

# CAARRAY 2.0

## *User's Guide*



NATIONAL<sup>®</sup>  
CANCER  
INSTITUTE

Center for Bioinformatics

*Revised December 21, 2007*



# CREDITS AND RESOURCES

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

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



# TABLE OF CONTENTS

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

## Chapter 4

### Creating and Managing Experiments .....27

Overview of an Experiment .....	27
Creating an Experiment .....	28
Steps for Creating an Experiment .....	29
Overview Tab .....	29
Contacts Tab .....	32
Annotations Tab .....	32
Data Tab .....	42
Publications Tab .....	42
Curation Tasks .....	43
Managing Array Designs .....	44
Managing Protocols .....	46
Managing [Controlled] Vocabulary [Terms] .....	50
Submitting an Experiment .....	52
Updating An Experiment Proposal .....	52
Editing an Experiment .....	52
Adding Data to an Experiment .....	53
Deleting an Experiment .....	53
Managing Experiment Visibility .....	53
Setting Public Visibility .....	54
Setting Collaboration Group Visibility .....	55

## Chapter 5

### Submitting Data to an Experiment .....57

Managing Data .....	57
Uploading Data Files .....	58
Validating Data Files .....	60
Importing Data .....	62
Supplemental Files .....	64
Downloading Files .....	65

## Chapter 6

### Extracting Data from caArray .....67

Downloading Data from caArray .....	67
Grid availability .....	68
Extracting Data Programmatically by API .....	68

## Chapter 7

### User Account Management .....69

Roles in caArray .....	69
------------------------	----

Managing User Accounts .....	70
Managing Collaboration Groups .....	70
Creating a Collaboration Group .....	70
Viewing Group Details .....	71
Editing Collaboration Group Details .....	72
<b>Appendix A</b>	
<b>caArray References .....</b>	<b>75</b>
<b>Appendix B</b>	
<b>caArray Glossary .....</b>	<b>77</b>
<b>Index .....</b>	<b>81</b>





# 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.


<b>Convention</b>	<b>Description</b>	<b>Example</b>
<b>Bold &amp; Capitalized Command</b> <b>Capitalized command &gt; Capitalized command</b>	Indicates a Menu command Indicates Sequential Menu commands	<b>Admin &gt; Refresh</b>
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
<b>Icon</b>	A toolbar button that you click	Click the <b>Paste</b> button (  ) to paste the copied text.
<b>Boldface type</b>	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 <b>Open</b> button.
<i>Italics</i>	Used to reference other documents, sections, figures, and tables.	<i>caCORE Software Development Kit 1.0 Programmer's Guide</i>
<b><i>Italic boldface monospace type</i></b>	Text that you type	In the New Subset text box, enter <b><i>Proprietary Proteins.</i></b>
<b>Note:</b>	Highlights a concept of particular interest	<b>Note:</b> This concept is used throughout the installation manual.

Table 2.1 caArray Guide Text Conventions

<b>Convention</b>	<b>Description</b>	<b>Example</b>
<b>Warning!</b>	Highlights information of which you should be particularly aware.	<b>Warning!</b> 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:

- Browsing and searching across experiments. See [Chapter 3 Navigating and Searching caArray](#)
- Creating and managing array experiments. See [Chapter 4 Creating and Managing Experiments](#)
- Annotating experiments. See [Chapter 5 Submitting Data to an Experiment](#)
- Uploading, validating, and importing array data. See [Chapter 5 Submitting Data to an Experiment](#)
- Extracting data from caArray. See [Chapter 6 Extracting Data from caArray](#)

- Managing collaboration groups. See [Chapter 7 User Account Management](#)

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 14


### caArray Fundamentals

---

**Note:** caArray 2.0 supports the following browsers:

- PC's running Microsoft XP  efox 2.0+ and Internet Explorer 6.0+
- Apple Macs running OS X  efox 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>.

---

**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 and Searching caArray

caArray allows a user to browse and search 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 and Searching Before Login](#), [Browsing and Searching After Login](#) and [Logging in](#).

### Browsing and Searching 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](#).

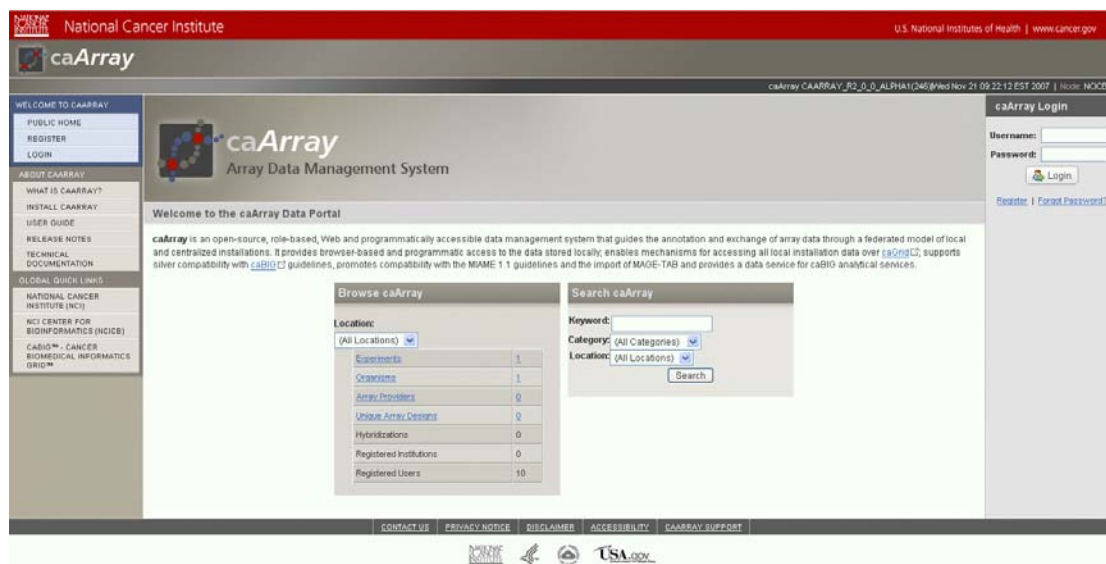


Figure 2.1 Prior to login, you can use these dialogs to browse or search the caArray database

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 the caArray Repository](#) 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 caArray 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.

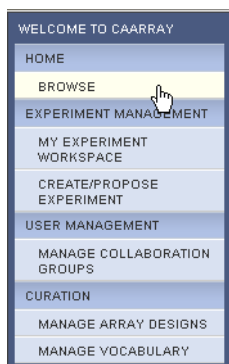
### *Navigating the caArray User Interface*

#### *Elements in the caArray User Interface*

### **Browsing and Searching After Login**

Once you have logged into caArray, the Browse and Search features are available from any page.

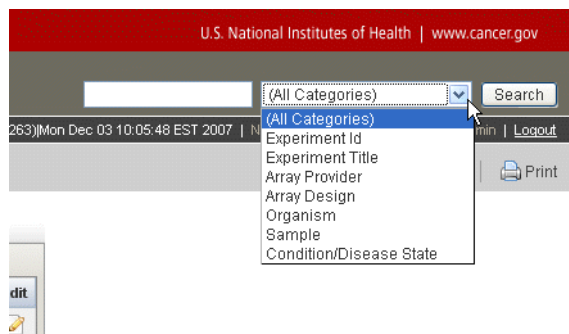
- To start the Browse function after login, click the **Browse** option on the left sidebar ([Figure 2.2](#)).



*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 and Searching Before Login](#).

- To start a Search after login, enter search query parameters in the Search text box in the upper right corner of the user interface ([Figure 2.3](#)).



*Figure 2.3 Search text box displays in every browser page*

For more information about launching a search, see [Chapter 3 Navigating and Searching caArray](#).

## 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 2.4](#)):

*Figure 2.4 caArray login page*

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

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 2.4](#)), 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 14.

**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 2.5*).

*Figure 2.5 New user account registration form*

3. In the Become a caArray User form, enter the appropriate information<sup>1</sup>.
  - **Security Information**
    - **Do you have an LDAP account** [a user profile with your institution] at [NCICB or your institution]?

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

**--Username\***

**--Password\***

1. Items with an asterisk or highlighted are required.

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

If your LDAP profile is not validated, 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.

If you select **No** [you do not have an LDAP account], the text boxes for entering the LDAP account information disappear. You should still indicate the role you will be assigned in caArray.

**Note:** Visibility/accessibility for Experiments is set per experiment. For more information see [Managing Experiment Visibility](#) on page 53.

To see a description of the roles in caArray 2.0, click the arrow to the right of this sentence.

<b>Role</b>	<b>Description</b>	<b>Permissible 2.0 Actions</b>
System Administrator	Person responsible for the effective operation of caArray	Manages users
Principal Investigator [PI]	Owens 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 abort the registration.

Once the request for a new account is sent to NCICB, it takes 24-72 hours to process. (The process time for a local installation at your institution may differ.) You will receive an email response when the account has been activated.

