

CAARRAY 2.0

User's Guide



NATIONAL[®]
CANCER
INSTITUTE

Center for Bioinformatics

Revised November 1, 2007

CREDITS AND RESOURCES

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

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

TABLE OF CONTENTS

| | |
|---|-----------|
| Credits and Resources | i |
| Using the caArray User's Guide | 1 |
| Introduction to the caArray User's Guide | 1 |
| Organization of this Guide | 1 |
| User's Guide Text Conventions | 2 |
| Chapter 1 | |
| About caArray | 5 |
| caArray Overview | 5 |
| Relationship of caArray to caBIG | 6 |
| caArray Standards | 8 |
| Security Structure Reword to reflect 2.0 | 9 |
| caArray Workflow | 10 |
| Chapter 2 | |
| Getting Started in caArray | 13 |
| Launching caArray | 13 |
| Browsing caArray Before Login | 14 |
| caArray New User Account and Login | 15 |
| Requesting a User Account | 16 |
| Using caArray Online Help | 18 |
| Navigating the caArray User Interface | 19 |
| The caArray provides a user-friendly interface to display options relating to login, accessing documentation or performing tasks for proposed or existing experiments. | 19 |
| Elements in the caArray User Interface | 19 |
| caArray Welcome Page Navigation Menu | 20 |
| Experiment Workspace | 21 |

Chapter 3

User Account Management25

| | |
|---|----|
| Initial questions | 25 |
| Roles in caArray | 26 |
| Permissions | 26 |
| Managing User Accounts | 27 |
| Registering a User Account | 28 |
| Creating a New User--this was written from the description of Registering a New User in the Getting Started chapter because...see next red text. | 28 |
| Editing User Accounts by an Administrator | 31 |
| Editing One's Own Account | 32 |
| Searching for a User | 32 |
| Viewing User Details | 34 |
| Managing Collaboration Groups | 37 |
| Creating a Group | 38 |
| Searching for a Group | 39 |
| Viewing Group Details | 40 |
| Modifying Collaboration Group Details | 40 |

Chapter 4

Navigating and Searching caArray43

| | |
|---------------------------------|----|
| Browsing caArray | 43 |
| Searching the Repository | 46 |
| Experiment Search Results | 47 |

Chapter 5

Creating and Managing Experiments49

| | |
|--|----|
| Overview of an Experiment | 49 |
| Creating an Experiment | 50 |
| Overview Tab | 51 |
| Contacts Tab | 53 |
| Annotations Tab | 54 |
| Data Tab | 63 |
| Publications | 63 |
| Submitting an Experiment | 64 |
| Editing an Experiment | 64 |
| Adding Controlled Vocabulary Terms | 64 |
| Managing Experiment Visibility | 65 |
| Managing Access to caArray Data | 65 |

This is from previous manual. When I get time to work through it, it will

| | |
|---|-----------|
| probably6 be removed. | 65 |
| Requesting Access to caArray Records | 65 |
| Granting Others Access to Your Work | 66 |
| Visibility | 66 |
| Updating an Experiment proposal | 67 |
| Experimental Design | 69 |
| Sources | 69 |
| Samples | 70 |
| Chapter 6 | |
| Submitting Data to an Experiment | 71 |
| Managing Data | 71 |
| Uploading Annotation and Array Content into a caArray project | 71 |
| Validating Data Files | 73 |
| Importing Array Designs | 74 |
| Files Types that can be Imported into caArray | 74 |
| Managing Associations | 76 |
| Chapter 7 | |
| Extracting Data from caArray | 77 |
| Downloading Data from caArray | 77 |
| Grid availability | 78 |
| Extracting Data Programmatically by API | 78 |
| Appendix A | |
| caArray References | 81 |
| Appendix B | |
| caArray Glossary | 83 |
| Index | 87 |

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](#):

Chapter 1 About caArray—This chapter provides an overview of the caArray infrastructure including a discussion on Model Driven Architecture, *n*-tier architecture with open APIs, use of controlled vocabularies, and registered metadata. It also includes a description about the workflow in caArray and the rationale for creating caArray objects, the Experiments comprised of those objects as well as data files uploaded to those Experiments.

Chapter 2 Getting Started in caArray—This chapter provides details about launching caArray, navigating the interface, accessing online help and performing basic functions in caArray such as searching caArray, submitting and modifying caArray objects.

Chapter 3 User Account Management—This chapter describes how to create and manage user accounts in caArray.

Chapter 4 Contacts—This chapter describes the process for creating and working with Contacts, namely individual and group contacts in caArray.

Chapter 5 Protocols—This chapter details instructions for creating and working with Protocols in caArray.

Chapter 6 Arrays and Array Designs—This chapter provides information about creating and working with Array Design models as well as Array objects in caArray.

Chapter 7 Controlled Vocabulary Terms—This chapter describes the means for accessing, using and adding standardized terms from Internet sources in the Controlled Vocabulary Terms section in caArray

Chapter 8 Biological Source Materials—This chapter describes the creation, annotation, and management of biological source materials in caArray.

Chapter 9 Experiments—This chapter describes the processes for using components created in caArray and data files uploaded into caArray to create a caArray object called an Experiment.

Appendix A—This appendix provides a glossary of words relative to microarray techniques in general and caArray in particular.

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

User's Guide Text Conventions

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


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

Table 2.1 caArray Guide Text Conventions

| Convention | Description | Example |
|--|---|--|
| <i>Italic boldface monospace type</i> | Text that you type | In the New Subset text box, enter <i>Proprietary Proteins.</i> |
| Note: | Highlights a concept of particular interest | Note: This concept is used throughout the installation manual. |
| Warning! | Highlights information of which you should be particularly aware. | Warning! Deleting an object will permanently delete it from the database. |
| { } | Curly brackets are used for replaceable items. | Replace {root directory} with its proper value, such as c:\cabio |

Table 2.1 caArray Guide Text Conventions (Continued)

CHAPTER 1

ABOUT CAARRAY

This [sectionchapter](#) describes briefly the microarray process as well as the caArray software.

Topics in this [sectionchapter](#) include:

- [caArray Overview on this page](#)
- [Relationship of caArray to caBIG on this page](#)
- [caArray Standards on page 8](#)
- [caArray Workflow on page 10](#)

caArray Overview

I've taken most of this from the Vision document. While I think it's good, I'd like to condense the text in this chapter. I doubt that readers of this user's guide will take the time to read this in its entirety the way it is.

If either of you can edit it, that would be great. With track changes on, you can just edit the text in the Vision doc (cause it's in Word :) and I'll make the appropriate changes here.

If neither of you have the time to do this, I could work it over. It will probably be easier for one of you because you know what is important enough to leave in and what can be excised.

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 shareable 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.

At the highest level, the following services are provided by caArray:

- Managing experiment annotation
- Array data upload, validation, and import
- Searching across experiments using a web browser
- Programmatic querying and retrieval of data over the grid or via a java API
- Managing of experiment security and access.

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. caBIG™ serves as the cornerstone of NCI's biomedical informatics efforts to transform cancer research into a more collaborative, efficient, and effective endeavor. In 2007, caBIG™ has moved from the pilot phase to the enterprise and is looking to evolve its set of maturing assets.

As the scientific community begins to better understand cancer at the molecular levels and personalized medicine is implemented in cancer patient care, researchers and clinicians will require more rapid access to—and easier methods to analyze—the multiple types of information involved. To this end, the vision of caBIG™ is a full cycle of integrated cancer research, extending from bench to bedside, and back again.

caArray, an integral component of the caBIG array of tools, has a committed following, years of user feedback and contribution on which to build from, sufficient resources, and an expansive vision to support integrative cancer research. It

Expression profiling is now a standard tool set with which to interrogate biological systems. Parallel advances in computing and new array technology provide an opportunity for collaboration and discovery within the scientific community and across traditional boundaries to reach clinicians and ultimately patients. The insistence on open source development provides the community with the greatest opportunity to gain access to the tools they desperately need to execute their respective mission. caArray was initially developed with expression profiling in mind, using the caBIG Compatibility Guidelines, as well as the Microarray Gene Expression Data (MGED) society standards for microarray data. Compatibility with these standards and guidelines was and remains required. However, the ability to add new standards that are developing is also necessary to facilitate data exchange and analysis across domains. A number of analytical tools and services that connect to caArray are already available - including geWorkbench and GenePattern - that provide a variety of analysis, visualization and annotation functions for microarray data.

The primary goal of caArray is to further translational cancer research through acquisition, dissemination and aggregation of high quality array data to support subsequent analysis. Initially envisioned to be the de facto repository for cancer related expression data, other applications – primarily the Gene Expression Omnibus (GEO) – has assumed the role. However, the quality of the expression data and the ability to

integrate the use of array technology into clinical research remains at issue. Further, the opportunity for caArray to evolve to handle other types of array data will provide greater impetus for use among the Cancer Centers and their collaborators which will ultimately benefit the cancer community. Data from expression, SNP, methylation, and protein arrays are anticipated for inclusion in caArray. Also under consideration are the burgeoning platforms of RNAi and CHip-chip data. Tissue microarrays represent another array-based technique that require a significantly different business process in their creation but are being considered for inclusion in the repository.

A significant challenge is to find common ground for the annotation, upload and extraction of these array data platforms to support meaningful analysis for cancer research. The need for logical and expedient approaches to storing, querying, retrieving and reporting on array-based data has increased over the last several years due to the contribution of the following factors:

- a significant increase in available array information (expression and others)
- decreases in the cost to generate array data
- improvements in the technology to generate this data
- advances in the approaches to analyze this data
- the need to increase the numbers of samples to enhance discovery and ensure the validity of results once found
- the need to ease the administration of the data and results by giving the community a single point of reference to perform the most essential tasks of array-based research.

The initial generation of caArray (1.0) and subsequent point releases (currently 1.5) have provided an interface for annotating scientifically significant meta-data along with the independent ability to upload and download data through an applet or accessing the MAGE-OM API over the grid service provided by caBIG. caArray 2.0 will improve upon this starting point by providing these features:

- Organizing the application around the natural workflow between investigators and the array labs that serve them
- Improve the user experience for storing and retrieving the data produced,
- Query the data through an easier to comprehend API
- Bridge the gap between the analysis tools in heavy use in the community today and the data they need to consume.

A simplified organizational structure for basic experiment relationships that fundamentally represent “annotation” versus the resulting array “data” is shown in Figure 1. This is an overall theme of the application.

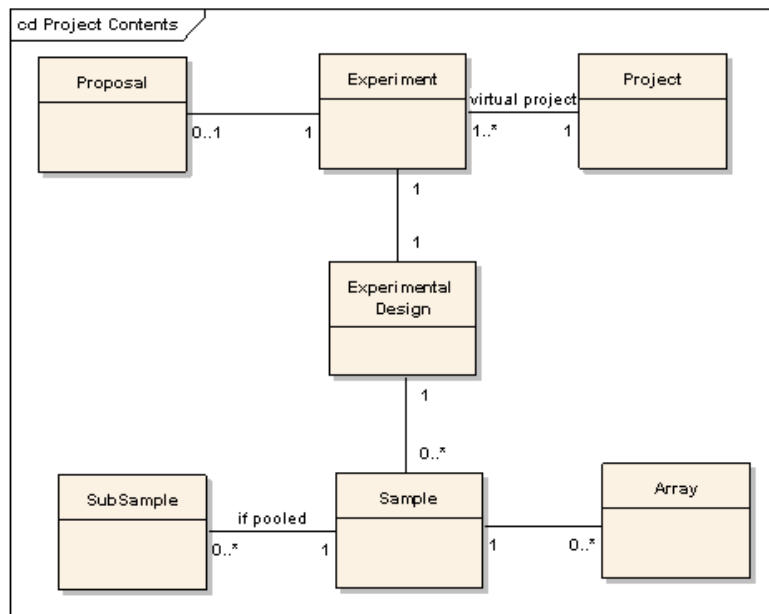


Figure 1.1 Basic Experiment relationships in caArray

caArray Workflow

caArray will look to perform intra-platform, intra-experiment queries for local adopters, via the NCICB instance and potentially across the grid collectively to increase the availability of quality data.

caArray will be built to scale 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 poised to exploit web services and/or databases from other components of caBIG when available and prudent. The desire to create an extensible array system that is non-platform-specific and potentially customizable is a theme that should influence the building of caArray.

caArray Overview

caArray Standards

Security Structure Reword to reflect 2.0

caArray Workflow

caArray Standards

As research and technology expand, it is of critical importance that established standards be used to ensure better and more accurate data collection and experiment results. Removing ambiguity by using standardized terminologies is one of the cornerstones of caArray. caArray is designed to support the international standard for microarray data and caBIG compatibility guidelines.

