

CAAMEL 1.6

User's Guide



Center for Bioinformatics

Revised September 14, 2007

CREDITS AND RESOURCES

caArray Development and Management Teams			
Development	Quality Assurance	Documentation	Project and Product Management
Eric Tavela ²	Ye Wu ⁴	Eric Tavela ²	Mervi Heiskanen ¹
Leslie Power ²	Scott Jiang ⁴	Leslie Power ²	Anand Basu ^{1c}
Steve Matyas ²		Rob Daly ²	Jerry Eads ⁵
Krishna Kanchinadam ²		Jill Hadfield ¹	Rob Daly ²
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 Terrapin Systems
⁴ Science Application International Corporation (SAIC)		⁵ Northern Taiga Ventures, Inc.	

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

TABLE OF CONTENTS

Using the caAMEL Guide	1
Introduction to the caAMEL User’s Guide	1
Organization of this Guide	1
Document Text Conventions	2
Chapter 1	
Getting Started with caAMEL	3
Overview of caAMEL and MAGE-ML	3
Objectives of caAMEL	4
Using caAMEL in a Nutshell	4
Launching caAMEL	5
caAMEL New User Account and Login	5
Getting Help	6
Quick Links	6
Chapter 2	
The caAMEL Workspaces	9
Overview of caAMEL Workspaces	9
Public Workspace	9
User Workspace	10
User Notification by Email	11
Chapter 3	
Working with MAGE-ML Documents	13
MAGE-ML Document Overview	13
caAMEL Workflow	14
Uploading Documents	15
Validating Documents	16
Downloading Documents	17
Deleting Documents	17
Document Sets	17

Creating a Document Set	18
Managing Document Sets	20
Importing Documents and Document Sets into caArray	20
Resolving Errors	22
Metadata Errors	22
Invalid Documents	22
MAGE-ML Document Errors	23
After caArray Import	23
Appendix A	
caArray Glossary	25
Index	27

USING THE CAAMEL GUIDE

This chapter introduces you to the *caAMEL 1.6 User's Guide* and suggests ways you can maximize its use.

Topics in this chapter include:

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

Introduction to the caAMEL User's Guide

The *caAMEL 1.6 User's Guide* is the companion documentation to the caAMEL software application. The *caAMEL User's Guide* includes information and instructions for the end user about using caAMEL.

Organization of this Guide

The *caAMEL 1.6 User's Guide* contains the following chapters:

Chapter 1 Getting Started with caAMEL—This chapter provides an overview of caAMEL, a brief description of MAGE-ML documents, objectives of caAMEL and instructions for launching the application. A workflow for using caAMEL completes this chapter. 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 caAMEL and the rationale for creating caAMEL objects, the Experiments comprised of those objects as well as data files uploaded to those Experiments.

Chapter 2 Navigating the caAMEL Workspaces—This chapter differentiates between the caAMEL workspaces and provides details about navigating the interface and performing basic functions in caAMEL,

Chapter 3 Working with MAGE-ML Documents—This task-based chapter describes functions that can be performed with MAGE-ML documents in the caAMEL workspaces. It includes instructions for uploading, validating, managing and importing documents into caArray.

Appendix A—This appendix includes a glossary for terms relevant to the caAMEL, MAGE-ML and caArray domains.

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

Document Text Conventions

Table 1.1 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
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 1.2 Programmer's Guide</i>
<i>Italic boldface monospace type</i>	Text that you type	In the New Subset text box, enter <i>Proprietary Proteins</i> .
Note:	Highlights a concept of particular interest	Note: This concept is used throughout the installation manual.
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 1.1 caAMEL 1.6 User's Guide text conventions

CHAPTER 1

GETTING STARTED WITH CAAMEL

This chapter outlines the purposes of the caAMEL application and describes MAGE-ML files.

Topics in this chapter include:

- *Overview of caAMEL and MAGE-ML* on this page
- *Objectives of caAMEL* on page 4
- *Using caAMEL in a Nutshell* on page 4
- *Launching caAMEL* on page 5
- *Getting Help* on page 6

Overview of caAMEL and MAGE-ML

caAMEL is a web-based application designed to work exclusively with MAGE-ML documents, uploading, validating and importing them to caArray, thus complementing the functionality of caArray.

The caAMEL project, as well as the caArray project that it is targeted to enhance, is sponsored by the National Cancer Institute Center for Bioinformatics (NCICB). caAMEL, like caArray, is an open source project, and the source code and APIs are available in the download page of the NCICB website <http://ncicb.nci.nih.gov/download/index.jsp>.

MAGE-ML, an acronym for Microarray Gene Expression Markup Language, is a standard data exchange format for microarray data. MAGE-ML uses eXtensible Markup Language (XML), which provides a consistent way to label data. MAGE-ML can be automatically derived from Microarray Gene Expression Object Model (MAGE-OM), which is developed and described in Unified Modeling Language (UML). A broad range of computer programs can process XML-labeled data, making it a valuable tool for both analysis and input. For more information about MAGE-ML, refer to the website: <http://www.mged.org/Workgroups/MAGE/introduction.html>.

Objectives of caAMEL

caAMEL was created to “fill a gap” in caArray 1.3.1, which is unable to perform MAGE-ML imports. Specifically, caAMEL was developed with the following primary objectives:

- To validate MAGE-ML documents
- To load valid MAGE-ML documents into a caArray repository

Using caAMEL in a Nutshell

If you want to import MAGE-ML into caArray, first make sure you are a registered caArray user. See [Registering](#) on page 5 for more information. Use that user account to login to caAMEL. If you are just going to use caAMEL to validate your MAGE-ML, you will not need an account, but can instead just work in the PUBLIC WORKSPACE. See the next section, [Launching caAMEL](#) for more information.

