HIEv User Manual

For software version 1.9

19 June 2013

There has been a significant rise in the number of sensors and sensor networks used in environmental research in recent years. This growth has brought with it the challenge of managing sensor infrastructure and the data produced by the increasing numbers of deployed sensors.

HIEv (usually pronounced "hive") was developed to address these challenges and specifically aims to:

- Ensure raw data is never lost
- Ensure that data can be used and interpreted in the future
- Allow researchers to make linkages between different types of data
- Make it easier for researchers to get access to the data they need
- Make it easier for technical officers to distribute the data
- Create entries in Research Data Australia (required by ANDS)

As a web application, HIEv is easily accessed from a variety of locations and platforms – all you need is a modern web browser and the URL of the server hosted by the research centre with which you are affiliated or collaborating.

The HEIv product was initially developed for the Hawkesbury Institute for the Environment (HIE) at the University of Western Sydney, with funding provided by ANDS under their Data Capture program.

The development name for HIEv was DC21, and this name still appears in some locations. In particular, it is the name of the open source GITHUB WIKI and repository for HIEv.

Contents

Contents 2

1 Overview 4

1.1 Installing HIEv 5

2 Glossary 6

3 Logging in to HIEv 7

3.1 Classes of Users 8

4 General Operation 9

4.1 The HIEv Main Screen 9

4.2 Entering Dates and Times 10

4.3 Signing Out 11

4.4 Changing Your User Settings 11

4.4.1 Overview Tab 11

4.4.2 Edit Details Tab 12

4.4.3 Change Password Tab 13

5 HIEv Data File Storage and Metadata 14

5.1 Basic Information 14

5.2 Summary Information Extracted from TOA5 Files 16

5.3 Column Information for TOA5 Files 16

6 Facilities and Experiments 18

6.1 Creating a Facility Entry 18

6.2 Editing a Facility Entry 20

6.3 Creating an Experiment Entry 21

6.4 Editing an Experiment Entry 24

6.5 Setting Up Experiment Parameters 24

7 Uploading Data files 26

7.1 Uploading RAW TOA5 data files 29

7.2 Manual Data Upload Action Summary 30

7.3 Automating the upload of data to HIEv 31

8 Managing Data Files 32

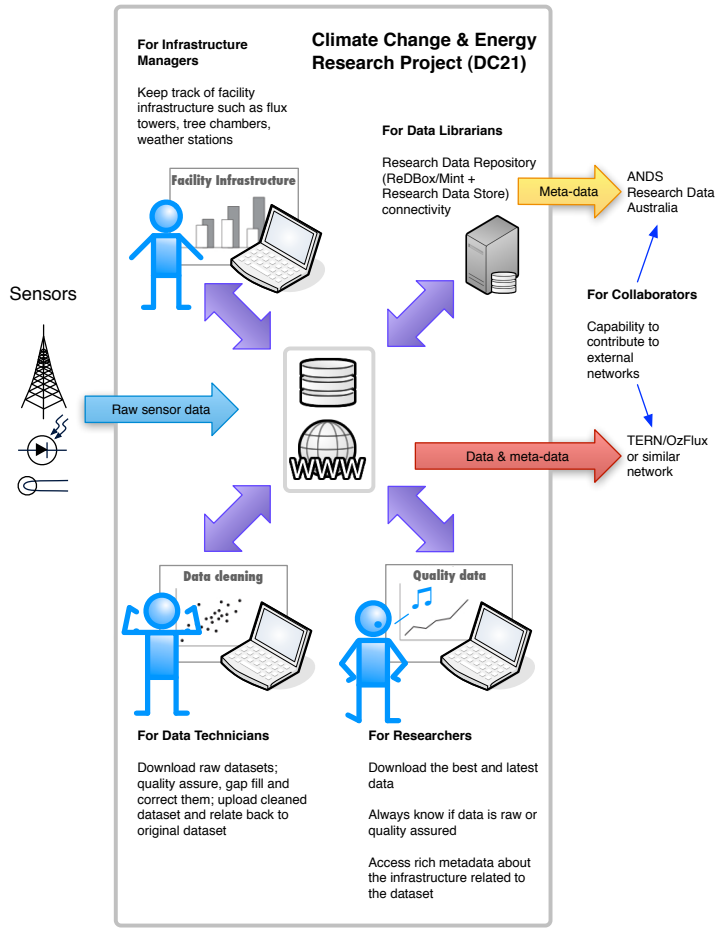
8.1 The Dashboard Tab 32

8.2 The Explore Data Tab and File Searching 32
8.2.1 Sorting 33
8.2.2 Searching 33
8.3 The Cart 41
8.3.1 Editing the Cart Contents 42
8.4 Viewing and Editing a File's Metadata 44
8.5 Deleting a Data File 45
9 Publishing Your Data 47
9.1 Creating a Package 48
9.2 Publishing a Package 49
9.3 Managing Published Packages 50
9.3.1 Publishing a second time 51
9.3.2 Deleting Published Packages 51
9.3.3 Editing Published Packages 51
9.3.4 Correcting Published Packages 51
9.4 Viewing Published data 52
10 Downloading files 53
11 HIEv Administration 54
11.1 Managing Users' Details 54
11.2 Authorising New Users – The Access Requests Tab 55
11.3 Managing Column Mappings 56
11.3.1 The Column Mappings tab 56
11.3.2 Updating from the Explore Data tab 57
11.4 Managing Background Tasks 58
12 Configuring Tags, Column Mappings and Experiment Parameters 60
13 Migrating data to a new system 62
14 Revision History 63
Appendix A - The Bagit format 64
README.HTML file 64
Appendix B - RIF-CS 65
Example RIF-CS file 65

1 Overview

HIEv is designed to act as a central repository for environmental research data. Technicians can configure field PCs to automatically push time-series data from sensors into HIEv. Time-series data and other data, such as photographs, videos, sound recordings, spreadsheets and other files can be manually uploaded. Researchers can use the system to discover and download the latest data available. Rich Metadata is stored for physical infrastructure ("Facilities"), for the Experiments that run at those Facilities, as well as for the individual data files to support discovery and interpretation.

This version of HIEv (Version 1.9) deals primarily with sensor networks which utilise Campbell Scientific data loggers, and provides rich metadata for data files which are uploaded in the TOA5 file format. (Because HIEv is open source, it would be possible for other users to add support for other file formats.)



All files in HIEv are grouped by Experiment, providing a convenient way to organise related files such as cleansed or gap-filled data, and analysis outputs.

Once finalised, Packages of data can be defined, described and published to ANDS as a RIF-CS data collection. This makes them available for discovery via the OAI-PMH protocol – see *Appendix B - RIF-CS* for more information. This enables researchers from outside the organisation that produced the data to discover it, and to request access to download a copy.

1.1 Installing HIEv

Instructions for installing and upgrading to a new version of the HIEv are in the Release Notes in the HIEv Version Documentation section of the GitHub wiki at https://github.com/IntersectAustralia/dc21/wiki#version-documentation.

The Release Notes will point to the Deployment Guide, also on the GitHub Wiki (https://github.com/IntersectAustralia/dc21/wiki), though these notes will also often contain some special instructions which need to be executed in conjunction with the deployment guide. Typically, a system administrator with Linux skills should do the installation.

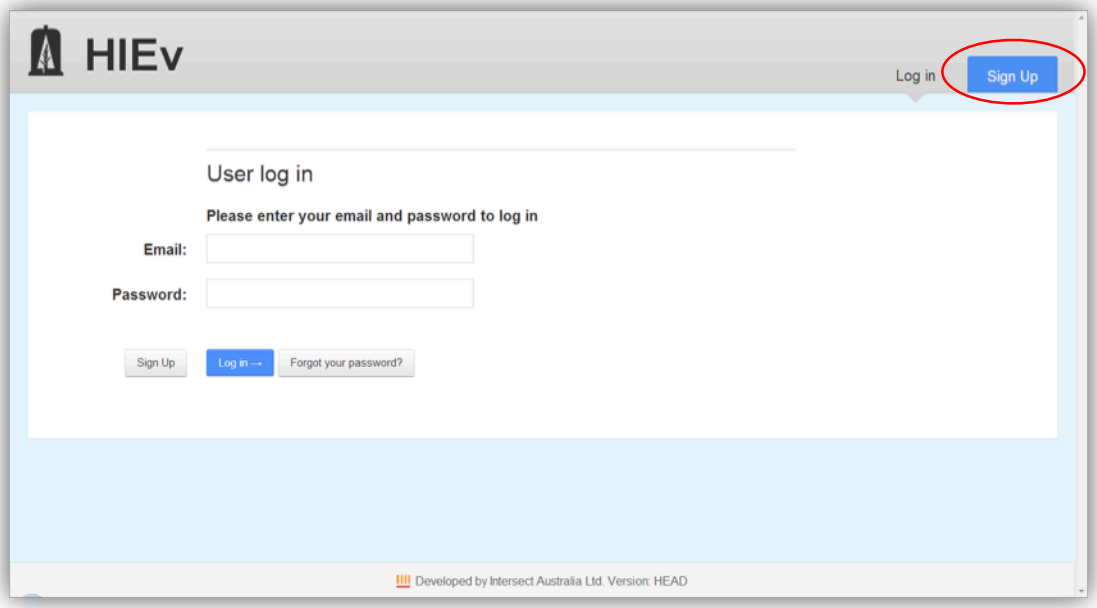
2 Glossary

ANDS	Australian National Data Service
API	Application Program Interface. For HIEv, this is an HTML interface which provides programmatic access to HIEv.
Bagit	A general purpose container file format. See *Appendix A - The Bagit format*.
Metadata	Data about a file of data. Typically, it may include information such as the date and time to which the data file relates, who collected it, where it was collected, why it was collected, explanation of columns in the data file, or any other information about the data in the file.
RDA	Research Data Australia. See http://researchdata.ands.org.au/.
RIF-CS	Registry Interchange Format - Collections and Services. See *Appendix B - RIF-CS*.
Ruby on Rails	The programming language in which HIEv is developed. See http://rubyonrails.org/ for more information.
TOA5	TOA5 format files are produced by the Campbell Scientific LoggerNet program. They are column formatted data files containing header information which describes the data in each column.

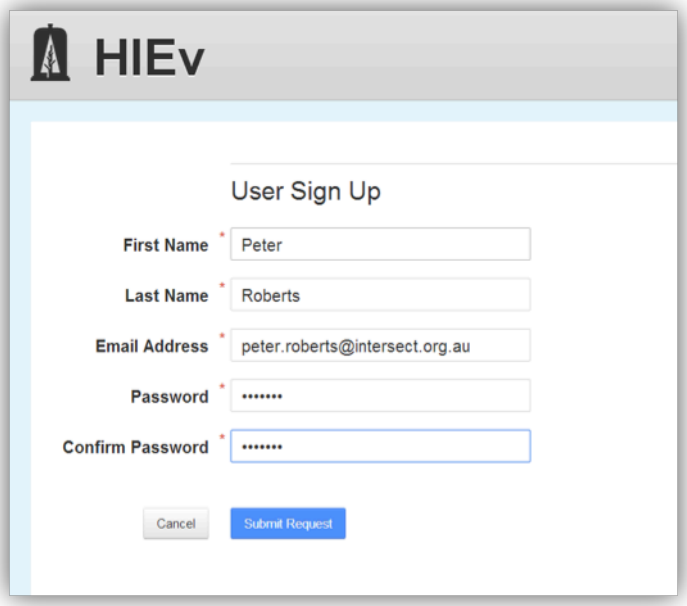
3 Logging in to HIEv

To begin using HIEv, enter the HIEv URL into your web browser. See your HIEv Application Administrator to find out your HIEv URL.

Before you can login you are required to have a HIEv account. You can apply for an account by clicking the blue **Sign Up** button on the top right of the screen.



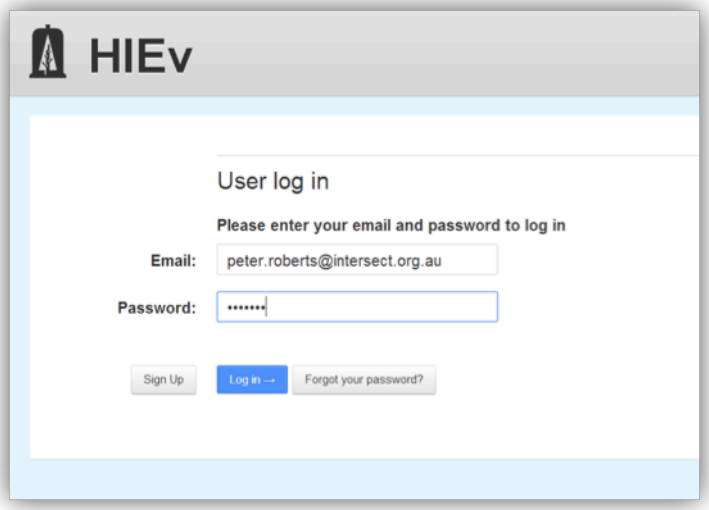
This will take you to a form where you will be requested to enter your first name, last name, email address and chosen password. Passwords must be between 6-20 characters and contain at least one each of: an uppercase letter; a lowercase letter; a digit and a symbol.



Once you have filled out the form and clicked **Submit Request** an email will be sent to the HIEv Application Administrator who will either approve or deny your request for access. Your

HIEv Application Administrator will typically be the HIEv Business Owner or Data Manager. This will probably not be your System Administrator, who would typically be a technical IT person. If your request is approved, you will receive an email informing you that you can now login using the password you entered on the original sign up form.

To see the login form, make sure you have the Log in tab selected on the top right. Next enter your Email address and password, and click the blue **Log in** button below.



Once you have logged in you will be taken to the main screen for the HIEv application.

3.1 Classes of Users

HIEv defines three classes of Users. You will be assigned to one of these classes by the person who authorises your request for Sign In to the system. The permissions of these three classes are restricted in the following ways:

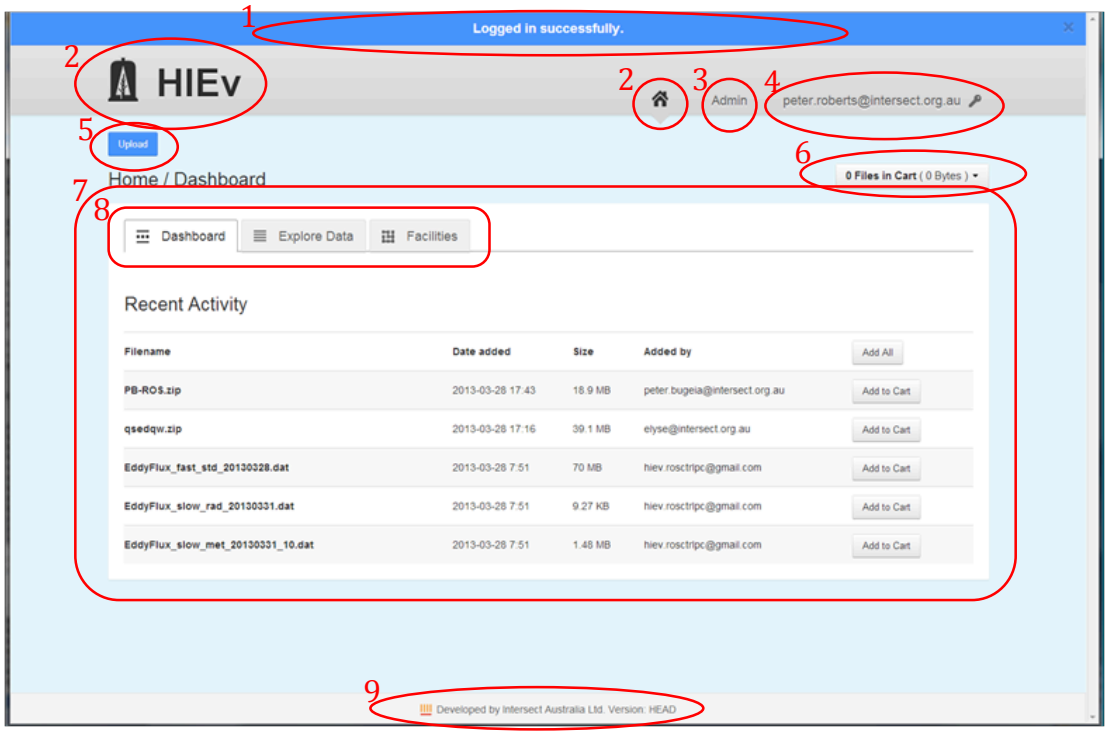
Researcher	No access to the Admin tab and its functions.
	Cannot edit the Metadata of files uploaded or created by other Users.
	Cannot delete files uploaded or created by other Users.
	Cannot use the Application Program Interface (API) functions, which are typically used to automatically upload and download data. (The API token can be generated but it cannot be used.)
	Cannot Publish files. (But can create Packages.)
API Uploader	No access to the Admin tab and its functions.
	Cannot edit the Metadata of files uploaded or created by other Users.
	Cannot delete files created by other Users.
	Cannot Publish files. (But can create Packages.)
Administrator	Permission to perform all functions in HIEv, including authorising new Users' requests for access to HIEv.

4 General Operation

This chapter describes aspects of the operation of HIEv which are common across a number of screens.

4.1 The HIEv Main Screen

The Main Screen consists of the following parts:



Most of these components are common to many of the screens in HIEv.

1	Message bar	This area will appear when HIEv displays an error or informative message.
2	Home button and HIEv Logo	Click either of these to show the Home/Dashboard (that is, the view shown above).
3	**Admin** button	Click to access HIEv administration functions (see *Chapter 11 HIEv Administration*). This button is only present if you have Administrator permissions.
4	Login ID	This is your login name. Click to open a dropdown menu of user operations. (See sections 4.2 and 4.4.)
5	Action button	In many screens, there is an action button at the top left corner. It is often an **Upload** button, which allows you to upload new data files to HIEv. (See *Chapter 7 Uploading Data files* for more information.) However, it may also be a button for another function which is more relevant to the data being displayed in the current view.
6	Cart status box	The HIEv web interface allows you to add files to a Cart, which operates like an e-Commerce shopping cart. Click in the Cart Status

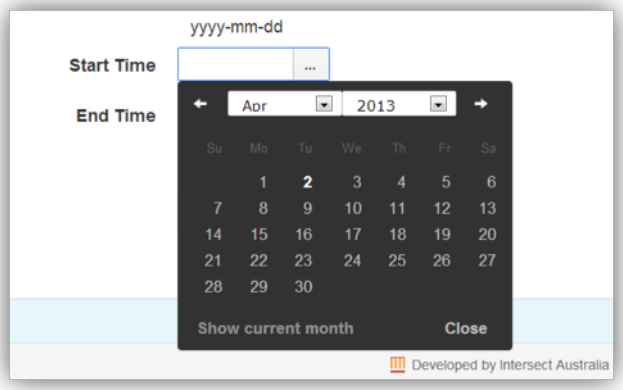
		to open a dropdown menu of Cart functions. (See section *8.3 The Cart)*
7	Working area	The content of this work area changes as you perform HIEv operations.
8	Tabs	Click on a tab to go to the default view for that tab. If it is the tab which is already displayed, it will still return to the default view for the tab.
9	Version indicator	This shows the version of HIEv which you are accessing.

4.2 Entering Dates and Times

There are a number of places where dates and times may be entered into the HIEv system.