MAGE-TAB—MAGE-TAB is a simple tab-delimited, spreadsheet-based format, which will be part of the MAGE microarray data standard. MAGE-TAB can be used for annotating and communicating microarray data in a MIAME compliant fashion.

MIAME 1.1 — Minimum Information About a Microarray Experiment describes the minimum information necessary to enable unambiguous interpretation of microarray experiments. Although details for particular experiments may differ, the MIAME objective is to define common elements among most experiments such as information on experiment and Array Design, samples, protocols and measurements. MIAME is not a formal specification, but a set of guidelines. For more information about MIAME 1.1, refer to the website: <http://www.mged.org/miame>

caBIG Silver-Level Compatibility — This term refers to the National Cancer Institute Bioinformatics (NCICB) Cancer Biomedical Informatics Grid (caBIG) Compatibility Guidelines. caBIG is a common, extensible informatics platform that integrates diverse data types and supports interoperable analytic tools. This platform allows research groups access to the extensible collection of emerging cancer research data, while supporting their individual research. For more information about caBIG, refer to the website: <https://cabig.nci.nih.gov/>. caBIG-Silver is a compatibility reference indicating that the software being used to submit and share cancer-related data meets or exceeds the Silver level of compliance. A description of the various caBIG guidelines and compliance levels such as bronze, silver and gold can be found at this website: https://cabig.nci.nih.gov/workspaces/VCDE/Documents/caBIG_Compatibility.pdf.

caArray Overview

Security Structure Reword to reflect 2.0

caArray Workflow

Security Structure Reword to reflect 2.0

caArray is designed to provide maximum flexibility in the collaborative sharing of microarray data, while allowing users and data managers to make choices regarding who can access the data submitted. Three levels of security and visibility are available in caArray:

- **Public** permissions or visibility gives anyone the ability to view the data, but with Read-Only access. The only person who can modify Public data is the owner of the data. Once your microarray experiment has been published, you may choose to make your experimental data public; anyone can view the data, but after publishing, ownership of the data is moved to the NCICB. Only the NCICB data curator can modify the data at that point.
- **Private** access or visibility restricts access to the data, allowing only the owner the ability to view the data.

Example: if you are in the initial information-gathering stage of a microarray experiment, you may choose to make the information available to only yourself.

Note: In caArray, if no visibility is selected when you are submitting or modifying an object, the default visibility is “Private”.

- **Group-level/Consortium** access or visibility allows only specified individuals or groups, as specified by the owner to view the microarray data. Groups have “read-only” privileges with exceptions granted to persons assigned “curation”

status; they are allowed “modify” privileges. Curation status must be requested through NCICB Application Support.

Example: If you are working on a collaborative effort with colleagues in your lab as well as with investigators at a different research facility, you may choose to make your specified data available to a group/consortium; the consortium has read-only access to the referenced objects. Only you and others in the group specified as Curators can modify the data.

Visibility can be designated with multiple choices. For example, if you assign Public visibility as well as a specific group for an object, the object is available for viewing to everyone, but only the members of the group assigned curation privileges can modify the object. The advantage to this arrangement is that a user with appropriate permissions can search for and display all objects assigned specified curation privileges.

caArray Overview

caArray will look to perform intra-platform, intra-experiment queries for local adopters, via the NCICB instance and potentially across the grid collectively to increase the availability of quality data.

caArray Standards

caArray Workflow

caArray Workflow

Important! Remove entire section?

After going through the use cases, I believe it would still be helpful to have a workflow section. I've left in some of the text. Could this be reworked to reflect 2.0, illustrating a suggested workflow? Perhaps just expand on the class diagram on page 8?

I would be willing to work on coming up with a workflow, but Brent, I'd like to talk it through with you first. Or if you come up with a workflow, I could draft the illustration.

Because there are dependencies between models of information in caArray that comprise an Experiment, it is important that you create objects in caArray in the order discussed and displayed in this section. [shown in Figure 1.2](#) If you do not do so, as you create an Experiment, you may find yourself thwarted by lack of a previously defined data element, and required to abandon partially filled out screens to work backwards.

Follow the numbered steps to create the listed objects. [ChaptersSections](#) in this caArray [User's Guide](#) online help correspond to each of these headings:

As you define the details of all of these components of a microarray experiment, associations between objects are identified and relationships are established. Then an Experiment object is created, incorporating the associated objects that comprise the experiment. These components of an Experiment are then used to annotate microarray

hybridization data and raw data files that are uploaded and parsed into the caArray database.

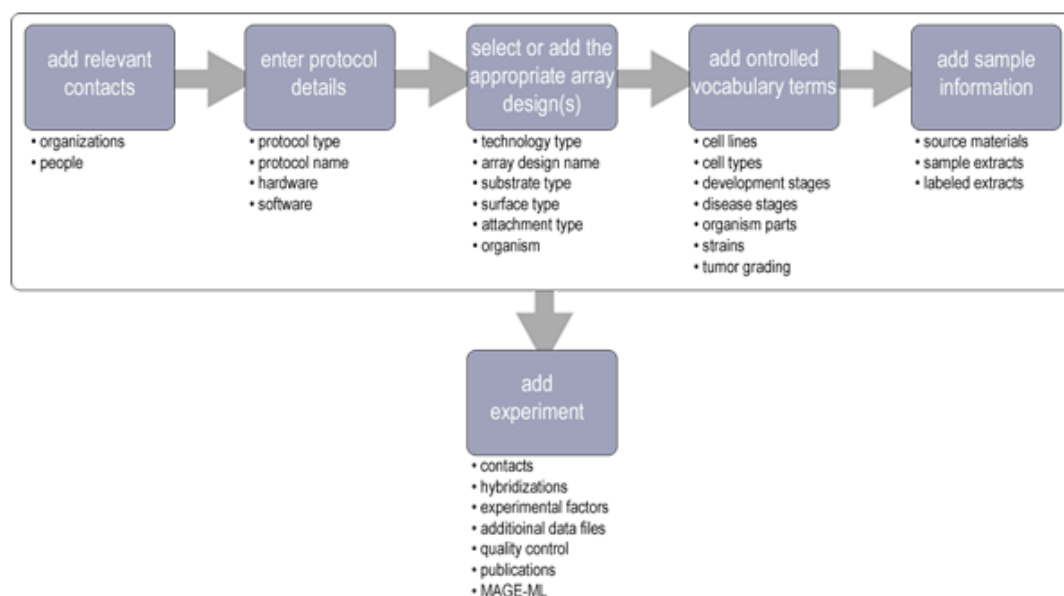


Figure 1.2 caArray Workflow. Each step in the workflow is described in its text and also serves as a [hypertext link](#) to a corresponding caArray user interface. Some steps in the workflow have dependencies on previous tasks. They are described, where appropriate, in this user's guide.

caArray Overview

caArray will look to perform intra-platform, intra-experiment queries for local adopters, via the NCICB instance and potentially across the grid collectively to increase the availability of quality data.

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 16
- *Using caArray Online Help* on page 18
- *Navigating the caArray User Interface* on page 19

Launching caArray

Note: caArray 2.0 supports the following browsers:

- PC's running Microsoft XP: Firefox 2.0+ and Internet Explorer 6.0+
- Apple Macs running OS X: ([TBD](#)) and Firefox 2.0+

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 caArray 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](#)).

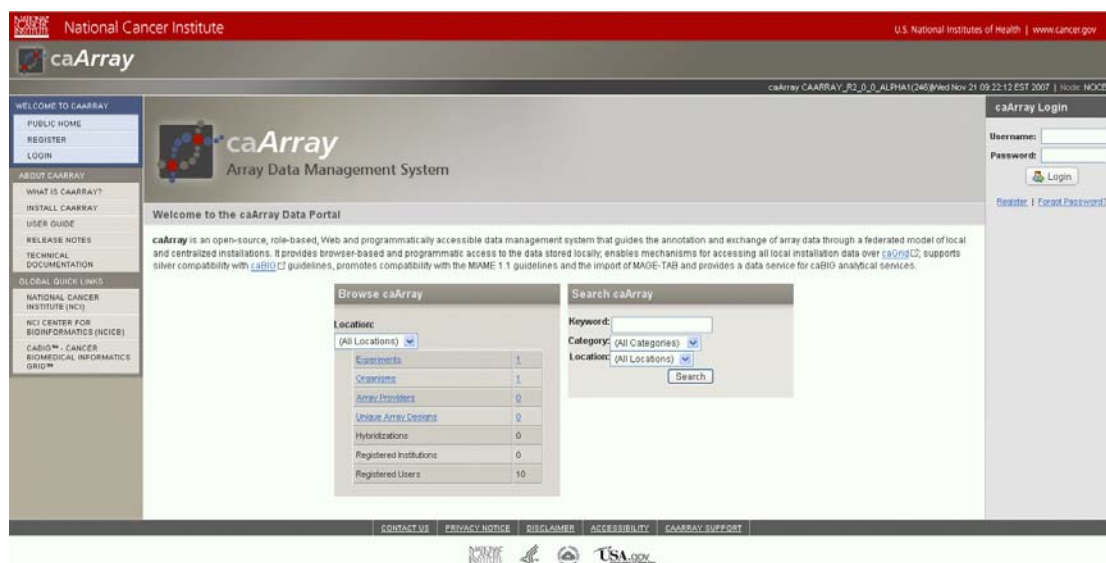


Figure 2.1 Prior to login, you can use these dialogs to browse or search the caArray database

The Browse caArray dialog box on the left center of the page lists the number of objects saved in the database, accessible to a public audience 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. For more information about browsing the caArray database, see *Browsing caArray* on page 43.

The Search caArray dialog on the right center of the page allows you to launch a search of the caArray database for public objects. For more information about executing a caArray search, see *Searching the Repository* on page 46.

After login, I can see no way to launch these features.

Menu options for launching searches are more extensive if you log in, because then you will have access to all objects to which you have been granted visibility. 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 16.

The caArray Interface

Elements of the caArray User Interface

caArray New User Account and Login

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

Figure 2.2 caArray login page

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

To log in, follow these steps:

1. Navigate to the caArray home page. Use the URL to the NCICB instance <http://array.nci.nih.gov> ([Figure 2.2](#)), 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 19.

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 2.3*).

Figure 2.3 New user account registration form

3. In the Become a caArray User form, enter the appropriate information¹.
 - **Security Information**
 - **Do you have an LCDAP account?**
If Yes, enter your username and password for the purposes of verifying that it is correct. If validated, you will be granted limited access to the application immediately as a “caArray User”; this will still require confirmation from the NCICB support teams.
If No then what?
 - **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 [Table 2.1](#). Any limits when person submitting their own request?. If what you said is true about permissions in 2.0, ("all roles (except Anon. User) mirroring capabilities of PI") then can every user do everything? Visibility/accessibility would just be set per experiment?

| Role | Description | Permissible Actions |
|-----------------------------|---|---|
| System Administrator | Person responsible for the effective operation of caArray | Manages users Manages user groups <u>Provides for overall management and quality of caArray ??</u> |
| Principal Investigator [PI] | Owns experiments and studies and/or projects | Submit data Review of experiment designs Submission of quality control metadata Upload of normalized data Annotation of parameters used |
| Lab Administrator | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Lab Scientist | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Biostatistician | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |

Table 2.1 caArray role descriptions

- Account Details
 - First Name*
 - Middle Initial
 - Last Name*
 - Email [address]*
 - Organization*
 - Address [Lines 1* and 2]
 - City Why is this not a required field?
 - State No longer listed--is this an oversight?
 - Postal [or Zip] Code Why is this not a required field?
 - Country*--Select from the drop-down list
 - Phone*
 - Fax

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

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

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

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

Using caArray Online Help




Verify this section after online help is implemented.

The 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 or menu option on the top menu bar or on left sidebar or both?? to open the complete online help project.

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

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

Note: caArray 2.0 does not have context sensitive help.

Navigating the caArray User Interface

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

Elements in the caArray User Interface

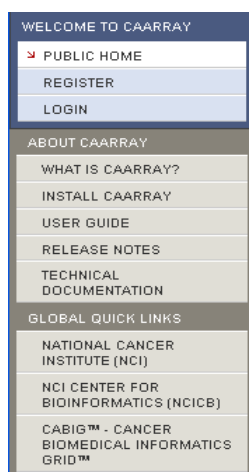
The terms in [the following table Table 2.2](#) identify and describe elements of the caArray user interface...

| Term | Definition |
|--|--|
| Left Vertical Navigation Menu | Hypertext links associated with the caArray application, caArray documentation and Global Quick Links. |
| Online help | caArray documentation accessed from your computer screen providing help and answers for questions about using the software. How implemented? Note: caArray 2.0 does not have context sensitive online help. You can open online help and use the TOC, index or perform a text search. |
| Browse caArray only before login? | The Browse dialog lists database object categories and the number of public objects in each categories. Click each hypertext link to browse details of the listed objects. |
| Search {caArray database objects} only before login? | The Search caArray dialog box allows you to launch a search of public objects from the Welcome page. Enter keywords or select a category or location to define the search at a basic level. Click the Advanced Search hypertext link to add further criteria to the search. For more information about caArray searches, see <i>Searching the Repository</i> on page 46. |
| Task tabs | Located across the top of some of the caArray user interface windows. Task tabs represent the tasks/annotations you create as components of an experiment |
| My Experiments/ Public Experiments | caArray task tabs. These comprise sets of Experiment information and annotations performed on or associated with gene expression or SNP microarrays. |

Table 2.2 *Elements of the caArray browser window*

caArray Welcome Page Navigation Menu

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



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

- **Public Home**--the Welcome/Login page If you click this after logging in, it appears to log you out without asking for confirmation. Is this what you intended?
- **Register**--the form for registering a new account
- **Login**--the Welcome/Login page If you click this after logging in, it appears to log you out without asking for confirmation. Is this what you intended?