---

**Note:** Once you register, you have immediate access  ibility 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 (  Help ) and on the left sidebar under the **About caArray** section.

Online help opens with two display panels:

1. The left panel displays the Table of Contents (TOC), and also offers access to the Index and Search features of online help. The TOC can be expanded. All topics listed in the TOC and index are hypertext links to the referenced topics.
2. The right panel displays the Welcome to caArray Online Help page and other topic contents.

The following features facilitate your navigation of online help:

- 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
<b>Left Vertical Navigation Menu</b>	Hypertext links associated with the caArray application, caArray documentation and Global Quick Links.
<b>[Online] Help</b>	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.  <b>Note:</b> 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.
<b>Browse caArray</b>	The <b>Browse</b> dialog lists database <del>object</del> categories and the number of public <del>objects</del> in each <del>categories</del> . Click each hypertext link to browse details of the listed <del>objects</del> .
<b>Search {caArray database objects}</b>	The Search caArray dialog box, available pre- and post login to all users, allows you to launch a search of pu  <del>objects</del> . Enter keywords or select a category or location to define the search at a basic level. For more information about caArray searches, see <a href="#">Searching the caArray Repository</a> on page 22.
<b>Task tabs</b>	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
<b>Work Queue/ Public Experiments</b>	caArray task tabs in the <i>My Experiment Workspace</i> <a href="#">My Experiment Workspace</a> . These comprise sets of Experiment information and annotations performed on or associated with <del>gene expression or SNP microarrays</del> . Experiments displayed on the Work Queue tab are those with which you are associated. Experiments displayed on the Public Experiments tab <del>have been assigned Public visibility</del> . For more information, see <a href="#">Managing Experiment Visibility</a> on page 53.

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.

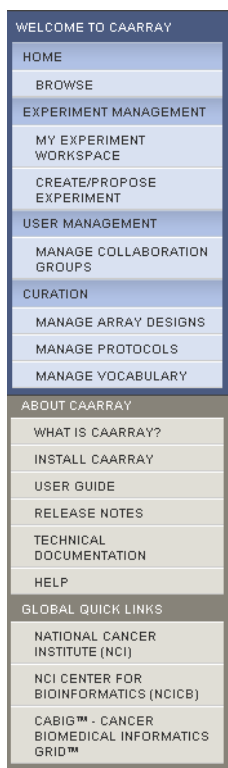


Figure 2.6 The caArray Welcome page navigation or left sidebar menu

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 requesting a new account
  - **Login**--the Welcome/Login page
- These four categories appear only after login
  - **Home**--Click **Browse** to display the Browse and Search dialog boxes
  - **Experiment Management**--Open the My Experiment Workspace and the page to begin creating an Experiment.
  - **User Management**--Allows you to create and manage collaboration users groups
 

**Note:** Only System Administrators can create users. For more information, see *Managing User Accounts* on page 70.
  - **Curation**--Includes functions related to managing Array Designs and Protocols and 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 pdf
- **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:

- **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~~ 



The screenshot shows the 'My Experiment Workspace' interface. At the top, there are tabs for 'Work Queue (4)', 'Completed Experiments (1)', and 'Public Experiments (0)'. Below the tabs, a table displays a list of experiments. The table has columns for Experiment Title, Array Type, Organism, Disease State, Status, Date Created, # Samples, Properties, Edit, and Delete. Four experiments are listed, all with a status of 'Draft' and a date created of '09/23/2007' or '10/01/2007'. The # Samples column shows '121' for each experiment. The table is paginated, showing 'Page 1' and navigation links for 'Back' and 'Next'.

Experiment Title	Array Type	Organism	Disease State	Status	Date Created	# Samples	Properties	Edit	Delete
<a href="#">armst-affy-arabi-316782</a>	Affymetrix	Mouse	Acute Lymphoblastic Leukemia, Adult	Draft	09/23/2007	121			
<a href="#">arnot-affy-chick-445892</a>	Affymetrix	Human	Cancer	Draft	10/01/2007	121			
<a href="#">frei-affy-human-91666</a>	Affymetrix	Mouse	Brain Stem Glioma, Childhood	Draft	09/23/2007	121			
<a href="#">prosc-affy-mouse-450860</a>	Affymetrix	Human	Cancer	Draft	10/01/2007	121			

Figure 2.7 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 53.

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 to the corresponding Experiment.~~ Click the icon to assign or modify the Experiment Permissions. See [Managing Experiment Visibility](#) on page 53.
- **Edit**--For Experiments with the appropriate permissions, allows the user to edit the Experiment. For more information, see [Editing an Experiment](#) on page 52.

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

## Browsing the caArray Repository

---

In caArray, you can browse the repository, 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 caArray Repository](#) on page 22. .

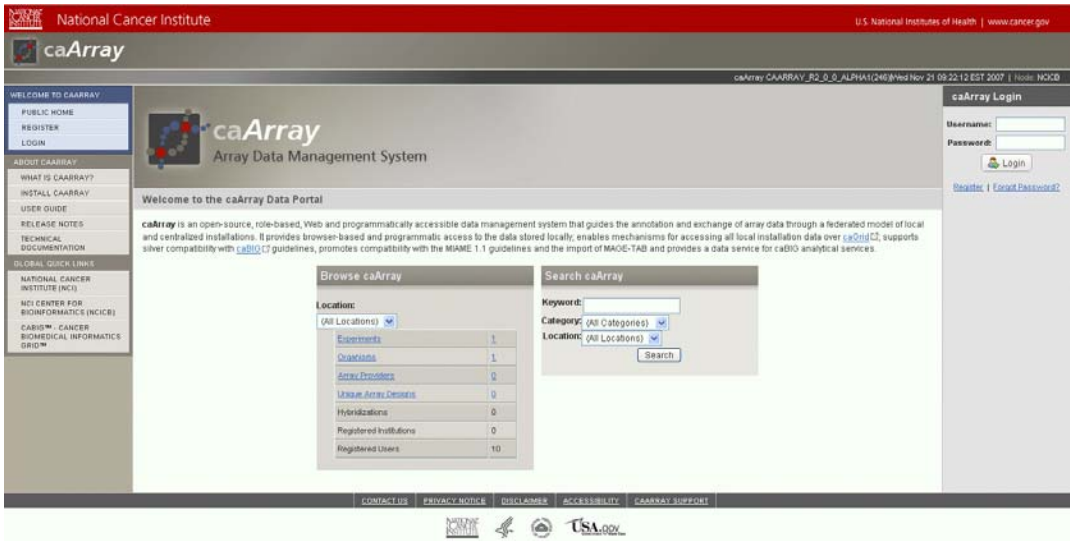


Figure 3.1 Prior to login, you can use these dialogs to browse or search the caArray database

To use the browse feature, follow these steps:

The following table [Table 3.1](#) lists the categories that display in the Browse dialog box.

1. Click any of the blue hypertext links in the Experiment properties category list of the Browse dialog box.

The page that opens depends on the category you selected. The descriptions in [Table 3.1](#) [this table](#) outline the detail that opens for each category that you click..

<i><b>Browse Dialog Box Category</b></i>	<i><b>Description</b></i>
<b>Experiments</b>	Both the experiments and corresponding number links open the Browse by Experiments page.
<b>Organisms</b>	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.

Table 3.1 Browse dialog box categories

<b>Browse Dialog Box Category</b>	<b>Description</b>
<b>Array Providers</b>	<p>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.</p> <p>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.</p>
<b>Unique Array Designs</b>	<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>
<b>Hybridizations</b>	The number of hybridizations in the repository is visible, for information only. You cannot open hybridizations from this page.
<b>Registered Users</b>	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 or at NCICB.

- Once the tab or page opens when you click any of these categories ([Table 3.1](#)), the same metadata displays on all pages for the list of Experiments located for that category ([Table 3.2](#)). ~~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.

<b><i>Experiment Category</i></b>	<b><i>Description</i></b>
<b><u>Experiment ID</u></b>	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.
<b><u>ID</u></b>	The Experiment title defined manually, naming and/or briefly describing the Experiment
<b><u>Assay Type</u></b>	<ul style="list-style-type: none"> <li>• <del>Gene Expression: experiment using microarrays intended to measure levels of transcribed genes</del></li> <li>• <del>SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA</del></li> </ul>
<b>Primary Contact</b>	<p>The person named as the Point of Contact for the Experiment.</p> <p><b>Note:</b> 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>
<b><u>Organism</u></b>	The organism that is the source of the sample biomaterials used in the Experiment
<b>Condition/Disease State</b>	The disease state of the source materials used in the Experiment
<b>Samples</b>	The number of samples identified in the Experiment. Click the hypertext link to open the Experiment to the Samples details page.
<b><u>Updated</u></b>	The date of the most recent update of the Experiment draft

Table 3.2 Experiment metadata categories

**Note:** Columns with underlined headings are sortable by clicking on the heading. caArray paginates the result sets in groups of 25.