Date and time entry fields appear with an ellipsis following the data entry boxes.

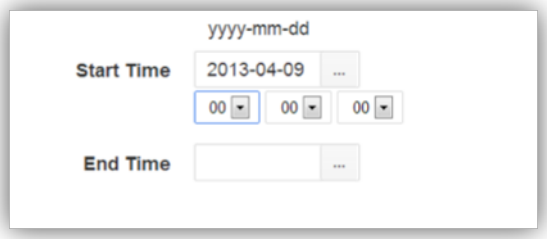
Click on the ellipsis to show a date picker dialog.



Click on the left and right arrows or Month and Year dropdown menus to select the required month. Click **Show current month** to display the calendar for the current month. Click on the date to select the date and close the date picker dialog. Clicking **Close** will close the date picker without selecting a date.

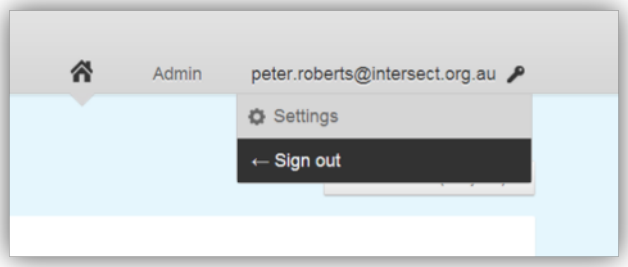
Alternatively, click in the date entry box and type in a date in YYYY-MM-DD format. Use the Tab key to move to the next field, or click in another field. Do not press Enter, as this will activate the data entry screen's Save button.

When the date has been selected, three extra fields will be displayed to permit entry of the Hour, Minute and Second for the time. Use the dropdown menus for these fields to enter the desired time. You cannot type the time in using your computer's the keypad.



4.3 Signing Out

Click on your login ID at the top right of the screen to see a dropdown menu.



Click on **Sign Out** to finish your HIEv session.

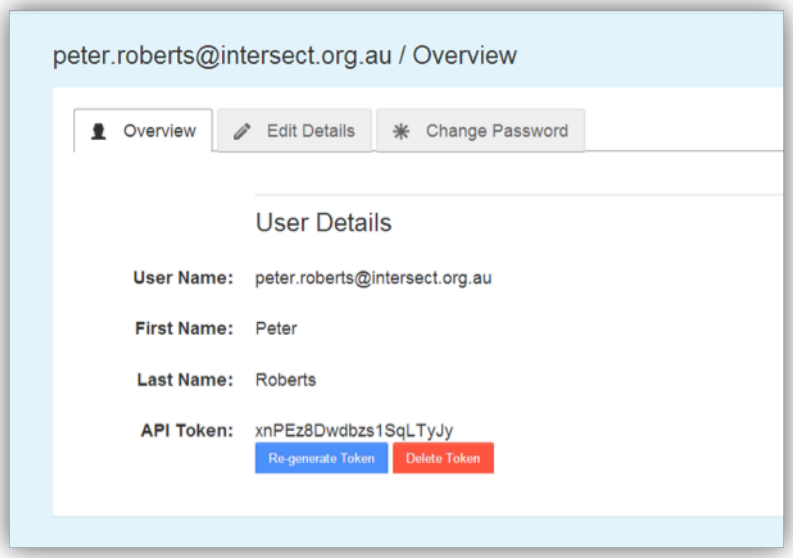
4.4 Changing Your User Settings

Click on your login ID at the top right of the screen to see a dropdown menu. Click on **Settings** to access the following three tabs.

When finished, click on the Home button to return to the HIEv Main Screen.

4.4.1 Overview Tab

This tab displays a summary of your User information.

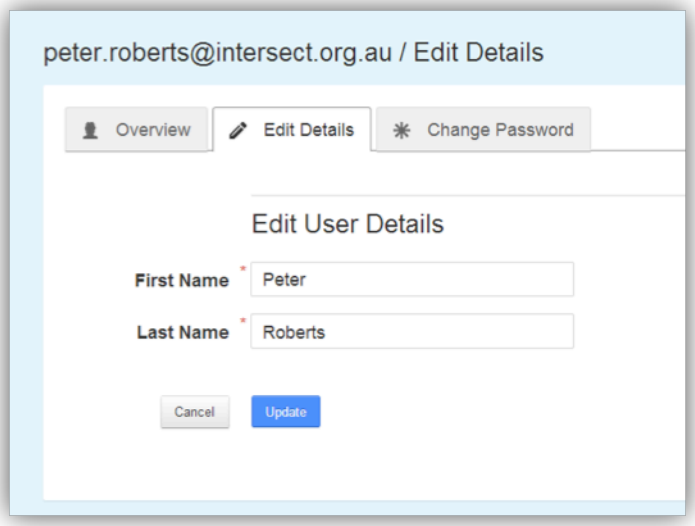


User Name Your valid email address which you use for logging on.

First Name Last Name	Your name.
API Token	A string of characters which you can use as an authorisation token in scripts which you write to make use of the HTML API for HIEv. The HTML API can be used to set up an automatic upload of data from a field PC. See the API definition on the GitHUB DC21/HIEv WIKI (https://github.com/IntersectAustralia/dc21/wiki/File-Upload-API) for instructions on using the HTML API. Initially, no token is displayed and only a single **Generate Token** button is displayed. Clicking on **Generate Token** will cause a token to be displayed in this field. Copy and paste it into the required place in your API scripts. Clicking on **Re-generate Token** will cause the current token to be invalidated and a new token to be generated and displayed. You must replace the token value in your API scripts with this new token so that your API scripts continue to work. This button is only displayed if a valid token is available. Note For security reasons, from time to time you should regenerate your Token and update it in your API scripts. Clicking on **Delete Token** will invalidate the displayed token. Your API scripts will no longer work. This button is only displayed if a valid token is available.

4.4.2 Edit Details Tab

The Edit Details tab allows you to update your First Name and Last Name.

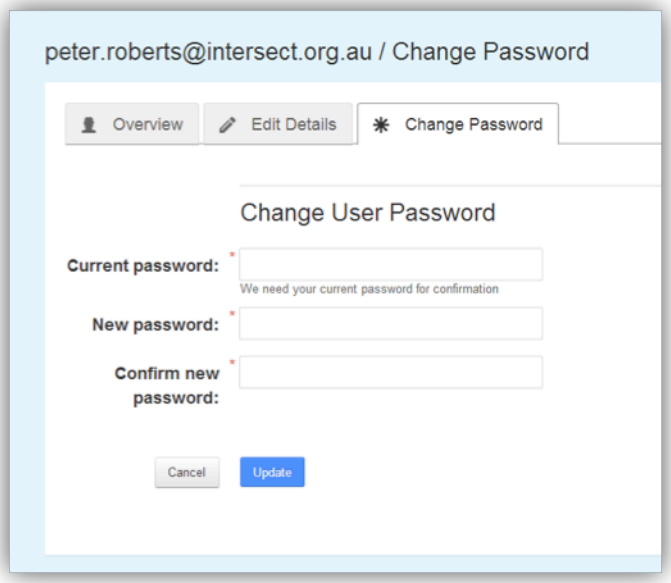


There is no function to change your email address. If you need to do this, you may need to create a new account, or approach your HIEv Application Administrator to change it for you.

Click **Cancel** to return to the Overview Tab without accepting changes, and click **Update** to store the changed values you've entered.

4.4.3 Change Password Tab

Use this tab to change your HIEv logon password.



Passwords must be between 6-20 characters and contain at least one each of: an uppercase letter; a lowercase letter; a digit and a symbol.

You must correctly enter your current password and the strings you enter for New password and Confirm new password must be identical for your password change request to be processed. Click **Cancel** to return to the Overview Tab without accepting changes, and click **Update** to store the changed password you've entered.

5 HIEv Data File Storage and Metadata

HIEv stores uploaded data files using a database structure. In addition to storing the files themselves, HIEv also stores information about each file, known as "Metadata". This Metadata falls into three categories.

5.1 Basic Information

The Basic Information is Metadata entered by the user when the data file is uploaded. It consists of the following fields.

Name — The Name to be used for the file which stores the data in HIEv.

Type — The Type of the data file is a single value that describes the data contained within the file. This value is chosen from a constrained list of possibilities defined by the System Administrator. The file's Type is generally aimed at tracking data through its various stages of processing.

There are three predefined data Types which are built into the system and cannot be changed:

RAW indicates a file that will be tested for known upload File Formats on file upload.

ERROR indicates a file that has failed to upload correctly. Such files are usually TOA5 files. (TOA5 format files are produced by the Campbell Scientific LoggerNet program. See *7.1 Uploading RAW TOA5 data files* for more information about TOA5 data files.)

PACKAGE indicates a file containing a collection of data files which is intended for Publishing. See *Chapter 9 Publishing Your Data* for more information.

If Type field has been set to one of the above predefined Types, it cannot be changed. If it has been set to one of the non-predefined Types, it can be changed, but only to another non-predefined Type.

File Format — Possible values for this field are:

TOA5 The file was inspected on upload and discovered to be TOA5 format. TOA5 files are processed differently on file upload. See section *7.1 Uploading RAW TOA5 data files* for more information.

BAGIT The file is a Package which is formatted as a BAGIT ZIP file. See *Appendix A - The Bagit format* for more information on BAGIT files.

UNKNOWN The file format is not known to HIEv.

The File Format field is set automatically by HIEv.

Description — The description entered by the User to describe the contents of this file.

Tags — The file Tags are a set of flags that have been given by the User to the file from a constrained list of possibilities defined by your HIEv Application Administrator. A file may be assigned multiple Tags or no Tags at all.

Experiment	This field indicates which experiment produced the file. A file can be associated with only one experiment. Any User with the appropriate permissions can create experiments. See *Chapter 6 Facilities and Experiments* for more information.
Facility	The Facility field indicates which facility the above experiment was run at. Each file must be associated with exactly one facility. Any User with the appropriate permissions can create facilities. The Facility field cannot be directly set. Its value is copied from the Facility field in the selected Experiment. See *Chapter 6 Facilities and Experiments* for more information.
Date added	The date on which this file was added to the HIEv database. For Packages, it's the date on which the package was created. This field is set automatically by the HIEv system.
Size	The size of this file. This field is set automatically by the HIEv system when a file is uploaded or a Package is created. This field also has a special function along with the Creation Status field, in that it tracks the progress of Package creation as it is being built. The Size field will show the progressive number of bytes which have been packaged. See sections *9.1 Creating a Package* for more information.
Creation Status	This field, along with the "Packaging progress" field, tracks the progress of Package creation as it is being built. This field tracks the overall creation status. Valid Status values are QUEUED, WORKING, COMPLETED and FAILED. See section *9.1 Creating a Package* for more information.
Packaging progress	This is a temporary field that is only available while a Package is being created. This field tracks the progress of Package creation by showing the progressive number of bytes that have been packaged so far. See section *9.1 Creating a Package* for more information.
File ID	File IDs are unique integers which are assigned and used internally by HIEv to identify files. File IDs cannot be changed by the User. In general, they will not change, but in the case of TOA5 files, may sometimes change after uploading further TOA5 data.
ID	This field provides the opportunity for Users to enter an additional external ID which has been used outside of HIEv to identify this data. It is a character string. For data files, the field is input by the user and it can be used to provide an ID to identify an individual data file. For HIEv Packages, the field is read-only and auto-generated on Package creation. The ID field, along with other Metadata, is copied into the RIF-CS file when the file is copied into a Package when it is created. No two files in HIEv can have the same non-null ID. See sections *9.1 Creating a Package* and *9.3* **Error! Reference source not found.** for more information about the use of this ID field.
Added by	This field indicates the User who uploaded the file to HIEv. For Packages, it's the User who Packaged it. This field is set automatically by HIEv.
Published	This applies to Packages only and indicates whether the Package has been Published or not. This field is set automatically by HIEv.

Published date	This applies only to Packages which have been Published and indicates whether the date on which the Package was Published. This field is set automatically by HIEv.
Start time End time	These fields only apply to Packages and non-TOA5 data files. They hold the dates and times which were manually entered when the file was uploaded or created. These times can be specified with a precision of one second and indicate the start and end times of the data in the file. (For TOA5 files, start and end times are automatically extracted and stored as part of the Summary Information.)

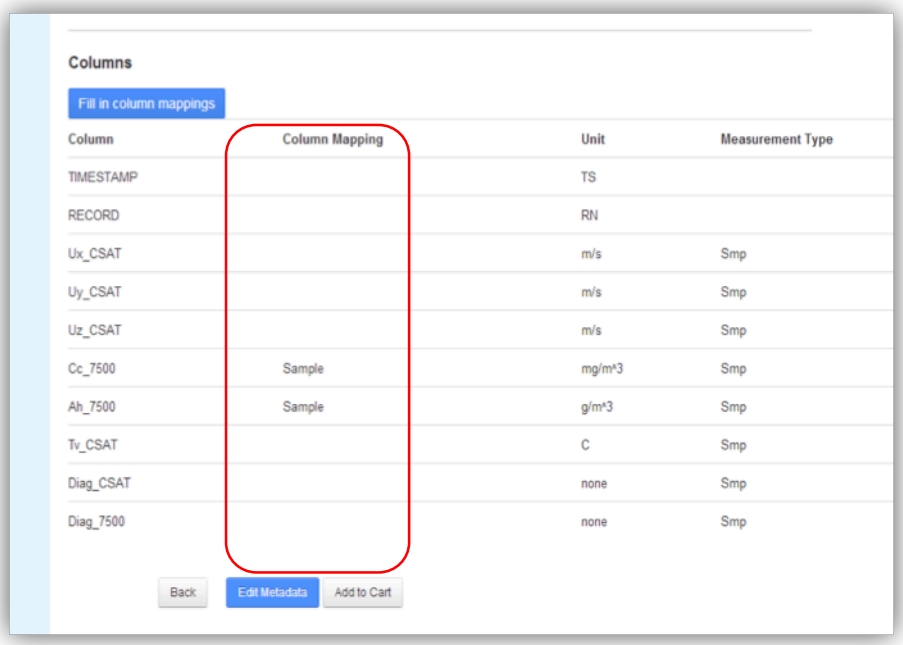
5.2 Summary Information Extracted from TOA5 Files

Summary Information is stored for TOA5 files only. It is collected automatically from the TOA5 data file and is not editable by HIEv Users.

Start time End time	The first and last times of the observations found in the TOA5 data file.
Sample interval field	The frequency of samples in the data file, if relevant.
Datalogger model	The model of data logger used to generate the TOA5 file.
Station name	Designates the Data Logger from which this data was generated and is checked during the upload of a TOA5 file to determine whether an existing TOA5 file or files needs to be replaced.
Serial number Os version Dld name Dld signature Table name	These five fields are also checked during the upload of a TOA5 file to determine whether an existing TOA5 file or files needs to be replaced.

5.3 Column Information for TOA5 Files

Column Information Metadata is collected automatically from TOA5 files when they are uploaded. This automatically collected Metadata for each TOA5 data file is displayed whenever the TOA5 data file's Metadata is displayed. It cannot be changed by HIEv Users.



In addition to the Column Information from the TOA5 file there is one extra column called Column Mapping, circled in red above. The value in this column is assigned by matching the value in the TOA5 column heading against the values in HIEv's Column Mapping Table. It is used to simplify searching by TOA5 column heading. See section *11.3 Managing Column Mappings* for more information about setting up the Column Mapping table.

6 Facilities and Experiments

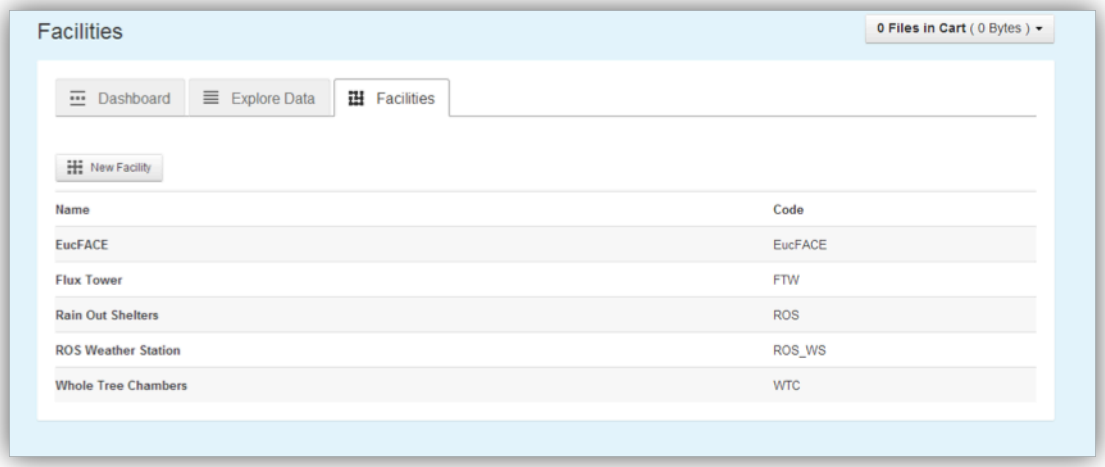
All data files uploaded to HIEv must be associated with an Experiment. In turn, each Experiment is associated with a Facility.

Therefore entries for Facilities and Experiments must be created before the associated data files are uploaded.

Facility — A Facility represents any instrument, or discrete set of instruments that are used in concert. This could be a multi-million dollar fixed Facility with hundreds of instruments and sensors, or a single piece of portable equipment that is taken out into the field.

Experiment — An Experiment is a set of related tests performed at a single Facility for which the data will be uploaded into HIEv.

Facility and Experiment entries are entered and updated using the Facilities tab on the main HIEv screen.



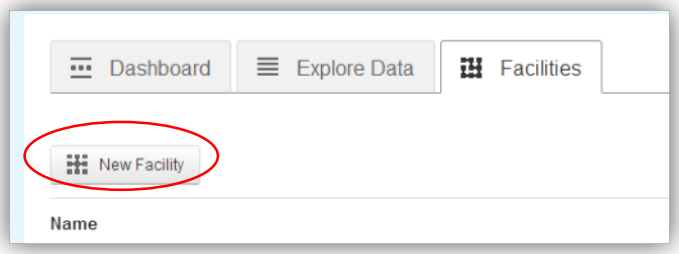
The Facilities tab lists all of the Facilities currently defined in HIEv.