Assuming you either:

- Logged in and are positioned at your user workspace, or
 - Navigated to the PUBLIC WORKSPACE,
1. In the Notification section, indicate whether you would like status notification to be sent to you by email. (This is a system email message that will update you on the status of your document or document set's validation or import after the task has completed.) See [User Notification by Email](#) on page 11 for more information.
 2. You can upload individual MAGE-ML documents in the MAGE-ML Documents section. Once a document is uploaded, you can validate, download, or delete it. Registered users can import it into caArray. The validation process may indicate that there were warnings, though the MAGE-ML was valid (e.g. a field value may be too large to import into caArray and will be truncated when imported). The "Import - Ignore Warnings" button allows you to import documents that generate these validation warnings. See [Uploading Documents](#) on page 15 for more information.

A document set is a group of potentially interrelated MAGE-ML documents. Document sets are validated as a group allowing for references between documents contained in the group. Open a document set to add additional MAGE-ML documents. Editing a document set allows you to change the name or description of the set. Registered users can import document sets to caArray. See [Document Sets](#) on page 17 for more information.

The recommended workflow in caAMEL is as follows:

1. Upload MAGE-ML documents (either as individual documents, or into a document set).
2. Validate the document(s).
3. Review the validation results when the job has completed.
4. Import the MAGE-ML content into caArray (if you are not simply validating the document[s]).

5. Review the import results.
6. Delete the document or document set in caAMEL when all tasks are complete.

Launching caAMEL

The caAMEL application is accessible at <https://caarraydb.nci.nih.gov/caamel/>.

caAMEL New User Account and Login

A user of caAMEL can work in either of two workspaces:

1. **Public Workspace**--The public workspace is available to anyone to validate MAGE-ML data, whether they are logged into caAMEL or simply visiting the site anonymously (without logging in). For more information, see [Public Workspace](#) on page 9.
2. **User Workspace**--If you want to import MAGE-ML documents into caArray, you must be a registered caArray user and currently logged into the caAMEL User Workspace. For more information, see the following section and [User Workspace](#) on page 10.

Registering

caAMEL does not offer its own user registration feature. Instead, if you wish to employ the User Workspace facilities of caAMEL, you must first register as a user of caArray. In fact, registered users of caArray are automatically considered users of caAMEL.

For more information about obtaining an account with caArray, refer to the caArray web page: <https://caarraydb.nci.nih.gov/caarray/>.

Logging Into caAMEL

When you click the URL to access caAMEL, the login page displays ([Figure 1.1](#)).

Figure 1.1 caAMEL login page

From any page of caAMEL, the left sidebar PUBLIC MENU displays **LOGIN**, if you are not currently logged into caAMEL.

- On the caAMEL login page, if you already have a caArray user account, enter the caArray account login information and click **Login**.
- If you do not have a caArray user account, refer to the caArray web page: <https://caarraydb.nci.nih.gov/caarray/> to apply for a user account. Then return to log into caAMEL.

Logging Out

Once logged in, to logout, select the **LOGOUT** option from the USER MENU. You can also log out by closing your browser window.

Getting Help

Each workspace page of caAMEL includes a basic summary that describes the use of that workspace page. If you need further assistance, click the USER'S GUIDE menu option on the left sidebar on any page in caAMEL to open this document. *Figure 1.2* illustrates where the USER'S GUIDE option can be found.

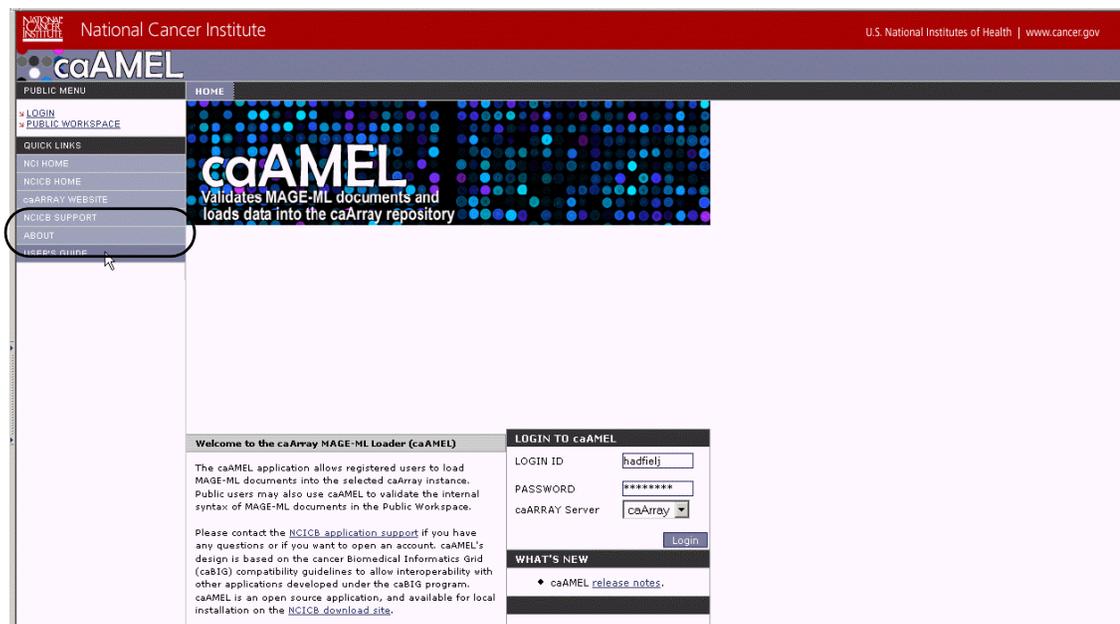


Figure 1.2 The left sidebar Quick Links menu; the USER'S GUIDE option is highlighted.

Quick Links

The **Quick Links** menu options on the left sidebar in most caAMEL viewers provide sources for caAMEL/caArray-related bioinformatics information on the Internet (*Figure 1.2*).

These include:

- **NCI Home** (<http://www.cancer.gov/>) is a link to the home page of the National Cancer Institute.

- **NCICB Home** (<http://ncicb.nci.nih.gov/>) is a link to the NCI Center for Bioinformatics home page.
- **caArray website** (<https://caarraydb.nci.nih.gov/caarray/>) is a link to the caArray portal, where you can access the caArray application and some documentation.
- **NCICB support** <http://ncicbsupport.nci.nih.gov/sw/> opens to the NCICB support page. NCICB Application Support subjects (with active hypertext links) display in the left sidebar of the browser window *Figure 1.3*.