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

- **What is caArray**--description
- **caArray Installation Guide**--opens the installation guide pdf
- **caArray 2.0 User's Guide**--opens the user's guide pdf
- **Release Notes**--opens release notes for caArray 2.0
- **Technical Documentation**--link to the caArray Technical Guide
- I think a "caArray Online Help" link should be added here.

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/>)
- **NCICB support** (<http://ncicbsupport.nci.nih.gov/sw/>) opens to a caArray project page; left in because workspace should probably have link to app support, though it currently does not. NCICB Application Support subjects (with active

hypertext links) display in the left-hand panel of the browser window (*Figure 2.4*).

National Cancer Institute

Center for Bioinformatics

HOME ABOUT NCICB INFRASTRUCTURE TOOLS PROJECTS DOWNLOADS TRAINING SUPPORT

Support

Welcome to the NCICB Application Support Site

The NCICB Application Support Group is dedicated to helping you use NCICB tools and resources to maximum benefit. We're here to help you find what you need, use and re-use our software, integrate data and applications, and facilitate the sharing of information and resources.

Telephone: 301-451-4384 or toll free: 888-478-4423
Email: ncicb@pop.nci.nih.gov

Telephone support is available Monday to Friday, 8 am – 8 pm Eastern Time, excluding government holidays. You may leave a message, send an email or submit a support request via the Web at any time.

When submitting support requests via email, please include:

- ♦ your contact information
- ♦ the name of the application/tool you are using
- ♦ the URL if it is a Web-based application
- ♦ a description of the problem and steps to recreate it
- ♦ the text of any error messages you have received

For more information on a specific application, follow the links below, or go to the [Supported Applications List](#) for a detailed breakout of all supported applications.

NCICB User Applications Manual

The NCICB User Applications Manual provides a thorough introduction to NCICB applications.

Web Self-Service

[Search FAQs](#)

Search our dynamic knowledgebase of Frequently Asked Questions (FAQs).

CONTENTS

| | |
|--|---|
| Cancer Centralized Clinical Database (C3D) | Centra |
| caArray | Clinical Research Information Exchange (CRIX) |
| caDSR Applications | DCP Enterprise System Knowledgebase (DESK) |
| Cancer Images Database (caIMAGE) | Enterprise Vocabulary Services (EVS) |
| Cancer Laboratory Information Management System (caLIMS) | Programming & API Support |
| Cancer Model Organisms Database (caMOD) | |

Figure 2.4 An example of the NCICB Quick Link help topic, in this case NCICB Application Support page.

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 4 Navigating and Searching caArray* and *Chapter 5 Creating and Managing Experiments* in this guide.

Upon login, in the user interface the 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)'. A button 'Add a New Experiment' is on the right. Below the tabs, it says 'Displaying 1-4 of 4 Total.' and 'Page 1 < Back | Next >'. The main table has columns: Experiment Title, Array Type, Organism, Disease State, Status, Date Created, # Samples, Properties, Edit, and Delete. The table contains four rows of experiment data.

| Experiment Title | Array Type | Organism | Disease State | Status | Date Created | # Samples | Properties | Edit | Delete |
|---|------------|----------|-------------------------------------|--------|--------------|---------------------|------------|------|--------|
| armst-affy-arabi-316282 | Affymetrix | Mouse | Acute Lymphoblastic Leukemia, Adult | Draft | 09/23/2007 | 121 | | | |
| arnol-affy-chick-445892 | Affymetrix | Human | Cancer | Draft | 10/01/2007 | 121 | | | |
| freij-affy-human-91666 | Affymetrix | Mouse | Brain Stem Glioma, Childhood | Draft | 09/23/2007 | 121 | | | |
| prosc-affy-mouse-450860 | Affymetrix | Human | Cancer | Draft | 10/01/2007 | 121 | | | |

Page 1 < Back | Next >

Figure 2.5 caArray Experiment Workspace

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)

The workspace displays three tabs:

1. **Work Queue**--proposed experiments in process of being created. This tab is now gone. Does it display only if user is in the process of creating an Experiment?
2. **My Experiments** - Proposed experiments that have been submitted for review
3. **Public Experiments** - Experiments that have been submitted and accepted with public visibility into the caArray database. For more information about visibility, see [cross ref.](#)

Will there be other tabs for people given private access to other Experiments? If so, what will they be called?

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 information for each corresponding experiment.

- **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 **Completed**
- **Permissions**--The accessibility given the corresponding Experiment. Click the icon to assign or modify the Experiment Permissions.
- **Edit**--For Experiments with the appropriate permissions, allows the user to edit the Experiment. What happens if user does not have permissions--no icon, different icon, or nothing happens when user clicks it?

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

Online help and additional resource links remain fixed through all the viewing windows.

CHAPTER 3

USER ACCOUNT MANAGEMENT

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

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

Topics in this [sectionchapter](#) include:

- [Roles in caArray on this page](#)
- [Searching for a User on page 32](#)
- [Creating a New User Account on page 33](#)
- [Managing a Group on page 38](#)
- [Creating a Group on page 37](#)
- [Granting Others Access to Your Work on page 41](#)
- [Visibility on page 41](#)

Initial questions

Do I need to address anything here about the System Admin needing to be set up to perform these functions?

The UC doc says “To execute this use case (Manage Users), a system admin. is logged in and can manage accounts of a given institution,...” This sounds like only the SA can perform these management tasks--or is it anyone with Admin permissions?

The UC doc indicates the installer determines whether “they want to use LDAP or the db authorization and authentication method”. If that is all determined ahead of time for a system, perhaps I should add qualifications when discussing filling out the new user form. I’d like to talk through this LDAP authorization so I understand it thoroughly. I

don't think we should assume that an admin user will automatically know about this for caArray.

Roles in caArray

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

| Role | Description | Permissible Actions |
|--|--|---|
| Anonymous User Note: Not listed under Roles | User without a caArray account or a non-logged in user | Browse and Search tasks |
| System Administrator | Person responsible for the effective operation of caArray | Manages users Manages user groups <u>Provides for overall management and quality of caArray ??</u> |
| Principal Investigator [PI] | Owns experiments and studies and/or projects | Submit data Review of experiment designs Submission of quality control metadata Upload of normalized data Annotation of parameters used |
| Lab Administrator | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Lab Scientist | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Biostatistician | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| External System Note: Not listed under Roles | Systems other than caArray from which caArray data can be extracted programmatically using an API. | For more informations, see <i>Extracting Data Programmatically by API</i> on page 78. |

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

Permissions

This seems to now be obsolete for 2.0. Revise it? Also, not sure where it belongs. The

| caArray | Roles (Actors) |
|---------|----------------|
|---------|----------------|

Table 1:

| | | Anonymous User | Principal Investigator | Research Scientist | Lab Administrator | Lab Scientist | Biostatistician | System Administrator | Collaborator | System Time | External System |
|-----------|---------------------------------|----------------|------------------------|--------------------|-------------------|---------------|-----------------|----------------------|--------------|-------------|-----------------|
| Use Cases | Acquire Experiment Data via API | | | | | | | | | | |
| | Manage Experiment Data Files | | | | | | | | | | |
| | Import Experiment Data | | | | | | | | | | |
| | Submit Experiment Proposal | | | | | | | | | | |
| | Download Experiment Data | | | | | | | | | | |
| | Manage Experiment Permissions | | | | | | | | | | |
| | Import Array Designs | | | | | | | | | | |
| | Login | | | | | | | | | | |
| | Manage Users | | | | | | | | | | |
| | Navigate Repository | | | | | | | | | | |
| | Search Repository | | | | | | | | | | |
| | Search Using Filters | | | | | | | | | | |
| | Manage Collaborators | | | | | | | | | | |
| | Register for an Account | | | | | | | | | | |
| | Report on Work Accomplished | | | | | | | | | | |
| | Update Experiment Proposal | | | | | | | | | | |
| | Import Data Files | | | | | | | | | | |
| | Validate Experiment Data Files | | | | | | | | | | |

Table 1:

information is important.

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

Managing user accounts includes creating, updating, resetting and disabling user accounts. All registered users can manage their own account data but not their role.

System admins can be assigned to "logical institutions" by CBIIT admins. What does this mean??

Registering a User Account

Covered in the Getting Started chapter. This has a brief explanation and a cross-reference to the other chapter. A person who wants to be caArray user can request a user account from the Welcome/Login page of caArray (see *Requesting a User Account* on page 16). A new user can also be added by a caArray user with Admin privileges. This section covers creating a new user by Admin user.

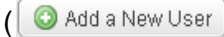
Before creating a new user account in caArray, you should perform a search to ascertain whether the user already has an account. I can't see an intuitive way to launch such a search. Is this true in 2.0?? Or just review names in user account table? For more information, see *Searching for a User* on page 32.

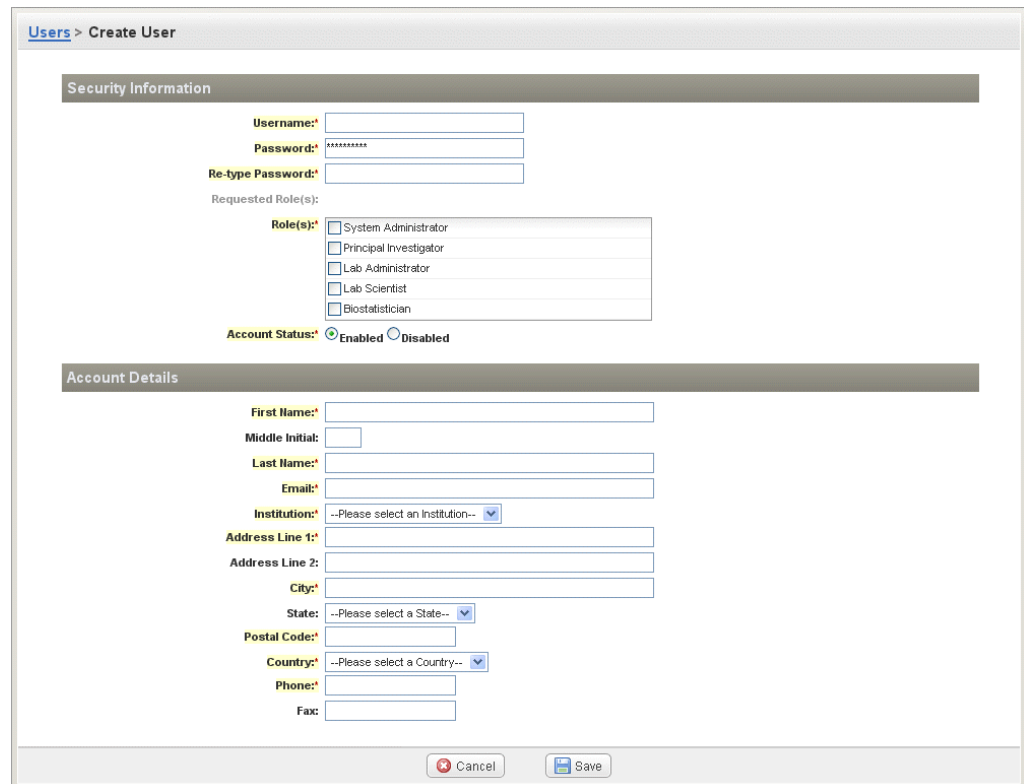
Creating a New User--this was written from the description of Registering a New User in the Getting Started chapter because...see next red text.

It is advisable, before creating a new user account in caArray, to perform a search to ascertain whether the user already has an account. I can't see an intuitive way to launch such a search. Is any or all of this paragraph true in 2.0??