6.1 Creating a Facility Entry

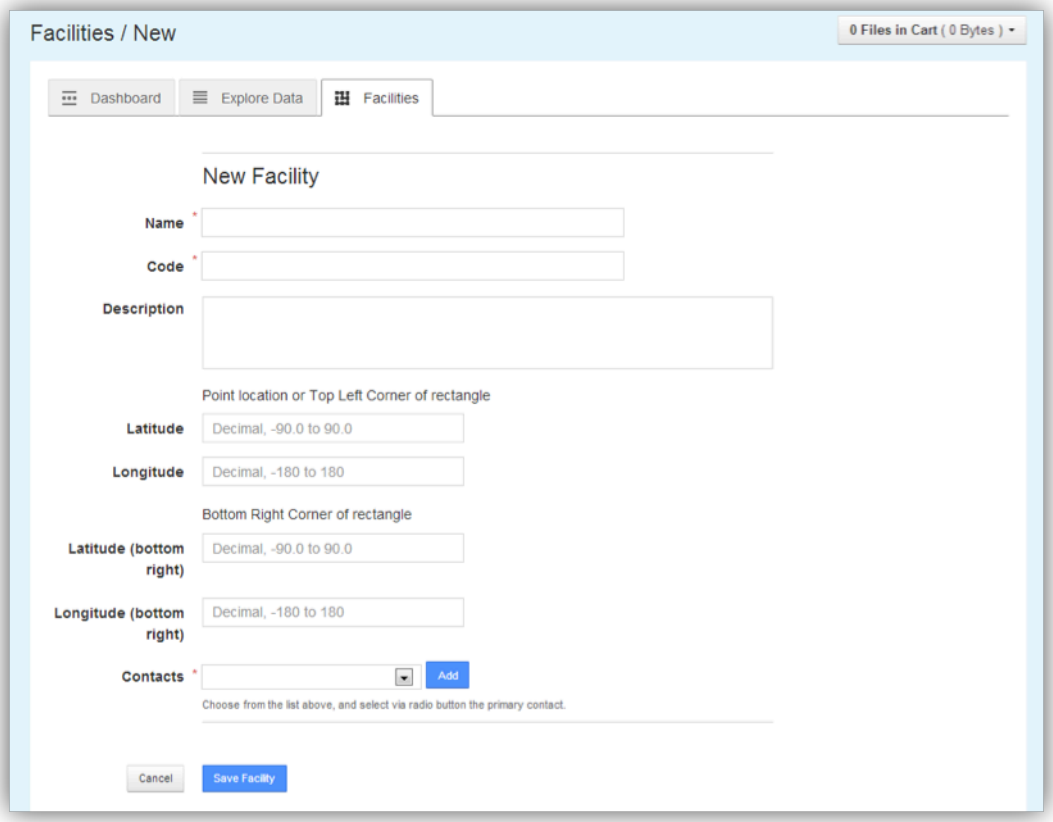
Note — Take care. Once created, a Facility entry cannot be deleted. This restriction prevents Experiments which reference a Facility from becoming invalid if the Facility were to be deleted.

To create a Facility entry:

- Click the **New Facility** button at the top-left of the Facilities tab.



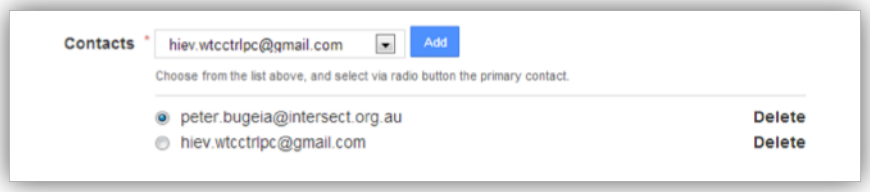
This will display a form that allows you to add details about the Facility you would like to create.



- Enter the details for the Facility.

Name	The Name for the Facility is a short, plain-English title that will be used in the application interface to refer to the Facility. The Name must be unique.
Code	The Code for the Facility is a short unique string. The Code must be unique.
Description	The Description of the facility should be as comprehensive as possible, describing details that would help a researcher both discover the facility when searching and assist that researcher in being able to interpret the data that is produced by the facility. These details would include things such as: • The purpose of the facility • Types of sensors installed at the facility

- Location of the sensors within the facility

Latitude and Longitude fields	The Latitude and Longitude for the facility are expressed in Decimal Degrees (http://en.wikipedia.org/wiki/Decimal_degrees) and can be taken directly from Google Maps. If a single set of co-ordinates is given, it is considered to be the central point for the Facility. Enter them only in the Latitude and Longitude fields, leaving the fields for the bottom-right corner empty. If two sets of co-ordinates are given they are considered to be a rectangle that bounds the facility. Enter the northwest corner in the first two fields and the southeast corner in the fields labelled "bottom right".
Contacts	The Contacts for the Facility are selected from the Users registered within HIEv. There must be at least one Contact for each Facility.  Select a Contact from the email addresses shown in the dropdown list and press the **Add** button to add it to the Contact list shown below this question. To remove an incorrectly added Contact, press the **Delete** button corresponding to that Contact's email address in the Contact list. Ensure that the primary Contact is highlighted correctly with the radio buttons to the left of the Contacts in the Contact list.

- Click on **Save Facility** to save the Facility's details and return to the Facilities list.

To abandon creating the Facility entry, click **Cancel**.

Once Facilities have been created they will appear in the list on the Facilities tab.

6.2 Editing a Facility Entry

To modify any of the parameters for a Facility entry:

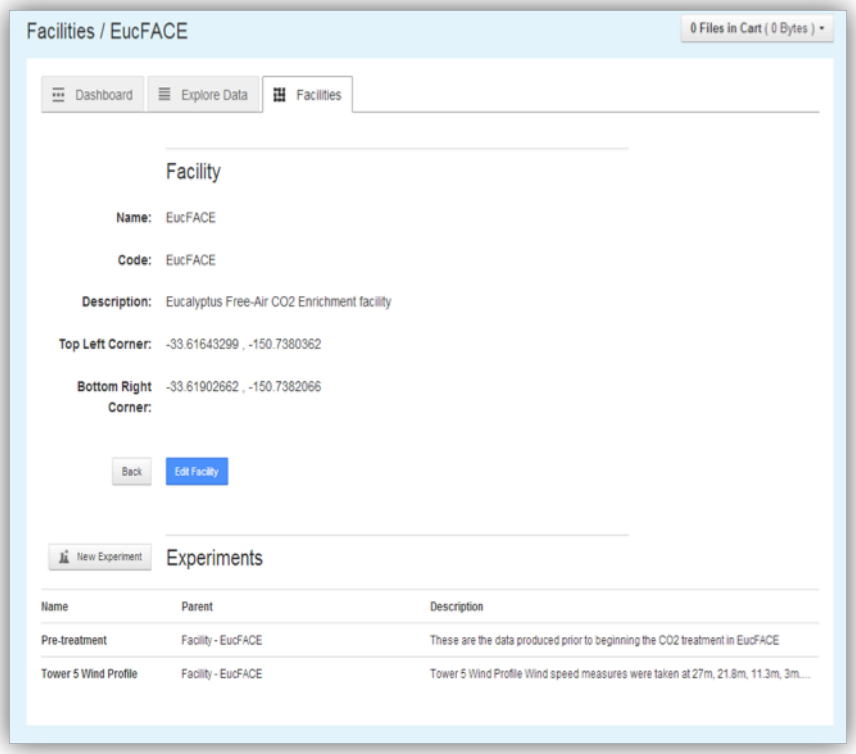
- From the Facility tab, click on the Facility Name of the entry you wish to edit. The Facility's details will be displayed.
- Click on **Edit Facility** to open the details edit screen, which is the same as the one described in *6.1 Creating a Facility Entry* above.
- Change the details as required.
- Click on the **Update** button so save your changes.

6.3 Creating an Experiment Entry

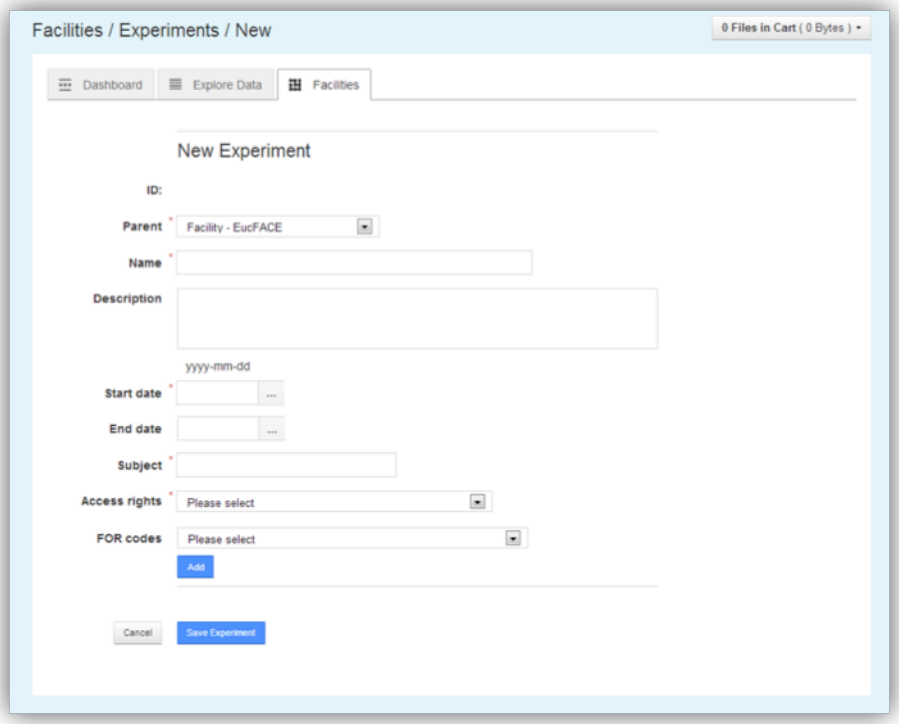
Note	Take care. Once created, an Experiment entry cannot be deleted. This prevents data files which reference this Experiment from becoming invalid if the Experiment were to be deleted.

To create an Experiment entry:

- On the Facilities tab, view the details of the Facility at which the Experiment will be run by clicking on the Facility name. The Facility details will be displayed and all of the Experiments registered at that Facility will be listed below the details of the Facility itself.



- Click on the **New Experiment** button at the top left of the Experiment list to show the Experiment parameters entry dialog.



- Enter the details for the Experiment.

Parent	The Parent for an Experiment is either the current Facility, or another Experiment running at that Facility. If an Experiment is selected, the new Experiment is considered a sub-experiment of the one selected.
Name	The Name for the Experiment should be short, but descriptive enough to uniquely identify the Experiment, including distinguishing an Experiment from those that are likely to come in the future. The Name should be unique for this Facility.
Description	The Description for the Experiment should describe the purpose of the Experiment and the techniques employed. Particular focus should be given to aspects of the Experiment that produce data that are stored in this system.
Start date	The Start date for the Experiment is the date that the Experiment was first considered to be active.
End date	The End date for the Experiment is the date that the Experiment concluded. This field should be left blank for Experiments that are currently active.
Subject	The Subject for the Experiment is a short phrase describing the Experiment's main research area. The Subject is primarily recorded to support publication to ANDS, who requests "A subject is a term, keyword, classification code or phrase representing the primary topic or topics covered by a registry object."
Access rights	The Access rights dropdown list box provides a selection of licences under which the data from this Experiment will be released. In Australia it is preferred that data is released under a

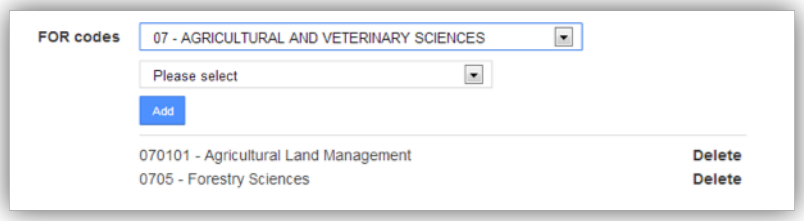
Creative Commons licence. The list of licences shown in the dropdown list is hard-coded in HIEv. More information about these licences can be found at http://creativecommons.org/licenses/.

FOR codes

Each experiment can have one or more Fields of Research (FOR) Codes. These FOR codes are the Australian and New Zealand Standard Research Classification (ANZSRC) codes. More information can be found at http://www.abs.gov.au/ausstats/abs@.nsf/Products/1297.0~2008~Main+Features~Chapter+3,Fields+of+Research?OpenDocument#11271429131099505 1, which says, in part,

> *The FOR is a hierarchical classification with three levels, namely Divisions (2 digits), Groups (4 digits) and Fields (6 digits). Each level is identified by a unique number.*
>
> *Each Division is based on a broad discipline. Groups within each Division are those which share the same broad methodology, techniques and/or perspective as others in the Division. Each Group is a collection of related Fields of research. Groups and Fields of research are categorised to the Divisions sharing the same methodology rather than the Division they support.*

A unique number identifies each level. In HIEv, FOR codes can be specified to Division, Group or Field of Research level.



The FOR codes are selected two digits at a time. After the Division is selected, a dropdown list box is displayed to allow selection from the Groups which are relevant to the selected Division. Similarly, after the Group is selected, a dropdown list box is shown to allow selection from the Fields of Research which are relevant to the selected Group.

Click the **Add** button to add the selected FOR code to the list for this Experiment, which will appear below the FOR code selection boxes.

You can click **Add** after selecting just the Division, the Division and Group, or all three levels, as appropriate for your experiment.

FOR codes can be removed from this Experiment by clicking the **Delete** button to the right of the FOR code in the list.

Note HIEv obtains FOR codes from a separate server, the details of which were set up during HIEv installation. If

this server is not available at the time you are creating or updating an experiment, you will not be able to select FOR codes. If this occurs, you should receive an error message and if so, you should report this to your system administrator. You should still be able to save your experiment and then later add FOR codes when the FOR server is available again.

- After you have correctly entered the Experiment details, click **Save Experiment** at the bottom of the page.

To abandon creating the Experiment, click **Cancel**.

6.4 Editing an Experiment Entry

To modify any of the details for an Experiment entry:

- From the Facility tab, click on the Facility name of the Facility which hosts the Experiment. The Facility details and Experiments for the Facility will be displayed.
- Click on the Experiment name for the Experiment you wish to modify. The Experiment details will be displayed.
- Click on **Edit Experiment** to open the Experiment detail entry screen, which is the same as the one described in *6.3 Creating an Experiment Entry* above.
- Change the details as required.
- Click on the **Update** button to save your changes.

6.5 Setting Up Experiment Parameters

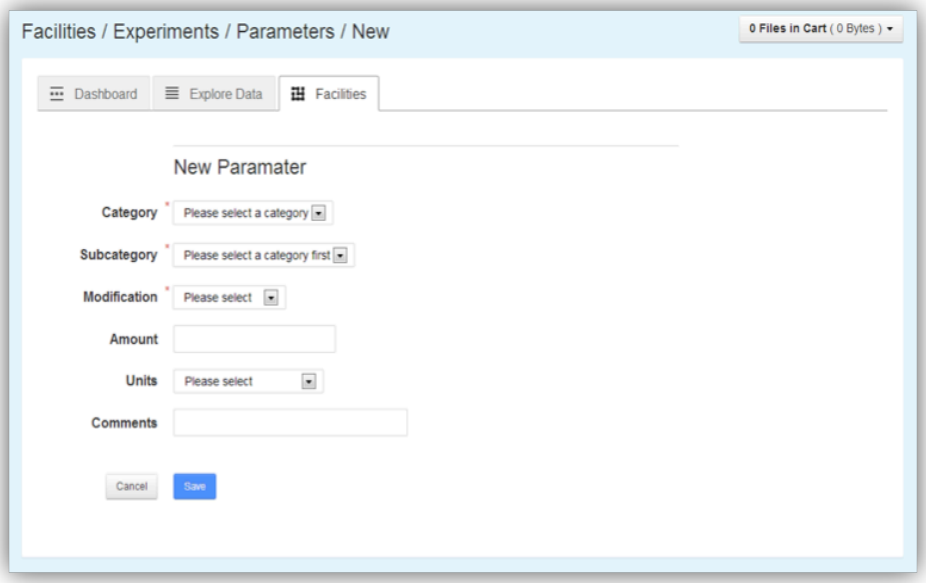
Experiments can optionally have one or more Experiment Parameters. These Parameters provide a structured way to describe Experimental treatments. They are documentary only. An Experiment can have multiple Parameters or none at all. By way of example, a Parameter may be the details of the raising of the CO2 levels within the tree chambers.

Because Parameters are not directly referenced by data files, the option to delete them is provided.

To add an Experiment Parameter:

- From the Facilities tab, click on the name of the Facility hosting the Experiment and then the name of the Experiment to which you wish to add a Parameter. The Experiment details will be shown.
- Click the **New Parameter** button directly below the Experiment details.

This button will display the form below.



- Enter the details of the Parameter. The System Administrator configures the values available in the dropdown list boxes for the questions in this form.

Category and Subcategory	These fields allow you to select the medium that is being modified from the dropdown lists. These fields are mandatory.
Modification	The Modification indicates the general way in which the medium has been modified. This field is mandatory.
Amount and Units	These optional fields allow more specific information to be recorded about the Modification.
Comments	This field can be used to record any unstructured, plain-text information you would like to record about the treatment.

- Click the **Save** button at the bottom of the form to save your Parameter details and return to the Experiment details screen.

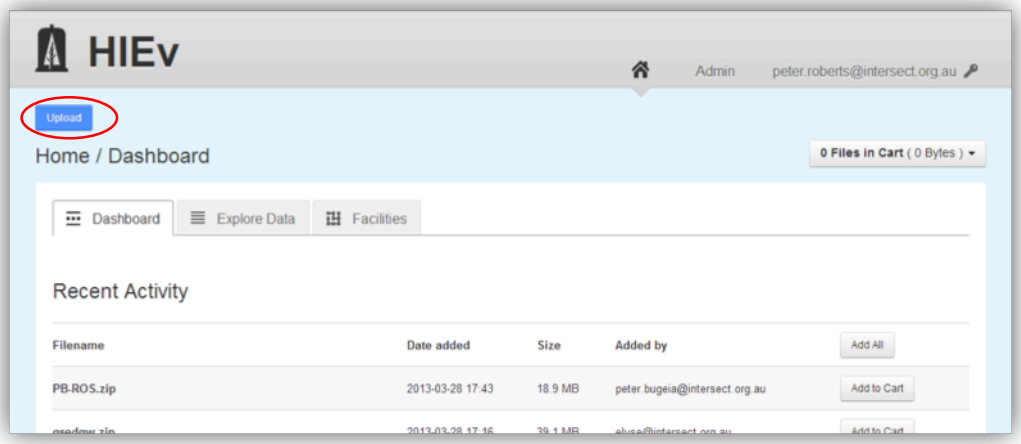
Once an Experiment Parameter has been created, it will appear in the Parameters list below the Experiment details.

Existing Experiment Parameters can be edited or deleted using the appropriate button to the right of the parameter in the **Actions** column.