The screenshot shows the NCICB Application Support Site. The header includes the National Cancer Institute logo and the Center for Bioinformatics. A navigation menu contains links for HOME, ABOUT NCICB, INFRASTRUCTURE, TOOLS, PROJECTS, DOWNLOADS, TRAINING, and SUPPORT. The main content area is titled "Support" and includes a welcome message, contact information (Telephone: 301-451-4384 or toll free: 888-478-4423; Email: ncicb@pop.nci.nih.gov), and a list of items to include in support requests. A sidebar on the left lists "RELATED LINKS" and "GLOBAL QUICK LINKS". Two additional boxes on the right promote the "NCICB User Applications Manual" and "Web Self-Service" with a "Search FAQs" button.

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.

RELATED LINKS

- caCORE Documentation
- NCICB Events
- Staff Directory

GLOBAL QUICK LINKS

- caCORE 3.1 now available
- caGrid 0.5 now available
- cancer Biomedical Informatics Grid™ (caBIG™)
- Open Development Initiative
- Mini-guide to this web site
- caCORE FTP site
- NCICB GForge Portal

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).

Figure 1.3 An example of the NCICB Quick Links help topic

CHAPTER 2

THE CAAMEL WORKSPACES

This chapter introduces you to the caAMEL workspaces.

Topics in this chapter include:

- [Overview of caAMEL Workspaces](#) on this page
- [Public Workspace](#) on this page
- [User Workspace](#) on page 10
- [User Notification by Email](#) on page 11

Overview of caAMEL Workspaces

A user of caAMEL can work in either of two workspaces:

- **Public Workspace**--The public workspace is available to anyone, whether they are logged into caAMEL or simply visiting the site anonymously (without logging in). For more information, see the following section.
- **User Workspace**--If you want to import MAGE-ML documents into caArray, you must be a registered caArray user and currently logged into the User Workspace. For more information, see [User Workspace](#) on page 10.

Public Workspace

The Public Workspace in caAMEL provides both registered users and anonymous users who are not logged into caAMEL a means to validate their MAGE-ML files. Both single documents and document sets are supported. Since the workspace is open to the public, you should be aware that other public users will also have access to the same files. If you are concerned about restricting access to view, validate, and potentially download a file, you are encouraged to use the User Workspace instead of the Public Workspace.

From any page in the caAMEL application, you can navigate to the Public Workspace. To do so, click the **PUBLIC WORKSPACE** link under **PUBLIC MENU** in the upper left sidebar (*Figure 2.1*).



Figure 2.1 The PUBLIC WORKSPACE link

From the Public Workspace (*Figure 2.2*) you can upload and validate documents..

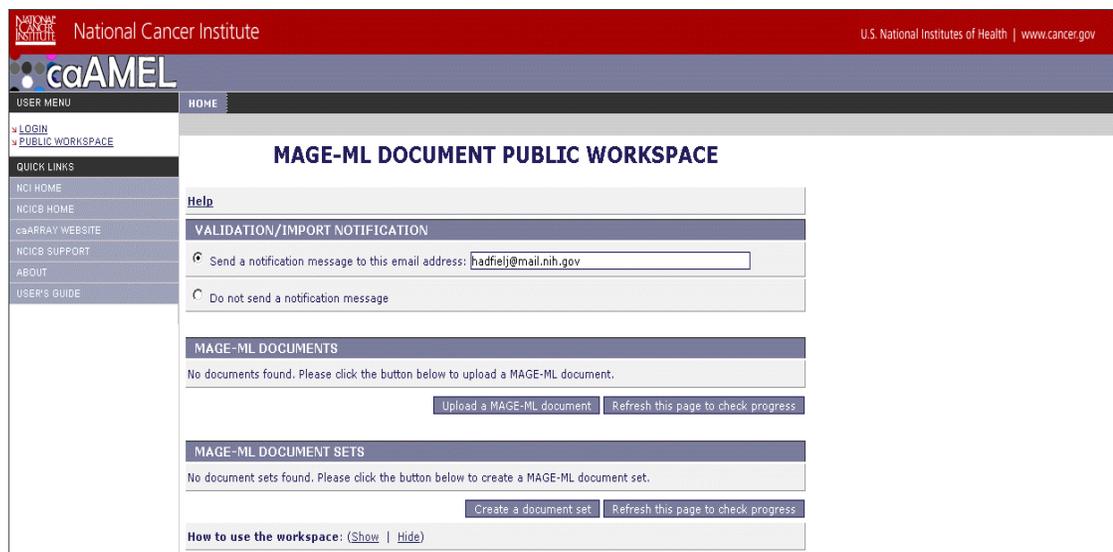


Figure 2.2 The Public Workspace

For more information about working with MAGE-ML documents in the Public Workspace, see *Chapter 3 Working with mage-ml Documents*.

User Workspace

A User Workspace is accessible only to users with a caArray user account who are logged into caAMEL. In most cases in caAMEL, it is recommended that you work in a User Workspace rather than the Public Workspace. This gives you a private place to work on your MAGE-ML documents and document sets. Additionally, you must be logged into caAMEL and working from a User Workspace to perform document imports into caArray.

- For more information about caAMEL login, see [caAMEL New User Account and Login](#) on page 5.
- For more information about working with MAGE-ML documents in the User Workspace, see [Chapter 3 Working with mage-ml Documents](#).

Once you have logged into caAMEL, click the User Workspace link that displays on the left sidebar from any caAMEL viewer window. This always returns you to the main User Workspace page ([Figure 2.3](#)).



Figure 2.3 User Workspace link

From the User Workspace ([Figure 2.4](#)) you can upload, validate and import documents and document sets.

