Flamingo User Guide

V 0.5 January 2017



# Introduction

## Oasis

The Oasis Platform can be described as follows:

Oasis is a calculation engine used to process Monte Carlo simulations of insurance companies’ exposures to catastrophic risk against synthetic event sets provided by model providers. In brief:

* Model providers build synthetic event sets to represent many years (100,000 years plus) of catastrophic events (say hurricanes) in a particular geography (say the East cost of the USA). This is necessary because the actual data that exists on historic events (~100 years) is too small to be statistically relevant and usable in estimating the probability of a particular event in the next year. These events are represented in the “hazard module” of oasis by storing the event and its severity in a particular area cell in tabular form. So, for example, they might say that a particular event (Event 2,465 say) has a severity of 90mph in area cell 1,456 (which is a 1km square on the coast of Florida in Dade County)
* Model providers also represent the “vulnerability” of categories of properties to particular severities. So, they might say that a 2 story, wood framed building in Florida when hit with a 90mph wind has vulnerability category 289 and will suffer 44% damage. (Note that this is vastly simplified and in reality, there will be a damage distribution rather than a point value and there would be many more parameters that drive the vulnerability – roof fixings, window type, foundations, local geography, etc.)
* Insurers then represent their exposures according to their location (i.e. which area cell is the property in) and their vulnerability category (i.e. 289 in the example above) and Oasis applies the damage quotient to the total value of the property (say 44% x $250,000 = $111,000). In reality, the distribution will be some variation around 44% to represent the secondary uncertainty and Oasis allows you to sample from that distribution to get a range of losses.

The Oasis platform has a four-tier architecture:

1. **Flamingo user interface:** Provides a reference user interface for modelling workflows. Extensible implementation using R-shiny framework. Runs in browser and no client-side installation required.
2. **Flamingo server:** Provides configurable business logic for modelling workflows. Provides an extensible exposure store based on a canonical data model, along with tools to transform exposure data from other common formats such as EDM and Cede. Implemented using Python (Flask), SQL server and embedded .NET.
3. **Mid-tier:** Provides a web service API for managing model data, running analyses and retrieving results. Implemented using Python (Flask).
4. **Calculation backend (ktools):** Calculation components implemented in C++. High performance, multi-threaded model execution and analytics. Provides a set of reference components for model execution.

## 

Figure - Oasis Architecture

## Flamingo

Flamingo is a business front end to Oasis and it has two main functions:

1. **Exposure management and transformations:** It allows a user to transform their exposure data into the formats required by both model providers in order to map the exposures into their particular area and vulnerability categorisations, and also to transform their data into the Oasis format required to run the oasis calculations.
2. **Running processes into the Oasis API layer and handling the outputs:** It provides a front end to the API interface with Oasis whereby you can present exposure data (pre-transformed into the correct format as above), choose your analysis settings (number of samples, output aggregation options, etc.), request an analysis, monitor the run and manage the output files.

### Exposure Management

Flamingo is designed to accept exposure data in many different formats and convert that data into a canonical data model format that is perfectly generic. This data can then be converted back out into many other data formats as required. The system utilises XSD and XSLT files to validate and transform the source exposure data into and out of the canonical data model. The three core conversions that Flamingo undertakes are:

1. Conversion from source to canonical format – this transformation takes data from a source format and converts into a cleansed version of that data ahead of loading into the canonical data model. This cleansing might include replacing empty values with a default value or filtering out data that is known not to be required in the modelling process (non-modelled building codes or geographies for example).
2. Conversion from canonical to model specific format – this conversion is used to get data into the format required by the model provider in order to assign the oasis keys specific to that model in an exterior lookup service.
3. Conversion from canonical to oasis format – this conversion takes the canonical exposure data, combines it with the model specific oasis keys that are returned from the lookup service and generates the abstract data files that are required to run models in oasis

The data conversions are all done within SQL Server using XSDs for data validation and XSLTs for transformation. These are files that define the underlying data schemas in the files (XSD) and the transformation rules to be applied (XSLT)