To create a new user account, as a user with Admin privileges, follow these steps:

1. Log in on caArray login page <http://array.nci.nih.gov>.
2. On the Welcome to caArray page, click **Manage Users** in the left sidebar. This option is no longer there. Can an admin add a new user? Looks like not.
No longer can see a way to open this page. The Users page that opens displays a listing of current users organized by last name, first name, organization, email address, account status (Enabled or Disabled) and the user's ability to Edit an Account. For more information about editing, see *Editing User Accounts by an Administrator* on page 31.

3. In the upper right corner of the page, select the **Add a New User** button (). This opens the user account registration form (*Figure 3.1*). =



4. In the Become a caArray User form, enter the appropriate information².

◦ Security Information

— **Do you have an LCDAP account?**

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

If No then what?

— **Username***

— **Password***

- **Requested role(s)*** --Select one or more of the roles. Roles are described in Table 3.1. Any limits when submitting this for a new user?. If what you said this week is true about permissions in 2.0, (“all roles

2. Items with an asterisk or highlighted are required.

(except Anon. User) mirroring capabilities of PI") then can every user do everything? Visibility/accessibility would just be set per experiment?

| Role | Description | Permissible Actions |
|-----------------------------|---|---|
| System Administrator | Person responsible for the effective operation of caArray | Manages users Manages user groups <u>Provides for overall management and quality of caArray ??</u> |
| Principal Investigator [PI] | Owns experiments and studies and/or projects | Submit data Review of experiment designs Submission of quality control metadata Upload of normalized data Annotation of parameters used |
| Lab Administrator | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Lab Scientist | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |
| Biostatistician | Same as PI in caArray 2.0 | Same as PI in caArray 2.0 |

Table 3.1 caArray role descriptions

- Account Details
 - First Name*
 - Middle Initial
 - Last Name*
 - Email [address]*
 - Organization*
 - Address [Lines 1* and 2]
 - City Why is this not a required field?
 - State No longer listed--is this an oversight?
 - Postal [or Zip] Code Why is this not a required field?
 - Country* --Select from the drop-down list
 - Phone*
 - Fax

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

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

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

6. Once all required information is entered, click **Save** Does Save submit the request?, or click **Cancel** to cancel the registration.

caArray validates the new user submission (including the LSAP, the username/ password uniqueness) and opens a confirmation page containing all data entered displays . Does this happen when new user requests or when Sys Admin requests, or both? Does this happens at same time system validates LSAP info? At this point, you as the administrator verify the information and submit it to the system. Is there a Submit button on that confirmation screen? Or how does SA submit?

caArray adds the new user and sends two distinct emails both to the new user?:


1. An email containing an introduction message containing the username, a description of the role(s) assigned the user and a link to the application's URL.
2. An email containing the password assigned to the new user. Can a new user reset the password?

When the account is registered, the UserID and password assigned to the user determines access rights for the software, based on roles assigned by the Admin User who created the account.

Editing User Accounts by an Administrator

Note: The options for performing user management tasks are visible in caArray after login, only if you have these Admin privileges. Is this still true?? I can't see any way to launch this, so I couldn't check it out. What will one see?

To modify the details for a registered caArray user, with Admin permissions, follow these steps:

1. Log in on caArray login page <http://array.nci.nih.gov>.
2. On the Welcome to caArray page, click **Manage Users** in the left sidebar.
3. Click the user name for the user of interest. The name is a hypertext link that opens the user profile.
4. Click the **Edit** button () at the bottom of the page. This opens the user details page.

The following fields are editable: In the forms that I opened, it looked like everything was editable. In 11/21 tag, none of this was available.

- a. Username and Password (The password is double entry and only available to Database management. This is not available for users with LDAP Authentication.) Are you saying the Editor cannot reset the password by a person with Admin permission?
- b. Current assigned role
- c. Organization Information
- d. Enable or Disable the account
5. Modify the editable information appropriately.
6. Click **Save**, and confirm the action at the prompt.

CaArray commits the change, and sends an email notification to the user of the edited account regarding the edit.

When a user without Admin privileges logs in, does he see the Users and Groups left sidebar menu option?

Editing One's Own Account

Could see no way of doing this in 11/21 tag. Perhaps does not belong in Admin chapter, but after Registration/Login section of the Getting Started chapter.

I'm confused about how and where a basic user launches this process. I take it that it's a different process than that launched through the User Management feature.

During the Proposal entry process or using the separate feature of a registered user's list of functions--where are these?, the ability to edit organization or username/password characteristics can be performed.

To edit your own user account, follow these steps:

1. Select **Edit Registration**. Where does this appear?
2. Some of the information that displays is not editable, Editable information is indicated by ???
 - a. For Database Authentication: you can edit the password (double entry)
 - b. For LDAP Authentication: neither username nor password are shown nor are they editable
 - c. Current and Desired Role(s) are editable (UC doc says only by SA)
 - d. Organization information is editable
3. Make the appropriate edits in the information.
4. Click WHAT?? to submit the changes or click **Cancel** to abort the edit process.
5. Verify the information in the confirmation screen that opens.
6. Confirm the changes by clicking WHAT?.

caArray commits the changes and sends an email notification to user of the edited account. It also notifies the installation system administrator, and displays a status of action in caArray?

Searching for a User

Viewing User Details

Creating a New User Account

Managing a Group

Viewing Group Details

Creating a Group

Managing a Group

Searching for a User

Can see no way to launch this in 11/21 tag.

In caArray 2.0, searching for a user involves the Filter feature on the Manage User's page. You can filter the user list based on organization, role, account status, or last name. If you submit the filter request with all defaults, all users are presented.

- Notes
- This search feature is a part of User Management, accessible only to Admin Users. If you do not have these privileges, you cannot see the options and features described in this section. For more information, see *User Management* on page 26.
 - This user search is to locate a person with an existing user account in caArray. This search is completely different than a Person search, described in *Contacts* on page 33, which searches for persons identified with key roles in Protocols, Arrays, Experiments, and so forth. How about these bullets?? Is any or all of this information true in 2.0??

To search for a user in caArray, follow these steps:

1. Click the **Manage Users** option on the left sidebar in the caArray portal.
2. The User form that opens (*Figure 3.1*), displays all valid caArray users by Last Name, Organization, and email address.

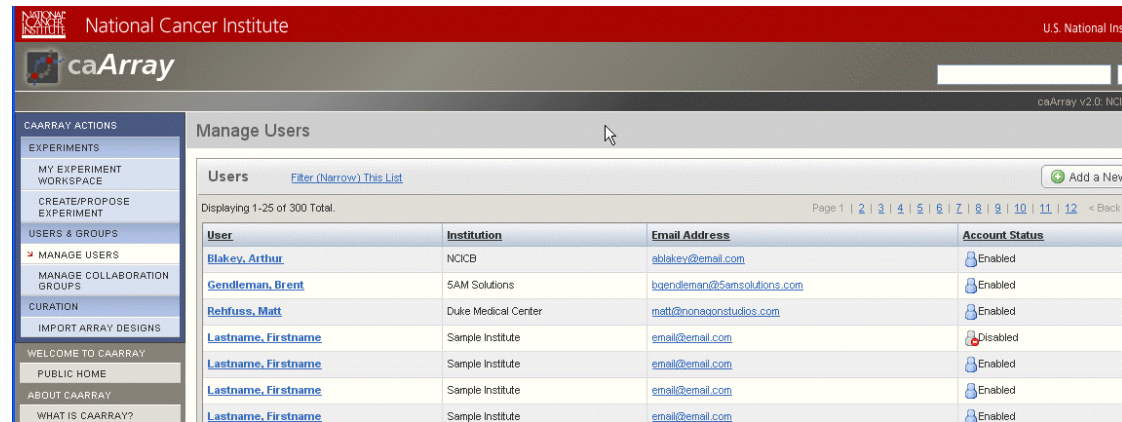


Figure 3.1 Users page

3. You can scroll through the list to find a user
- or
4. Click the **Filter (Narrow) This List** hypertext link at the top of the page.

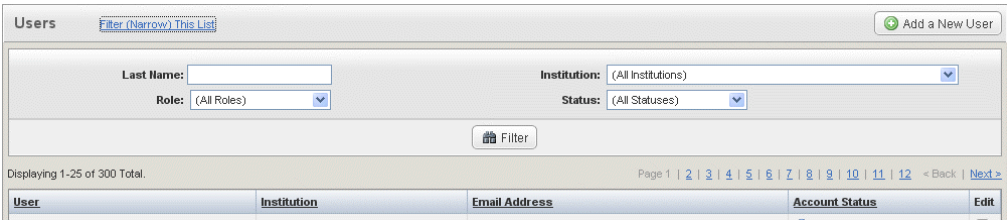



Figure 3.2 Filter users section of the Users page


5. In the new section of the page that opens (*Figure 3.2*), enter or select the following fields that identify the user for whom you are searching:
 - **Last Name**

- **Role**
- **Institution**
- **Email address**

Note: The search is case insensitive; wild cards are always implied on both sides of the query string. [True?](#)

6. Click the **Filter** button  **Filter** .

Note: By clicking the **Filter** button without setting parameters, the search is unrestricted; all Users then display on the **Users** page.

Search results display on the same page. [For more information see User Search Results](#). If you do not locate the User you want, you can click the **Add a New User** button ( Add a New User) on the upper right of the page to create a new user account. For more information, see [Creating a New User Account](#) on page 33.

[User Search Results](#)

[Viewing User Details](#)


[Creating a New User Account](#)

[Disabling a User Account](#)

[Activating a User](#)

User Search Results

User search results display a list of users meeting the search criteria on the Users page. If no criteria were entered, a list of all the users in the system is returned. The initial list is sorted in ascending order of users' last names. All columns in the search results are sortable. [Both last 2 statements true?](#)

- To view details for a filtered result, click the hypertext link in the **User** column.
- To send an email to the user, click the corresponding email address.
- If the appropriate user is not found in the search, you as an Admin User can create a new user by clicking the **Add a New User** button  Add a New User in the upper right corner of the page. For more information, see [Creating a New User Account](#) on page 33.

[Searching for a User](#)

[Viewing User Details](#)

[Disabling a User Account](#)

[Activating a User](#)

Viewing User Details

From a list of users on the User page, a logged in Admin User can view any user details. To do so, click the hypertext link in the **User** column, [as described in the User Search Results](#).

The User Details page displays user information in read-only format ([Figure 3.3](#)).

[Users](#) > Arthur Blakey

Security Information

Authentication Method: LDAP Account [NCICB]
 Username: ablakey
 Password: *****
 Requested Role(s): Principal Investigator
 Current Role(s): Principal Investigator
 Account Status: Enabled

Account Details

First Name: Arthur
 Middle Initial:
 Last Name: Blakey
 Email: ablakey@email.com
 Institution: NCICB
 Address Line 1: 2617 Wooster ST
 Address Line 2:
 City: New York
 State: NY
 Postal Code: 33322
 Country: United States
 Phone: 111-222-3333
 Fax:

Figure 3.3 *User details page*

7. Click the **Back** key in your browser to close the page.

Searching for a User

User Search Results

Activating a UserRequesting an Account Through NCICB Application Support

Creating a New User Account

Disabling a User Account

Activating a UserRequesting an Account Through NCICB Application Support

Viewing User Details

Activating a UserRequesting an Account Through NCICB Application Support

Disabling a User Account

Activating a User

Requesting an Account Through NCICB Application Support

Is this still possible in 2.0? To request a caArray user account from NCICB Application Support, complete one of the following steps:

1. Contact NCICB Application Support:
 - NCICB@pop.nci.nih.gov
 - Phone at 888-478-4423 (toll-free) or 301-451-4384 (local)

OR

2. Go to the NCICB caArray login page <https://caarraydb.nci.nih.gov/caarray/> and click the **Register** button. Enter the requested information, as displayed in

Figure 3.4 and itemized as follows below the figure. Follow the screen prompts to complete the registration.

Figure 3.4 User account registry form


3. You must provide the following information in order to obtain a caArray account³:
 - *Institution
 - *First and Last Names
 - *Complete Address, including country
 - *Telephone number
 - *Email address

Disabling a User Account

An Admin User can disable a caArray user account. This does not remove the underlying record, but after performing this task, the user account status will display on the Modify User Details page as disabled. true??

To disable a user, follow these steps:

1. Perform a search to locate the user. For more information, see *Searching for a User* on page 32.
-
3. Items with an asterisk are required fields.

2. Click the user name link in the search results to open the User Details page.
3. Click the **Edit** button () in the lower page center to open the User Details page.
4. Click the **Disabled** button next to **Account Status**.
5. Click **Save**. Upon a successful save, the user account is disabled immediately.

Searching for a User

User Search Results

Viewing User Details

Activating a User


Creating a New User Account

Activating a User

Activating a User

An Admin User can activate a caArray user account. The user account is noted on the Modify User Details page as **Active**.