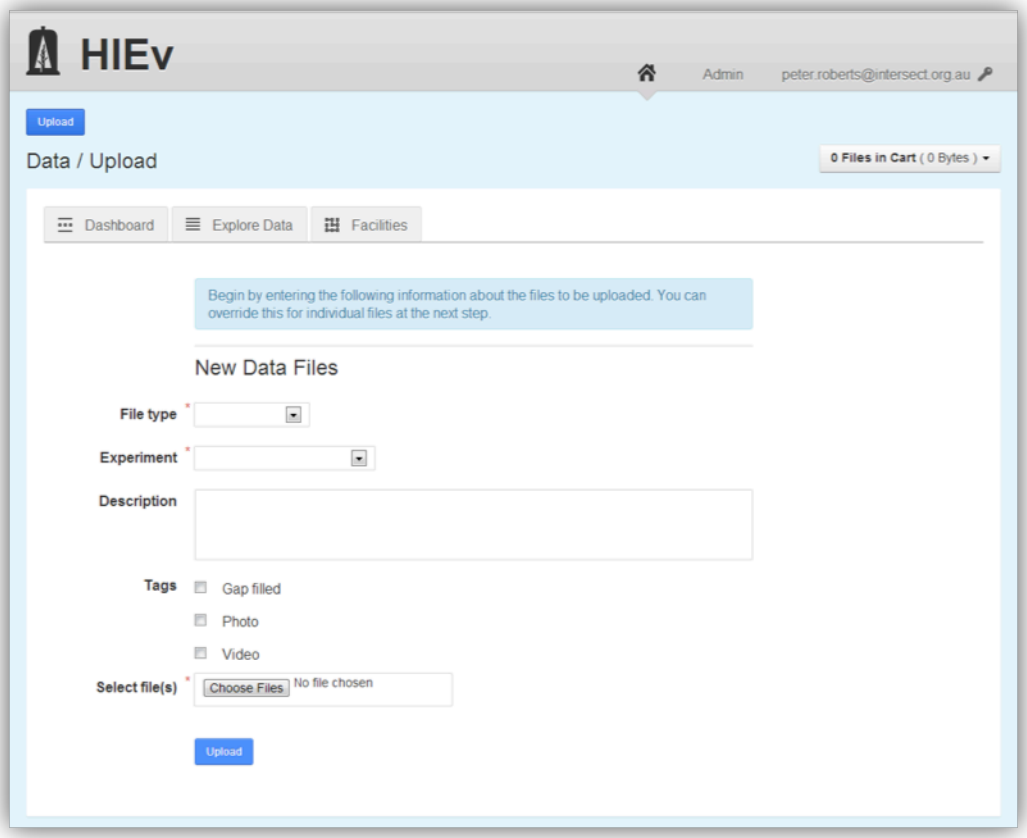
7 Uploading Data files

To upload one or more new data files:

- Click the **Upload** button which is displayed at the top left of the main screen.



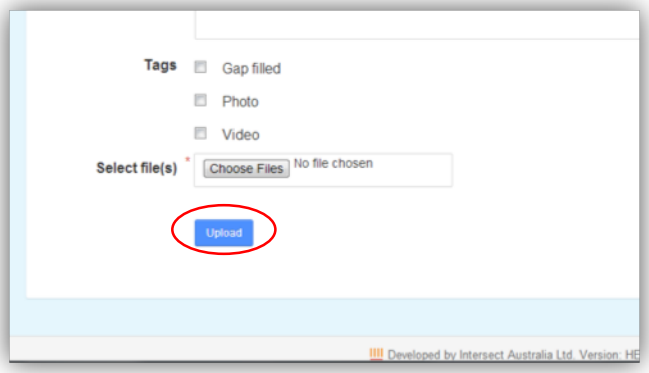
Clicking this button will take you to the New Data Files form.



- Enter all the Metadata that you want applied to the new file or files. See the section *5.1 Basic Information* for information about the meaning of Metadata fields.

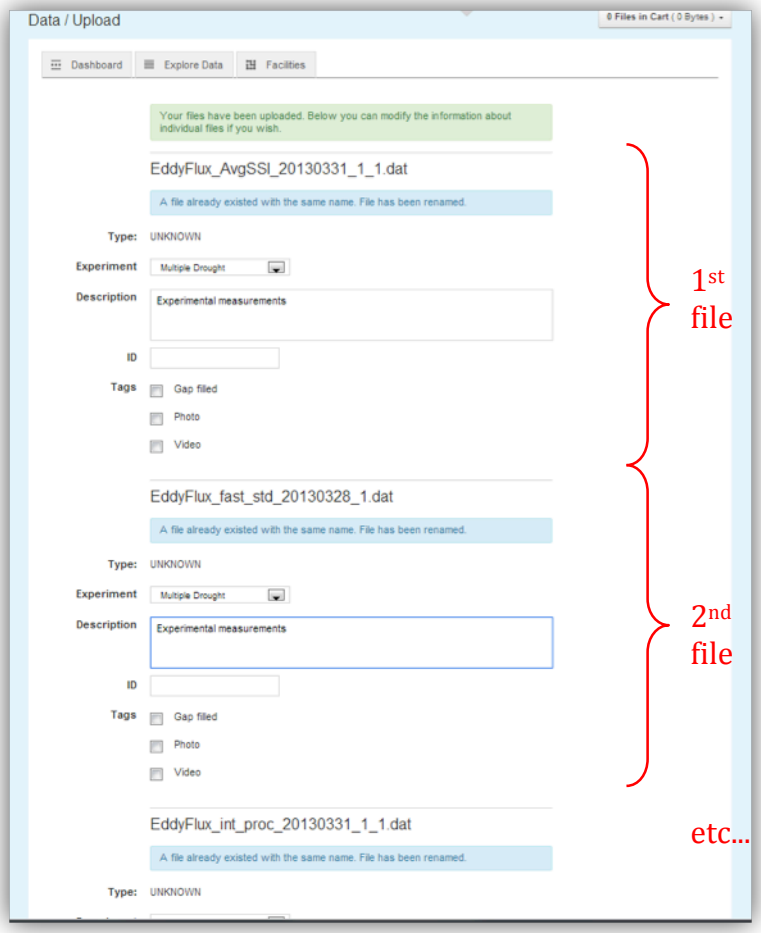
Note Files that are uploaded with a File Type of **RAW** and are also detected as containing valid TOA5 header information are treated as a special case. See section *7.1 Uploading RAW TOA5 data files* for more information.

- Click on **Choose Files** and use the file dialog which opens to select one or more files to be uploaded. When you confirm your file choice and return to the New Data Files dialog, a second **Choose Files** button will appear below the first.
- If you require more files to be uploaded, click on the new **Choose Files** button and use the file dialog to select one or more further files to be uploaded. Repeat this step using the new **Choose File** button which appears after each file selection dialog is closed until you have selected all the files you require to be uploaded.
- Click on the **Upload** button at the bottom of the screen to cause the files to be uploaded. (Do not click on the **Upload** button at the top of the screen. This button is the one you clicked to start this process. It will restart the file upload process and you will have to repeat the steps you have already done.)



After the file or files have successfully uploaded, a screen will be displayed with a separate section for each uploaded file displaying the Metadata applied to that file.

- For each file shown on this screen, adjust the Metadata for that file as you require.



  In particular, if the start and end dates and times for the data were not automatically extracted for the file, enter them now if required. See *4.2 Entering Dates and Times* for instructions on entering dates and times.

- Click on the **Update** button at the bottom of this screen to apply the modified Metadata to the files and complete the upload process.

If an uploaded file has the same filename as another data file that already exists within the system, HIEv will automatically suffix a unique number to the end of the original filename, before the file extension, to avoid the conflict.

7.1 Uploading RAW TOA5 data files

When a TOA5 CSV file is uploaded and the **Type** is specified as RAW, it is considered to become part of the canonical stream of data for that data logger. As a result, there will only ever be a single file with a **Type** of RAW that contains any given sample from a TOA5 data logger.

The system automatically determines the start time, end time and data logger that generated the data from the TOA5's header information and uses it to control processing of this file. Other metadata is also extracted from the TOA5's header information.

As a result of this processing:

1. If a TOA5 file is uploaded with a Type of RAW, and the file being uploaded is a complete superset of another file (or files) that are also RAW TOA5 files from the same data logger, the subset files will be replaced with this new file, regardless of the file names.
2. If a TOA5 file is uploaded with a Type of RAW that only partially overlaps an existing file of RAW data from the same data logger, the file will be uploaded, but its Type changed to ERROR and the original file(s) left in place.
3. If a TOA5 file is uploaded with a Type of RAW that overlaps an existing file of RAW data, but does not pass a sample-by-sample comparison with the original file(s), the file will be uploaded, but its Type changed to ERROR and the original file(s) left in place.

The Start and End Dates for the observations in a TOA5 file are extracted and stored as Metadata for the file. Therefore, when TOA5 files are uploaded as RAW data, HIEv does not permit the User to enter Start and End Dates.

7.2 Manual Data Upload Action Summary

The following table summarises the upload processing that will be done and the results for various combinations of file type and existing data.

IF...				**THEN...**		
File Type selected is	**File Format is**	**File Overlap is**	**And File**	**Resulting type will be**	**Resulting file name will be**	**Resulting messages to user**
RAW	TOA5	None	does not exist	RAW	as per uploaded	success
RAW	TOA5	None	exists	RAW	suffixed - see (1)	filename change (3)
RAW	TOA5	Safe	does not exist	RAW	as per uploaded	safe replacement (2)
RAW	TOA5	Safe	same as file being replaced	RAW	as per uploaded	safe replacement (2)
RAW	TOA5	Safe	exists (but is not the file being replaced)	RAW	suffixed - see (1)	safe replacement (2), filename change (3)
RAW	TOA5	Unsafe	does not exist	ERROR	as per uploaded	bad overlap (4)
RAW	TOA5	Unsafe	exists	ERROR	suffixed - see (1)	bad overlap (4), filename change (3)
RAW	Non-TOA5	N/A - only for TOA5	does not exist	RAW	as per uploaded	success
RAW	Non-TOA5	N/A - only for TOA5	exists	RAW	suffixed - see (1)	filename change (3)
Not RAW	TOA5	N/A - we don't check unless RAW	does not exist	as specified	as per uploaded	success
Not RAW	Non-TOA5	N/A - only for TOA5	exists	as specified	suffixed - see (1)	filename change (3)

Notes:

(1) suffixed means appending _1 (or the next available number) - e.g. blah.dat becomes blah_1.dat (or blah_2.dat if blah_1.dat already exists)

(2) MESSAGE: The file replaced one or more other files with similar data. Replaced files: <filenames here>

(3) MESSAGE: A file already existed with the same name. File has been renamed.

(4) MESSAGE: File cannot safely replace existing files. File has been saved with type ERROR. Overlaps with <filenames here>

7.3 Automating the upload of data to HIEv

As well as via the web interface, data can be uploaded to HIEv using an HTTP-based API. The upload of data into the system is facilitated through a Ruby script. On Windows, there is also a .BAT script which wraps this Ruby script.

This aspect of HIEv operation is beyond the scope of this manual. Instructions and downloadable scripts for Windows can be found in the HIEv WIKI documentation on GitHub at https://github.com/IntersectAustralia/dc21/wiki/Setting-Up-Automated-Load-From-PC.

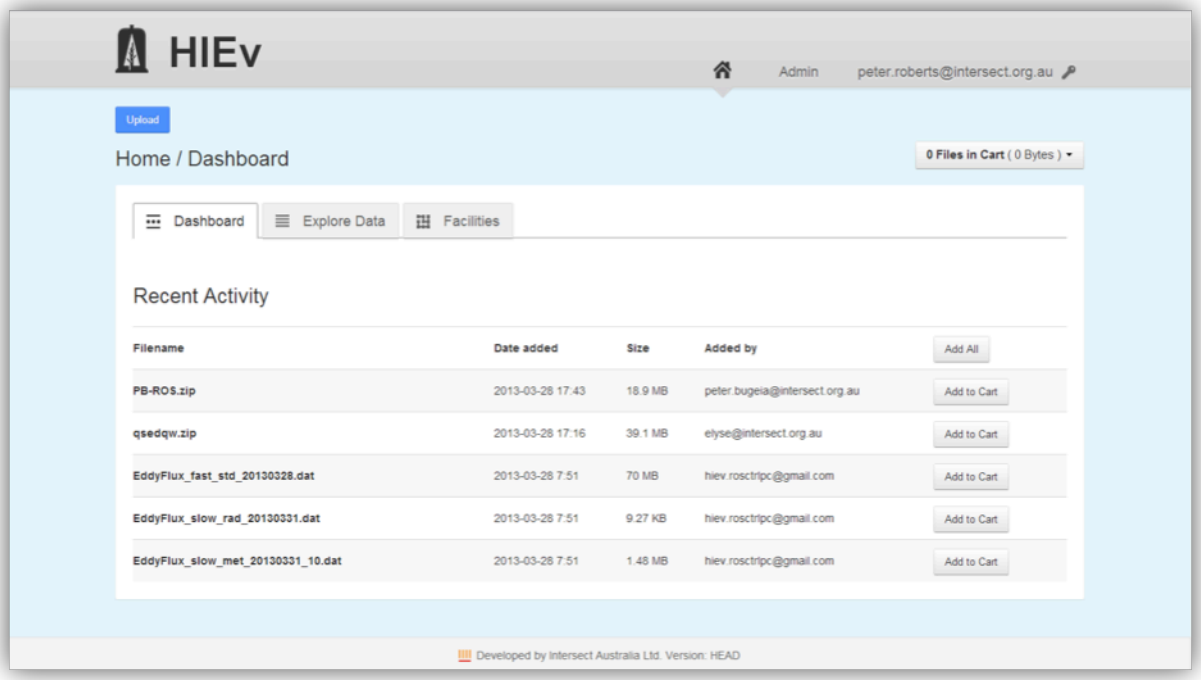
8 Managing Data Files

The key views of the files uploaded into the HIEv database are the Dashboard and Explore Data views. These views allow you to perform the key functions of HIEv.

Access these views by clicking on their respective tabs on the HIEv Home Screen.

8.1 The Dashboard Tab

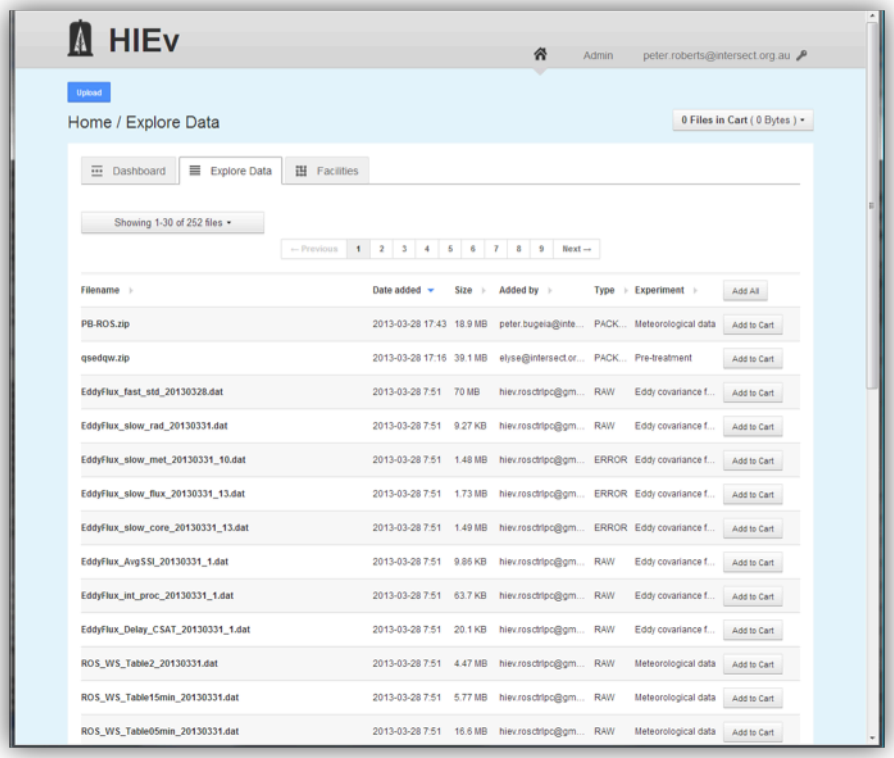
The default tab on the Home Screen is the Dashboard tab. It shows a list of the five files which have been uploaded or packaged most recently by all Users of HIEv.



The operation of the Dashboard tab is similar to the operation of the Explore Data tab, except that searching and sorting functions are not supported. See the description of the Explore Data tab below for more information.

8.2 The Explore Data Tab and File Searching

The Explore Data tab provides the main data management functions of HIEv. The initial view shows all data files which have been uploaded. If there are more files than fit on one screen, only the first 30 files will be shown, and the subsequent files can be shown by paging through the data using the page number buttons.



8.2.1 Sorting

Click on the heading of any column in the file list in this view to sort the files into increasing order by that column. Click again to reverse the sort order.

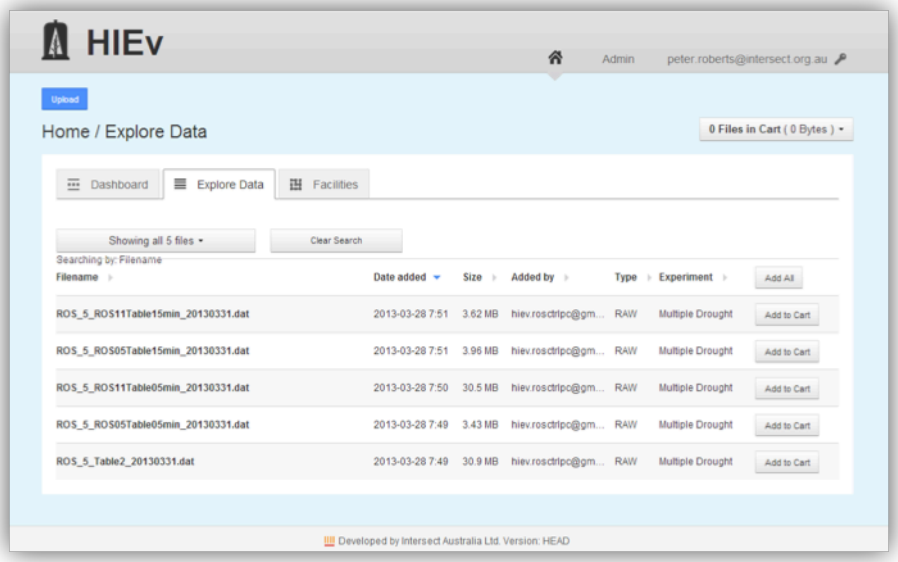
A triangular up or down arrow to the right of any column heading indicates the active sort order. A grey right-pointing arrow indicates that the file list is not sorted by that column.

Re-sorting the list always resets the display to the first page of the file list.

Sorting can be done by only one column at a time.

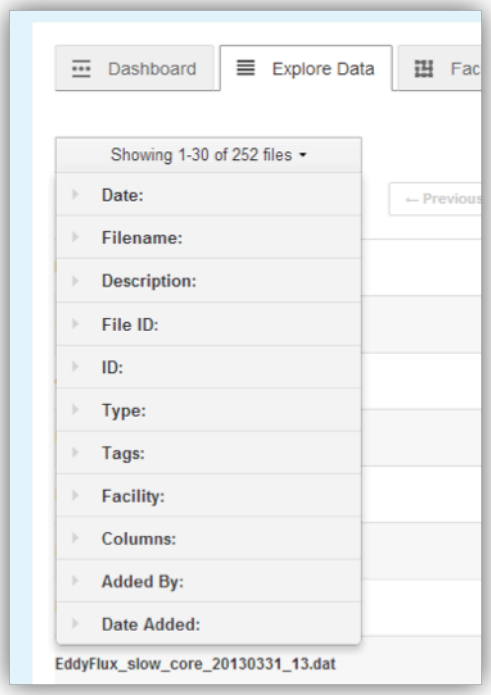
8.2.2 Searching

When the number of files uploaded to your system becomes large, finding the files you are interested in may become difficult. The Search facility is provided to assist. It supports searching for data files using the Metadata that was supplied at the time each file was uploaded.



The screen above shows the Explore Data tab when a search is active, restricting the number of files shown. Note the **Clear Search** button, which is only present when a search is active. Click this button to return to displaying all files.

To set or change the search conditions, click on the **Showing...files** button, which will show a dropdown list of search parameters. The exact text of this button changes, depending on how many files are presently shown in the file list.



If there is a search active, one or more of the search categories in this menu will be expanded when you first display it. These are the search categories which have active search data.