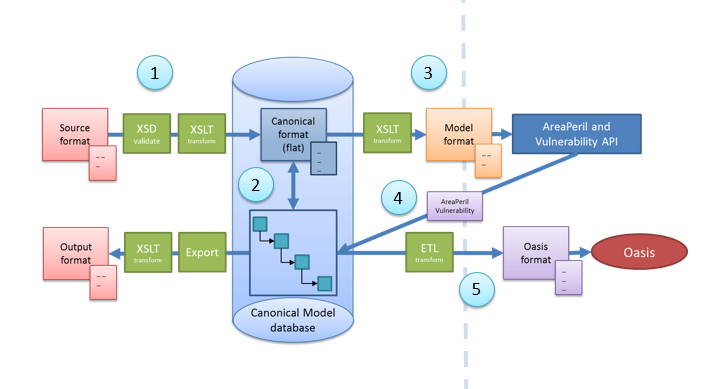


Figure – Data Transformations

Figure 1 illustrates the steps are taken when converting data in Flamingo via the following steps

1. Source Location and Contract data files conversion to Canonical Location and Contract data files
2. Canonical Location and Contract data files to Canonical Data Model
3. Canonical Location file to Model Location file conversion
4. Model specific oasis key files (matching and non-matching location files)
5. Canonical data model to Oasis file format

### Canonical Data Model

The canonical data model is Flamingo’s way of storing exposure data. It is a generic data model that uses a general hierarchy and “key value pairs” in combinations with “profiles” to define the data formats. The hierarchy uses generalised terms to describe the relationships between levels in any class of business, with the mapping to those levels being specific to the class. Each level in the hierarchy then has a related values table which stores the key value pairs, and a profile to describe the meta data associated with the keys.

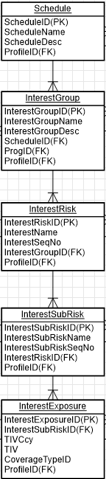


Figure – General Hierarchy

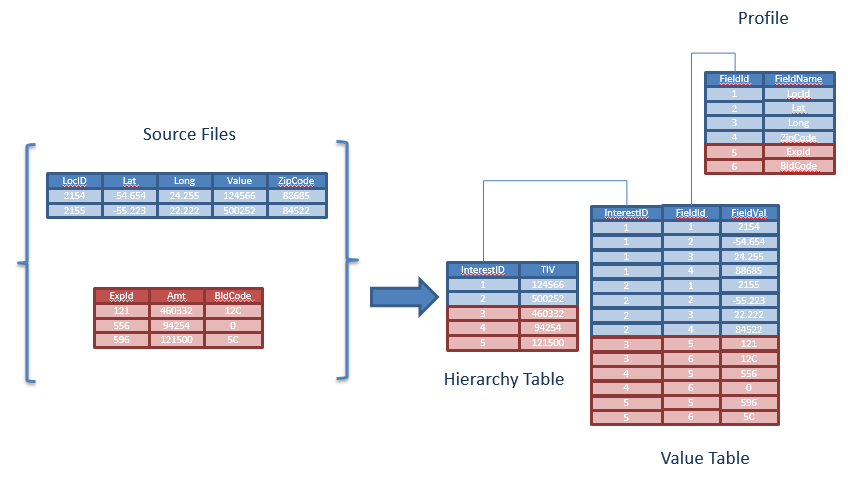


Figure - Key Value Pairs

### Processes

Processes in Flamingo are ordered lists of API instructions to the Oasis mid-tier.

There are 4 basic API calls that are executed in any one process, and these are:

1. Post Exposures – this sends up the exposure files as generated in the exposure management in Flamingo and receives confirmation of the location of the exposure files in the Oaiss data store
2. Post Analysis – this sends up a reference to the location of the exposures and the requested analysis settings – i.e. which analyses are required. It receives back a queue resource form the Oasis mid-tier.
3. Get analysis status – here Flamingo polls the Oasis mid-tier with the queue resource request and receives back the status of that analysis: either in progress, failed or completed. Once the analysis is completed, the Oasis mid-tier also returns the location of the output files that have been generated
4. Get outputs – this call passes the output location and receives back the files themselves from the oasis data store