See also *Searching the caArray Repository*

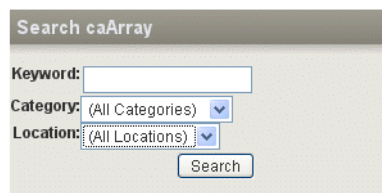
## Searching the caArray Repository

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. *Before login*, from the caArray Portal Welcome page, locate the Search dialog box on the right center of the page.

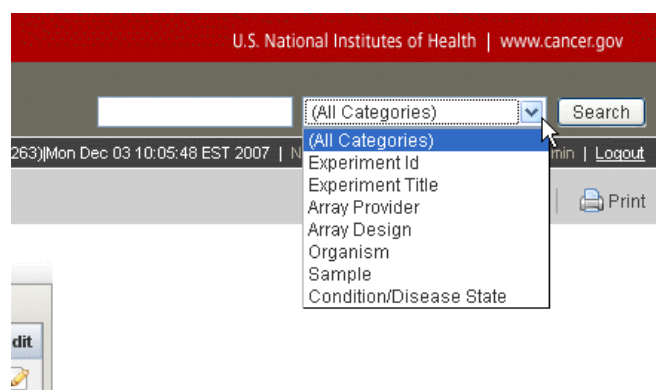


The image shows a 'Search caArray' dialog box. It contains a 'Keyword:' text input field, a 'Category:' dropdown menu currently set to '(All Categories)', and a 'Location:' dropdown menu currently set to '(All Locations)'. A 'Search' button is located at the bottom right of the dialog.

Figure 3.2 Search dialog box

OR

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



The image shows the upper right-hand corner of the caArray portal page. It features a search bar with a dropdown menu for categories. The dropdown menu is open, showing options: '(All Categories)', 'Experiment Id', 'Experiment Title', 'Array Provider', 'Array Design', 'Organism', 'Sample', and 'Condition/Disease State'. A 'Search' button is to the right of the dropdown. The page header includes 'U.S. National Institutes of Health | www.cancer.gov' and a timestamp '263) Mon Dec 03 10:05:48 EST 2007 | N...'. There are also links for 'min | Logout' and a 'Print' button.

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"> <li>• <del>All of the below</del></li> <li>• Experiment ID</li> <li>• Experiment Title</li> <li>• Assay Provider</li> <li>• Array Design</li> <li>• Organism</li> <li>• Sample</li> <li>• Disease State</li> </ul> <p>Only experiments in the category selected will be searched. If you do not choose a category, <del>All of the Below</del> (default) remains selected, and caArray will search all Experiments.</p>

Table 3.3 Search criteria options


Search Option	Description
<b>Keyword</b>	In the text box, enter one or more words, separated by spaces. Example: <b>breast cancer</b> <b>Note:</b> Queries are case insensitive; wild cards are 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.
<b>Location</b>	Search this instance of caArray at your institution or at NCICB. 

Table 3.3 Search criteria options

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

**Note:** If you click **Search** without defining query parameters, the search is unrestricted, and all Experiments in caArray that have not been explicitly removed from visibility display on the Search Results page.

See also [Experiment Search Results](#).

## Experiment Search Results

For launching a caArray search, see [Searching the caArray Repository](#).

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 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 3.4the following table](#). Most of these properties were identified when the Experiment was created or edited.

---

**Note:** Columns with underlined headings are sortable by clicking on the heading. caArray paginates the result sets in groups of 25.

---

Search Results Properties	Search Results Fields Descriptions
<u><b>Experiment ID</b></u>	The autogenerated identification assigned by caArray. Click the hypertext link to open the corresponding experiment tabs which contain all current Experiment information.
<u><b>ID</b></u> 	The Experiment title defined manually, naming and/or briefly describing the Experiment
<u><b>Assay Type</b></u>	<del>Gene Expression: experiment using microarrays intended to measure levels of transcribed genes</del> <del>SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA</del> 





<b>Search Results Properties</b>	<b>Search Results Fields Descriptions</b>
<b>Primary Contact</b>	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.
<b><u>Organism</u></b>	The organism that is the source of the sample biomaterials used in the Experiment
<b>Disease State</b>	The disease state of the source materials used in the Experiment
<b>Samples</b>	The number of samples identified in the Experiment. Click the hypertext link to open the Experiment to the Samples details page.
<b><u>Updated</u></b>	The date of the most recent update of the Experiment draft

Table 3.4 Experiment metadata categories

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 52.
- For information about contacting the Experiment POC, see [Primary Contact](#) in [Table 3.4](#).
- For information about extracting data from an Experiment, see [Downloading Data from caArray](#) on page 67.

**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, including all components.

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

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

*[Creating an Experiment on page 28](#)*

*[Managing Array Designs on page 44](#)*

*[Managing Protocols on page 46](#)*

*[Managing \[Controlled\] Vocabulary \[Terms\] on page 50](#)*

*[Updating An Experiment Proposal on page 52](#)*

*[Managing Experiment Visibility on page 53](#)*

### 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, protocols, quality control and data processing steps, and so forth. You then upload research data files related to the hybridizations performed in the Experiment, and link the sample information you submitted to the hybridizations.

Once you log into caArray, the My Experiments Workspace displays by default. The basic elements of a caArray Experiment, shown on the Experiments page, are described in [the following table Table 4.1](#):

<b>Term</b>	<b>Definition</b>
<b>Experiment Overview</b>	Basic information about an experiment such as IDs, service and assay types, provider of array and array designs, source of biomaterials, and disease state
<b>Contacts</b>	Principal Investigator and/or Point of Contact for the Experiment
<b>Annotations</b>	Experimental factors and design, biomaterials used, hybridizations
<b>Data</b>	Experimental data files uploaded and imported into caArray; imported data available for extraction from caArray; Supplemental data
<b>Publications</b>	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 to the caArray repository. 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.

### *Creating an 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** (*Figure 4.1*). This opens the Overview tab for entering overall characteristics for the Experiment..



Figure 4.1 Create/Propose Experiment on left sidebar

3. Proceed to the *Overview Tab*.

### 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 Characteristics<sup>1</sup> as described in the [the following table Table 4.2](#).

Experiment Details

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: test1

Status: Draft

Experiment Identifier:

Service Type\*: Full

Assay Type\*: Gene Expression

Provider\*: Affymetrix

Array Designs: DrosGenome1  
GenomeWideSNP\_6.Full  
hg-ut33\_plus\_2  
HT\_HG-U133A  
Mapping250K\_Nsp  
Mapping250K\_Sty  
Test3

Organism\*: Homo sapiens

Save

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	<b>Draft</b> displays by default before the Experiment is formally submitted.
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 <b>l</b> ast name followed by a 5 character number. <i>Example:</i> klemm-90765. After the Experiment has been saved or submitted, the Experiment ID displays <del>generally</del> as 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
<b>Service Type</b>	<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"> <li><b>Full</b>--typical relationship between PI and lab; with this selection, the Experiment is created in caArray and “submitted to a lab” using the standard workflow.</li> </ul> <p><del>Note: Full Service Experiments are placed in Public workspaces</del></p> <ul style="list-style-type: none"> <li><b>Publish</b>--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.</li> </ul> <p><del>Note: Publish Service Experiments remain in the PI's Experiment queue for making public when desired.</del></p>
<b>Assay Type</b>	<p>Select from the drop-down menu the appropriate assay type. Options are the following:</p> <ul style="list-style-type: none"> <li><b>Gene Expression</b>--experiment using microarrays intended to measure levels of transcribed genes</li> <li><b>SNP</b>--experiment using microarrays intended to detect nucleotide changes in chromosomal DNA</li> <li><b>aCGH</b>--<u>define</u></li> <li><b>exon</b>--<u>define is this as in “exon” or “intron”. Why would this generic term be used here?</u></li> </ul>
<b>Provider</b>	<p>Select from the drop-down menu the provider of the array.</p> <p><b>Note:</b> Only Affymetrix, Illumina and GenePix are supported in caArray 3.0. <u>If this is true, why are so many “providers” listed here?</u></p> <p>Once selected, caArray automatically loads a corresponding list of Array Designs (next field)</p>
<b>Array Designs</b>	<p>Select one or more array design from the automatically-generated list of array designs corresponding to the Provider you selected. <del>A significant number of Array Designs have been imported into caArray.</del> or you can choose to import those of your choice. For mor information, see <a href="#">Managing Array Designs</a> on page 44.</p>
<b>Organism</b>	<p>Select from the drop-down menu the organism that is the source of the sample biomaterial used in the Experiment.</p>

Table 4.2 Fields for Overall Experiment Characteristics

- 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 <b>Yes</b> or <b>No</b> . If answer is <b>No</b> , caArray opens additional text boxes where you can add contact information for the project Point of Contact (POC). <b>Note:</b> 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 4.4<sup>1</sup>](#).