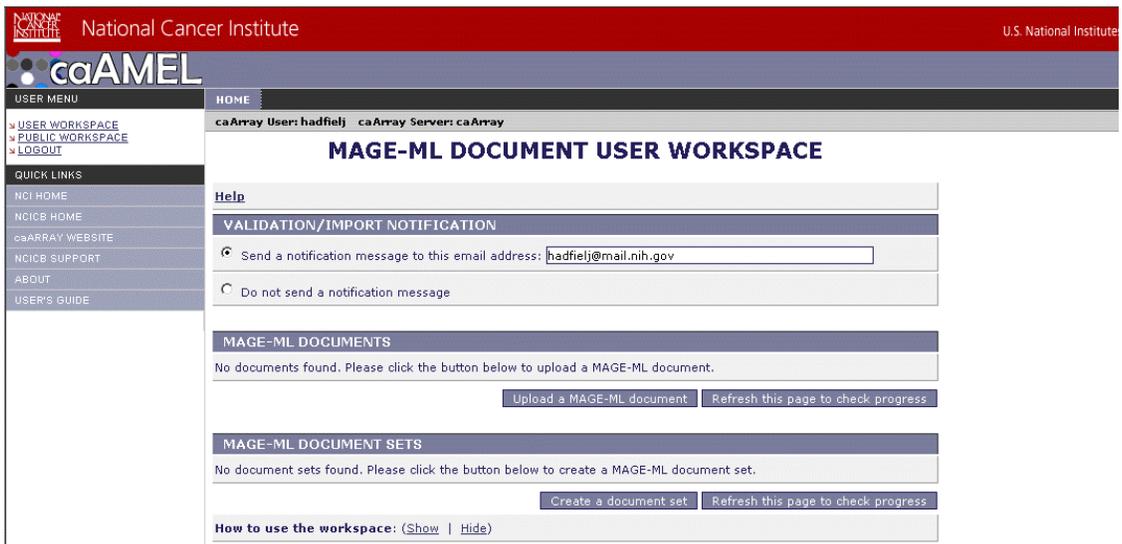


Figure 2.4 User Workspace

User Notification by Email

Tasks such as validating and importing data can be somewhat time consuming. To relieve you of the burden of waiting for long-running tasks to complete, caAMEL performs many tasks asynchronously. You can choose to periodically “refresh” a page to get an updated status of percentage complete, or you can indicate you would like to be notified by email that your document or document set’s validation or import has been completed. This frees you for other tasks.

The user notification feature is active in either the Public or User Workspaces. If you are logged into caAMEL, the email address you specified in your profile in your caArray account is used by default, but you can change it prior to launching any operation that may result in a notification.

To indicate that you want to be notified by email, follow these steps:

1. In the Notification section at the top of a caAMEL workspace, select the option **Send a notification message to this email address** ([Figure 2.5](#)).
2. Enter the appropriate email address or, if you are logged in, verify the email address.
3. The notification feature will be activated once you begin the upload, validation or import task.



VALIDATION/IMPORT NOTIFICATION	
<input checked="" type="radio"/>	Send a notification message to this email address: <input type="text" value="moreau@islandresearch.com"/>
<input type="radio"/>	Do not send a notification message

Figure 2.5 Specifying Email Notification

CHAPTER 3

WORKING WITH MAGE-ML DOCUMENTS

This task-oriented chapter outlines the procedures for working with MAGE-ML documents in caAMEL and importing them into caArray.

Topics in this chapter include:

- *MAGE-ML Document Overview* on this page
- *Document Sets* on page 17
- *Importing Documents and Document Sets into caArray* on page 20
- *Resolving Errors* on page 22
- *After caArray Import* on page 23

MAGE-ML Document Overview

The primary purpose of caAMEL is to facilitate importing MAGE-ML documents into caArray. MAGE-ML documents are XML files that adhere to the MAGE-ML 1.1 or MAGE-ML 1.0 DTD.

Users familiar with the structure of MAGE, or who have used caArray can anticipate the components that can be represented in a MAGE-ML document, including such things

as organizations, persons/contacts, protocols, experiments, and biosequences. [Figure 3.1](#) displays a segment of a typical MAGE-ML document.

```

1  <?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
2  <!DOCTYPE MAGE-ML PUBLIC "-//OMG//DTD MAGE-ML 1.1//EN" "MAGE-ML.dtd">
3  <MAGE-ML identifier="amoreau-contact.xml">
4    <AuditAndSecurity_package>
5      <Contact_assnlist>
6        <Person identifier="caamel.nci.nih.gov:person:1">
7          address="5 Monkey Lane, Suite 5A, Island Bay, 03020"
8          email="moreau@islandresearch.edu"
9          fax="1-800-555-1212"
10         firstName="Alphonse"
11         lastName="Moreau"
12         midInitials=""
13         name="Alphonse Moreau"
14         phone="1-342-555-1212"
15         tollFreePhone="1-866-555-1212"
16         URI="http://www.islandresearch.edu">
17         <Affiliation_assnref>
18           <Organization_ref identifier="caamel.nci.nih.gov:organization:1" />
19         </Affiliation_assnref>
20       </Person>
21     </Contact_assnlist>
22   </AuditAndSecurity_package>
23 </MAGE-ML>

```

Figure 3.1 Example of a simple MAGE-ML Document

Once a document is uploaded into caAMEL, you can validate, download, or delete it. Registered users can import it into caArray. The validation process may indicate that there are “warnings” for the document, even though the MAGE-ML is valid. For example, a field value may be too large to import into caArray and will be truncated when imported. The **Import - Ignore Warnings** option allows you to import documents in spite of these validation warnings.

A document set is a group of potentially interrelated MAGE-ML documents and data files. Document sets are validated as a group, allowing for references between documents contained in the group. You can add additional MAGE-ML documents or referenced data files to a document set. Editing a document set allows you to change the name or description of the set. Users logged into caAMEL can import the MAGE-ML content and data files contained in document sets to caArray.

caAMEL Workflow

The recommended workflow in caAMEL is as follows:

1. Upload MAGE-ML documents (either as individual documents, or into a document set).
2. Validate the document(s).
3. Review the validation results when the job has completed.
4. Import the MAGE-ML content into caArray (if you are not simply validating the document[s]).
5. Review the import results.
6. Delete the document or document set in caAMEL when all tasks are complete.