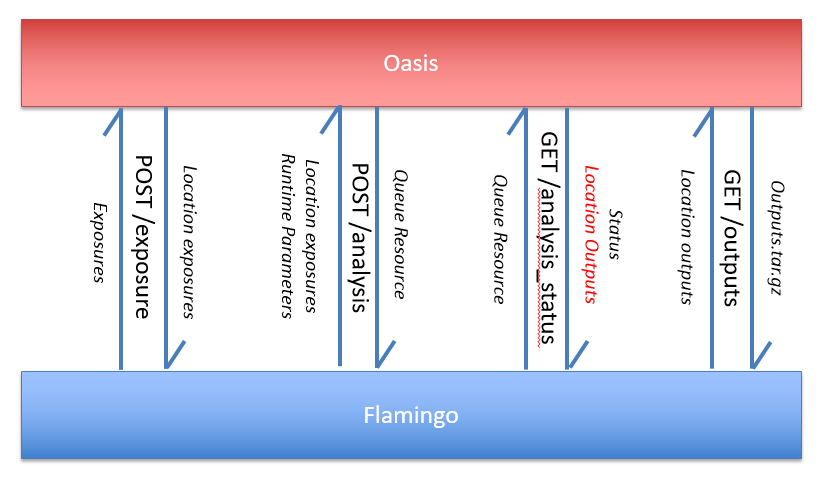
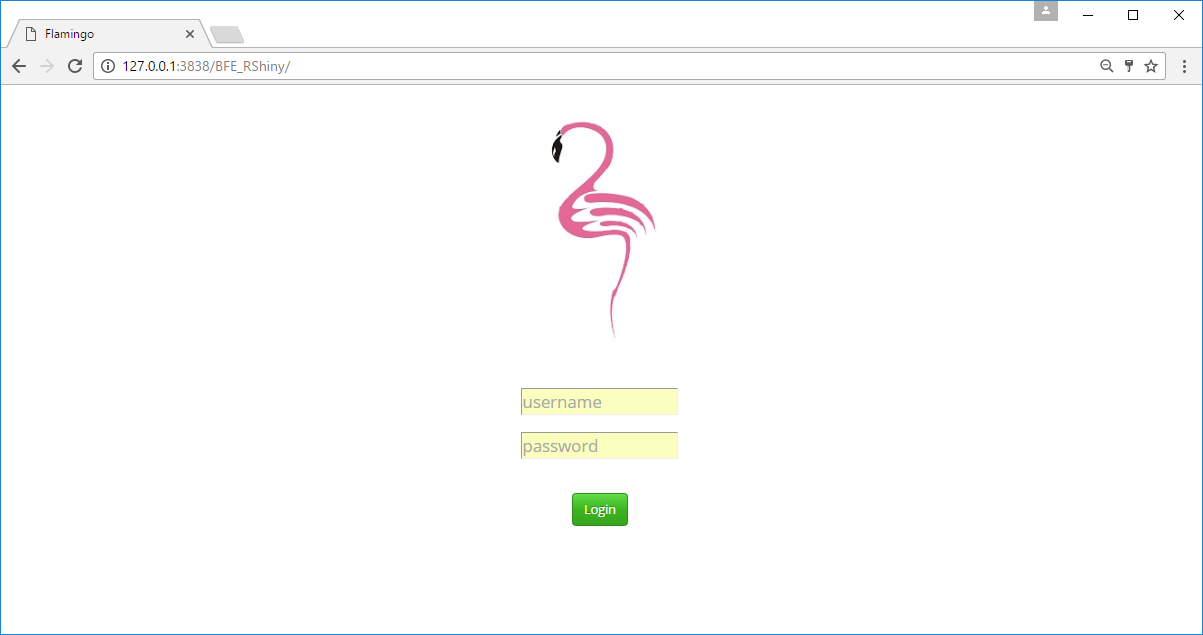