<b>Experimental Design Fields</b>	<b>Description</b>
<b>Experimental Design Type*</b>	Select the Experimental Design type among the options in the drop-down menu.
<b>Experimental Design Description</b>	Enter a description for the Experimental Design used for the Experiment.
<b>Quality Control Types</b>	Select the QC type in the displayed list.
<b>Quality Control Description</b>	Enter a description for the Quality Control used for the Experiment.
<b>Replicate Types</b>	Select one or more Replicate Types <del>2-kinds</del> from the displayed list. Replicates can be either technical (arrays) or biological (laboratory animals or samples, etc.)
<b>Replicate Description</b>	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** to save the draft. ~~The name of the factor now displays on the Experimental Factors subtab.~~ Click **Cancel** to return to the subtab without adding the design.
3. Proceed to the ~~Experimental Factors~~**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 4.5<sup>2</sup>](#).

<b>Experimental Factors Fields</b>	<b>Description</b>
<b>Factor Name*</b>	Enter a name for the Experimental Factor.

Table 4.5 Experimental Factor fields

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


<b>Experimental Factors Fields</b>	<b>Description</b>
<b>Description</b>	Enter a description for the Experimental Factor.
<b>Category</b> 	Select the appropriate category for the Experimental Factor in the displayed list.

Table 4.5 Experimental Factor fields

- Click **Save** to save the draft. ~~The name of the factor now displays on the Experimental Factors subtab.~~ Click **Cancel** to return to the subtab without adding the factor.
- Repeat steps 1 - 3 as often as needed to enter all the Experimental Factors for this Experiment.
- Click **Save** or **Cancel** to abort the action.
- Proceed to the **Sources Tab** Sources 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 *Managing [Controlled] Vocabulary [Terms]* on page 50.



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 acidson the array. *Example:* The Extract RNA is labeled with a fluourescent dye, yielding the labeled extract RNA.

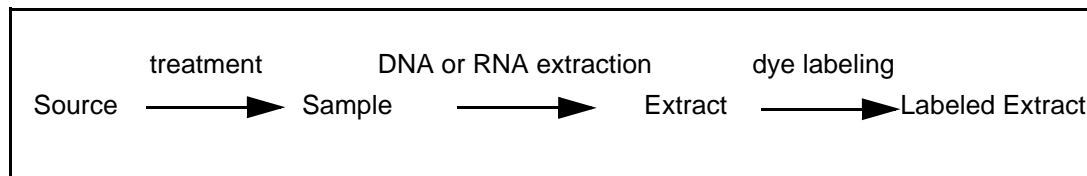


Figure 4.3 Biomaterials components and their relationship in caArray

The nature of each type of biomaterial is defined in varying dialog boxes in caArray.

[Sources Tab](#)

[Samples Tab](#)

[Extracts Tab](#)

[Labeled Extracts Tab](#)

## Sources Tab

A 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.

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

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 4.6](#).<sup>1</sup>

Source Fields	Description
<b>Source Name*</b>	Name assigned to the source
<b>Description</b>	Description of the source
<b>Tissue Site*</b>	<b>Tissue Site</b> is the site from which the Source material was obtained. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.
<b>Material Type</b>	<b>Material Type</b> is the descriptor for the type of material gleaned from the tissue site. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.

Table 4.6 [Fields for documenting a Source](#)




1. Fields with a red asterisk are required.

Source Fields	Description
<b>Cell Type</b>	Cell Type is the descriptor for the category of cells used for the Source material. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.
<b>Disease State</b>	Disease State is the descriptor used to describe the disease condition when the Source was obtained. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.

Table 4.6 *Fields for documenting a Source*

- Click **Save** to save the Source to the Experiment. ~~The name of the Source now displays on the Sources subtab.~~ Click **Cancel** to return to the subtab without adding the Source.

**Note:** When you create Samples for this Experiment, you must specify the Source(s) for the Samples. The Samples column on this Sources tab will then be populated.

- Repeat steps 1-3 as often as necessary to add all appropriate Sources to the Experiment.
- 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.
- To edit the new Source, click the corresponding **Edit** button (  Edit ) and edit the data. Click **Save** to save the edits.
- To delete a Source, click the Delete icon (  ) in the corresponding row.
- Proceed to the [Samples Tab](#) *Samples Tab*.

*Biological Source Material*

*Extracts Tab*

*Labeled Extracts Tab*

## Samples Tab

A Sample is the original source biomaterial after initial treatment events. *Example:* A tissue or biopsy material treated to create a cell lysate. The Samples subtab (under the Annotations Tab) displays any previous Samples that may have been added previously to the Experiment you are creating.

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

- On the Samples tab, click the **Add a New Sample** button to add a new Sample.

2. In the Sample form that opens, enter the information described in [the following table Table 4.7](#).<sup>1</sup>


<b>Samples Fields</b>	<b>Description</b>
<b>Sample Name*</b>	Enter a name for the sample.
<b>Description</b>	Enter a description of the sample.
<b>External ID</b>	Enter an identification value given to the Sample outside of caArray, for example, in the lab that did the work
<b>Source(s)*</b>	Sources must already have been saved to caArray. Select one or more <b>Sources</b> from which the Sample was derived. As you do so the selected source(s) move into the <b>Selected Sources</b> panel.  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. See <a href="#">Sources Tab</a> on page 35 for more information.
<b>Material Type</b>	<b>Material Type</b> is the descriptor for the type of material gleaned from the tissue site. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.
<b>Protocol Type</b>	Select from the drop-down menu the descriptor for the category of protocol used to prepare the Sample.
<b>Protocol</b>	Select from the drop-down list the protocol used to prepare the sample. If the appropriate protocol has not been entered into the system, click <b>Add</b> to open the page where you can add a new protocol. For more information, see <a href="#">Creating a Protocol</a> on page 48.

Table 4.7 Fields for documenting Samples


3. Click **Save**. The name of the Sample now displays on the Samples subtab. Click **Cancel** to return to the subtab without adding the Sample.

**Note:** When you create Extracts for this Experiment, you must specify the Samples for the Extracts. The Extracts column on this Samples tab will then be populated.

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

4. Repeat steps 1 - 3 as often as needed to enter all the Samples used in this Experiment.
5. 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.

1. Fields with a red asterisk are required.

- 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*.

### *Biological Source Material*

#### *Sources Tab*

#### *Labeled Extracts Tab*

## Extracts Tab

An Extract is a Sample after a treatment event in which DNA or RNA is extracted<sup>1</sup> 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.

- On the Extracts subtab, 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 4.8<sup>1</sup>](#).

<i>Extracts Fields</i>	<i>Description</i>
<b>Extract Name*</b>	Name assigned to the Extract
<b>Description</b>	Description of the extract
<b>Samples*</b>	<p>Samples must already have been saved to caArray. Select one or more <b>Samples</b> from which the Extract was derived. As you do so, the selected sample(s) move into the Selected Samples panel.</p> <p>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. See <a href="#">Samples Tab</a> on page 36 for more information.</p>
<b>Material Type</b>	<b>Material Type</b> is the descriptor for the type of material gleaned from the tissue site. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.
<b>Protocol Type</b>	Select from the drop-down menu the descriptor for the category of protocol used to prepare the Extract..

Table 4.8 Fields for documenting an Extract




1. Fields with a red asterisk are required.

<b>Extracts Fields</b>	<b>Description</b>
<b>Protocol</b>	Select from the drop-down list the protocol used to prepare the Extract. If the appropriate protocol has not been entered into the system, click <b>Add</b> to open the page where you can add a new protocol. For more information, see <a href="#">Creating a Protocol</a> on page 48.

Table 4.8 Fields for documenting an Extract

- Click **Save**. The name of the Extract now displays on the Extracts subtab. Click **Cancel** to return to the subtab without adding the Extract.

**Note:** When you create Labeled Extracts for this Experiment, you must specify the Extracts for the Labeled Extracts. The Labeled Extracts column on this Extracts tab will then be populated.

- 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](#).

### Biological Source Material

#### Sources Tab

#### Samples Tab

### Labeled Extracts Tab

A Labeled Extract is an Extract that has been labeled for detection of the nucleic acids on the array. *Example:* The Extract RNA is labeled with a fluorescent dye, yielding the labeled extract RNA.

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 4.9<sup>1</sup>](#).

<b>Labeled Extracts Fields</b>	<b>Description</b>
<b>Labeled Extract Name*</b>	Name assigned to the Extract
<b>Description</b>	Description of the extract




Table 4.9 Fields for documenting a Labeled Extract

<b>Labeled Extracts Fields</b>	<b>Description</b>
<b>Extracts*</b>	Extract(s) from which the Labeled Extract was derived. Extracts must already have been saved to caArray. Select one or more <b>Extract</b> from which the Labeled Extract was derived. As you do so, the selected Extract(s) move into the Selected Extracts panel. 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. See <a href="#">Extracts Tab</a> on page 38 for more information.
<b>Material Type</b>	<b>Material Type</b> is the descriptor for the type of material gleaned from the tissue site. There are three ways you can enter terms for annotating this attribute. See <a href="#">Managing [Controlled] Vocabulary [Terms]</a> on page 50 for more information about using this feature.
<b>Protocol Type</b>	Select from the drop-down menu the descriptor for the category of protocol used to prepare the Labeled Extract..
<b>Protocol</b>	Select from the drop-down list the protocol used to prepare the Labeled Extract. If the appropriate protocol has not been entered into the system, click <b>Add</b> to open the page where you can add a new protocol. For more information, see <a href="#">Creating a Protocol</a> on page 48.