To activate a user, follow these steps:

1. Perform a search to locate the user. For more information, see *Searching for a User* on page 32.
2. Click the user name link in the search results to open the User Details page.
3. Click the **Edit** button () to open the User details page.
4. Click the **Enabled** button next to Account Status.
5. Click **Save**. Upon a successful save, the user account is immediately activated, indicated on the User Details page by ???.

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 (define i.e. the user who created the object, project, samples, etc??) to a project, set of projects or particular samples and their underlying array data. To be assigned to a collaboration group, a user must have a caArray user account.

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

Move somewhere else: A good rule of thumb to follow is if you can see an **Edit** button or **Delete** option, then you can modify or delete an object. If you cannot see those options, then you cannot edit or delete the object. Is this true in 2.0??

Note: To add a collaborator (user) to a group, the user must have a valid caArray user account.

Managing a Group

[Group Search Results](#)

[Creating a Group](#)

[Viewing Group Details](#)

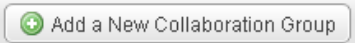

[Managing a Group](#)

Creating a Group

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

If you are an Admin User, before creating a new group in caArray, you should perform a search to ascertain whether the group already exists. For more information, see [Managing a Group](#) on page 38. Still true in 2.0??

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

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

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

[Group Search Results](#)

[Viewing Group Details](#)

[Modifying Collaboration Group Details](#)

[Creating a Group](#)

[Managing a Group](#)

Searching for a Group


Could see no way to launch such a search in 11/21 tag.

This search feature is a part of User Management, accessible only to Admin Users.

To ascertain whether a particular collaboration group already exists in caArray, follow these steps:

1. Click the **Manage Collaboration Groups** option on the left sidebar, under **Users & Groups**.
2. On the Collaboration Groups page that opens (*Figure 3.5*), click the hypertext link, **Filter (Narrow) this List**.
3. In the new section of the page that opens, enter the name of the group for which you are searching, the group owner's last name or a group member's last name.

Note: The search is case insensitive; wild cards are always implied on both sides of the query string. True?

4. Click the **Filter** button ( Filter).



| Manage Collaboration Groups | | | |
|---------------------------------|---|---|---|
| Collaboration Groups | | Filter (Narrow) This List | Add a New Collaboration Group |
| Displaying 1-3 of 3 Total. | | Page 1 < Back Next > | |
| Collaboration Group Name | Group Members | Edit | Delete |
| Collaborators A | Arthur Blakey, Brent Gendleman, Matt Rehfuess, Firstname Lastname, Firstname Lastname (View All 65) |  |  |
| Collaborators B | Arthur Blakey, Phillip Joe Jones, Jonathan Bonham | | |
| Collaborators C | Joseph Addai, Marion Barber, Ronnie Brown, LaMont Jordan, LaDainian Tomlinson (View All 35) | | |
| Displaying 1-3 of 3 Total. | | Page 1 < Back Next > | |

Figure 3.5 Search Results display on the same page as the Search Groups form

Search results display on the same page. Anything else I should add about how they display? For more information see *Group Search Results*. If you do not locate the group you want, you can click the **Add a New Collaboration Group** hyperlink on the upper right corner of the page to create a new group.

Group Search Results

Viewing Group Details

Creating a Group

Managing a Group

Group Search Results

Group search results display below the Search Groups form, displaying a list of groups meeting the filter criteria (*Figure 3.5*). Anything else I should add about how they display? If no criteria were entered, a list of all the groups in the system is returned. The initial list is sorted in ascending order of group names. Columns are sortable.

To view details for a search hit, click the hypertext link in the **User Name** or **Login Name** columns corresponding to the result.

If the appropriate group is not found in the search, you as an Admin User can create a new group by clicking the **New Group** link on the left navigation sidebar. For more information, see [Creating a Group](#) on page 37.

[Managing a Group](#)


[Viewing Group Details](#)

[Managing a Group](#)

Viewing Group Details

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

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

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

[Managing a Group](#)

[Group Search Results](#)


[Creating a Group](#)

[Modifying Collaboration Group Details](#)



[Managing a Group](#)

Modifying Collaboration Group Details

To modify collaboration group details, follow these steps:

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

4. To perform these edits, do the following: I may change this format later. Please check for accuracy though.

| | |
|-----------------------------|--|
| Edit the group name | Enter new name in the Group Name text box. Note: The new member must already have a valid caArray user account. |
| Add a new group member | Note: The new member must already have a valid caArray user account. Click the Add a New Group Member button at the top right of the page. In the new section of the page that opens, you can search for the group member using one or more criteria. <u>true??</u> Enter the last name, select the Role category, the Institution, and the Status to be searched. Click the Filter button. |
| Remove a group member | On the Collaboration Group page, in the Remove column, click the icon () that corresponds to the group member. |
| 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. To edit <i>individual</i> user details, you must return to the Manage Users feature. (See <i>Editing User Accounts by an Administrator</i> on page 31) |
| Delete the group | To delete the entire Collaboration Group, click the Delete button ( Delete) at the bottom center of the page. |

Managing a Group

Group Search Results

Creating a Group

Viewing Group Details

Managing a Group

CHAPTER 4

NAVIGATING AND SEARCHING CAARRAY

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

Topics in this [sectionchapter](#) include:

- [Browsing caArray on this page](#)
- [Searching the Repository](#) on page 46

Browsing caArray

In caArray, you can browse the application, a feature that can be launched before you login or by a non-registered “Anonymous User”. [What about after login?](#)

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, accessible to a public audience 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 4.1](#)).

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

From this page, you can also launch a search of the caArray repository. For more information, see *Searching the Repository* on page 46. .

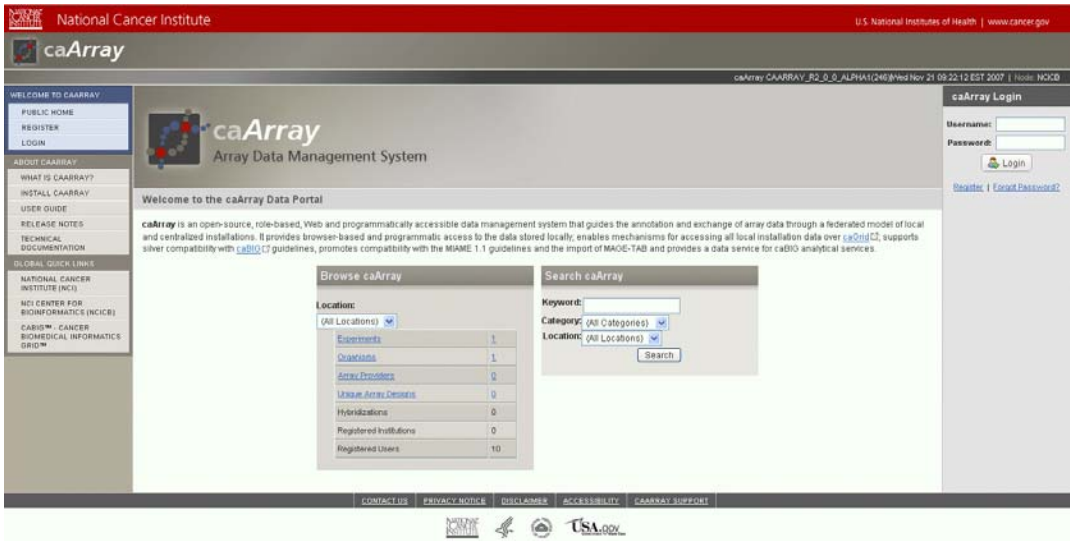


Figure 4.1 Prior to login, you can use these dialogs to browse or search the caArray database

To use the browse feature, follow these steps:

1. Click any of the blue hypertext link in the Experiment properties category list of the Browse dialog box. The page that opens depends on the category you selected.

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

Table 4.1 Browse dialog box categories

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

Table 4.1 Browse dialog box categories

Note: The ability to change the location for browsing is deferred to a future release of caArray.

- Once the tab or page opens when you click any of these categories ([Table 4.1](#)), the same metadata displays on all pages for the list of Experiments located for that category ([Table 4.2](#)). Most of these properties were identified when the Experiment was created or edited.

| | |
|------------------------|---|
| | |
| Experiment ID | The autogenerated identification assigned by caArray. Click the hypertext link to open the corresponding experiment tabs which contain all current Experiment information. Only public display? Or private display as well, but no links? |
| ID | The Experiment title defined manually, naming and/or briefly describing the Experiment |
| Assay Type | <ul style="list-style-type: none"> Gene Expression: experiment using microarrays intended to measure levels of transcribed genes SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA |
| Primary Contact | 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. or anyone else for that matter. |
| Organism | The organism that is the source of the sample biomaterials used in the Experiment |

| | |
|----------------------|--|
| | |
| Disease State | The disease state of the source materials used in the Experiment |
| Samples | The number of samples identified in the Experiment. Click the hypertext link to open the Experiment to the Samples details page. |
| Updated | The date of the most recent update of the Experiment draft |

Table 4.2 Experiment metadata categories

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

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

Searching the Repository

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

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

Searches can be initiated by entering criteria of two different types

1. Basic searches that search a limited set of often-referenced values,
2. Advanced criteria such as field-specific searches based on specific values and ranges of values (such as certain genes with a minimal intensity value, or specific probe sets, etc.). [Oh, really?? How enter these criteria?](#)

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

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

Search caArray

Keyword:

Category: (All Categories) ▼

Location: (All Locations) ▼

2. Define the search criteria by using the search options described in the following table [Table 4.3](#):

| Search Option | Description |
|---------------|---|
| Category | Select one of the Experiment properties categories listed. Only experiments in the category selected will be searched. If you do not choose a category, All Categories (default) remains selected, and caArray will search all Experiments. |
| Keyword | Enter one or more words, separated by spaces. . Example: breast cancer. Note: No logic statements, such as AND or OR or SQL statements are supported in these search features. <u>Case insensitive; wild cards implied on both sides of the query string??</u> |
| Location | Search by location is not implemented in caArray 2.0 |

Table 4.3 Search criteria options

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

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

Experiment Search Results

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

Search results are listed in table format, with columns displaying properties for each Experiment; fields are described in [Table 4.4](#). Most of these properties were identified when the Experiment was created or edited.

| <u>Experiment ID</u> | The autogenerated identification assigned by caArray. Click the hypertext link to open the corresponding experiment tabs which contain all current Experiment information. <u>Only public display? Or private display as well, but no links?</u> |
|-----------------------------|---|
| ID | The Experiment title defined manually, naming and/or briefly describing the Experiment |
| <u>Assay Type</u> | <ul style="list-style-type: none"> • Gene Expression: experiment using microarrays intended to measure levels of transcribed genes • SNP: experiment using microarrays intended to detect nucleotide changes in chromosomal DNA |
| Primary Contact | The person named as the Point of Contact for the Experiment. Click the hypertext link or the envelope icon (✉) to open an email form where you can draft an email to this contact, if named. <u>or anyone else for that matter.</u> |

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

Table 4.4 Experiment metadata categories

The details for categories underlined in [Table 4.4](#) can be sorted by clicking column headings. caArray paginates the result sets in groups of 25.

Underlining.

Can anonymous user see only public Experiments or even private ones that cannot be opened?

You can open any Experiment to which your assigned permissions grant you access. Private Experiments can be listed in the results, but can be viewed only if you have appropriate permissions. Only the ability to contact PI 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 PI for the Experiment, or with appropriate permissions, edit it or extract.

- For information about editing an Experiment, see *Editing an Experiment* on page 64.
- For information about contacting the Experiment PI, see *Primary Contact* in [Table 4.4](#).
- For information about extracting data from an Experiment, see *Downloading Data from caArray* on page 77.

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 5

CREATING AND MANAGING EXPERIMENTS

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

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

Overview of an Experiment on this page

Managing Experiment Visibility on page 65

Updating an Experiment proposal on page 67

Overview of an Experiment

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

The basic elements of a caArray Experiment are described in [the following table Table 5.1](#):

| Term | Definition |
|----------------------------|--|
| Experiment Overview | Basic information about an experiment such as IDs, service and assay types, provider of array and array designs, source of biomaterials, and disease state |
| Contacts | Principal Investigator and/or Point of Contact for the Experiment |

Table 5.1 *Elements of a caArray Experiment*

| Term | Definition |
|---------------------|--|
| Annotations | Experimental factors and design, biomaterials used, hybridizations |
| Data | Experimental data files uploaded and imported into caArray; imported data available for extraction from caArray; Supplemental data |
| Publications | Publications associated with the Experiment, primarily, journal articles |

Table 5.1 *Elements of a caArray Experiment*

Adding an Experiment in caArray involves two main segments:]

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

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

Move these somewhere else:.

| | |
|-------------------------------|---|
| Organisms | Select from the drop-down menu the organism that is the source of the sample biomaterial |
| Array Providers | <u>The company that manufactured the array used in the Experiment</u> |
| Unique Array Designs | Descriptions of microarrays that include such information as array layout and design and the technology used to create a microarray. |
| Hybridizations | The process of incubating one or more labeled extracts with an array. Hybridization methods form the basis for microarray techniques. |
| Registered Institution | <u>Cancer center, academic institution, laboratory or commercial vendor with an established caArray account??</u> |
| Registered Users | <u>Person with a validated caArray User Account. May or may not be associated with a Registered Institution.</u> |

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.