Figure - Oasis APIs

# User Guide

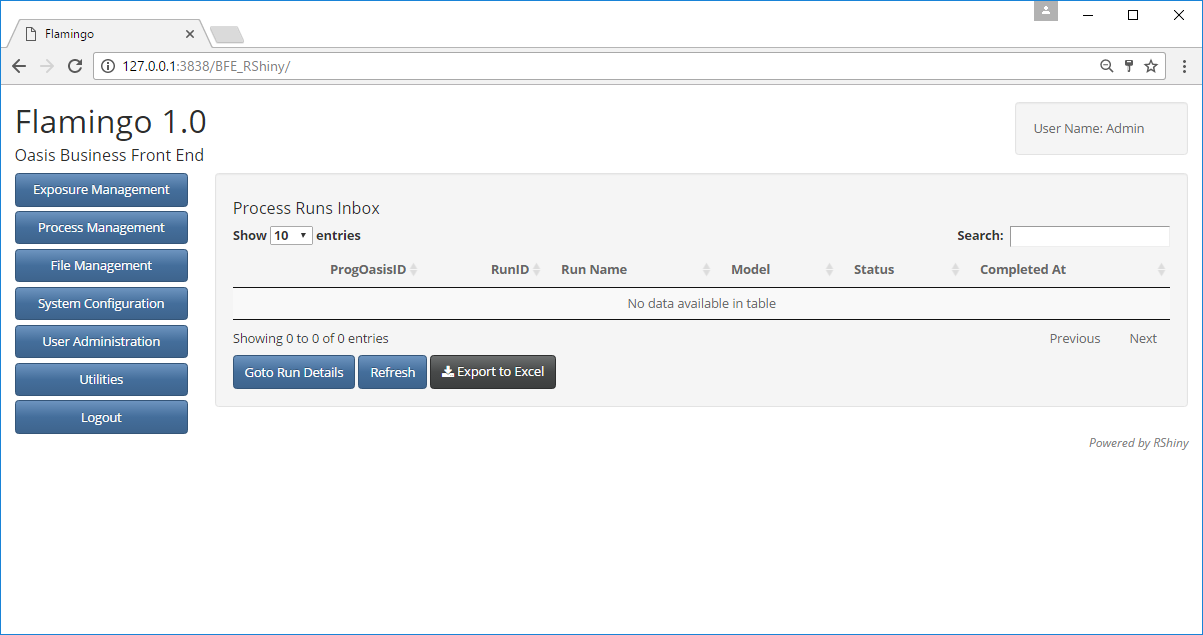
## Login

The login screen is where you should enter your Flamingo username and password



## Landing Page

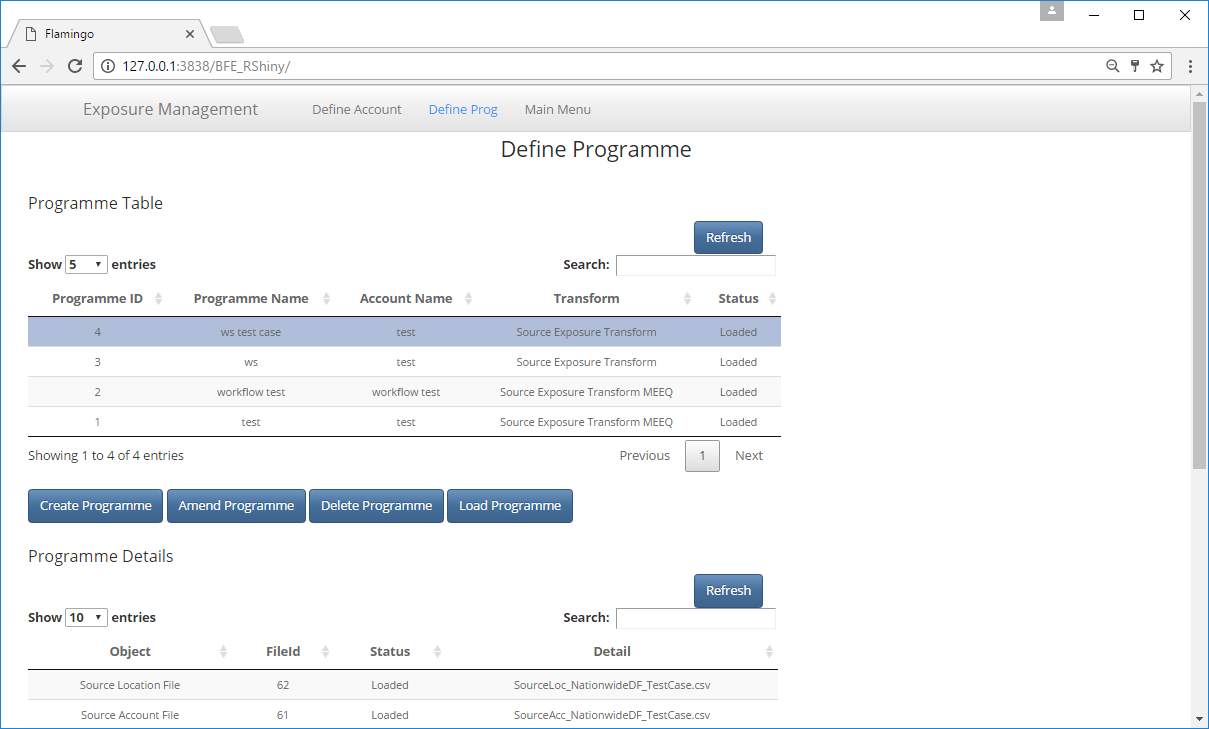
The Landing page displays a number of areas of Flamingo functionality on the left hand side and a list of previously requested processes in the main section with the current status.