Table 4.9 Fields for documenting a Labeled Extract

- Click **Save**. ~~The name of the Labeled Extract now display on the Extracts subtab.~~ Click **Cancel** to return to the subtab without adding the Labeled Extract.

**Note:** When you create Hybridizations for this Experiment, you must specify the Labeled Extracts for the Hybridizations. The Hybridizations column on this Labeled Extracts tab will then be populated.

- Repeat steps 1 - 3 as often as needed to enter all the Labeled Extracts used in this Experiment.
- To copy the Labeled Extract, click the **Copy** icon (  ) in the **Copy** column. caArray copies the Labeled 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~~.

### Biological Source Material

- Fields with a red asterisk are required.



[Sources Tab](#)

[Samples Tab](#)

[Extracts 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.

1. 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 4.10](#)<sup>1</sup>.

<b>Hybridizations Fields</b>	<b>Description</b>
<b>Hybridization Name*</b>	Name assigned to the Extract
<b>Description</b>	Description of the hybridization
<b>Labeled Extracts*</b>	Labeled Extract(s) used in the hybridization protocol. Select one or more listed Labeled Extracts. As you do so, the selected Labeled Extract(s) move into the Selected Labeled Extracts panel.  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.  See <a href="#">Labeled Extracts Tab</a> on page 39 for more information.

Table 4.10 Fields for documenting a hybridization

3. Click **Save**. The new Hybridization object displays on the Hybridization tab. Click **Cancel** to return to the subtab without adding the Hybridization.
- Note:** On the saved draft, the uncompressed size of the hybridization file is defined by caArray when it becomes available.
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 (  Edit ) 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](#).

[Biological Source Material](#)

[Sources Tab](#)

[Samples Tab](#)

1. Fields with a red asterisk are required.

*Extracts Tab**Labeled Extracts 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 4.11](#).

<i><b>Data Tabs</b></i>	<i><b>Description</b></i>
<b>Manage Data</b>	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.
<b>Imported Data</b>	This subtab list all files that have been imported into caArray.
<b>Supplemental Files</b>	This tab lists <a href="#">files</a> and documents that have been uploaded to caArray.
<b>Download Data</b>	From this tab, you can download (extract) data that has been imported into caArray.

*Table 4.11 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 4.12](#)).<sup>1</sup>

<i><b>Publications Fields</b></i>	<i><b>Description</b></i>
<b>Title*</b>	Title of the publication

*Table 4.12 Fields for documenting Publications*

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

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

Table 4.12 Fields for documenting Publications

- 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.

## Curation Tasks

Curation Tasks in caArray are those which can be managed by the Curation options on the left sidebar ([Figure 4.4](#)):

*Managing Array Designs*

*Managing Protocols*

*Managing Array Designs*

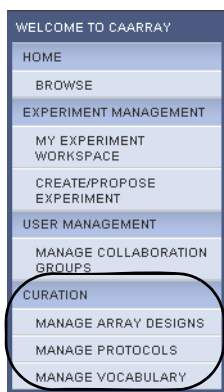


Figure 4.4 Curation options display in the left sidebar

## Managing Array Designs

A significant number of Array Designs are already pre-loaded into caArray. Or you can upload, validate and import Array Designs of your choice through this Curation tool. You can also view Array Designs, or edit those for which you have permissions.

[Viewing Array Designs](#)

[Adding an Array Design](#)

[Editing an Array Design](#)

### Viewing Array Designs

To view Array Designs in the system, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Array Designs**.

The Import Array Designs page that opens displays all Array Designs that have been pre-loaded or imported later into caArray (see [Importing Data](#) on page 62). Properties corresponding to those array designs are described in [Table 4.13](#) the following table.

**Note:** Columns with underlined headings are sortable by clicking on the heading.


<b><u>Array Designs Properties</u></b>	<b>Description</b>
<b><u>Array Design Name</u></b>	Name assigned the Array Design
<b><u>Provider</u></b>	The provider of the array design. This is generally the company or group that manufactured the array design.
<b><u>Assay Type</u></b>	The assay type used for the Array Design. <ul style="list-style-type: none"> <li>• <b>Gene Expression</b>--experiment using microarrays intended to measure levels of transcribed genes</li> <li>• <b>SNP</b>--experiment using microarrays intended to detect nucleotide changes in chromosomal DNA</li> </ul>
<b><u>Version Number</u></b>	The version number of the Array Design
<b><u>Feature Type</u></b>	The technology type or platform of the reporters on the array.
<b><u>Organism</u></b>	The organism used for the Array described by the Array Design.
<b>Edit</b>	If you do not have permissions to Edit this Array Design, this icon is not visible. If it is, click the <b>Edit</b> icon (  ) to open the Array Designs details page where you can edit the data. For more information, see <a href="#">Editing an Array Design</a> the following section.
<b>Status</b>	<b>Imported</b> <u>is this the only possible status?</u>

Table 4.13 Array Designs properties

If you click the Array Design Name in the Import Array Designs page, the details page that opens displays the name including file type extension of the uploaded/imported Array Design file.

## Adding an Array Design

### Editing an Array Design

## Adding an Array Design

**Note:** Array Design files must be in one of the following formats to be successfully imported into caArray:

- Affymetrix Array Designs are limited to `.cdf` files.
- Illumina Array Designs are limited to `.csv` files.
- Genepix Array Designs are either `.gal` (the actual design file) or derived from a `.gpr` file

To add an Array Design to caArray, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Array Designs**.

The Import Array Designs page that opens displays all Array Designs that have been pre-loaded or imported at a later time into caArray (see previous topic). [Importing Data](#)).

2. To import a new Array Design, click **Import a New Array Design** in the upper right corner.
3. On the form that opens, enter the appropriate information for the Array Design in the Array Design Details fields provided (and described in [the following table Table 4.14](#)).<sup>1</sup>

<b>Array Designs Details Properties</b>	<b>Description</b>
<b>Array Design Name</b>	Name assigned the Array Design. You should be sure to enter the accurate Array Design name.
<b>Assay Type</b>	The assay type used for the Array Design. <ul style="list-style-type: none"> <li>• <b>Gene Expression</b>--experiment using microarrays intended to measure levels of transcribed genes</li> <li>• <b>SNP</b>--experiment using microarrays intended to detect nucleotide changes in chromosomal DNA</li> <li>• <u><b>aCGH</b></u></li> <li>• <u><b>exon</b></u></li> </ul>
<b>Provider</b>	The provider of the Array Design. This is generally the company or group that manufactured the array design.
<b>Version Number</b>	The version number of the Array Design
<b>Feature Type</b>	The technology type or platform of the reporters on the array.

Table 4.14 Array Designs properties

1. Fields with a red asterisk \* are required.

<b>Array Designs Details Properties</b>	<b>Description</b>
<b>Organism</b>	The organism used for the Array described by the Array Design.

Table 4.14 Array Designs properties



4. In the Upload Array Design File section, click the **Browse** button to navigate to the file.
5. Click **Save** to launch the Array Design import process.  
The process includes uploading the file, validating it and importing it into the system. You should not leave the user interface once this process is underway, or you may have to start all over.

### Viewing Array Designs

### Editing an Array Design

#### Editing an Array Design

To edit an Array Design, follow these steps

1. On the row corresponding to the Array Design, click the **Edit** icon (  ).
- OR
2. Open the Array Design by clicking on its name, and click the **Edit** button (  Edit ) at the bottom of the details page.  
All information is editable.
3. From the Edit page, you can initiate uploading of an Array Design file. The newly uploaded file will replace the existng Array Design file.
4. Save any edit by clicking the **Save** button.

### Managing Array Designs

### Viewing Array Designs

### Adding an Array Design

## Managing Protocols

In caArray, you can create and manage protocol(s) for referencing in an Experiment. A protocol provides detailed documentation about the precise actions taken in any procedure that might be part of an experiment. For example, a Protocol could describe the steps a laboratory used for any kind of process used in an Experiment, such as the way a source material or sample is derived, the method used for labeling an extract or the methods used for running a hybridization or creating an image file of array results.

A protocol can be created independently of a specific experiment, or added during the course of creating an experiment. A protocol can be used by any caArray user but it can only modified by the owner of the protocol or another user with assigned permissions.