Uploading Documents

Document upload can be performed in either the Public Workspace or in the User Workspace. For more information about these interfaces, see [Chapter 2 The caAMEL Workspaces](#). The procedure for uploading documents is the same in both workspaces.

Before you can validate a MAGE-ML document, it needs to be uploaded to caAMEL. If you have ever uploaded a file using a web browser, this process should not be unfamiliar to you. If you have a good connection speed, and a moderately sized file, the process will go quickly. However, if you are on a dial-up connection and uploading a massive MAGE-ML document, such as over 1 GB in size, the upload might be quite time consuming.

To upload one or more documents, follow these steps:

1. From the workspace, choose the **Upload a MAGE-ML document** button ([Figure 3.2](#)).

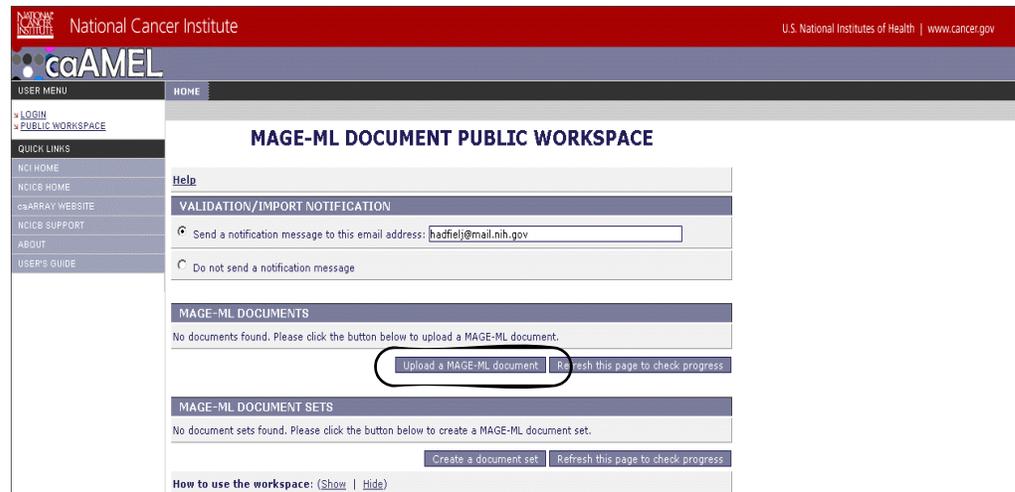


Figure 3.2 Upload a MAGE-ML document button

2. On the Upload MAGE-ML page that opens, browse for and select up to four different files to upload ([Figure 3.3](#)).



Figure 3.3 Upload MAGE-ML page

3. Press the **Upload** button.

When the upload completes, an updated view of the workspace displays, as shown in [Figure 3.4](#).

MAGE-ML DOCUMENT USER WORKSPACE

[Help](#)

VALIDATION/IMPORT NOTIFICATION

Send a notification message to this email address:

Do not send a notification message

MAGE-ML DOCUMENTS

Name	Size	Status	Uploaded	Validated	Imported
repository_MAGE.xml	1.47 Kb	Not Validated	01/10/07 06:51 PM		

MAGE-ML DOCUMENT SETS

No document sets found. Please click the button below to create a MAGE-ML document set.

How to use the workspace: ([Show](#) | [Hide](#))

Figure 3.4 Workspace after file has been uploaded

The document's filename, its size and status display in the middle of the page (circled). With each document contained in the workspace, there is an associated number of action options corresponding to that file, including **Validate** and **Delete**. Those operations are depicted with icons and corresponding pop-up text. The available icons that display vary according to the workspace you are working in.

Note: For efficient file transfer, caAMEL can receive uploaded .ZIP files and unzip them upon arrival. This permits users who have large MAGE-ML files to zip them first, then upload them. Choosing a zip file and uploading it is performed in the same way as described in [Uploading Documents](#) on page 15.

Validating Documents

Note: Document validation can be performed in either the Public Workspace or in the User Workspace. For more information about these interfaces, see [Chapter 2 The caAMEL Workspaces](#). The procedure for validating documents is the same in both workspaces.

Once a file has been uploaded, it can be validated to ensure it is both well-formed XML and also meets the semantic requirements of valid MAGE-ML documents.

To validate any uploaded file, click the Validate button () that corresponds to the document you wish to validate.

The validation process can be somewhat time consuming. If the file is fairly small, you can choose to monitor the process yourself, using the **Refresh the page to check progress** button to update the status field to report on progress (for example, "Validating (45%)..."). Alternatively, caAMEL permits you to specify an email address to which notification of completion can be sent once the process is completed. For more information, see [User Notification by Email](#) on page 11.

Once the validation process is completed, the **Status** will either state **Valid** or **Invalid**. The date and time validation was completed displays in the **Validated** column. If the document was deemed **Invalid**, click the hypertext link to view the reason(s).

MAGE-ML DOCUMENTS						
Name	Size	Status	Uploaded	Validated	Imported	
simple.xml	554 bytes	Valid	01/18/07 12:26 PM	01/18/07 12:26 PM		   
repository_MAGE.xml	1.47 Kb	Validation Error	01/10/07 06:51 PM			   
biosequence.xml	32.43 Kb	Valid	01/18/07 12:26 PM	01/18/07 12:27 PM		   
array_and_design.xml	2.87 Kb	Validating (0%)	01/18/07 12:26 PM			   
valid.xml	1.05 Kb	Not Validated	01/18/07 12:26 PM			   

Figure 3.5 The MAGE-ML Documents table displays the status of validation results

Downloading Documents

Note: Document download can be performed in either the Public Workspace or in the User Workspace. For more information about these interfaces, see [Chapter 2 The caAMEL Workspaces](#). The procedure for downloading documents is the same in both workspaces.

Any document, once uploaded, can be downloaded. To do so, click the corresponding **Download** () button. In the dialog box that opens, browse for the destination for the document or specify to open the document. Click **OK** to execute your choice.