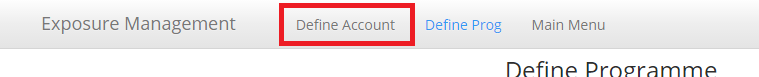
## Exposure Management

Once you’re logged in navigate to “Exposure Management” on the top left

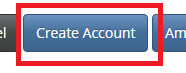




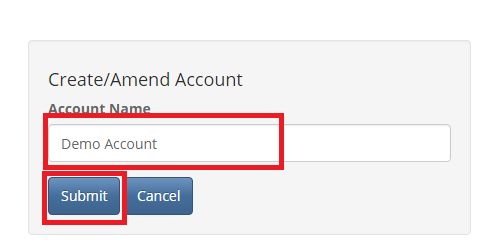
Click on “Define Account” on the top menu bar



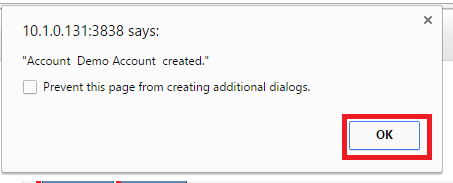
Click on “Create Account”



Name the Account and click on “Submit”



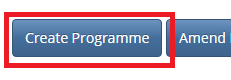
Click “OK” on the pop up box that appears



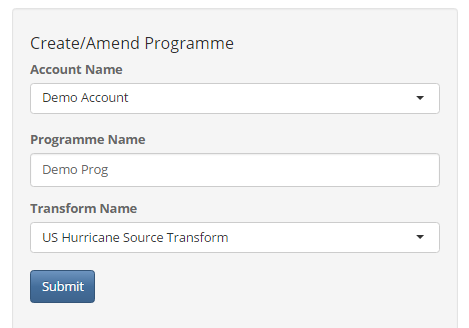
Navigate to “Define Prog” on the menu bar at the top



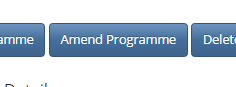
Click on “Create Programme”



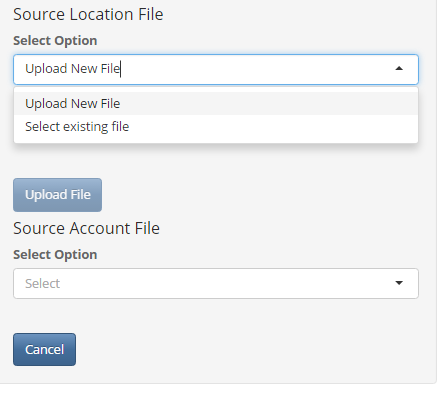
Select your account, name the programme and select a transform, then click “submit”



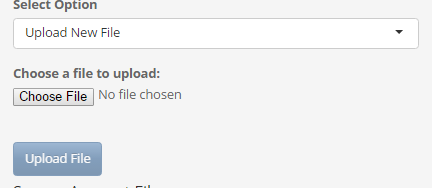
Select the programme that you have just created and click on “Amend Programme”