Click on the Metadata field by which you wish to search to expand it.

If you specify more than one search condition using more than one Metadata field, the file list will display only those files which satisfy *all* of the conditions you specify.

Click on any **Update Search Results** button to cause HIEv to redo the search based on all the criteria you have entered. These buttons will appear when you open any of the Metadata search fields.

8.2.2.1 Regular Expressions

Regular expressions are used for searching for specific substrings in general text. They are used widely across many computer systems. HIEv uses regular expressions to provide comprehensive search functionality for Filenames, Descriptions and IDs.

A few of the more useful functions of regular expressions are described briefly below. However, a comprehensive description of regular expressions is beyond the scope of this manual. Users can read a thorough description at http://www.regular-expressions.info/reference.html.

HIEv's use of regular expressions is not case sensitive. Therefore, you can enter either upper or lower case characters and get the same result.

The following few examples are provided as a quick start to using regular expressions.

Simple string		A simple string containing no special characters will match that string, regardless of where within the target string it occurs. **Example:** `mpl` will match the string `example` at the fourth character.
Start of string	^	The character ^ will match the beginning of the string. **Example:** `^exa` will match `example`, but `^xa` will not.
End of string	$	The character $ will match the end of the string. **Example:** `le$` will match `example`, but `pl$` will not.
Any character	.	The period character will match any character. **Example:** `a.c` will match `abc`, `aac`, `adc`, `a7c` and `a-c` but it will not match `ac` or `abbc`.
Repeated character	*	Asterisk causes matching to zero or more repetitions of the preceding character. **Example:** `ab*c` will match `ac`, `abc`, `abbc` or `abbbc`, but will not match `a7c` or `ahc`. It will match `aac` at the second character and `acc` at the first character, because there are zero characters between the `a` and `c` in those strings.
Repeated characters	+	The plus sign causes matching to one or more repetitions of the previous character. **Example:** `ab+c` will match `abc`, `abbc` or `abbbc` but will not match `ac`.

Alternate characters	[ ]	Strings enclosed within square brackets will match any one of the characters within the brackets. **Example:** `a[123]b` will match `a1b`, `a2b` or `a3b` only. It will not match `ab` or any other substring.
Character ranges	[-]	Use – between `[]` to match one of range of characters. **Examples:** `[0-9]` matches any digit. `[a-z]` matches any letter. `[a-z0-9]` matches any digit or letter.

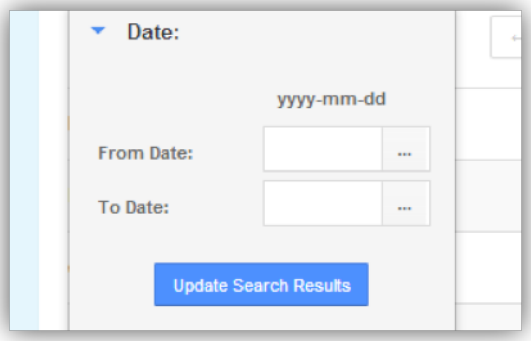
Escape character	\	In order to match a special character, precede it with the backslash character.
		Special characters are `[\^$.|?*+(){}`
		Putting `\` before other characters often has a special meaning, so should be avoided.
		Examples: `\\` will match `\`
		`\.` will match `.`
		`\*` will match `*`
		`\[` will match `[`

Combinations		Any of the above search methods can be combined. **Examples:** `^.c` will match any string with `c` as its second character. `^abc$` will match the string `abc` only. `abcd` or `aabc` will not be matched. `1[abcd]+2` will match any combination of the characters `a`, `b`, `c` or `d` which occurs between the digits `1` and `2`. `[\[\]]` will match either `[` or `]`. `\.+` will match any run of periods. `[0-9]+` will match any integer number. `[0-9]+\.[0-9]*` will match any number with a decimal point.

If an invalid regular expression is entered, HIEv will place an error message at the top of the screen, clear the search field and ignore the regular expression.

8.2.2.2 Restricting by Data Date

The **Date** field allows you to search for files based on the start and end date specified in the file's Metadata.



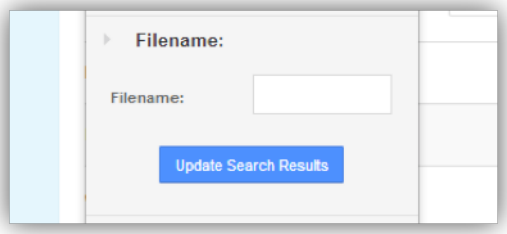
Dates can be entered in either the **From Date**, **To Date** or both. If only a **From Date** is specified, all files containing data on or after that date will be included. If only a **To Date** is specified, all files containing data for before or on that date will be included. See *4.2 Entering Dates and Times* for instructions on entering dates.

If you restrict by date, files which have no dates in their Metadata will not be displayed.

For TOA5 files, this search option checks the Start and End Dates in the Information from the File (see section *5.2 Summary Information Extracted from TOA5* Files) and for all other files, this search option checks the Basic Metadata Information (see section *5.1 Basic Information*).

8.2.2.3 Restricting by Filename substring

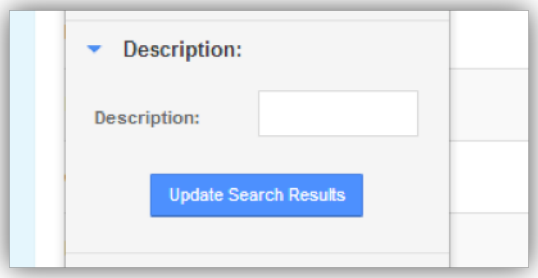
The **Filename** field allows you to search for files based on their filename.



HIEv treats the search string as a regular expression. See section *8.2.2.1 Regular Expressions for more information.*

8.2.2.4 Restricting by Description substring

The **Description** field allows you to search for files based on their free-form text descriptions.

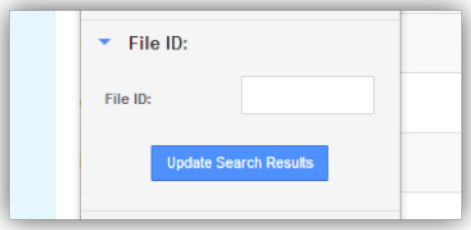


HIEv treats the search string as a regular expression. See section *8.2.2.1 Regular Expressions* for more information.

Use the search string ^$ to search for files without any Description.

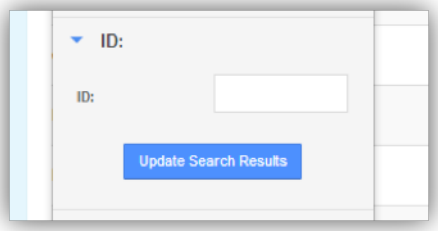
8.2.2.5 Restricting by File ID

Enter an integer number into the File ID search field to display only the file with that File ID. It is not usual to use this search method with any other search criterion, as this method will always display exactly one file, or no files if there is no file with the entered File ID.



8.2.2.6 Restricting by ID

The **ID** field allows you to search for files based on their entered ID.

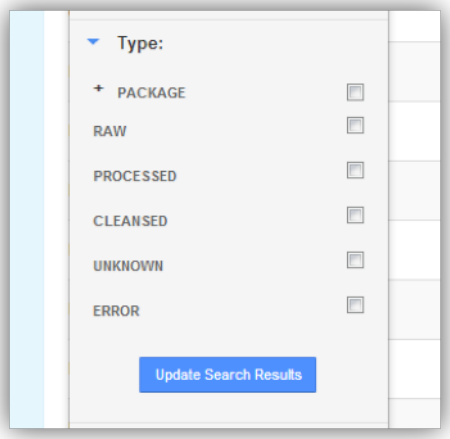


HIEv treats the search string as a regular expression. See section *8.2.2.1 Regular Expressions* for more information.

Use the search string `^$` to search for files without any ID, or `^ *$` to search for files with an ID consisting of just spaces.

8.2.2.7 Restricting by File Type

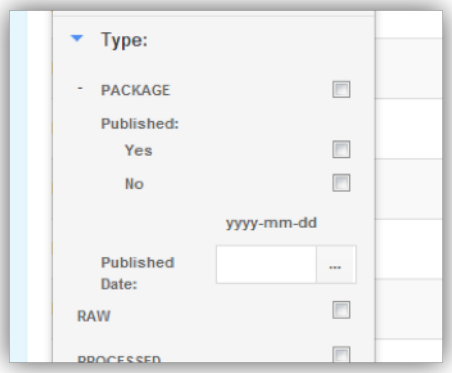
The **Type** search parameters allow you to search for files based on their specified type.



The set of possible types is displayed as a list of checkboxes. Selecting none of the checkboxes is the same as selecting them all – that is, files will not be filtered based on their

Type. When at least one checkbox has been selected, only files of that type will be returned in the search results. More than one type can be selected.

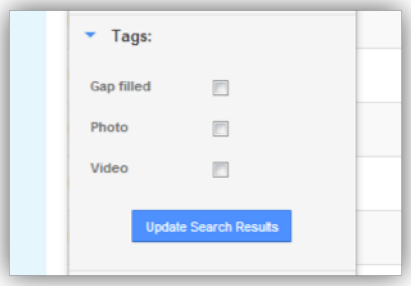
Click on the **+** sign to the left of the **Package** option to open further search conditions for Packages. (Clicking on the word **Package** itself sets its checkbox.) You can click on the minus sign to close it again.



Select the **Yes** or **No** checkbox to display only files which are, or are not, Published. To show a Package file regardless of its Published status, leave both checkboxes unchecked. If a **Published Date** is entered, then only files Published on that date will be displayed. See *4.2 Entering Dates and Times* for instructions on entering dates.

8.2.2.8 Restricting by Tags

The **Tags** interface allows you to search based on the tags that have been assigned to a file.



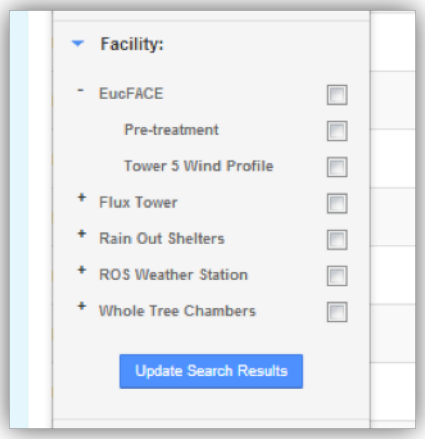
Like the Type interface, selecting none of the available checkboxes means that files will be returned in the search results regardless of the tags they have. Once a checkbox is selected, only files that have the corresponding tag will be listed.

If more than one Tag is selected, all files with any of those tags will be listed.

It is not possible to search for files which do not have a specific tag.

8.2.2.9 Restricting by Facility

The **Facility** interface allows you to search for files based on the Facility or Experiment that produced the file.

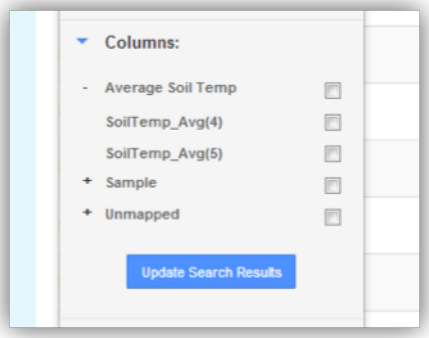


The interface is a two-level hierarchy of checkboxes. The top level shows all the Facilities defined in the system and the second level shows the Experiments that are defined for those Facilities. Selecting a Facility selects all of the Experiments for that facility. If only specific Experiments are required, clicking on the **+** sign to the left of a Facility will expand the hierarchy and allow individual Experiments to be selected or deselected.

Selecting more than one Facility or Experiment causes all files for any of those Facilities or Experiments to be listed.

8.2.2.10 Restricting by Data File Columns

The **Columns** interface allows you to search for TOA5 format data files that contain specified columns.



Like the Facility interface, this shows a two-level hierarchy of checkboxes. The top level contains all the Names that are defined in the Columns Mapping table. When these Names are expanded by clicking the **+** sign next to them, they show all of the TOA5 column headings which map to these Names. The last top level group is an extra group called **Unmapped** that contains all the TOA5 column headings that appear in any file in the entire HIEv data store but are not mapped to a standard Name in the Column Mappings table. See *11.3 Managing Column Mappings* for more information.

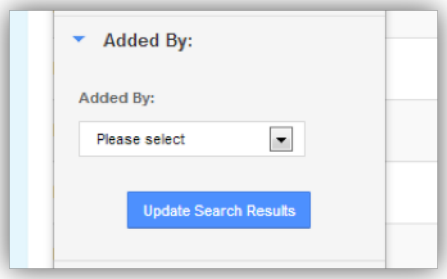
Selecting the checkbox for a standardised top-level Name will select all the TOA5 column headings that are mapped to it. Clicking the **+** sign to the left of the top-level Name will allow you to select them individually.

Selecting more than one checkbox will cause any file which has any of those corresponding TOA5 column headings to be listed.

If you set any checkbox in this search function, only TOA5 data files will be listed.

8.2.2.11 Restricting by Person Who Added the File

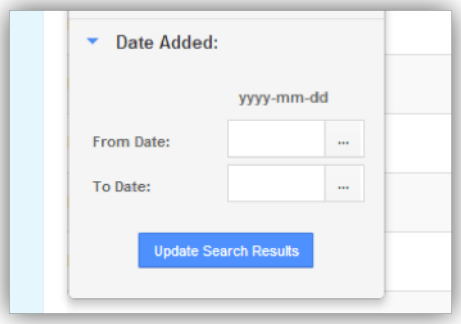
The **Added By** field allows you to search for files that were uploaded by a specific User.



The User must be selected from the list of a Users registered in the system.

8.2.2.12 Restricting by Upload Date

The **Date Added** interface allows you to search for files that were uploaded between a specified set of dates.



Like the **Date** field, the **Date Added** interface allows you to select a **From date** and a **To Date**. If both dates are specified only files uploaded between those dates will be included in the search results. If only a **From Date** is specified, all files uploaded after that date will be included. If only a **To Date** is specified, all uploaded before that date will be included. See *4.2 Entering Dates and Times* for instructions on entering dates.

8.3 The Cart

The Cart operates like an e-Commerce shopping cart. HIEv provides functions for adding files to the Cart and for doing operations, such as Downloading and Publishing, on all files in the Cart. See more information about Downloading in *Chapter 10 Downloading files* and Publishing in *Chapter 9 Publishing Your Data*.

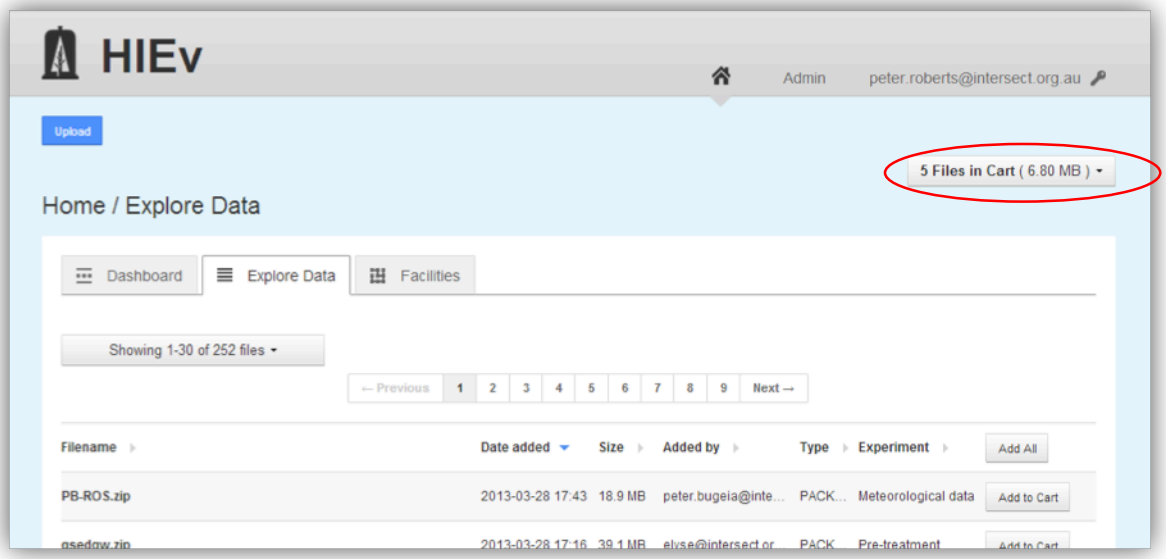
Unlike most e-Commerce systems, the content of the Cart persists between login sessions.

If a file is deleted from the system or is replaced by a new upload, that file will disappear from all Users' Carts.

Note	There is only one Cart per User account. If two people simultaneously use the same login account, the results can be unpredictable.

Add a file to the Cart by clicking on any **Add to Cart** button for that file. There are **Add to Cart** buttons in multiple places, including the Dashboard file list and the Explore Data file list.

The Cart status box, circled in red below, shows the number of files in your Cart and the total size of these files. It appears at the top right of many HIEv screens.

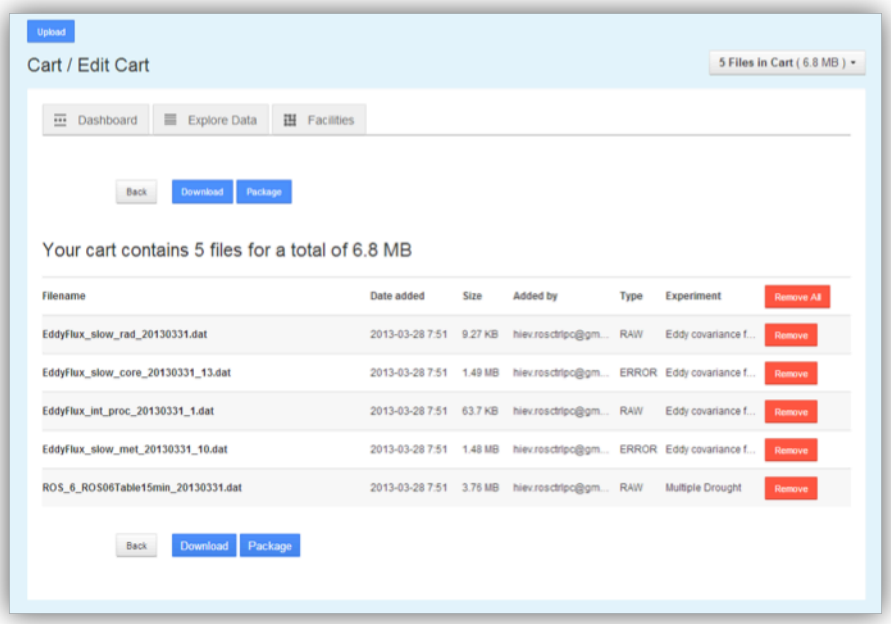


Click on the Cart status box to show a dropdown menu of operations which can be performed on the files in the Cart. These operations are:

Download	Click on this option to download data files to your local computer. See *Chapter 10 Downloading files* for instructions on using this feature.
Package	Click on this option to create a publishable Package containing all files in the Cart. See section *9.1 Creating a Package* for instructions on using this feature.
Clear cart	Click on this option to remove all files from the Cart. It does not delete the files themselves.
Edit cart	Click on this option to view a list of the Cart contents, remove individual files from the Cart, download all files in the Cart or create a publishable Package using all files in the Cart. See section *8.3.1 Editing the Cart* for more details.