**Note:** Protocols do not include procedures relating to the statistical analysis of array results.

[Viewing Protocols](#)

[Viewing Protocol Types](#)

[Creating a Protocol](#)

[Editing a Protocol](#)

[Editing a Protocol Type](#)

## Viewing Protocols

To view existing protocols in caArray, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Protocols**.

The **Protocols** tab displays by default. All Protocols that have been created in caArray display on this tab. Properties corresponding to those Protocols are described in [Table 4.15](#)the following table.


<b><i>Protocol Properties</i></b>	<b><i>Description</i></b>
<b><u>Name</u></b>	Name assigned the Protocol
<b><u>Type</u></b>	Descriptor of the protocol type, such as labeling or hybridization.
<b><u>Description</u></b>	Description of the protocol procedure. Include any and all appropriate details, such as the detailed steps taken in a laboratory procedure. <i>Example:</i> Description of a procedure for labeling RNA with fluorescent tabs to be used in a hybridization procedure.
<b><u>Contact</u></b>	The name of the person to contact for information about the protocol.
<b><u>URL</u></b>	Link to a source of external documentation related to the Protocol
<b>Edit</b>	Click the <b>Edit</b> icon (  ) to open the Array Designs details page where you can edit the data. For more information, see <a href="#">Editing a Protocol</a> on page 49.

Table 4.15 Protocol properties

**Note:** Columns with underlined headings are sortable by clicking on the heading.

2. To view details of a protocol, click its name.

**Note:** Any protocols you did not create are in read-only mode.

[Managing Protocols](#)

[Viewing Protocol Types](#)

[Creating a Protocol](#)

[Editing a Protocol](#)

## Editing a Protocol Type

### Viewing Protocol Types

**Note:** Protocol Types are not created specifically as separate entities in caArray. They are identified when a user creates a protocol. For more information, see [Creating a Protocol](#) on page 48.

To view existing Protocol Types in caArray, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Protocols**.
2. By default, the Manage Protocols page displays. To view Protocol Types, click the **Protocol Types** tab.

The **Protocol Types** tab displays all Protocol Types that have been created in caArray. Properties corresponding to those Protocol Types are described in [Table 4.15](#) the following table.


<b>Protocol Type Properties</b>	<b>Description</b>
<b>Value</b>	The descriptor of the protocol type, such as labeling or hybridization.
<b>Description</b>	The description of the Protocol Type.
<b>Source</b>	The controlled vocabulary that is the source for the descriptor term value for the Protocol Type. The Source name is a hypertext link that takes you to the website for the source.
<b>Edit</b>	Click the <b>Edit</b> icon (  ) to open the Protocol Type details page where you can edit the data. For more information, see <a href="#">Editing a Protocol Type</a> on page 50.

Table 4.16 Protocol properties

3. To view details of a Protocol Type, click its **Value**.

## Managing Protocols

### Viewing Protocols

### Creating a Protocol

### Editing a Protocol

### Editing a Protocol Type

### Creating a Protocol

To create a protocol, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Protocols**.

The Manage Protocols page that opens displays all Protocols that have been created in caArray.



2. To create a new Protocol, click the **Add Protocols** button in the upper right hand corner of the page.
3. In the Manage Protocols form that opens, enter the appropriate information for the new protocol. Fields are described in [Table 4.17](#) the following table.<sup>1</sup>

<b>Protocol Properties</b>	<b>Description</b>
<b>Name*</b>	Name assigned the Array Design
<b>Description</b>	Description of the protocol procedure. Include any and all appropriate details, such as the detailed steps taken in a laboratory procedure. <i>Example:</i> Description of a procedure for labeling RNA with fluorescent tabs to be used in a hybridization procedure.
<b>Type*</b>	Descriptor of the protocol type such as “labeling” or “hybridization” from a controlled vocabulary, for example MGED.  Select a listed type or if the appropriate value is not displayed, begin typing a term in the <b>Filter</b> 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 protocol type cannot be found.
<b>Contact</b>	The name of the person to contact for information about the protocol.
<b>Software</b>	Name of software used in the Protocol. <i>Example:</i> GenePix Pro 3.0.1.22
<b>Hardware</b>	Name of hardware used in the Protocol. <i>Example:</i> GeneChip(R) Fluidics Station 450 <sup>®</sup>
<b>URL</b>	Link to a source of external documentation related to the Protocol

Table 4.17 Protocol fields

4. Click **Save** to save the protocol. Click **Cancel** to halt the action. In both cases, you are returned to the Manage Protocols page. The protocol you just added is listed first in the list of protocols.

[Managing Protocols](#)

[Viewing Protocols](#)

[Viewing Protocol Types](#)

[Editing a Protocol](#)

[Editing a Protocol Type](#)


## Editing a Protocol

**Note:** A protocol can be edited only by the owner of the protocol or by another user with assigned permissions.


---

1. Items with an asterisk are required.

To edit a protocol, follow these steps:

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Protocols**.
2. On the row corresponding to the Protocol, click the **Edit** icon (  ).

OR

3. Open the Protocol by clicking on its name, and click the **Edit** button (  Edit ) at the bottom of the details page that opens.
4. All information for a Protocol is editable. Make the appropriate edits on the form that opens.
5. Save any edits by clicking the **Save** button. To abort the edit, click the **Cancel** button. This returns you to the Manage Protocols page.

*Managing Protocols*


*Viewing Protocols*

*Viewing Protocol Types*


*Creating a Protocol*

*Editing a Protocol Type*

### Editing a Protocol Type

1. After logging into caArray, on the left sidebar under **Curation**, click **Manage Protocols**.
2. On the Manage Protocols page, click the **Protocol Types** tab.
3. On the row corresponding to the Protocol Type, click the **Edit** icon (  ).

OR

4. Open the Protocol Type by clicking on its name, and click the **Edit** button (  Edit ) at the bottom of the details page.
5. All information for a Protocol Type is editable. Make the appropriate edits on the form that opens.
6. Save any edits by clicking the **Save** button. To abort the edit, click the **Cancel** button. This returns you to the Manage Protocols page.

*Managing Protocols*

*Viewing Protocols*

*Viewing Protocol Types*

*Creating a Protocol*



*Editing a Protocol*

### 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.
3. To add a new term, click the **Add** button (  Add ). 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.

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

Table 4.18 Fields for entering a new vocabulary term

4. Once you have entered the appropriate information, click **Save**.

This returns you to the original tab, where you can continue defining the Experiment attribute.

## 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 53.

---

## 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 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 the caArray Repository](#) on page 19 or [Searching the caArray 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 or In Process status (having been submitted), 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 the caArray Repository](#) on page 19 or [Searching the caArray 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 the caArray Repository](#) on page 19 or [Searching the caArray 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, using the Permissions feature of caArray. Public access to the Experiment can be configured, and well as Collaboration Group access.

## Setting Public Visibility

## Setting Collaboration Group Visibility

### Setting Public Visibility

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 the caArray Repository](#) on page 19 or [Searching the caArray Repository](#) on page 22.
2. Click the **Permissions** icon (  ) in the row listing the Experiment.
3. The Experiment Permissions page that opens displays a panel that describes the visibility policies of caArray and allows you to set visibility for your Experiment.

The Visibility options are:

- **Visible**--applied by default to each experiment upon creation, this profile exposes summary information without access to annotation and array data.
- **Read**--grants read access to the experiment as a whole, providing a preview into its content
- **Read Selective**--grants selective access to specific sample annotation and data
- **No Visibility**--completely removes the Experiment summary information from view

**Note:** Your ability to set Visibility options varies according to the draft/submission status of the Experiment. See below.

*For an Experiment in Draft status:*

**Note:** Only **Visible** and **No Visibility** options are available for an Experiment in Draft status.

- a. To control public access, under **The Public**, click the **Edit Access Control** button.
- b. In the Control Access to Specific Content panel that displays on the right, select the visibility option from the drop-down list. The only available options for and Experiment draft) are: **Visible** and **No Visibility**.
- c. The option at the bottom of the left hand panel relates to The Cancer Genome Atlas (TCGA) policy (for human SNP data), which is that only four fields of annotation can be provided to the open public:
  - Clinical Diagnosis
  - Histologic Diagnosis
  - Tissue Anatomic Site
  - Pathologic Status

Click this option for these limits to apply to your Experiment.

**Note:** The Collaboration Groups visibility option is not available to an Experiment in Draft status.

- d. Click **Save** under the TCGA option to execute the Public access choices

*For a “submitted” Experiment:*

**Note:** Four visibility options plus Collaboration Group Visibility are available for Experiments that have been submitted to the caArray repository.

- a. To control public access, under **The Public**, click the **Edit Access Control** button.
- b. In the Control Access to Specific Content panel that displays on the right, select the visibility option from the drop-down list. The four available options are those described above: **Visible**, **Read**, **Read Selective** and **No Visibility**.
- c. The option at the bottom of the left hand panel relates to The Cancer Genome Atlas (TCGA) policy (for human SNP data), which is that only four fields of annotation can be provided to the open public:
  - Clinical Diagnosis
  - Histologic Diagnosis
  - Tissue Anatomic Site
  - Pathologic Status

Click this option for these limits to apply to your Experiment.