- On the left sidebar, click **Create/Propose Experiment**. This opens a tab for entering overall characteristics for the Experiment.



Figure 5.1 Create/Propose Experiment on left sidebar

Overview Tab

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

To complete the Overview tab, follow these steps:

- On the Overview tab, enter the appropriate information for Overall Experiment Characteristics¹ as described in the [the following table Table 5.2](#).

Experiment: New Experiment

Overview

Overall Experiment Characteristics

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

Experiment Title*:

Status: Draft

Experiment Identifier:

Service Type*: --Select a Service Type--

Assay Type*: --Select an Assay Type--

Provider*: --Select a Provider--

Array Designs:

Organism*: --Select an Organism--

Tissue Sites*: Filter: Add

- Leg
- Brain
- Cell Line
- Feed

Selected Tissue Sites

Tissue Types*: Filter: Add

- whole_organism
- cell lysate
- organism_part
- DNA
- total_RNA

Selected Tissue Types

Cell Types: Filter: Add

- Chromaphine Cells
- Brain Tissue

Selected Cell Types

Conditions: Filter: Add

- Adrenocortical Carcinoma

Selected Conditions

Save

| Overview Tab Field | Description |
|-----------------------|--|
| Experiment Title | Title designate by the PI or other user creating the Experiment |
| Status | Draft or Completed? |
| Experiment Identifier | This project Identifier is autogenerated by caArray upon the initial save of the Experiment. The Experiment Identifier is not editable. The ID is generated using the PIs last name followed by a 5 character number. Example: klemm-90765. The Experiment ID is geneally a hypertext link that opens the Experiment. |
| Service Type | <p>Service Type is <u>describe..</u>. Select from the drop-down menu the appropriate service type. <u>What happens when these options are selected? I see in UC with Publish, caArray messages PI that "submission to a lab is removed."</u></p> <p>Options are: <u>descriptions of each</u></p> <ul style="list-style-type: none"> • Full <p>Note: Full Service Experiments are placed in Public workspaces</p> <p><u>How are private indicated??</u></p> <ul style="list-style-type: none"> • Publish--for bulk <u>upload of data typically after a significant amount of data has been collected without caArray. It is intended to provide a means to annotate this data but the higher priority is to make data available to the caArray community and give opportunity for further annotation as time permits. Turns off data validation feature.</u> <p>Note: Publish Service Experiments remain in the PI's Experiment queue for making public when desired.</p> |
| Assay Type | <p>Select from the drop-down menu the appropriate assay type.</p> <p>Options are the following:</p> <ul style="list-style-type: none"> • Gene Expression--experiment using microarrays intended to measure levels of transcribed genes • SNP--experiment using microarrays intended to detect nucleotide changes in chromosomal DNA |
| Provider | <p>Select from the drop-down menu the provider of the array.</p> <p>Note: Only Affymetrix , Illumina and GenePix are supported in caArray 2.0.</p> <p>Once selected, caArray automatically loads a corresponding list of Array Designs (next field)</p> |
| Array Designs | Select one or more array design from the automatically-generated list of array designs corresponding to the Provider you selected. |
| Organism | Select from the drop-down menu the organism that is the source of the sample biomaterial |

Table 5.2 Fields for Overall Experiment Characteristics

1. Fields with a red asterisk * are required.

| Overview Tab Field | Description |
|---------------------|---|
| Tissue Sites | Click the source(s) of the tissue(s) used in the array. If the appropriate site is not displayed, to find a term of interest, begin typing a term in the Filter test box; the selections in the <u>box</u> will be limited according to the text you enter. To add a new term, click Add. This take you to the Manage Vocabularies page where you can do so. <u>relationship to adding in Filter text box or adding in Manage Vocabularies section.</u> <u>Jill, add Manage Vocabularies section for adding terms to these 4 fields. ??</u> |
| Tissue Types | Click the tissue type used in the array. If the appropriate site is not displayed, to find a term of interest, begin typing a term in the Filter test box; the selections in the <u>box</u> will be limited according to the text you enter. To add a new term, click Add. This take you to the Manage Vocabularies page where you can do so. |
| Cell Types | Click the type of cells used in the array. If the appropriate cell type is not displayed, to find a term of interest, begin typing a term in the Filter test box; the selections in the <u>box</u> will be limited according to the text you enter. To add a new term, click Add . This take you to the Manage Vocabularies page where you can do so. |
| Conditions | <u>Do you mean type of cancer? That's what it looks like from some samples in the app.</u> |

Table 5.2 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. 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 5.3](#). How is this information “made available automatically.” (from UC doc)? There are no drop down menus for selecting pre-entered information.

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

Table 5.3 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 for each of the seven subtabs.

Experimental Design

1. Enter on the Experimental Design subtab (under the Annotations Tab) the appropriate information as described in [the following table Table 5.4](#)¹.

| Experimental Design Fields | Description |
|---|---|
| Experimental Design Type* | Select the Experimental Design type among the options in the drop-down menu. |
| Experimental Design | <u>Is this field for a description? Then say so (like the next field).</u> |
| Experimental Design Quality Control Description | Enter a description for the Quality Control used for the Experiment. |
| Quality Control Types | Select the QC type in the scroll down list. |
| Replicate Description | If there are replicates used in the experiment, describe the number of replications and how the replicates were generated. <u>Define in context of caArray?</u> |
| Replicate Types | Select one or more Replicate Types from the scroll down list. |

Table 5.4 Experimental Design fields

1. Fields with a red asterisk * are required.

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

Experimental Factors

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

1. Click the **Add a New Experimental Factor** button to add a new Experimental Factor.
2. In the form that opens, enter a name for the new factor. This is a required field, as indicated by the red asterisk^{*1}.
3. Click **Save**. The name of the factor now displays on the Experimental Factors subtab.
4. Repeat steps 1 & 2 as often as needed to enter all the Experimental Factors for this Experiment.
5. Click **Save**.

How is the category (field for a saved Exp. Factor) determined or specified?

6. Proceed to the Sources subtab.

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 [Chapter 9 Controlled Vocabulary Terms](#).

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 5.2](#) 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.

1. Fields with a red asterisk * are required.

- **Labeled Extracts** are Extracts that have been labeled for detection of the nucleic acids. *Example:* The Extract RNA is labeled with a fluorescent dye, yielding the labeled extract RNA.

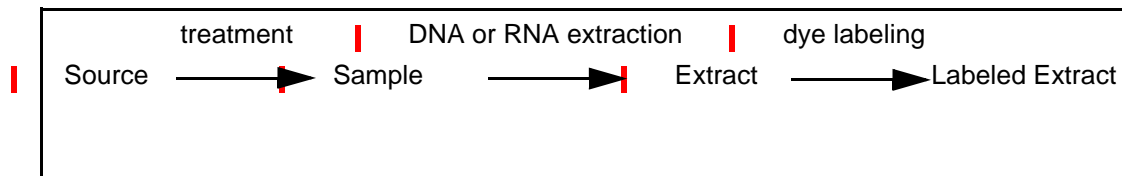


Figure 5.2 Biomaterials components and their relationship in caArray

The nature of each type of biomaterial is defined in varying dialog boxes in caArray. The relevant fields for each is described in [the associated topics Table 5.5, Table 5.6, and Table 5.7](#).

Sources

Samples

Extracts

Labeled Extracts

Sources

The term Source refers specifically to original source material before it has been subjected to any treatment. The nature of Sources in caArray is described in [the following table Table 5.5¹](#).

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

Table 5.5 Fields for documenting Sources

Samples

Extracts

Labeled Extracts

1. Items with an asterisk are required fields.

Samples

The term Sample refers specifically to original source material after it has been initially treated. The nature of Samples in caArray is described in fields in [the following table Table 5.6](#)¹:

| Sample Fields | | |
|---------------|--|------------------------------|
| Field | Description | Example |
| Sample Name | Name given to the Sample Note: It is a good idea to clearly name Samples | |
| Description | Description of the Sample. | |
| *Organism | The species from which the Sample was obtained | <i>Homo sapiens</i> |
| Tissue Site | The name of the tissue site from which the Sample was obtained | NIH Clinical Center |
| Source(s) | The name of the Source from which the Sample was derived | Lung Biopsy tissue--PatientX |
| Extract(s) | Extract(s) that have been identified in caArray as being derived from this Sample | Private |

Table 5.6 *Fields for documenting Samples*

Sources

Extracts

Labeled Extracts

Extracts

An Extract represents the nucleic acid extracted from a Sample for use in microarray experiments. The nature of Extracts in caArray is described in fields in [the following table Table 5.7](#):

| Extract Fields | | |
|-------------------|--|---------------------|
| Field | Description | Example |
| Extract Name | Name given to the Extract | |
| Description | Description of the Extract | cDNA, total RNA |
| Nucleic Acid Type | The type of nucleic acid extract: DNA or RNA <u>cDNA??</u> | |
| Related Sample(s) | The name of the Sample from which the Extract was derived | <i>Homo sapiens</i> |

Table 5.7 *Fields for documenting Samples*

1. Items with an asterisk are required fields when adding a new Sample. See *Sources Tab* on page 58.

| Extract Fields | | |
|------------------|--|---------|
| Field | Description | Example |
| Labeled Extracts | Labeled Extract(s) that have been identified in caArray as being derived from this Extract | |

Table 5.7 *Fields for documenting Samples*

Sources

Samples

Labeled Extracts

Labeled Extracts

A Labeled Extract is an Extract that has been labeled with signal emitters such as fluorescent materials for nucleic acid detection. Labeled Extracts are detected by automated equipment. The nature of Labeled Extracts in caArray is described in the fields in [the following table Table 5.8](#):

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

Table 5.8 *Fields for documenting Labeled Extracts*

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

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



Sources Tab

To add a Source, follow these steps:

1. On the Sources tab, click the **Add a new Source** button in the upper right-hand corner of the tab.
2. In the Sources form that opens, enter the information as described in the following table [Table 5.9](#).

| Source Fields | Description |
|---------------|---|
| Source Name | Name assigned to the source |
| Description | Description of the source |
| Tissue Site | Site where the source tissue was derived. Select or enter a new value in the Filter text box and click Add . |

Table 5.9 Fields for documenting a Source

3. Click **Save** to save the Source to the Experiment.
 - Once the draft is saved, the Organism you specified when you began creating the Experiment displays in the **Organism** field.
 - If you have already specified Samples for this Experiment, you specified the Source for the Samples. They now display in the **Related Samples** column.
4. Repeat steps 1-3 as often as necessary to add all appropriate Sources to the Experiment.
5. To copy the Source, click the Copy icon () in the Copy column. caArray copies the Source attributes, renames it using the existing source name and adding an incremental number. The copied source now displays under the existing one. To edit the new source name, click the hypertext Name link, click the Edit button () and rename the Source. You can recopy the Source as many times as you like.

Note: To delete a Source, click the Delete icon () in the corresponding row.

6. Proceed to the [Samples Tab](#) [Samples Tab](#).

Samples Tab

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

1. 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 5.10](#).

| Samples Fields | Description |
|----------------|-----------------------------|
| Sample Name | Name assigned to the sample |
| Description | Description of the sample |

Table 5.10 Fields for documenting Samples



| Samples Fields | Description |
|-----------------------|---|
| Source(s) | Click the Source of the material used in the array, if it displays in the Source list. If the appropriate source is not displayed, to find a term of interest, begin typing a term in the Filter test box; the selections in the <u>box</u> will be limited according to the text you enter. To add a new term, click <u>Add</u> . This takes you to the Manage Vocabularies page where you can do so. |

Table 5.10 Fields for documenting Samples

3. Click **Save**.

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

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

4. Repeat steps 1 - 3 as often as needed to enter all the Samples used in this Experiment.
5. To copy the Sample, click the **Copy** icon () in the Copy column. caArray copies the Sample attributes, renames it using the existing Sample name and adding an incremental number. The copied Sample now displays under the existing one. To edit the new Sample name, click the hypertext **Name** link, click the **Edit** button ( Edit) and rename the Sample. You can recopy the Sample as many times as you like.

Note: To delete a Sample, click the Delete icon () in the corresponding row.