Deleting Documents

Note: Document deletion can be performed in either the Public Workspace or in the User Workspace. For more information about these interfaces, see [Chapter 2 The caAMEL Workspaces](#). The procedure for deleting documents is the same in both workspaces.

To delete an uploaded document, click on the corresponding **Delete** () button. When the confirmation prompt displays, click **OK** to confirm the deletion or **Cancel** to halt the deletion. Once completed, the row representing the document concerned no longer appears, confirming the deletion of the document.

Document Sets

caAMEL provides for the management of document sets.

A document set is a group of potentially semantically meaningful, interrelated, MAGE-ML documents and data files. Creating a document set provides a means to organize your documents with common metadata, such as a document set name and description.

Document set components are validated as a group, thus allowing for references between documents contained in the group.

Note: Because the documents are managed as a group, if one of the documents within a document set is invalid, the document set is considered invalid.

You can add additional MAGE-ML documents or referenced data files to a document set. Editing a document set allows you to change the name or description of the set. Users logged into caAMEL can import the MAGE-ML content and data files contained in document sets to caArray.

- Notes:**
- Document sets can be created and managed in either the Public Workspace or in the User Workspace. For more information about these interfaces, see [Chapter 2 The caAMEL Workspaces](#). The procedure for working with document sets is the same in both workspaces, except for the import function which is only available in the User Workspace.
 - You can use document sets even for sets containing only one document. This would provide the advantage of being able to name and describe the document, for convenience.

Creating a Document Set

The first step to employing document sets is to create an empty one. To do so, follow these steps:

1. Click the **Create a document set** button ([Figure 3.6](#)).

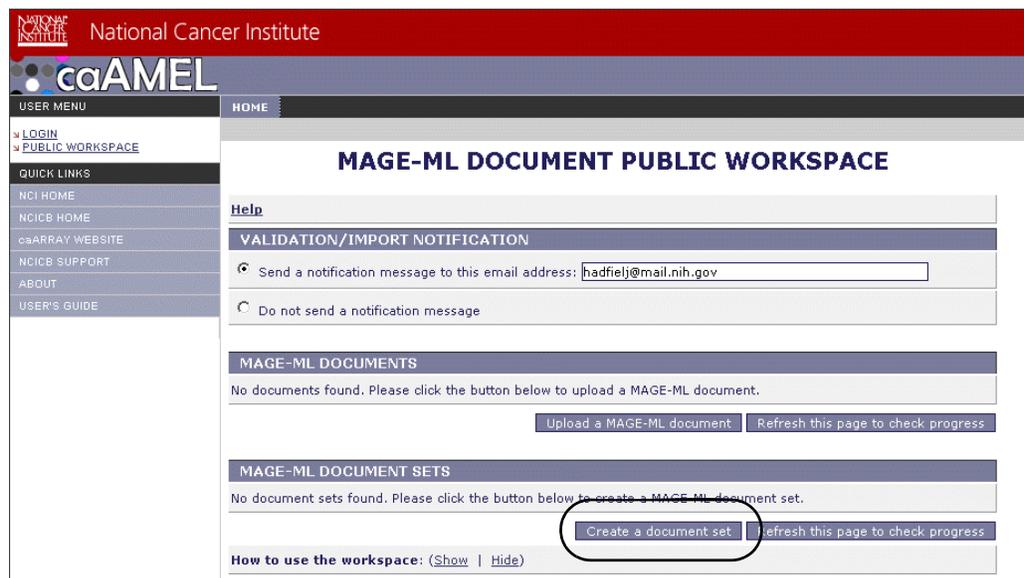


Figure 3.6 Create document set button

2. On the page that opens, enter a name and description for the documentation set (*Figure 3.7*).

CREATE DOCUMENT SET

How to use the workspace: A document set is a group of potentially interrelated MAGE-ML documents. Document sets are validated as a group, allowing for references between documents contained in the group. Create the document set on this page and click **Save**. On the next page, upload documents to the document set.

DOCUMENT SET PROPERTIES

Document Set Name: Brazilian monkey Exp

Description: These files represent the collection of MAGE-ML data from Dr. Moreau's Brazilian Monkdy experiment. These are the experiments that resulted in the discovery of the five mutant....

Save Cancel

Figure 3.7 Specifying name and description for a new document set

3. Click **Save**. This opens the new, empty document set, ready to have MAGE-ML documents uploaded to it (*Figure 3.8*).

MAGE-ML DOCUMENT SET

How to use the workspace: Note that only document sets are validated and imported; individual documents cannot be individually validated. Click *Return to Workspace* to perform actions on the document set. Otherwise, use this form to upload additional documents into your set, or download or delete individual documents.

Note: Document sets may also contain data files that are referenced in MAGE-ML documents (i.e. in <DataExternal> elements). Any referenced data files must be uploaded as part of the document set or the document set will be considered invalid. On import, the data files will also be transferred into caArray.

DOCUMENT SET
Brazilian monkey Exp

No documents found. Please click the button below to upload a MAGE-ML document.

Upload a MAGE-ML document Return to Workspace Refresh this page to check progress

Figure 3.8 The resulting empty document set, ready to have MAGE-ML documents uploaded to it

4. To upload documents into a document set, click the **Upload a MAGE-ML document** button in the MAGE-ML DOCUMENT SET page. Once one or more documents are uploaded to the set, each displays in the MAGE-DOCUMENT Set page (*Figure 3.9*).

MAGE-ML DOCUMENT SET

How to use the workspace: Note that only document sets are validated and imported; individual documents cannot be individually validated. Click *Return to Workspace* to perform actions on the document set. Otherwise, use this form to upload additional documents into your set, or download or delete individual documents.

Note: Document sets may also contain data files that are referenced in MAGE-ML documents (i.e. in <DataExternal> elements). Any referenced data files must be uploaded as part of the document set or the document set will be considered invalid. On import, the data files will also be transferred into caArray.