- d. Click **Save** under the TCGA option to execute the Public access choices.

### *Managing Experiment Visibility*

### *Setting Collaboration Group Visibility*

## Setting Collaboration Group Visibility

A section at the bottom of the left panel of the Experiment Permissions page allows you to set Experiment visibility for Collaboration Groups. You can use an existing group, or with proper permissions, you can create the group from this page. For information about how to create collaboration groups through the User Management features, see [Managing Collaboration Groups](#) on page 70.

To configure Experiment visibility for a collaboration group, follow these steps:

1. Perform either a browse or search to locate the Experiment. For more information, see [Browsing the caArray Repository](#) on page 19 or [Searching the caArray Repository](#) on page 22.


2. Click the **Permissions** icon (  ) in the row listing the Experiment.

The Experiment Permissions page that opens displays a panel that describes the visibility policies of caArray and allows you to set visibility for your Experiment. For information about setting Public visibility, see [Managing Experiment Visibility](#) on page 53.

*If the Collaboration Group already exists:*

1. At the bottom of the left panel, under **Collaboration Groups**, select in the drop-down list the group to which you want to assign visibility.
2. Click the **Edit Access Control** button.
3. In the Control Access to Specific Content to [Group] panel that displays on the right, select the visibility option for the group from the Experiment Access drop-down list. The five available options are: **None**, **Read**, **Read/Write** and **Read/Write/Selective**.
4. Click **Save** in that panel to execute the Collaboration Group visibility choices.

*If the Collaboration Group must be created:*

1. If the group to which you want to assign Experiment visibility does not yet exist, click the **Add New Group** button.
2. In the Add a New Collaboration Group form, enter the name of the group and click **Save**.
3. The Manage Collaboration Groups page that opens display all groups in the system, including the one you just created. Click the Edit icon (  ) to specify members for the group. For more information about creating and working with a Collaboration Group, see [Managing Collaboration Groups](#) on page 70.
4. Proceed from step
5. Click **Save** in that panel to execute the Collaboration Group visibility choices.

See also [Setting Public Visibility](#).



## 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 58
- [Validating Data Files](#) on page 60
- [Importing Data](#) on page 62
- [Supplemental Files](#) on page 64
- [Downloading Files](#) on page 65

## 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](#).

<b>Data Tabs</b>	<b>Description</b>
<b>Manage Data</b>	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.
<b>Imported Data</b>	This subtab list all files that have been imported into caArray.
<b>Supplemental Files</b>	This tab lists files and documents that have been uploaded to caArray, and identified as supplementary (reference) files.
<b>Download Data</b>	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.

- Notes:**
- To perform tasks on the Data tab, your Experiment must be in Edit mode.
  - 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.
  - Importing Array Design files is performed through the Curation tool in caArray, not on the Data tab. For more information, see [Managing Array Designs](#) on page 44.

---

**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 62.

---

## 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.

[Steps for Uploading Data](#)

[Deleting a File](#)

[Validating Data Files](#)


[Importing Data](#)

[Supplemental Files](#)

### Steps for Uploading Data

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 the caArray Repository](#) on page 19 or [Searching the caArray Repository](#) on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (  ) corresponding to the Experiment to which you want to upload data. 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, click **Browse** to navigate to 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](#)).

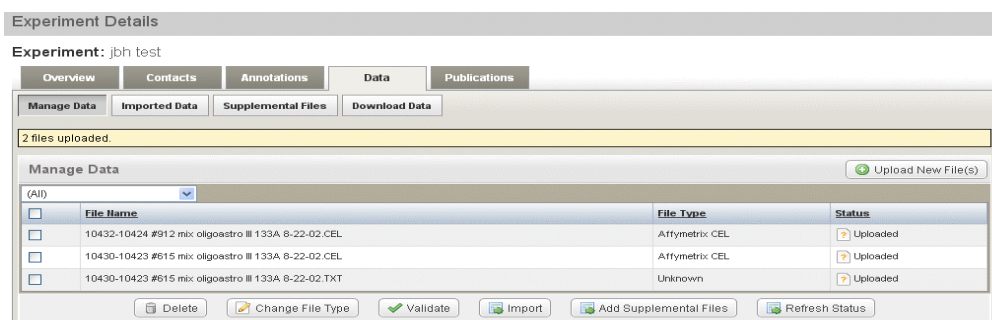


Figure 5.1 Manage Data subtab displays files in caArray and their status

**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.

[Uploading Data Files](#)

[Deleting a File](#)

[Validating Data Files](#)

[Importing Data](#)

## Supplemental 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.

### Uploading Data Files

#### Steps for Uploading Data

#### Validating Data Files

#### Importing Data

### Supplemental Files

## 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 4.2the 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 64.

- Word documents
- Excel spreadsheets
- PowerPoint files
- PDFs

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

The 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 62.

- **Illumina:** Gene expression and SNP

---

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

---

### Uploading Data Files

#### Steps for Validating Data


[Validation Errors](#)

[Importing Data](#)

[Supplemental Files](#)

## Steps for Validating Data

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 the caArray Repository](#) on page 19 or [Searching the caArray 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 64

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

[Uploading Data Files](#)

[Validating Data Files](#)

[Validation Errors](#)

[Importing Data](#)

[Supplemental Files](#)

## 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
- 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

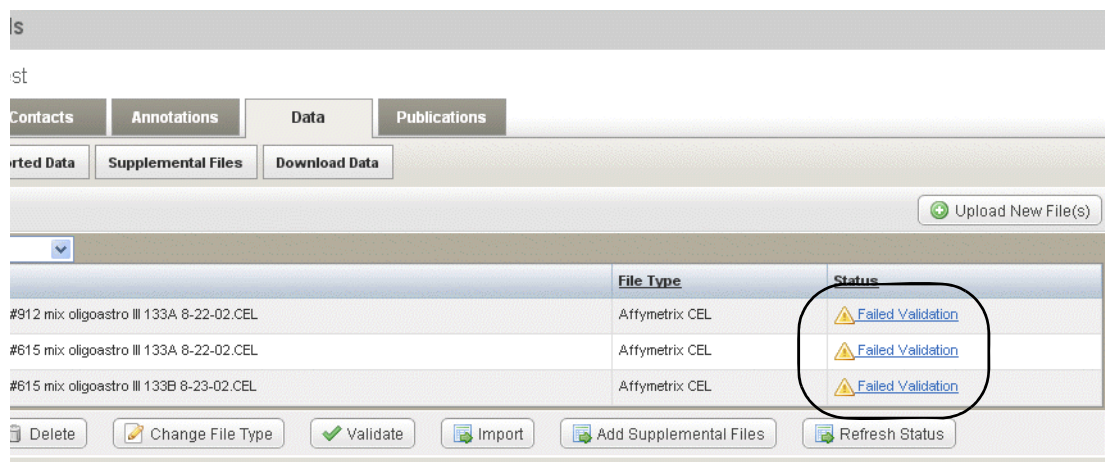


Figure 5.2 Validation failures display on the Manage Data table.

[Uploading Data Files](#)

[Validating Data Files](#)

[Steps for Validating Data](#)

[Importing Data](#)

[Supplemental Files](#)

## 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.

---

[Uploading Data Files](#)

[Validating Data Files](#)

[Files Types that can be Imported into caArray](#)

[Steps for Importing Data](#)

[Supplemental Files](#)

## 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"> <li>• Affymetrix .CEL</li> <li>• Affymetrix .CHP</li> <li>• Affymetrix.DAT (can be imported even though they cannot be validated in caArray)</li> <li>• Affymetrix.EXP (can be imported even though they cannot be validated in caArray)</li> <li>• GenePix .GPR</li> <li>• Illumina data files</li> </ul>
Array Design files	<ul style="list-style-type: none"> <li>• .CDF</li> <li>• MAGE-TAB ADF (Array Design Format)</li> <li>• Illumina .CSV and .txt</li> <li>• GenePix GAL</li> </ul> <p><b>Note:</b> These can be uploaded, validated and imported only through the Manage Array Design feature described in <a href="#">Managing Array Designs</a> on page 44.</p>
MAGE-TAB files	<ul style="list-style-type: none"> <li>• MAGE-TAB with single SDRF (Sample and Data Relationship Format)</li> <li>• MAGE-TAB with multiple SDRFs</li> <li>• IDF (Investigation Description Format) only, no referenced SDRFs</li> <li>• SDRF with only source &gt; Hybridizations</li> <li>• ADF (Array Design Format)</li> <li>• MAGE-TAB with existing samples in caArray.</li> </ul>

Table 5.2 File types that can be imported into caArray

[Uploading Data Files](#)

[Validating Data Files](#)

[Importing Data](#)


[Steps for Importing Data](#)

[Supplemental Files](#)

## 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 the caArray Repository](#) on page 19 or [Searching the caArray 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 60.
  - 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 the file set is **Imported**.
7. After a successful import, you can delete the file set. See [Deleting a File](#) on page 60.