6. Proceed to the **Extracts tab** *Extracts Tab*.

Extracts Tab

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

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



1. Click the **Add a New Extract** button to add a new Extract.
2. In the Extract form that opens, enter the information described in *the following table Table 5.11*.

| Extracts Fields | Description |
|------------------------|------------------------------|
| Extract Name* | Name assigned to the Extract |
| Description | Description of the extract |

Table 5.11 Fields for documenting an Extract

| Extracts Fields | Description |
|------------------------|--|
| Sample* | Sample from which the Extract was derived. Select or enter a new value in the Filter text box and click Add . |

Table 5.11 Fields for documenting an Extract

- Click **Save**. The name of the Extract now display on the Extracts subtab.
- Note:** When other data about the new Extract has already been entered in this Experiment, the fields that display on the Extracts tab will be populated - **Nucleic Acid Type**, and **Labeled Extracts derived from the Extracts**. If those fields are not currently populated, they will become such when the appropriate information is entered.
- Repeat steps 1 - 3 as often as needed to enter all the Extracts used in this Experiment.
 - To copy the Extract, click the **Copy** icon () in the Copy column. caArray copies the Extract attributes, renames it using the existing Extract name and adding an incremental number. The copied Extract now displays under the existing one. To edit the new Extract name, click the hypertext **Name** link, click the **Edit** button ( Edit) and rename the Extract. You can recopy the Extract as many times as you like.

Note: To delete an Extract, click the **Delete** icon () in the corresponding row.

- Proceed to the [Labeled Extracts Tab](#)**Labeled Extracts tab**.

Labeled Extracts Tab

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

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



- Click the **Add a New Labeled Extract** button to add a new Labeled Extract.
- In the Labeled Extract form that opens, enter the information described in [the following table Table 5.12](#) .

| Labeled Extracts Fields | Description |
|--------------------------------|--|
| Labeled Extract Name* | Name assigned to the Extract |
| Description | Description of the extract |
| Extracts* | Extract from which the Labeled Extract was derived. Select or enter a new value in the Filter text box and click Add . <u>??</u> <u>no Add button.</u> |

Table 5.12 Fields for documenting a Labeled Extract

- Click **Save**. The name of the Labeled Extract now display on the Extracts subtab.

Note: When other data about the new Labeled Extract have already been entered in this Experiment, the fields that display on the Extracts tab will be populated - Labels Used, Amount, Protocol. NEXT BUILD If those fields are not currently populated, they will become such when the appropriate information is entered.

- Repeat steps 1 - 3 as often as needed to enter all the Extracts used in this Experiment.
- To copy the Extract, click the **Copy** icon () in the Copy column. caArray copies the Extract attributes, renames it using the existing Extract name and adding an incremental number. The copied Extract now displays under the existing one. To edit the new Extract name, click the hypertext **Name** link, click the **Edit** button () and rename the Extract. You can recopy the Extract as many times as you like.

Note: To delete a Labeled Extract, click the **Delete** icon () in the corresponding row.

- Proceed to the **Hybridizations Tab**.

Hybridizations Tab

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

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

- Click the **Add a New Hybridization** button to add a new Hybridization.
- In the Hybridizations form that opens, enter the information described in the following table *Table 5.13*.

| <i>Hybridizations Fields</i> | <i>Description</i> |
|------------------------------|--|
| Hybridization Name* | Name assigned to the Extract |
| Labeled Extract* | Labeled Extract used in the hybridization protocol. Select or enter a new value in the Filter text box and click Add . <u>?? no Add button.</u> |

Table 5.13 Fields for documenting a hybridization

- Click **Save**. The name of the Extract now display on the Extracts subtab.

Note: On the saved draft, the uncompressed size of the hybridization file is defined by caArray; this information displays on the Labeled Extract tab.

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

- Repeat steps 1 - 3 as often as needed to enter all the hybridization data used in this Experiment.

- Click **Save**. The name of the Hybridization now displays on the Hybridizations subtab.

Note: To delete a Hybridization, click the **Delete** icon () in the corresponding row.

- Proceed to the Data tab.

Data Tab

The Data tab is the vehicle for uploading, validating, importing and downloading data relating to caArray Experiments. When you click on the Data tab, four subtabs where you initiate data-related tasks display. They are described in [the following table Table 5.14](#).

| Data Tabs | Description |
|--------------------|--------------------|
| Manage Data | |
| Imported Data | Define |
| Supplemental Files | Define |
| Download Data | |

Table 5.14 Tabs for performing data-related tasks

[Flesh out this section.](#)

[For more information about submitting data to an Experiment, see Chapter 6 Submitting Data to an Experiment.](#)

Publications

This 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 reviews, books and journal articles.

To add publication associations, follow these steps:

- Locate the Experiment for which you want to associate publications.
- Click the **Publications** tab.
- Click the **Add a New Publication** button.
- 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 5.15](#)).¹

| Publications Fields | Description |
|----------------------------|----------------------------------|
| Title* | Title of the publication |
| Authors | Author(s) of the publication |
| URL | URL for locating the publication |

Table 5.15 Fields for documenting Publications

- Fields with a red asterisk * are required.

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

Table 5.15 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 and returns you to your Experiment workspace, which is now updated with the state of the project.

Submitting an Experiment

Once you decide the Experiment is completely finished, you can choose to submit it for review. Click the Submit Experiment button. At that point...

Editing an Experiment

At any point, after an Experiment draft has been saved, with appropriate permissions, you can open the draft and edit it.

To edit an Experiment, follow these steps:

- Perform either a browse or search to locate the project you want to edit. For more information, see *Browsing caArray* on page 43 or *Searching the Repository* on page 46.
- On the row corresponding to the Experiment you want to edit, click the Edit button. this opens the Experiment details tabs.

The information that is editable is:

o

- Make the appropriate changes in the Experiment
- Click **Save** to save your changes to the draft.

Adding Controlled Vocabulary Terms

In caArray 2.0, many attributes are available for the entry of terms for creating Experiments and storing data in the caArray repository. ??

Managing Experiment Visibility

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

To assign or modify Experiment visibility, follow these steps:

1. Click on the icon in the Permissions column that corresponds to the Experiment.
2. The Experiment Permissions page that opens displays two visibility categories:
 - a. Browsability--The default setting is "Not Browsable. To change the setting to Browsable, click the Toggle button.
 - b. Policies--The only option here is to click the **Use TCGA Policy** check box. The Cancer Genome Atlas (TCGA) has established a policy in which only 4 fields of annotation can be provided to the open public:
 - Clinical Diagnosis
 - Histologic Diagnosis
 - Tissue Anatomic Site
 - Pathologic Status

If you check the box, the same access standard will apply to your Experiment.Visibility.

Note: At this point, visibility cannot be assigned to Collaboration Groups.

Online Help

Revise: Easily accessible and prominent informational text indicating the distinction between public, institutional and collaboration group access as well as the type of access control available (Read and Read Selective; Edit and Edit Selective; and None) is vital to understanding the nuances of this use case.

Managing Access to caArray Data

This is from previous manual. When I get time to work through it, it will probably be removed.

Requesting Access to caArray Records

Much of the information provided by others is proprietary, and not yet public. The owner of private data determines who has access to proprietary data.

To request access to caArray records submitted or "owned" by existing caArray Contacts, whether persons or organizations, you must contact NCICB Application Support, ncicb@pop.nci.nih.gov. In your contact email, you must include the following information:

- Your name
- Your email address
- Name of the owner of the information you wish to access (person or organization)

- The specific information you wish to access, indicated by accession numbers or other unique identifying information available in the caArray records.
- A concise but complete statement of the reason you wish to be granted access to this information.

Granting Others Access to Your Work

Visibility

Security Structure Reword to reflect 2.0

Granting Others Access to Your Work

You can choose to grant others access to your work so that they can review your experiments, compare experimental design, compare results, etc. To assign access to others, you must contact [NCICB Application Support](#). In your contact email, you need to include all of the following information:

- Your name
- Your email address
- Your telephone number
- The name(s) of the person or persons, and/or the organization or organizations, to which you wish to grant access and the name of the group with which the new users should be associated
- The contact telephone numbers and email addresses for all persons and organizations to be granted access
- The level of access that you wish to grant to each person/group (**Read only, Read and Modify**).

Note: If only certain records are to be accessed, this must be explicitly indicated by accession numbers or other unique information that is part of the caArray records or class of records in question.

Work on redundancy. Don't know if this still applies or not. Put in creating and managing experiments??

Visibility

Security Structure Reword to reflect 2.0

Visibility

In caArray, the term “visibility” is used to describe the access to objects granted to logged in users. *The following table* [Table 5.16](#) describes visibility options and their interpretation.

Note: When specifying visibility, to make multiple selections, press the CTRL key and click on all items you want to select. To deselect a visibility option, press CTRL + click on the item.

| Term | Definition |
|-----------------------|--|
| Public | Anyone can view all records, with Read-Only access. |
| Private | Can be viewed only by submitter ("owner") of object. Private is assigned by default if no other Visibility option(s) are selected when an object is created. |
| Public + Group | Anyone can view all records, with Read-Only access. Modify privileges are assigned to members of the specified group(s) who are assigned "Curation" privileges. Note: The advantage to assigning a group access to data that is already Public, is in being able to isolate objects with that group's visibility. For example, the user (with appropriate visibility) could locate from the Public objects, all Experiments tagged with the group's visibility. |
| Read-Only | Permission only to view records created by other persons or organizations |
| Delete | Permission to delete records created by other persons or organizations |
| Read/Modify | Permission to view and/or modify records created by other persons or organizations |

Table 5.16 *caArray visibility and permission settings*

The visibility of items is relevant when a user performs a search. In the list of search results only items are displayed which are "visible" for the user who performed the search.

Security Structure Reword to reflect 2.0

Work on redundancy. Don't know if this still applies or not. Put in creating and managing experiments??

Granting Others Access to Your Work

From here on, to be reworked:


Updating an Experiment proposal

Once an Experiment proposal has been created in caArray, with the appropriate permissions, you can edit a proposed project to to update a proposal in draft status or to reflect appropriate changes.



Note: To update an Experiment proposal, the Experiment must be in "draft" or "returned for revision--in 2.0?" status.

To update an Experiment proposal, follow these steps:

1. Log in to your available instance of caArray.

- ° The caArray user interface displays the My Experiment workspace, where all Experiments to which you have visibility are listed. Must be logged in to edit? How access Browse and Search when logged in?
 - 2. Browse or search for the Experiment you want to edit. For more information about these functions, see *Browsing caArray* on page 43 or *Searching the Repository* on page 46.
 - 3. When you have located the experiment you want to edit, click the **Edit** icon () on the row that corresponds to the Experiment.
 - 4. On the Experiment pages that open, modify the Experiment information appropriately. Required fields are indicated with a red asterisk*. For more information about the Experiment fields, see *Creating an Experiment* on page 50.
- If an entry is not valid, caArray displays a message indicating type of information needed.
- 5. If you are not finished entering or modifying data for the Experiment, click **Save**. If you have finished added information, click **Submit Experiment Proposal**.
 - 6. How soon does the the Status of an Experiment become visibly updated after submission?

Can import source and sample annotations? When?? or manually enter source and sample annotations. Cannot overwrite or version related to asynchronous importing of annotations.

- **Permissions**--click the icon () to view the permissions assigned to this Experiment draft--
- **Edit**--the Edit icon () indicates the Experiment draft is editable (what indicates not editable?)

Delete--what does gray trash can indicate? How interpret this column?

