
 - adding new 37
 - description 34
 - fields 35
 - tab 34
- Source Materials
 - annotations for 34
 - characteristics 34
 - definition 34
- Sources tab 37
- submitting Experiment 45
- Supplemental Files 55

T

- Technical Support i, 16
- text conventions in user guide 2

U

- uploading data to Experiment 50

user's manual conventions [2](#)

user account, new [62](#)

user groups

 managing [62](#)

users group

 editing [64](#)

V

validating data file [51](#)

validation errors [52](#)

viewing group details [63](#)

visibility, setting Experiment [47](#)