DOCUMENT SET
Brazilian monkey Exp

Name	Size	Status	Uploaded	Validated	Imported
biosequence.xml	32.43 Kb	Not Validated	01/10/07 07:19 PM		
simple.xml	554 bytes	Not Validated	01/10/07 07:19 PM		
valid.xml	1.05 Kb	Not Validated	01/10/07 07:19 PM		

Upload a MAGE-ML document Return to Workspace Refresh this page to check progress

Figure 3.9 Populated document set

Managing Document Sets

Once created, document sets can have documents uploaded to them (described in the previous section), they can be validated (✓), the metadata can be edited (✎), and the document set can be opened (📁) to facilitate exploring validation results, downloading and/or deleting documents that are in the set, and adding new documents to the set. All of these procedures are initiated by clicking the appropriate button that corresponds to the database set (Figure 3.10). The tasks are completed as described for working with single documents elsewhere in this chapter.

- [Validating Documents](#) on page 16
- [Metadata Errors](#) on page 22

The screenshot displays two tables. The top table, titled 'DOCUMENTS', lists individual documents with columns for Size, Status, Uploaded, Validated, and Imported. The bottom table, titled 'DOCUMENT SETS', lists document sets with columns for Description, # of Files, Total Size, Status, Validated, and Imported. In the 'DOCUMENT SETS' table, the 'Validated' and 'Imported' columns for the first set are circled, and the corresponding icons (a green checkmark and a green folder) are also circled. Below the tables are buttons for 'Upload a MAGE-ML document' and 'Refresh this page to check progress'.

DOCUMENTS	Size	Status	Uploaded	Validated	Imported				
	554 bytes	Valid	01/18/07 12:26 PM	01/18/07 12:26 PM		✓	📁	📁	📁
.xml	1.47 Kb	Validation Error	01/10/07 06:51 PM			✓	📁	📁	📁
	32.43 Kb	Valid	01/18/07 12:26 PM	01/18/07 12:27 PM		✓	📁	📁	📁
.xml	2.87 Kb	Validating (0%)	01/18/07 12:26 PM			✓	📁	📁	📁
	1.05 Kb	Not Validated	01/18/07 12:26 PM			✓	📁	📁	📁

Upload a MAGE-ML document Refresh this page to check progress

DOCUMENT SETS	Description	# of Files	Total Size	Status	Validated	Imported			
	7 files represent the portion of MAGE-ML data from Dr. Moreau's Brazilian fly experiment. These are experiments that led to the discovery of the mutant....	7	40.06 Kb	Validating (0%)			✓	📁	📁

Create a document set Refresh this page to check progress

workspace: (Show | Hide)

Figure 3.10 Attributes and options available for document sets

Importing Documents and Document Sets into caArray

Note: Importing documents and document sets can be performed only from the User Workspace. You must be logged into caAMEL to work from the User Workspace. For more information, see [User Workspace](#) on page 10.

After uploading a document to your User Workspace, you can import a single document or a document set into caArray. A MAGE-ML document or document set must be considered valid before caAMEL permits them to be imported into caArray. You can specify validation before you attempt import or you can proceed directly to import. If you go right to import, caAMEL performs a validation implicitly before the import begins.

Note: Import can be performed under two options: **Import** (✓) and **Import (Ignore Warnings)** (📁).

Import—Import is initiated when you click this button. If the document is judged to be valid, the import proceeds without a problem. If you have not previously validated the

document, and caAMEL judges the document to be invalid, the import will be aborted. An error message displays in the User Workspace and/or by email notifying you of the termination of the process (*Figure 3.11*).

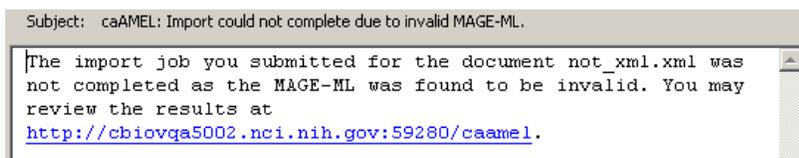


Figure 3.11 Email notice after an attempt to import an invalid document

Import (Ignore Warnings)—When you click this button, import proceeds in spite of warnings that may be generated by the document validation. Such warnings reflect the strictness of MAGE-ML and the peculiarities of caArray. In short, by ignoring warnings, even if there are some generated by the validation portion, the import process can continue to completion rather than being cancelled without your knowledge. This warning could signal, for example, that a text field in the document exceeds field limits in caArray and will be truncated on import.

To perform an import into caArray, follow these steps:

1. Click the Import () or Import (Ignore Warnings) () button that corresponds to the document or document set to be imported.

Note: These two buttons are not present in the viewer if you are not logged into caAMEL.

The process may be somewhat time consuming. You can monitor the import status by clicking the **Refresh this page...** button or waiting to receive notice by email that the process has finished.

At that point, the Imported column in User Workspace displays the date and time of the successful import, if it completes successfully (*Figure 3.12*).

MAGE-ML DOCUMENTS										
Name	Size	Status	Uploaded	Validated	Imported					
simple.xml	554 bytes	Imported	01/18/07 12:26 PM	01/18/07 12:39 PM	01/18/07 12:39 PM					
repository_MAGE.xml	1.47 Kb	Validation Error	01/10/07 06:51 PM							
biosequence.xml	32.43 Kb	Importing (97%)	01/18/07 12:26 PM	01/18/07 12:39 PM						
array_and_design.xml	2.87 Kb	Validating (0%)	01/18/07 12:26 PM							

Figure 3.12 caAMEL viewer indicating a successful document import into caArray

Note: If your MAGE-ML document(s) refer to external data files in a <DataExternal> tag (e.g. .cel files, .chp files, .txt files, etc.), you must import the MAGE documents and related data files as part of a document set. caAMEL will not consider data files or MAGE documents that refer to them as valid unless they are part of the same document set.

Resolving Errors

Metadata Errors

Typographical errors in document set information are simple to correct. Click the **Edit** () button for the document set to be edited. In the Edit Document Set dialog box that opens, make the correction(s) to the name or description.