Project Characteristics

Represent minimum set of attributes required to submit an Exp for review.

| Screen Name | Value | Field Type | Required |
|-----------------------|---|------------|----------|
| Experiment Title | Alpha/Numeric (100 characters) | Text Box | Yes |
| Experiment Identifier | Auto-generated once saved | Hidden | No |
| Payment Mechanism | Grant, Purchase Order, Other | List Box | Yes |
| Payment Number | Alpha/Numeric (100 characters) | Text Box | No |
| Service Type | Full, Partial, Analysis | List Box | Yes |
| Platform Type* | Gene Expression, SNP, DNA aCGH | List Box | Yes |
| Manufacturer | Affymetrix, Agilent, GenePix, Illumina, Imagene, UCSF Spot, Custom* | List Box | No |

| | | | |
|-------------------|--|--------------|-----|
| Array Design | Derived List of Available Designs from Manufacturer | List Box | No |
| Source | List of available Organisms | Multi-Select | Yes |
| Tissue Site | List of Available Terms based on Organism | Multi-Select | Yes |
| Tissue Types | List of Tissue Types (e.g., frozen_tissue_cells) | Multi-Select | Yes |
| Cell Type | List of Available Terms | Multi-Select | No |
| Disease/Condition | List of Available Terms | Multi-Select | Yes |
| Pooled Samples? | Yes/No | Radio | Yes |
| Sample Types | List of Sample Types based on Platform Type (e.g., DNA for SNP, RNA for Gene Expression) | Multi-Select | Yes |

Experimental Design

| Screen Name | Value | Field Type | Required |
|--|--|--|----------|
| Experiment Design Type | List of Available Types (e.g., time series, pharmacogenic) | List Box | Yes |
| Experimental Design | Alpha/Numeric Text (2000 Characters) | Text Area | Yes |
| Experimental Quality Control Description | Alpha/Numeric Text (2000 Characters) | Text Area | No |
| Quality Control Types | List of Available Types (biological replicates, peer review) | List Box | No |
| Replicate Description | Alpha/Numeric Text (2000 Characters) | Text Area | No |
| Replicate Types | List of Available Types (biological or technical replicates) | List Box | No |
| Experimental Factors | Name, Description and Category (list) | Mixed with Multiple Factors able to be added | Yes |

Sources

| Screen Name | Value | Field Type | Required |
|--------------------|---|------------|----------|
| Source Name | Alpha/Numeric Text | Text Box | Yes |
| Source Description | Alpha/Numeric Text (500 characters) | Text Area | No |
| Organism | Displayed from Project Properties | Read-only | N/A |
| Organism Part | Displayed from Project Properties – Unless more than one selected, then display the option(s) | Select | Yes |
| Cell Type | | | |
| Strain | | | |

| | | | |
|-----------------------------|--|--|--|
| Developmental Stage | | | |
| Disease State | | | |
| Genetic Variation | | | |
| Genotype or Mutation Status | | | |
| Age | Numeric value (including 0 and decimal amounts) and a qualifying list (minutes, hours, days, weeks, months, years) | | |

Samples

| Screen Name | Value | Field Type | Required |
|-----------------------------|--|------------|----------|
| Source Name | Alpha/Numeric Text | Text Box | Yes |
| Source Description | Alpha/Numeric Text (500 characters) | Text Area | No |
| Organism | Displayed from Project Properties | Read-only | N/A |
| Organism Part | Displayed from Project Properties – Unless more than one selected, then display the option(s) | Select | Yes |
| Cell Type | | | |
| Strain | | | |
| Developmental Stage | | | |
| Disease State | | | |
| Genetic Variation | | | |
| Genotype or Mutation Status | | | |
| Age | Numeric value (including 0 and decimal amounts) and a qualifying list (minutes, hours, days, weeks, months, years) | | |

CHAPTER 6

SUBMITTING DATA TO AN EXPERIMENT

This [section](#)[chapter](#) describes the processes for submitting data to caArray Experiments by performing these tasks:

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 following topics are part of this chapter:

- [Uploading Annotation and Array Content into a caArray project on this page](#)
- [Validating Data Files](#) on page 73
- [Importing Array Designs](#) on page 74
- [Managing Associations](#) on page 76
- [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 designs?](#)

Managing Data

the ability to import MAGE-TAB files that contain source and sample information is available in 2.0 through the Manage Data Use Case.

Uploading Annotation and Array Content into a caArray project

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

Due to large file size of array data and to lesser extent, the annotation, the upload may take minutes or even hours to complete. So the upload can be run in the background.

Data owner must be logged in. (described previously).

At least one project is available to have data uploaded, validated and imported into it.

Service type must be known (complete partial, or analysis) and assay type (gene expression or SNP).

.zip files can be uploaded.

Files types that can be uploaded?

To upload data files associated with an experiment, follow these steps:

1. Go to the NCICB caArray login page <http://array.nci.nih.gov> and log in.

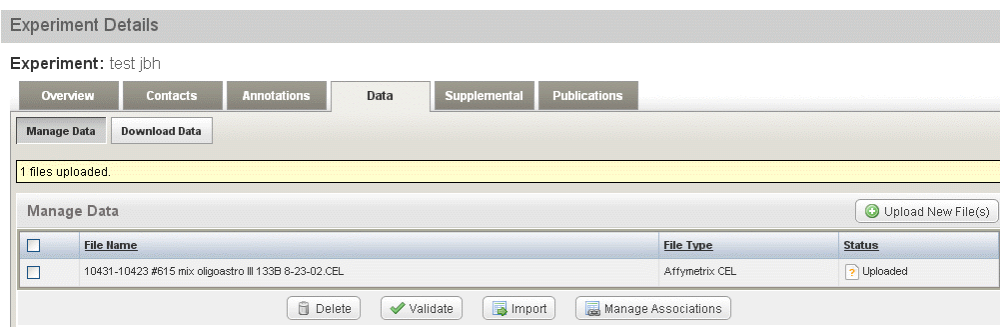
Once you are on the caArray Portal Welcome login page, the browser displays the Experiment workspace.

1. Locate and open the Experiment for which you want to upload files.
2. Select the **Data** tab, and **Manage Data** subtab.
3. Click the **Upload New Files** button.
4. In the form that opens, browse for the file you want to upload and 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.

caArray launches the upload process.

When the upload has finished, caArray displays the list of files, their status (**Uploading**, **Uploaded**, **Validating**, **Validated**, **Importing** and **Imported**), the type of file it is and the ability to select or deselect any one or all of the files to support the following actions: If the file is compressed, the system extracts the file, discards the original and presents each of the files in the zip to the user, indicating that each one has been uploaded. Can upload more than one file at a time? I couldn't select more than one.



The ability to see the total number of files selected for upload and how many have made it numerically will be available – e.g., 1 of 4, 2 of 4, 3 of 4, 4 of 4.

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 delete files.

Deleting a File

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

Deleting a file

Choose file(s) to be deleted, and clicks delete. User confirms the deletion. File is physically deleted. and returns owner to

Validating Data Files

Once files have been uploaded to caArray, they are available for validation and importing into the application for annotation of array data. Content validation does not determine the accuracy of information from scientific viewpoint, but it can ensure that the uploaded data files, even those previously non-validated, are sound for importing into the system.

A user associated with the PI's research lab can perform the validation.

One project with at least one complete file is available to a user. Files that were attempted to be uploaded but incomplete are not available for this task. Type of array (gene expression or SNOP) must be known.

Following file types can be validated:

MGED: IDF, SDRF, ADF

Affymetrix: .dat, .cel, .chp, exp

Illumina: Gene expression and SNP

To validate uploaded data files in caArray, follow these steps:

1. Go to the NCICB caArray login page <http://array.nci.nih.gov> and log in.

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

2. Locate and open the Experiment for which you want to validate files.
3. Select the **Data** tab, and **Manage Data** subtab.

4. 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.

5. From the UC: For files where the type cannot be inferred, the Data Owner is prompted to select the file type before validation can proceed. True?

caArray performs structural and then content validation against each file you have selected, updating the status of each file periodically (10 seconds) until all

files display the validation status in the Status column, Validated or Validation Errors. If validation fails, the file cannot be imported.

A validation error can be structural or content-based. To view a validation error description, click the Status ???.

“Unexpected Error has occurred” when tried to validate a bad file.

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

Should I include a section on MAGE-TAB validation failure (UC p. 6)

Importing Array Designs

Separate chapter??

Important of annotation and array data from previously uploaded files. The import feature makes validated data available for download through the user interface and APIs.

To perform an import, the user must be logged in and have appropriate permissions for data access. The data to be imported must be successfully uploaded and available for import. If the data has not been previously validated, when you launch the import caArray will take the data through the validation process. Once it is successfully validated, caArray will automatically continue with the import.

Files Types that can be Imported into caArray

The following file types can be imported into caArray: (anything in red I'm not sure about)

- Raw/processed data files:
 - Affymetrix .CEL
 - Affymetrix .CHP
 - GenePix .GPR
 - Illumina data files
- Array design files
 - .CDF
 - MAGE-TAB ADF (Array Design Format)
 - Illumina .CSV and .txt
 - GenePix GAL
- Affymetrix .DAT
- Affymetrix .EXP

- Array data file without annotations. Without a current sample, create sample first. User must be able to select sample associated with data. Select multiple samples?
- MAGE-TAB files
 - MAGE-TAB with single SDRF (Sample and Data Relationship Format)
 - MAGE-TAB with multiple SDRFs
 - IDF (Investigation Description Format) only, no referenced SDRFs
 - SDRF with only source > Hybridizations
 - ADF (Array Design Format)
 - MAGE-TAB with existing samples in system. Warn that data already exists for samples. Allow update, no update, or file association only
- Other file types

The following file types support information sharing, but the actual validation is minimal. What does this mean? I uploaded a Word file (.doc) and when I tried to validate, got an error message: “An unexpected error has occurred.”: Word docs, Excel spreadsheets, PowerPoint files and pdfs. They can be uploaded and imported, though the only validation is that they are present. I got the same error message when I tried importing the Word doc. These files are not associated with any samples, but are associated at an Experiment level.

After a successful import, data contained in the files is persisted in caArray. Status of file set is “Imported”.

Service type (complete, partial, or analysis) and type of array (gene expression or SNP) must be known so system knows what type of data it is about to receive. Please clarify this sentence.

To import data, follow these steps:

1. Go to the NCICB caArray login page <http://array.nci.nih.gov> and log in.
Once you are on the caArray Portal Welcome login page, the browser displays the Experiment workspace.
2. Locate and open the Experiment for which you want to import files.
3. Select the **Data** tab, and **Manage Data** subtab.
4. 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 73.
- 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.

5. After a successful import, you can delete the file set. See *Deleting a File* on page 73.

Managing Associations

CHAPTER 7

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 78

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 associated with private data for which (s)he has the appropriate permissions. [UC says only the public data port can be downloaded.](#)

must be available in a project without restriction (public data) or a private project with some public data. Only public data can be downloaded.

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

1. Go to the NCICB caArray login page.
On the caArray Portal Welcome login page, using either the Browse or the Search feature, locate the Experiment of your choice. For more information, see [Browsing caArray](#) on page 43 or [Searching the Repository](#) on page 46.
2. To open the Experiment, click the Experiment ID.
Experiment details display on each of the tabs associated with the Experiment.
3. Click the **Data** tab, then click the **Download Data** subtab. caArray displays a list of all downloadable data by File Name, File Type and file sizes.

Note: You can sort the columns of the list by clicking on the column headers.

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

5. How are the following indicated?

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

With the additional ability to

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

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

Grid availability

Implemented in 2.0?? 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.

Extracting Data Programmatically by API

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

- geWorkBench
- caIntegrator
- WebGenome
- Grid client through a caGrid service

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

Import Data

1. Ability to import source or sample annotation directly into the UI not supported for 2.0. But the ability to import MAGE-TAB files that contain source and sample information is available in 2.0 through manage data UC.

[Experiment characteristics in table 1 in Submit Experiment Proposal UC p. 7.](#)

When PI has experiment of hundreds of disease/conditions, must enter on a sample by sample basis.

Use user-controlled vocabulary for non-NCI Thesaurus values.

A P P E N D I X



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

CAARRAY GLOSSARY

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

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

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

Table B.1 caArray Glossary

INDEX

A

- account, creating new user 16
- activating user 37
- adding
 - Group 38

B

- biomaterials
 - relationship between 56

C

- caArray
 - common uses of 19
 - creating user account 16
 - dependencies between objects 10
 - glossary 83
 - interface 19
 - interface description 19
 - launching 13
 - login 14, 43, 46, 50, 72, 73, 75, 77
 - objectives 5
 - overview 5
 - permissions 9
 - recommended workflow 10
 - requesting access to caArray
 - records 65
 - security structure 9
 - standardized terminologies 8
 - standards 8
 - task tabs 21
 - visibility 9
- caArray User's Guide
 - introduction 1
 - organization 1
 - text conventions 2
- caBIG Silver-Level Compatibility 9
- Consortium visibility, description 9
- creating
 - group 38
 - user's group 38

D

- disabling a user 36

E

- Experiments
 - overview 49

G

- Glossary 83
- group
 - creating 38
 - modifying details page 40
 - search results 39
 - viewing details page 40
- group-level visibility, description 9

H

- help in caArray 18

L

- Labeled Extract
 - definition 56
 - fields, descriptions, examples 58
- launching caArray 13

M

- MIAME 1.1, description 9
- modifying
 - group details 40

N

- NCICB Application Support 35

O

- online help
 - icons 18
 - Quick Links 20
 - using 18

overview, chapters in guide 1

description of permissions in
caArray 9, 66

P

private visibility, description 9

public visibility, description 9

R

records, requesting access to 65

registering new user 16

S

Sample Extract

definition 55

fields, descriptions, examples 57

searching

caArray User 32

from caArray login screen 14, 43, 46,
50, 72, 73, 75, 77

user's groups 39

search results

group 39

user 34

Security in caArray 9

Source Material

fields, descriptions, examples 57

Source Materials

annotations for 55

characteristics 55

definition 55

T

text conventions in user guide 2

U

user

activating a 37

creating account 16

disabling 36

requesting new account from
NCICB 35

searching for 32

search results 34

viewing details 34

user's manual conventions 2

V

viewing

group details 40

user details 34

visibility