8.3.1 Editing the Cart Contents

Selecting the Edit Cart option will display the following screen.

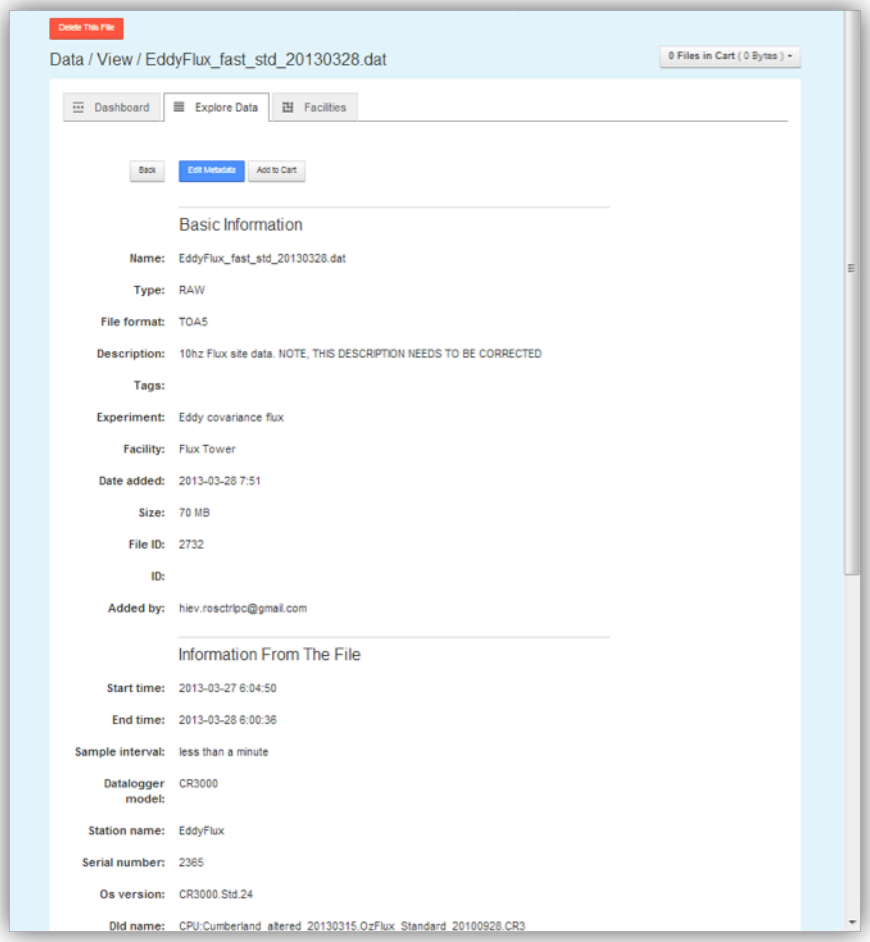


There is one line in the table on this screen for each file in your Cart. These files can be removed individually by clicking on the **Remove** button on the relevant line. If the **Remove All** button is clicked, the Cart will be emptied and further Cart operations cannot be performed until files are again added to the Cart.

Package and Download buttons are also available on this screen. See section *9.1 Creating a Package* and *Chapter 10 Downloading files* for more information about the operations initiated by these buttons.

8.4 Viewing and Editing a File's Metadata

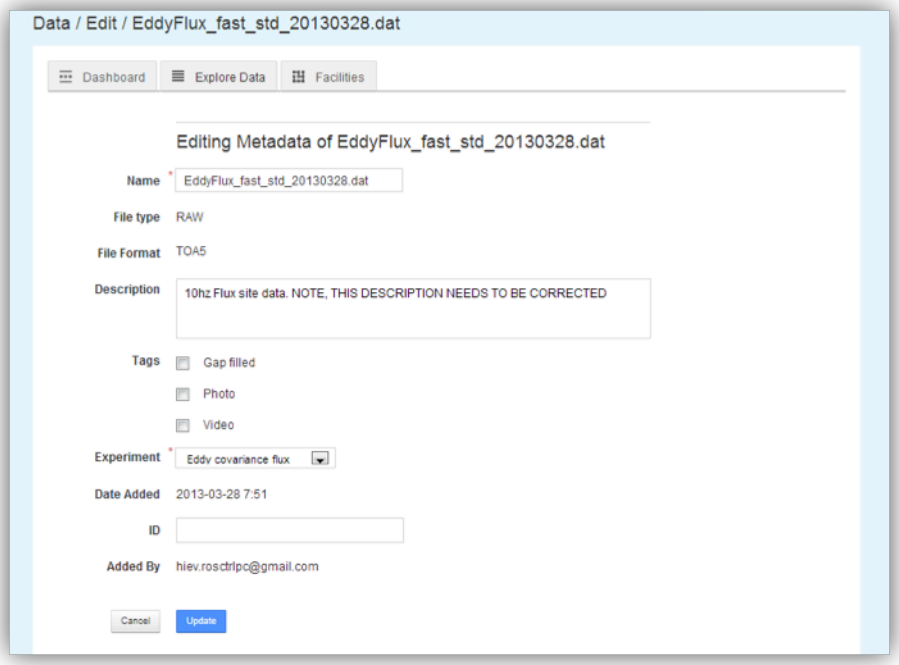
Clicking on any filename in the Filename column of the Dashboard tab, Explore Data tab or Edit Cart view will display the Metadata for that file in a screen similar to the following.



Files of all types will display the Basic Information section showing the Basic Metadata, although some fields are not shown for some file types. TOA5 files also show **Information From The File** and **Columns** sections. See *Chapter 5 HIEv Data File Storage and Metadata* for more information.

There is a button which says **Add to Cart**. Click this button to add this file to your Cart. If the file is already in your Cart, then the button will change to say **Remove from Cart**, and you can use it to remove the file from your Cart.

There is also a button to edit the file's Metadata. Clicking this button will take you to a form that allows you to modify the file's Metadata.



Once you have finished editing the Metadata, click the **Update** button to save your changes. If all your changes are valid, the Metadata will be updated and you will be returned to the Metadata view.

Note	You will only have permission to edit the Metadata of this file if you are logged in with the User credentials which were used when the file was uploaded, or you have Administrator permission.

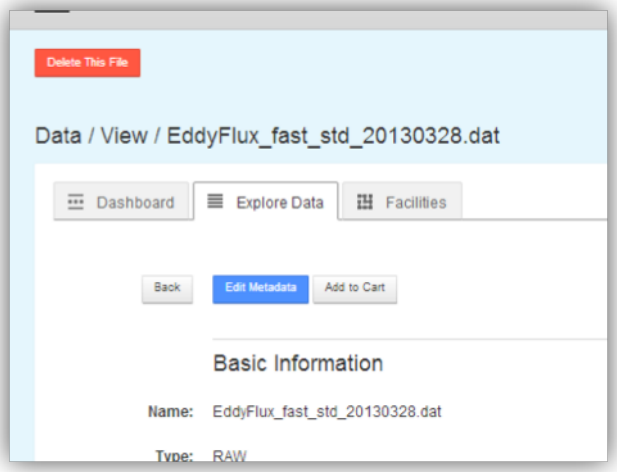
8.5 Deleting a Data File

Note	Deleting files removes the file completely from HIEv. *It is irreversible.*

In order to delete a file from the HIEv database, first view the Metadata for that file by clicking on the file name in either the Dashboard or Explore Data views. If you have permission to delete that file, a red **Delete This File** button will appear at the top of the Metadata View screen.

Note	You will only have permission to delete this file if you are logged in with the User credentials which were used when the file was uploaded, or you have Administrator permission.

You cannot delete a file which has a non-null ID. In order to delete it, you must clear the ID first. This restriction makes it more difficult to inadvertently delete a Package which has been Published and has had its Published ID entered.



When you click on the **Delete This File** button, you will be asked to confirm that you do wish to delete the file.

9 Publishing Your Data

When enough data has been produced from a facility to warrant publishing (for example, six months of weather station data) or a researcher wants to publish their research paper and site their research data package, the data collected can be Published to the Australian Research Data Commons.

Before Publishing, the data files to be published must be combined into a single Package file. Package files are ZIP files which use the Bagit format, which is described in *Appendix A -The Bagit format*.

Once created, Package files are shown sorted with the other data files in the file list which is displayed on the Explore Data tab.

Exactly which files should be included in a Package intended for Publishing is largely dependent on what is meaningful for the data and research discipline in question. It is entirely valid to have the same data appear in multiple Package files if that will aid discovery and reuse for other researchers. For example, a large set of data could be divided into two smaller, but overlapping, sets of data that represent different lenses (research problems) that the data could be seen through.

Once a Package has been published the Metadata describing the Package will be made available for harvesting via the OAI-PMH protocol. After this has occurred, it will become discoverable in Research Data Australia.

The process of Publishing involves a few steps:

- The User creates a Package which contains the data files to be Published using the HIEv Publish function. This step copies the selected data files' Metadata into a RIF-CS file and copies the data files themselves into a Package ZIP file. These two files are logically linked together.
- The User Publishes the Package using HIEv's Publish function. This copies the RIF-CS file into a specific location so that it can be harvested by the OAI-PHM harvester. It also sets the Published flag and Published Date field in the ZIP file's Metadata.
- At some subsequent time, the OAI-PHM harvester will discover the RIF-CS file. The harvester copies the RIF-CS file and the Packaged ZIP file it refers to into the Published data store. There is no indication in the Package's HIEv Metadata when or if this has occurred.

Note	Files in a Packaged ZIP file are copies of the files in HIEv. Changing the files in HIEv will not affect the files in the Packaged ZIP file.
Note	The exact details of the harvesting operation may vary from site to site. At some sites, the Packaged ZIP file may not be copied during harvesting and may be accessed directly by external data users.
Note	The RIF-CS file is created when the Package is created, and the RIF-CS file is the file that is Published. Therefore, changes to the Package Metadata after creation do not affect what is Published.

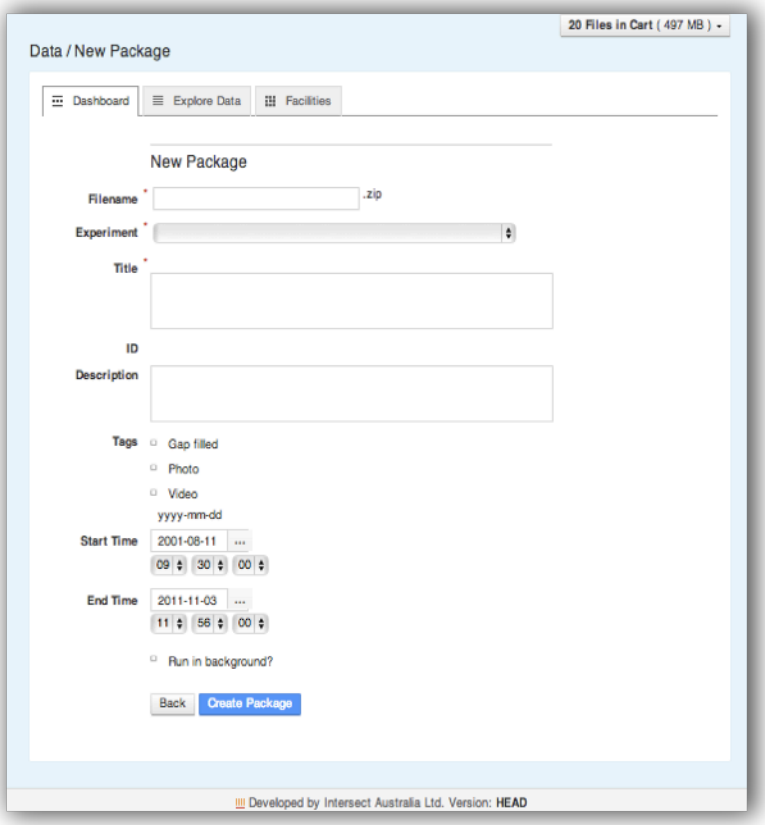
9.1 Creating a Package

Creating a Package creates two related components:

Bagit ZIP File	This ZIP file contains a snapshot of all the data for this Package. This includes copies of the data files and a Readme.HTML file which contains a copy of the all of the data files' Metadata. This Readme.HTML file is both human readable and machine readable. See *Appendix A - The Bagit format* for more information.
Matching RIF-CS File	This file contains a copy of the Package Metadata which is entered at the time the Package is created. Only selected parts of the Metadata from the associated Experiments and Facilities are copied into the RIF-CS file. See *Appendix B - RIF-CS* for more information.

To create a Package containing one or more files:

- Determine the external Package ID that will be assigned to this Package. See the HIEv Application Manager for information on how this ID is determined on your site.
- Add the required file or files to your Cart, ensuring that the Cart contains only those files you wish to include in your Package. See *Chapter 8.3 The Cart* for instructions on using the HIEv Cart.
- Click on the Cart status box to open the Cart dropdown menu.
- Select **Package** from the dropdown menu. Alternatively, selected **Edit Cart** from the dropdown menu and then click on the **Package** button on that screen, as shown below. If this method is used, the Cart can be reviewed prior to Packaging.
- The New Package screen will be displayed.



- Enter the Metadata to be associated with your Package. See section *5.1 Basic Information* for details of this Metadata. Note that you cannot enter a value for the ID field – it will be automatically generated when the package is created.
- Choose whether you want to run the package creation as a background task by checking the "Run in background?" checkbox (this is the default). You should run the package creation as a background task if you think it might take a long time to create the package. You can uncheck the checkbox but this will mean you will have to wait until the package is created until you can continue to use the application.
- Click on **Create Package** to cause your Package file to be created and saved. The ID field will be automatically generated as part of the package creation process. The Package file can now be viewed in the Explore Data tab. Note that if you click on **Back**, you will be returned to the Explore Data tab and the Package will not be created.

Note It is very important to check the Package Metadata **Create Package** This Metadata, including selected Metadata from the included data files and the Facilities and Experiments they reference, is copied into the RIF-CS file immediately after pressing this button. The data in the RIF-CS file cannot be edited. If it is wrong, the Package must be deleted and re-created.

Note If "Run in background?" was checked when you pressed **Create Package,** the package creation is queued to be run in the background and you can continue to use the HIEv application while it is being created. You can monitor the progress of the creation by refreshing the browser window and looking at the "Creation Status" and "Package progress" fields. See section *5.1 Basic Information* for details of these fields.

Note If a background task fails to complete, there is an administration function available to see more information and to kill it if necessary. See section *11.4 Managing Background Tasks* for details.

It is possible to create a Package which contains other Packages. There may be circumstances when this is meaningful.

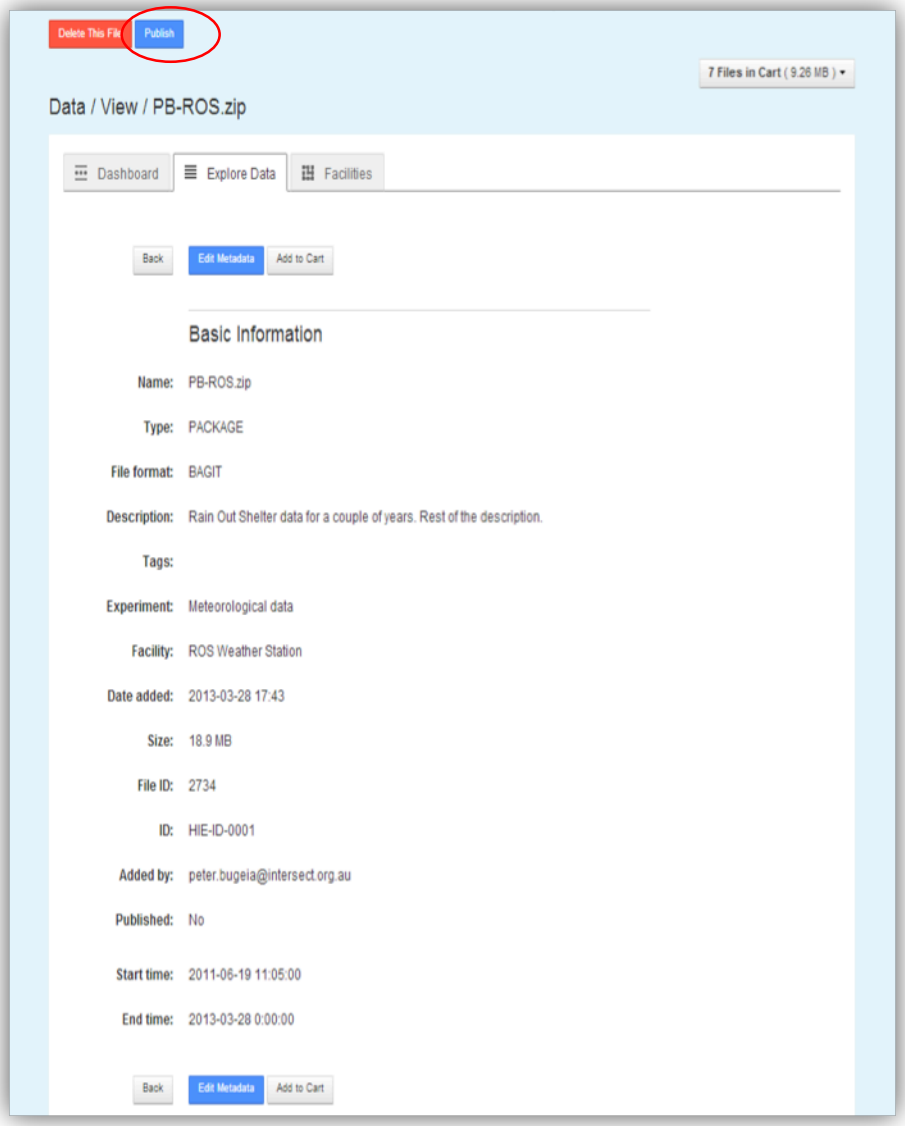
9.2 Publishing a Package

When a Package has been created and its Metadata is correct, it can be Published.

Publishing can be done by Users with Administrator permissions only.

To Publish a Package:

- View the Package's Metadata by clicking on its filename on either the Dashboard tab or Explore Data tab. Review its Metadata to ensure you have selected the correct Package file and that it is ready to Publish.



- Click on the **Publish** action button at the top of the Metadata screen. This **Publish** button will not appear if the Package has already been Published. A Package cannot be Published twice.
- A dialog box is shown for you to indicate that you are sure you wish to proceed. Click on **OK**. If you click on the **Cancel** button, you will be returned to the Package Metadata screen.

9.3 Managing Published Packages

Note	Considerable care must be taken when managing Published Packages. You should have a thorough understanding of the way the harvesting and storage of Published Packages is configured on your system before using the functions described in this section.

Once Packages are Published using the HIEv Publish function, they are available for harvesting, but they are not necessarily harvested promptly. Depending on how your system

is configured, it may take some time for them to be harvested, perhaps even days. In addition, you cannot tell from HIEv if the Package has already been harvested or not.

9.3.1 Publishing a second time

HIEv prevents you Publishing a Package more than once. However, Publishing a second time is possible if a new package with the same data is created. Such Publishing is generally harmless, even if the original version has already been harvested. The exact results will depend on the way your system is configured.