EDIT DOCUMENT SET

DOCUMENT SET PROPERTIES	
Document Set Name	Brazilian monkey Exp
Description	These files represent the collection of MAGE-ML data from Dr. Moreau's Brazilian Monkey experiment. These are the experiments that resulted in the discovery of the five mutant....
<input type="button" value="Save"/> <input type="button" value="Cancel"/>	

Figure 3.13 Editing document set meta-data

Invalid Documents

If you are not sure why a document or document set has been deemed “Invalid”, click the **Invalid** hypertext link in the document or document set's **Status** field. For example, in the case of a Brazilian Monkey3 Experiment document set, following the **Invalid** link opens a page in which one document in the set is considered invalid, as illustrated in *Figure 3.14*.

MAGE-ML DOCUMENT SET

How to use the workspace: Note that only document *sets* are validated and imported; individual documents cannot be individually validated. Click *Return to Workspace* to perform actions on the document set. Otherwise, use this form to upload additional documents into your set, or download or delete individual documents.

Note: Document sets may also contain data files that are referenced in MAGE-ML documents (i.e. in <DataExternal> elements). Any referenced data files must be uploaded as part of the document set or the document set will be considered invalid. On import, the data files will also be transferred into caArray.

DOCUMENT SET Brazilian Monkey3						
Name	Size	Status	Uploaded	Validated	Imported	
valid.xml	1.05 Kb	Valid	01/18/07 01:59 PM	01/18/07 01:59 PM		 
invalid_attributes.xml	1.03 Kb	Invalid	01/18/07 01:59 PM	01/18/07 01:59 PM		 
biosequence.xml	32.43 Kb	Valid	01/18/07 01:59 PM	01/18/07 01:59 PM		 

Figure 3.14 Examining which documents in a document set are invalid

To further explore the errors in an invalid document, click the **Invalid** link in its **Status** field. The table that opens displays details covering individual errors in the invalid document (*Figure 3.15*).

MAGE-ML DOCUMENT VALIDATION RESULTS						
VALIDATION RESULTS FOR invalid_attributes.xml						
Type	Line Number	Name	Identifier	Element	Description	Suggested Fixes
Error	6		arraydesign	ArrayDesign.numberOfFeatures	An invalid value was provided for an attribute of type Integer. The value provided was five	Edit the MAGE-ML document to provide a valid value
Error	11		array	Array.arrayXOrigin	An invalid value was provided for an attribute of type Float. The value provided was five and a half	Edit the MAGE-ML document to provide a valid value
Error	11		array	Array.arrayYOrigin	An invalid value was provided for an attribute of type Float. The value provided was	Edit the MAGE-ML document to provide a valid value
Error	22		arraymanufacture	ArrayManufacture.manufacturingDate	An invalid value was provided for an attribute of type Date. The value provided was not a date	Edit the MAGE-ML document to provide a valid value
Warning	6		arraydesign	ArrayDesign.name	This MAGE element of type ArrayDesign is unnamed. In order to be accessible in caArray MAGE elements of this type must include a name. By default, caAMEL will use the element's identifier (arraydesign) as its name.	You may either accept the generated name or edit the file to provide a name for this element.
Warning	11		array	Array.name	This MAGE element of type Array is unnamed. In order to be accessible in caArray MAGE elements of this type must include a name. By default, caAMEL will use the element's identifier (array) as its name.	You may either accept the generated name or edit the file to provide a name for this element.
<< Back						

Figure 3.15 Examining validation results for a documentation set

MAGE-ML Document Errors

Once errors in a MAGE-ML document are discovered and understood, it is up to you to resolve them. While caAMEL can point out where it finds errors, caAMEL cannot resolve them. You must return to the MAGE-ML source documents to resolve the reported errors.

Once resolved, the updated document can be uploaded into caAMEL again, and the process of validation and/or import can be re-attempted.

After caArray Import

Once a document or document set is successfully imported into caArray, you may find it unnecessary to keep the original MAGE-ML documents in caAMEL. If that is the case, you should “clean up after yourself” to limit disk space consumption by deleting documents and document sets with which they are finished.

A P P E N D I X



CAARRAY GLOSSARY

Acronyms, objects, tools and other terms referred to in the chapters or appendixes of this caAMEL User's Guide are described in this glossary.

<i>Term</i>	<i>Definition</i>
caBIG	cancer Biomedical Informatics Grid
CGH	Comparative Genomic Hybridization
MAGE 1.1	MAGE 1.1 is a widely-used microarray data standard or guideline
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
NCI	National Cancer Institute
NCICB	National Cancer Institute Center for Bioinformatics
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

INDEX

C

- caAMEL
 - glossary 25
 - help 6
 - launching 5
 - login 5
 - logout 6
 - objectives 4
 - overview 3
 - Public Workspace 9
 - quick links 6
 - user account 5
 - User Workspace 10
 - workflow 4, 14
- caAMEL User's Guide
 - introduction 1
 - organization 1
 - text conventions 2
- caArray
 - after importing to 23
 - importing document to 20
 - user account for caAMEL 5
- creating document set 18

D

- deleting documents 17
- document
 - deleting 17
 - downloading 17
 - importing to caArray 20
 - invalid 22
 - MAGE-ML, errors 23
 - MAGE-ML description 13
 - MAGE-ML example 14
 - uploading 15
 - validating 16
- document conventions 2
- document set
 - creating 17
 - description 14, 17
 - managing 20

- downloading documents 17

E

- editing metadata 22
- email notification 11
- errors
 - in MAGE-ML documents 23
 - invalid document 22
 - metadata 22

G

- Glossary 25

H

- help in caAMEL 6

I

- import (ignore warnings) 21
- invalid document 22

L

- launching caAMEL 5
- login 5
- logout 6

M

- MAGE-ML
 - description 3
 - document description 13
 - document errors 23
 - document example 14
 - document handling in caAMEL 14
 - overview 3
- managing document sets 20
- metadata errors 22

O

- overview, chapters in guide 1

P

Public Workspace 9

Q

Quick Links 6

T

task

- completion 11

- notification 11

text conventions, user guide 2

U

uploading

- document 15

- zip file 16

user

- notification 11

- registering 5

User Workspace 10

V

validating documents 16

validation, time 16

W

workflow in caAMEL 4, 14

Z

zip file upload 16