*Uploading Data Files*

*Validating Data Files*

*Importing Data*

*Files Types that can be Imported into caArray*

*Supplemental Files*

## 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.



*Uploading Data Files*

*Validating Data Files*

*Importing Data*

*Files Types that can be Imported into caArray*

*Steps for Importing Data*

## 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 68


### 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 the caArray Repository](#) on page 19 or [Searching the caArray Repository](#) on page 22.
3. In the My Experiment workspace or the search results page, click the **Edit** button (  ) corresponding to the Experiment from which you want to download files. This opens the Experiment to the Overview tab; you are in **Edit** mode.
4. Select the **Data** tab, and the **Download Data** subtab.

**Note:** You can sort the columns of the list by clicking on the column headers. All files that are part of this Experiment display on this tab.

5. Identify one or more file(s) to be downloaded. Click the plus icon (  ) to the left of the **File Name** column on the row of the file(s) you select. That places the data set in the download queue, visible in the right panel on the tab. The identifying information for the file, including the total file size displays there.
6. Click the **Launch Download Job** button to initiate the download process. The job proceeds until all designated files are downloaded.
  - The length of time for the download is dependent upon the file size.
  - You can continue to work in caArray during the download process.

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.

7. To remove selected files from the queue, click the **Remove** icon(s) (  ) corresponding to the data file or click the **Cancel** button.

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. For more information, see [SSS](#)

---

**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
- calIntegrator
- 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 [section](#)[chapter](#) 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.

Topics in this [section](#)[chapter](#) include:

- [Roles in caArray on this page](#)
- [Managing User Accounts](#) on page 70
- [Managing Collaboration Groups](#) on page 70

## 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><b>Role</b></i>	<i><b>Description</b></i>	<i><b>Permissible Actions</b></i>
Anonymous User <b>Note:</b> 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
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

<b>Role</b>	<b>Description</b>	<b>Permissible Actions</b>
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 <b>Note:</b> Not listed under Roles	Systems other than caArray from which caArray data can be extracted programmatically using an API.	For more informations, see <a href="#">Extracting Data Programmatically by API</a> on page 68.

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](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.

[Viewing Group Details](#)

[Creating a Collaboration Group](#)

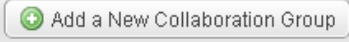




[Editing Collaboration Group Details](#)

### Creating a Collaboration Group

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

For more information, see [Viewing Group Details](#) on page 71. 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. Click Save. This returns you to the Manage Collaboration Groups page that lists groups that have been created in the system. When a group you create first displays here, it is empty.
  5. To add users to a Collaboration Group, click the **Edit** icon (  ) corresponding to the group.
  6. The group details page that opens lists current members of the group. Click the **Add a New Group Member** button (  ).
  7. In the form that opens, enter the users last name, first name, and organization, and click the Filter button (  ) to find the user. The user displays in the Member Name column.
  8. For each member to be added to the group, click the **Add** icon (  ) in the far right column of the screen.
  9. Click the **Save** button to save the group information. or click **Cancel** to abandon creation of the group.
- I'm not sure I have this description correct because no users are in the system that I can add to a group so I can check it out.
10. On the Manage Collaboration Groups page, 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*

*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 **Edit** icon (  ) corresponding to the group. The group details lists all users 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.

*Creating a Collaboration Group*


[Viewing Group Details](#)



[Viewing Group Details](#)

[Editing Collaboration Group Details](#)

## Editing Collaboration Group Details

To edit collaboration group details, follow these steps [check details](#):

1. From the Manage Collaboration Groups page, click the **Edit** icon (  ) corresponding to the group you select.
2. 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. [check this out](#)
  - Delete the group.
3. 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.
Add a new group member	<p><b>Note:</b> The new member must already have a valid caArray user account.</p> <p>Click the <b>Add a New Group Member</b> 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. <a href="#">true??</a> Enter the last name, select the Role category, the Institution, and the Status to be searched. Click the Filter button.</p>
Remove a group member	On the Collaboration Group page, in the <b>Remove</b> column, click the icon (  ) that corresponds to the group member.
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 <b>Delete</b> button (  Delete ) at the bottom center of the page.

[Creating a Collaboration 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](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.

<b>Term</b>	<b>Definition</b>
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

<b>Term</b>	<b>Definition</b>
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><a href="#">UC summary p. 4 for more</a></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 href="#">a set of annotation and array data that describe the experimental design, the samples used and the array data produced</a></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><a href="#">the material used for extraction onto an array</a></u>
Source	<u><a href="#">the material – typically tissue or cell lines - from which a sample is taken</a></u>
URI	Uniform Resource Identifier
URL	Uniform Resource Locators
XML	Extensible Markup Language ( <a href="http://www.w3.org/TR/REC-xml/">http://www.w3.org/TR/REC-xml/</a> ) - 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

- Array Designs 45
- Experiment data 53
- Group 70
- Publication 42
- vocabulary term 51

Annotations tab 32

Application Support i, 16

Array Designs

- adding 45
- editing 46
- managing 44
- viewing 44

## B

biological source materials

- description 34
- relationship between 35

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 77
- interface description 14
- launching 7
- login 8
- My Experiment Workspace 16
- objectives 5
- overview 5
- references 75
- 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 70

- editing 72

- viewing details 71

collaboration groups

- managing 70

- setting Experiment visibility 55

Contact tab 32

creating

- collaboration group 70

- Experiment 29

- group 70

- protocol 48

- user account 70

curation tasks 43

## D

Data tab

- deleting a file 60

- description 42

- downloading data files 67

- importing files 62

- Manage Data subtab 57

- Supplemental Files 64

- uploading data 58

- validating data 60

deleting

- data file 60

- Experiment 53

downloading data files 67

draft status, Experiment visibility in [54](#)

## E

editing

- Array Designs [46](#)
- collaboration group [72](#)
- Experiment [52](#)
- Protocol [49](#)
- Protocol Typer [50](#)

Experiment

- adding data to [53](#)
- Annotation tab [32](#)
- Contact tab [32](#)
- creating [29](#)
- deleting [53](#)
- editing [52](#)
- managing data submission [57](#)
- managing visibility [53](#)
- overview tab [29](#)
- submitting [52](#)
- updating [52](#)
- uploading data to [58](#)

Experimental Factors tab [33](#)

Experiments, overview [27](#)

Extract

- adding new [38](#)
- description [34](#)
- tab [34](#)

extracting caArray files [67](#)

Extracts tab [38](#)

## F

file types, importable [63](#)

## G

Global Quick Links [16](#)

Glossary [77](#)

group

- creating [70](#)
- viewing details page [71](#)

## H

help in caArray [13](#)

Hybridization, adding new [41](#)

Hybridizations tab [41](#)

## I

importable data file types [63](#)

importing data files [62](#)

## L

Labeled Extract

- adding new [39](#)
- description [35](#)
- tab [35](#)

Labeled Extracts tab [39](#)

launching caArray [7](#)

## M

managing

- Array Designs [44](#)
- controlled vocabulary terms [36](#)
- data [57](#)
- user accounts [70](#)
- user groups [70](#)
- vocabulary terms [50](#)

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

Protocol

- creating [48](#)
- editing [49](#)

Protocols

- description [46](#)
- viewing [47](#)

Protocol Types

- editing [50](#)
- viewing [48](#)

Publication, adding [42](#)

Publications tab [42](#)

public visibility, Experiment [53](#)

## R

references [75](#)

registering new user [11](#)

## S

Sample

- adding new [36](#)
- description [34](#)
- tab [34](#)

- Samples tab [36](#)
- searching
  - caArray repository [22](#)
  - from caArray login screen [8](#)
- search results [24](#)
- Source
  - adding new [35](#)
  - description [34](#)
  - tab [34](#)
- Source Materials
  - annotations for [34](#)
  - characteristics [34](#)
  - definition [34](#)
- Sources tab [35](#)
- submitted status, visibility in [55](#)
- submitting Experiment [52](#)
- Supplemental Files [64](#)

## T

- Technical Support [i, 16](#)
- text conventions in user guide [2](#)

## U

- uploading data to Experiment [58](#)
- user's manual conventions [2](#)
- user account, new [70](#)
- user groups
  - managing [70](#)
- users groups
  - editing [72](#)

## V

- validating data file [60](#)
- validation errors [61](#)
- viewing
  - Array Designs [44](#)
  - collaboration group details [71](#)
  - Protocols [47](#)
  - Protocol Types [48](#)
- visibility
  - collaboration groups, setting [55](#)
  - Experiment in draft status [54](#)
  - Experiment in submitted status [55](#)
  - setting public [53](#)
- vocabulary terms
  - adding [51](#)
- vocabulary terms, managing [50](#)