9.3.2 Deleting Published Packages

It is possible to use the Delete function on any Package file that you have created. (If you have administration privileges, you can also delete the Package files created by others.) However, the actual package file and associated RIF-CS, while no longer visible through the HIEv user interface, are not physically deleted from the system. Instead, they are moved to an archive directory on the HIEv server. If necessary, a system administrator can recover these files for you.

If you delete a Package before it is harvested, the Package file and corresponding RIF-CS will be archived and no longer available for harvesting. This effectively undoes the Publish function and the Package will not be harvested.

If you delete a Package after it has been harvested, the Package file and corresponding RIF-CS will be archived, bit it will not affect any already harvested version of the RIF-CS and Package. If you subsequently don't want the deleted package to be available externally, you will need to have it manually removed from the application which harvested it, eg: a local metadata store and/or the RDA website.

Note	On sites where the harvesting process copies only the RIF-CS file and external users reference the Bagit ZIP directly on HIEv, you must be wary of deleting a Published Package as it will invalidate the links in the harvested RIF-CS file.

9.3.3 Editing Published Packages

As explained above, at the time a Package is created its matching RIF-CS file is also created. (The action of Publishing only copies that RIF-CS file to a discoverable location for harvesting.)

Therefore, even though editing the Metadata of a Package is possible, it will not alter the Metadata already stored in the RIF-CS file. For this reason, editing a Package's Metadata is not recommended.

The only effective way to change the Metadata of a Package is to delete the Package and recreate it. Prior to Publishing, this is always a safe and reliable way to update a Package's Metadata.

After Publishing, the effect of Package deletion and recreation depends on whether harvesting has already occurred (see previous section).

9.3.4 Correcting Published Packages

If a Package is Published incorrectly and you know it has not been harvested, deleting it is sufficient to avoid the Publishing of the incorrect data.

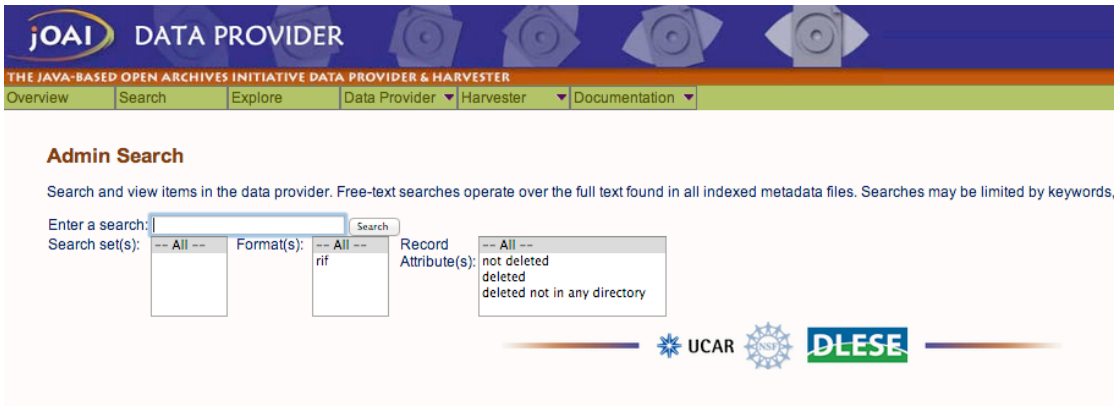
If it has been harvested, deleting it will have no downstream effect. It will be necessary for you to contact the administrator of your Published data store and ask for its removal.

If the Metadata of a Published Package is not correct, you can delete the Package, recreate it with correct Metadata and re-Publish. If it has already been harvested, some systems may correctly overwrite the old data with the new when it is harvested. However, it is generally best to contact the administrator of your data store and ask for the removal of the original package.

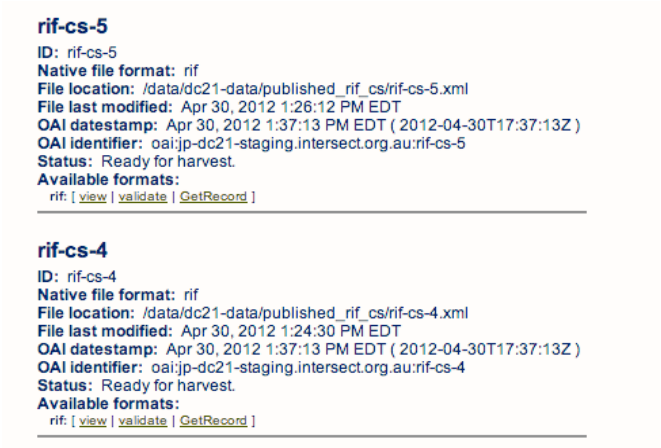
9.4 Viewing Published data

HIEv implementations use the Java Open Archives Initiative (jOAI) to harvest Published data. This tool can be configured in many ways, so it is best to understand its operation specific to your site. See you System Administrator for these details and also the URL of your HIEv server.

The descriptions of published Packages can be viewed by going to the jOAI web interface at **http://<*your.HIEv.server*>/oai/admin/query.do** and performing a search.



Clicking the **Search** button with the search field blank will show all published Packages.



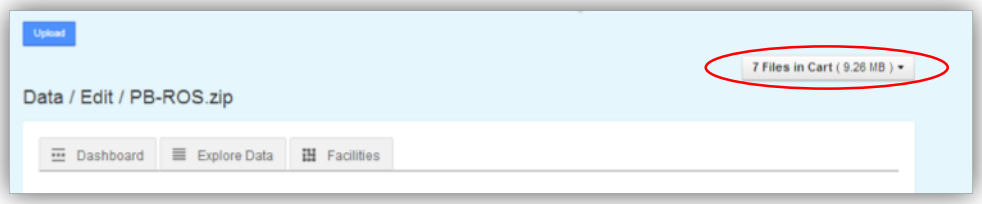
10 Downloading files

HIEv allows you to download any data file, or multiple data files, to your local computer.

If you download a single file, it will be saved on your computer in its usual format. If you simultaneously download more than one file, the files will be combined into a ZIP file and that ZIP file will be downloaded to your computer. (This ZIP file is *not* a Bagit format ZIP file.)

To download one or more files:

- Add those files to your Cart, ensuring that the Cart contains only those files you wish to download. See section *Chapter 8.3 The Cart* for instructions on how to add data files to your Cart. Take care not to select too much data for one download. Depending on the speed of your data link, it could take a very long time to download.
- Click on the Cart status box to open the Cart dropdown menu.



- Select **Download** from the dropdown menu. Alternatively, select **Edit Cart** from the dropdown menu and then click on the **Package** button on that screen, as shown below. If this method is used, the Cart can be reviewed prior to Packaging.
- A file dialog box will open. Navigate to the sub-directory on your local computer into which you wish to save the downloaded data and select the name you wish to use for the downloaded data file. The file dialog will be for the one data file if only one file is in your Cart, or it will be for one ZIP file if multiple files are in your Cart.

Note When downloading data files, only the data files themselves are downloaded. Metadata is not downloaded.

If you download a Packaged ZIP file, you can access as much of the Metadata for each of the files as is included in that Packaged ZIP file. See section *9.1 Creating a Package* for instructions on creating a Packaged ZIP file and *Appendix A - The Bagit format* for details of the Bagit format, which is used for Packaged ZIP files.

11 HIEv Administration

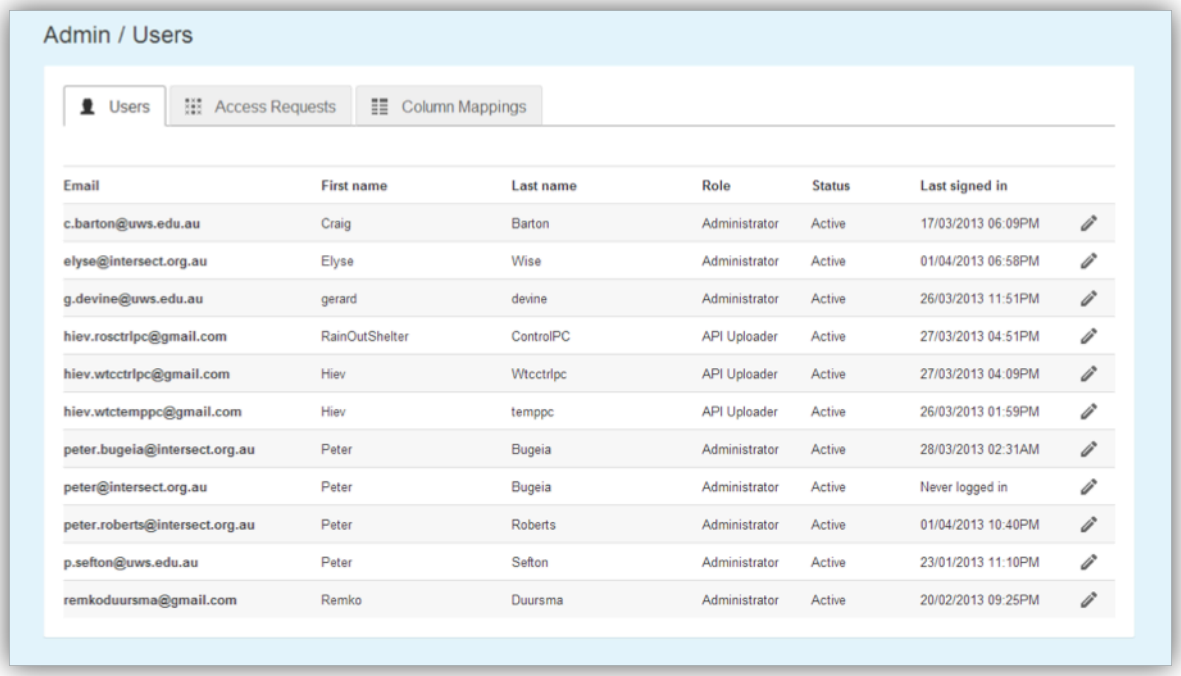
When a User account is created it is associated with a role within HIEv. This role dictates what permissions the User has within the system. The most powerful role a User can be given within the system is that of the Administrator.

Only Administrators have access to the **Admin** link at the top right of the main screen. It accesses the Admin section, which has three tabs:

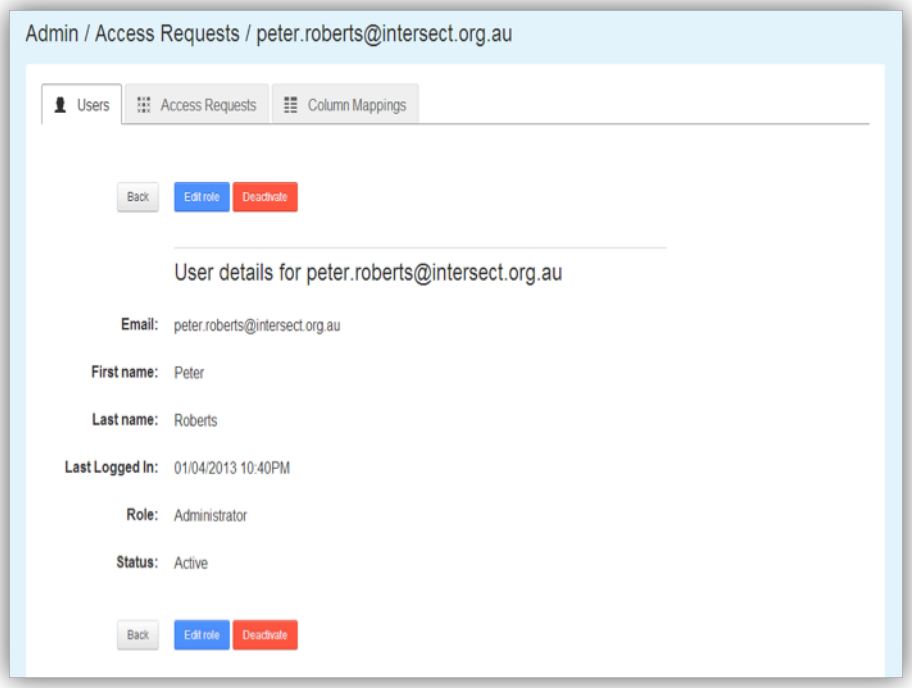
Users tab	This tab is used for managing the details of the Users who can access the system.
Access Requests tabs	This tab is used for processing requests by access the system made by new Users.
Column Mappings tab	This tab is one method of managing the list of defined Column Mappings.

11.1 Managing Users' Details

The Users tab lists all the Users that are registered within the system.



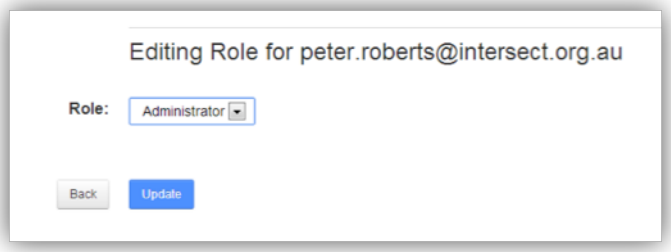
Click on a User's email address to open a screen showing that User's details.



Two functions are provided:

Deactivate Disable the account from being used to login to the system. No data uploaded by the User will be deleted.

Edit role Change the role that will be assigned to the User for future logins. Clicking this button opens the following screen that allows the User's role to be changed. See section *3.1 Classes of Users* for information on roles.

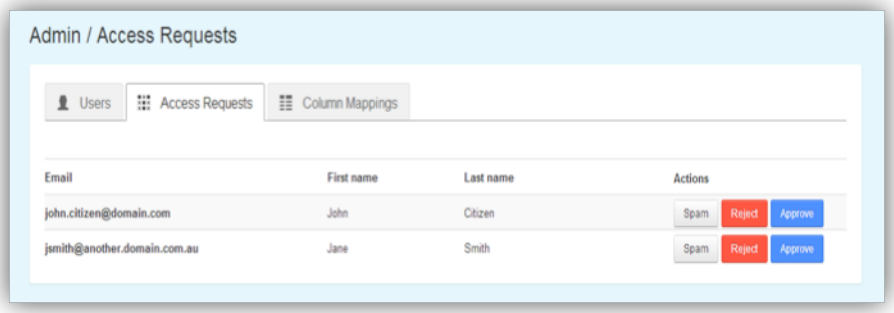


You can click on the pen ✎ at the right of each User's entry on the Users list to jump directly to this dialog.

It is not possible to delete a User's entry. This ensures that historical information relating to that User remains meaningful. Instead of deletion, a User's login account should be Deactivated.

11.2 Authorising New Users – The Access Requests Tab

The Access Requests tab is where an administrator can approve or deny requests for a User account in the system.



Each access request line in this table has three buttons:

Spam Click on this button to ignore the account request and remove it totally from the system.

Reject Click on this button to reject the access request. An email informing the User that his or her request for an account has been rejected will be sent to the email address listed.

Approve Click on this button to accept the User's access request. It will take you to a screen where you must select a role for the User in the system. Completing this approval process results in a confirmation email being sent to the User. See section *3.1 Classes of Users* for information on the permissions of the three available roles.

Clicking any of these three buttons will remove the request from the Access Requests table.

11.3 Managing Column Mappings

Column Mappings are a way of defining a relationship between the column headings in TOA5 data files (the Code part of the mapping) to a standard name from a defined ontology (the Name part of the mapping.)

The Column Mappings are stored once for the whole system and all Users share the one set of mappings.

A basic set of Column Mappings is defined as part of the configuration of HIEv at installation. In addition, further mappings can be added as they are needed.

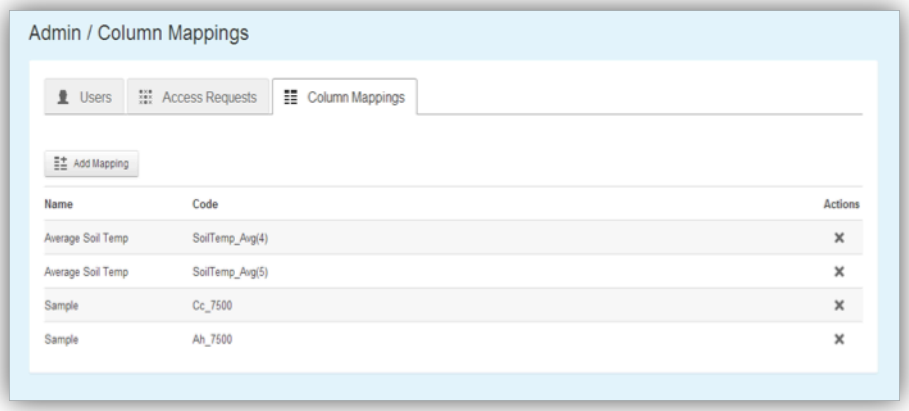
These Columns Mappings are used by HIEv for two purposes:

- When the Metadata for a TOA5 file is displayed, the Column Mappings table is checked against the column names in the TOA5 file. If any match is found, the Name from any matching entry in the Columns Mapping table is shown in the Metadata display in the Name column of the Columns information for that file.
- When searching for data files by Columns, the options in the search parameters are the column headings from all TOA5 data files stored in HIEv. They are grouped and sorted by their matching Name from this table. Those without matching Names are shown in an Unmapped list as the final item.

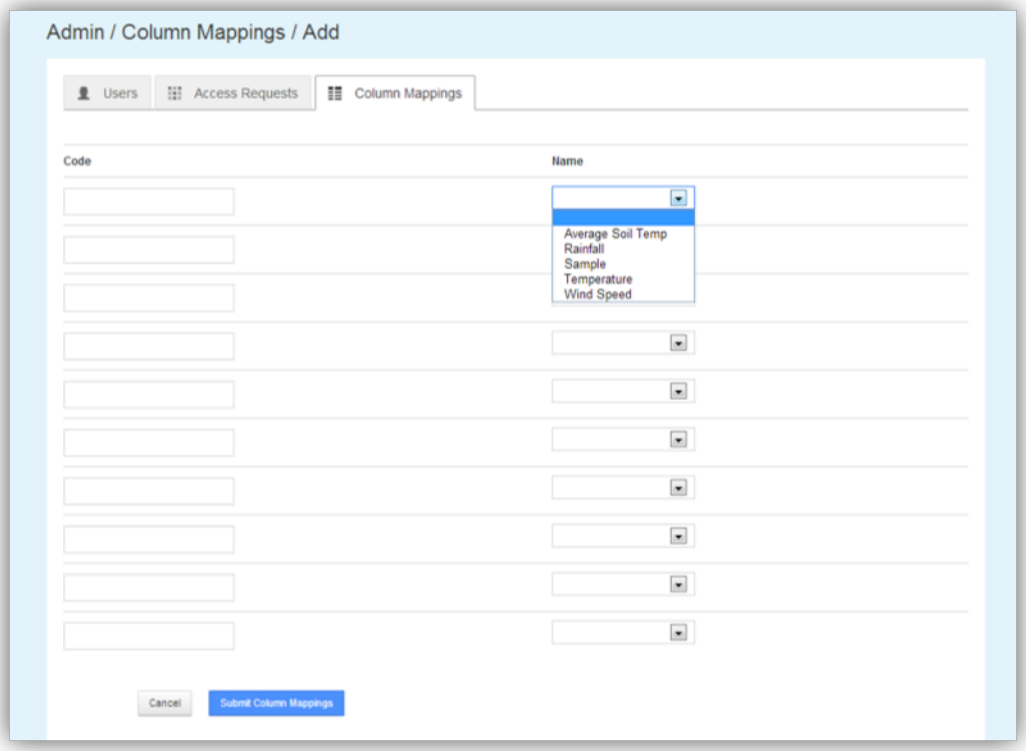
11.3.1 The Column Mappings tab

This tab allows Users with Administrator permission to add and delete column mappings.

To delete an existing mapping, click the **X** in the far right **Actions** column of the table for the mapping you wish to delete.



To **add** more mappings click the **Add Mappings** button at the top left of the tab. This will display a form where the mapping pairs can be defined.



On the left of each row the code from a TOA5 column header can be entered and on the right the standard Name to map to can be selected from a drop-down list.

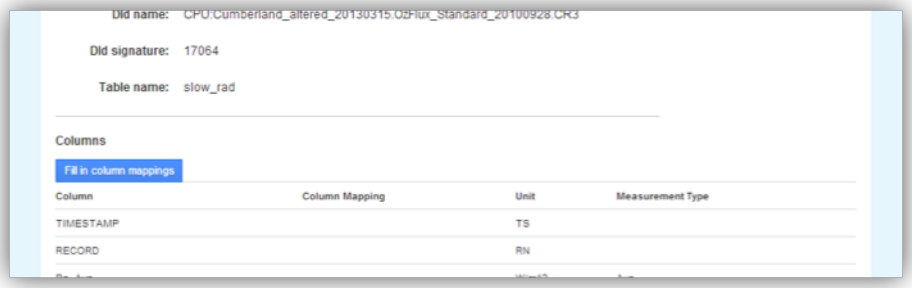
Once the mappings are defined, click the **Submit Column Mappings** button.

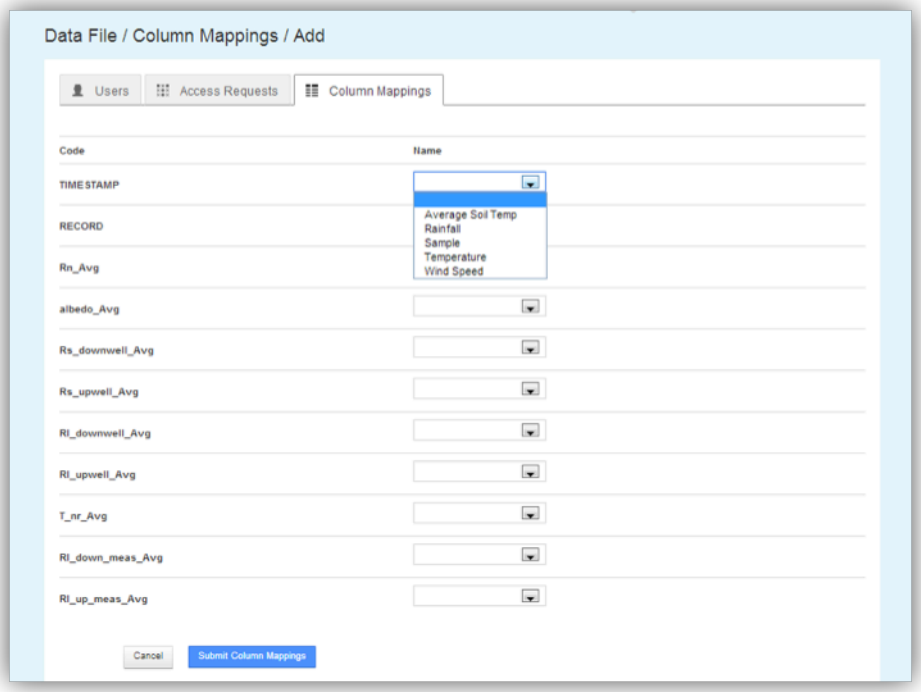
11.3.2 Updating from the Explore Data tab

Column mappings may also be defined using the **Fill In Column Mappings** button on the **View Metadata** page for a TOA5 file.

This is the preferred method for updating Column Mappings, as it avoids the need to manually type the column headers into the Code fields. However, it permits setting mappings relevant to the columns in this TOA5 data file only.

All Users can perform this function.





The list of unmapped column headings in this TOA5 data file is displayed in the left hand column and dropdown boxes are shown in the right hand column. Use the dropdown boxes to set the equivalent names for any of the column headings shown. Click on **Submit Column Mappings** to cause these new mappings to be added to the system-wide Column Mapping table. You will observe that these new mappings are now listed in the Column Mapping column of the Columns section of the Metadata display for that TOA5 file.

11.4 Managing Background Tasks

If you are a HIEv administration user and you are currently logged into HIEv, you can access the Ruby background task administration screen by creating a new browser tab and entering the following:

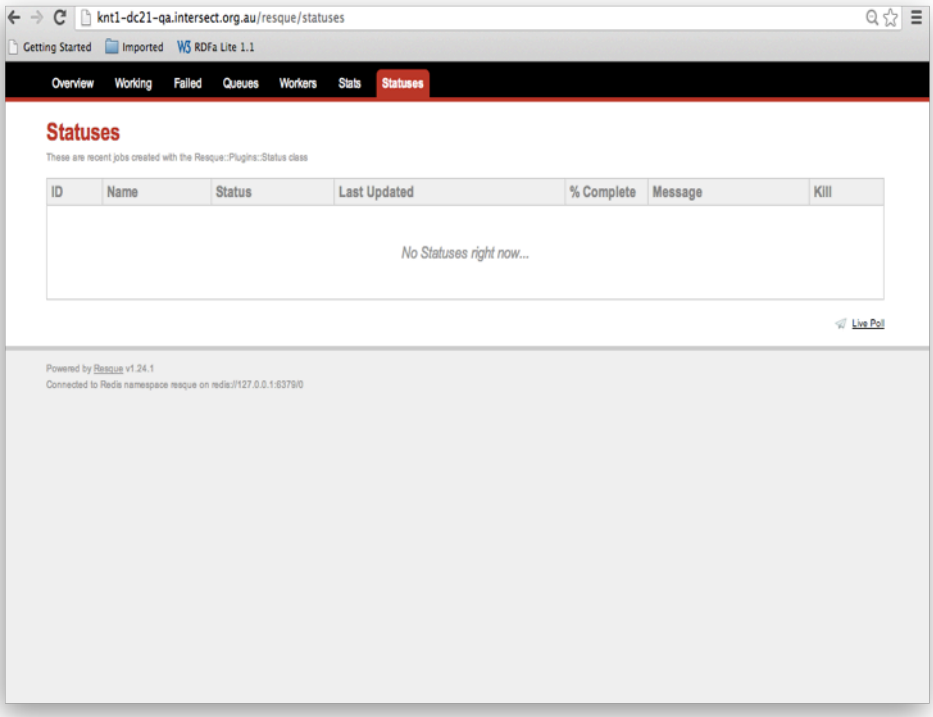
Hiev_url/resque/statuses

where HIEv_url is the url of your HIEv application.

The following screen will be displayed in the new tab.

You can use this screen to view the status of currently running packaging tasks, and to delete tasks (via the Kill function) if they aren't working properly. Killing a background task

may or may not delete the partially created HIEv Package. If the Package still exists after the background task has been killed, you should manually delete the package.



Note	If the background task detects an error during processing it will attempt to set the Creation Status to FAILED. In this case, the job entry for this task should provide an indication as to why the failure occurred.

12 Configuring Tags, Column Mappings and Experiment Parameters

When HIEv is first installed, the available Tags, Column Mappings and Experiment Parameters are populated in the database from the configuration file dc21app_extra_config.yml. Once the system has been installed, the Experiment Parameters and Tags can be modified by modifying the lookup tables that store these values. This can be done directly using a tool like PSQL. Another convenient way to do this is to use the Rails console. The instructions below use the Rails console to add rows to the tables.

If you wish to delete or modify existing rows in these tables, make sure you maintain referential integrity with existing records.

To begin, ssh to the server on which HIEv is running on and determine the Rails Environment the system is running as. This will generally be "production", but you can check by looking for a RailsEnv line in the Apache Rails configuration (e.g. <root>/etc/httpd/conf.d/rails_dc21app.conf). This will look something like the following.

```
RailsEnv production
```

In this case, the Rails Environment is "`production`". Once the value of this setting has been determined for your system, navigate to the location of the application installation (e.g. <root>/home/devel/dc21app/current) and enter the directory "dc21app/current". From here you can start the Rails Console using this command.

```
RAILS_ENV=<RailsEnv> bundle exec rails console
```

e.g.

```
RAILS_ENV= production bundle exec rails console
```

This will give you a prompt similar to:

```
Loading production environment (Rails 3.1.1)
1.9.2p290 :001 >
```

From this prompt you can issue commands to add Tags and Experiment Parameters.

To add a Tag use this command.

```
Tag.create!(name: '<Tag name>')
```

e.g.

```
Tag.create!(name: 'Analysed')
```

This will result in output similar to:

```
(0.3ms)  BEGIN
(1.3ms)  SELECT 1 FROM "tags" WHERE LOWER("tags"."name") = LOWER('Analysed') LIMIT 1
SQL (8.7ms)  INSERT INTO "tags" ("created_at", "name", "updated_at") VALUES ($1, $2, $3)
      RETURNING "id"  [["created_at", Fri, 14 Sep 2012 10:55:24 EST +10:00], ["name",
      "Analysed"], ["updated_at", Fri, 14 Sep 2012 10:55:24 EST +10:00]]
(0.6ms)  COMMIT
=> #<Tag id: 6, name: "Analysed", created_at: "2012-09-14 00:55:24", updated_at: "2012-09-14
      00:55:24">
```

To add a Modification or a Unit for an Experiment Parameter, use the commands:

```
ParameterModification.create!(name: 'Above average')
ParameterUnit.create!(name: 'PSI')
```

Parameter Categories and Sub Categories require an extra step to define the relationship between the two.

```
parameter_category = ParameterCategory.create(name: 'Light')
parameter_category.parameter_sub_categories <<
ParameterSubCategory.create(name: 'Brightness')
```

This will result in output similar to:

```
1.9.2p290 :001 > parameter_category = ParameterCategory.create(name: 'Light')
(0.4ms)  BEGIN
SQL (121.5ms)  INSERT INTO "parameter_categories" ("created_at", "name", "updated_at") VALUES
      ($1, $2, $3) RETURNING "id"  [["created_at", Fri, 14 Sep 2012 16:14:26 EST +10:00],
      ["name", "Light"], ["updated_at", Fri, 14 Sep 2012 16:14:26 EST +10:00]]
(0.5ms)  COMMIT
 => #<ParameterCategory id: 8, name: "Light", created_at: "2012-09-14 06:14:26", updated_at:
      "2012-09-14 06:14:26">
1.9.2p290 :002 > parameter_category.parameter_sub_categories <<
1.9.2p290 :003 >   ParameterSubCategory.create(name: 'Brightness')
(0.3ms)  BEGIN
(0.3ms)  ROLLBACK
(0.2ms)  BEGIN
ParameterCategory Load (0.7ms)  SELECT "parameter_categories".* FROM "parameter_categories"
      WHERE "parameter_categories"."id" = 8 LIMIT 1
SQL (1.2ms)  INSERT INTO "parameter_sub_categories" ("created_at", "name",
      "parameter_category_id", "updated_at") VALUES ($1, $2, $3, $4) RETURNING "id"
      [["created_at", Fri, 14 Sep 2012 16:14:27 EST +10:00], ["name", "Brightness"],
      ["parameter_category_id", 8], ["updated_at", Fri, 14 Sep 2012 16:14:27 EST +10:00]]
(0.5ms)  COMMIT
ParameterSubCategory Load (0.7ms)  SELECT "parameter_sub_categories".* FROM
      "parameter_sub_categories" WHERE "parameter_sub_categories"."parameter_category_id" = 8
      ORDER BY "parameter_sub_categories"."name"
=> [#<ParameterSubCategory id: 27, name: "Brightness", parameter_category_id: 8, created_at:
      "2012-09-14 06:14:27", updated_at: "2012-09-14 06:14:27">]
```

To add a Column Mapping name use this command.

```
ColumnMapping.create!(code:'<Code>', name:'<Name>')
```

This will result in output similar to:

```
1.9.2p290 :001 > ColumnMapping.create!(code:'VOL', name:'Volume')
   (0.1ms) BEGIN
   (1.0ms) SELECT 1 FROM "column_mappings" WHERE LOWER("column_mappings"."code") =
      LOWER('VOL') LIMIT 1
  SQL (8.5ms)  INSERT INTO "column_mappings" ("code", "created_at", "name", "updated_at")
      VALUES ($1, $2, $3, $4) RETURNING "id"  [["code", "VOL"], ["created_at", Wed, 31 Oct
      2012 14:18:53 EST +11:00], ["name", "Volume"], ["updated_at", Wed, 31 Oct 2012 14:18:53
      EST +11:00]]
   (0.9ms)  COMMIT
 => #<ColumnMapping id: 6, code: "VOL", name: "Volume", created_at: "2012-10-31 03:18:53",
      updated_at: "2012-10-31 03:18:53">
```

13 Migrating data to a new system

To restore a **pg_dump** you pass the file to psql with an empty database. If you have an existing database with the same name, you need to drop it first and recreate it.

The command to drop the database is **dropdb**. So you 'su' to the **postgres** user and run these commands.

```
$ sudo su - postgres
$ dropdb <database name>
$ createdb <database name>
```

Once you have done that, you can exit the **postgres** user, and restore the database dump.

```
$ exit
$ psql -U <user> <database name> < sql_dump.sql
```

To restore the data, you need to untar it into your root directory. It is likely that your permission system won't allow you to create a directory under root, so you should create it manually, and assign the right permissions to it.

```
$ sudo mkdir /data
$ sudo chown <user>:<group> /data
$ cd /
$ tar xvf <tar file>
```

14 Revision History

Version No.	Revision Date	Summary of Changes	Revised by
V1.8	4 Apr 2013	Initial HIEv version. (Copied and substantially updated from User Manual Version 1.1 for DC21 V1.6b.)	Peter Roberts
V1.9	19 Jun 2013	Updates for version release 1.9.	Peter Bugeia

Appendix A - The Bagit format

BagIt is currently defined in an Internet Engineering Task Force (IETF) internet draft.

Quoting from the preamble of the Bagit entry on Wikipedia:

> *BagIt is a hierarchical file packaging format designed to support disk-based storage and network transfer of arbitrary digital content. A "bag" consists of a "payload" (the arbitrary content) and "tags", which are Metadata files intended to document the storage and transfer of the bag. A required tag file contains a manifest listing every file in the payload together with its corresponding checksum. The name, BagIt, is inspired by the "enclose and deposit" method,[1] sometimes referred to as "bag it and tag it".*
>
> *Bags are ideal for digital content normally kept as a collection of files. They are also well-suited to the export, for archival purposes, of content normally kept in database structures that receiving parties are unlikely to support. Relying on cross-platform (Windows and Unix) filesystem naming conventions, a bag's payload may include any number of directories and sub-directories (folders and sub-folders). A bag can specify payload content indirectly via a "fetch.txt" file that lists URLs for content that can be fetched over the network to complete the bag; simple parallelization (e.g., running 10 instances of "wget") can exploit this feature to transfer large bags very quickly. Benefits of bags include*
>
> - *Wide adoption in digital libraries (e.g., the Library of Congress).*
> - *Easy to implement using ubiquitous and ordinary filesystem tools.*
> - *Content that originates as files need only be copied to the payload directory.*
> - *Compared to XML wrapping, content need not be encoded, saving time and storage space.*
> - *Received content is ready-to-go in a familiar filesystem tree.*
> - *Easy to implement fast network transfer by running ordinary transfer tools in parallel.*

Further information about the Bagit hierarchical file packaging format can be found at various places on the Internet, including:

Internet Engineering Task Force – http://www.ietf.org

Wikipedia – http://en.wikipedia.org/wiki/BagIt

Version 0.97 of the Bagit specification - http://tools.ietf.org/html/draft-kunze-bagit-08

README.HTML file

Within the data subdirectory of the Bagit ZIP file there is a README.HTML file, which is intended for both human and machine reading. Viewing this file in a web browser will summarise the Package contents, including its Metadata, list of its data files, and Metadata for the included data files.

The machine readable parts of this file conform to the RDFa Lite (Resource Description Framework in attributes Lite) specification described at http://www.w3.org/TR/rdfa-lite/. This semantic information can be parsed by a program which can then build discoverable facets of information for a search engine.

Appendix B - RIF-CS

Quoting from the Global Registries website (http://globalregistries.org/rifcs.html):

> *The **Registry Interchange Format - Collections and Services** (RIF-CS) Schema was developed as a data interchange format for supporting the submission of Metadata to a collections service registry. It is based on ISO2146 but only includes elements needed for a collection service registry and so is not a full binding to the standard.*
>
> *A collection in the RIF-CS Schema context could be a repository, a registry, a collective work or an index/database. There are no hard and fast rules about what constitutes a collection and it is up to the data providers to consider what their collections are and what Metadata should be provided. The RIF-CS schema also supports other registry object types, namely services, activities and parties. Any or all of these along with their relations to each other are able to be expressed in RIF-CS format.*

The Australian National Data Service (ANDS – http:// http://www.ands.org.au) uses the RIF-CS standard for management of data in the Australian Research Data Commons. It provides a training resource for RIF-CS at http://www.ands.org.au/training/rif-cs/index.html.

ANDS uses the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH – see http://www.openarchives.org/pmh/tools/tools.php) to collect RIF-CS data.

Example RIF-CS file

```
<?xml version="1.0" encoding="UTF-8"?>
<registryObjects xmlns="http://ands.org.au/standards/rif-cs/registryObjects"
      xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:schemaLocation="http://ands.org.au/standards/rif-cs/registryObjects
      http://services.ands.org.au/documentation/rifcs/1.3/schema/registryObjects.xsd">
  <registryObject group="University of Western Sydney">
    <key>/data_files/2734</key>
    <originatingSource>https://jp-dc21-staging.intersect.org.au/</originatingSource>
    <collection type="dataset">
      <name type="primary">
        <namePart>PB-ROS.zip</namePart>
      </name>
      <location>
        <address>
          <electronic type="url">
            <value>https://jp-dc21-staging.intersect.org.au/data_files/2734/download</value>
          </electronic>
        </address>
      </location>
      <subject type="local" xml:lang="en">Meteorological data</subject>
      <subject type="anzsrc-for">05</subject>
      <subject type="anzsrc-for">0502</subject>
      <description type="brief">Rain Out Shelter data for a couple of years. Rest of the
      description.</description>
      <description type="rights">http://creativecommons.org/licenses/by-nc-
      nd/3.0/au</description>
      <coverage>
        <temporal>
          <date type="dateFrom" date_format="W3CDTF">2011-06-19T21:05:00+10:00</date>
          <date type="dateTo" date_format="W3CDTF">2013-03-28T11:00:00+11:00</date>
        </temporal>
      </coverage>
      <coverage>
        <spatial type="gmlKmlPolyCoords">150.73946,-33.61006</spatial>
      </coverage>
      <relatedInfo>
        <notes>Published by Peter Bugeia (peter.bugeia@intersect.org.au)</notes>
      </relatedInfo>
      <relatedInfo>
        <notes>Unique ID: HIE-ID-0001</notes>
      </relatedInfo>
      <relatedInfo>
```

```
        <notes>Primary contact for ROS Weather Station is Craig Barton
      (c.barton@uws.edu.au)</notes>
      </relatedInfo>
    </collection>
  </registryObject>
</registryObjects>
```