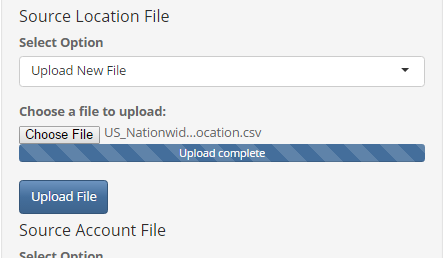
Select the “Upload New File” Option for the Source Location File



Click “Choose File” and select your source location file form your file system

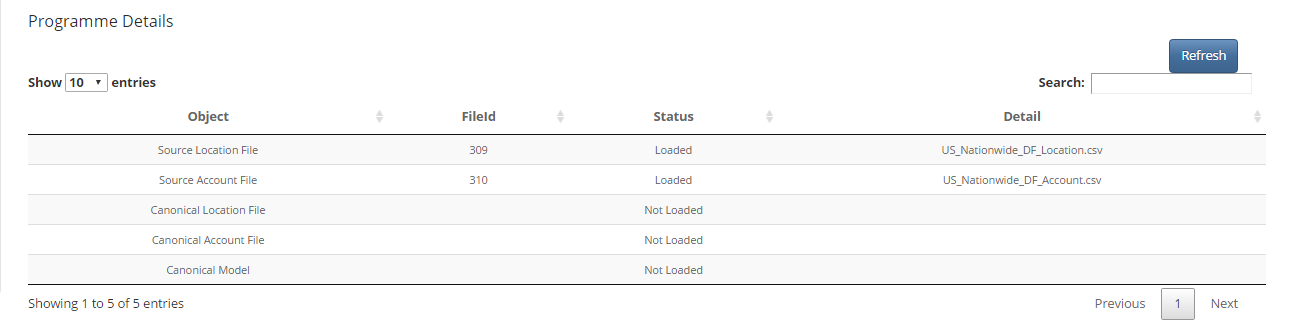


Once the upload bar is complete, click on “Upload File”



Repeat Source Location File steps for Source Account File

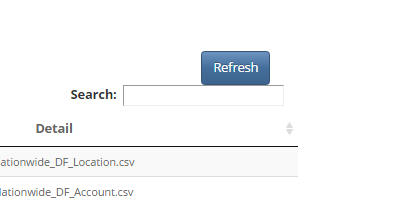
Once both files are uploaded you should see the files listed in the “Programme Details” Table



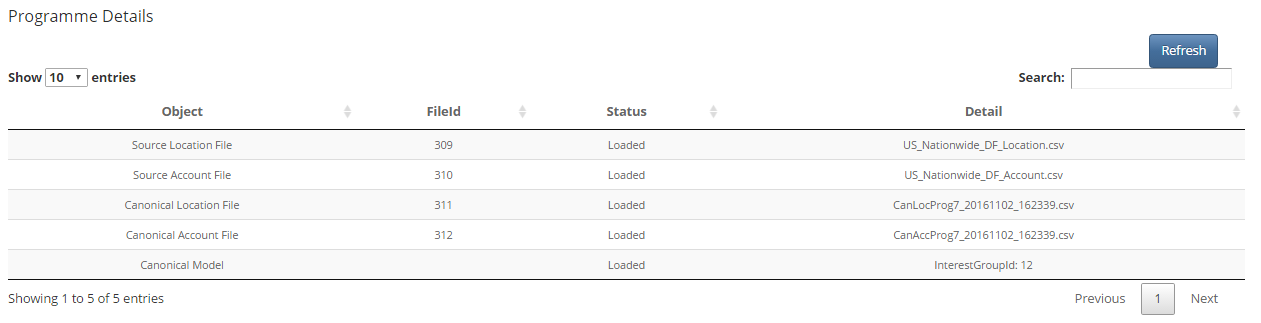
Now click on the “Load Programme” button



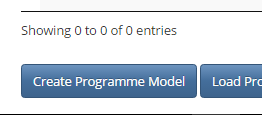
Click on the “Refresh” button above the “Programme Details” table to refresh the details



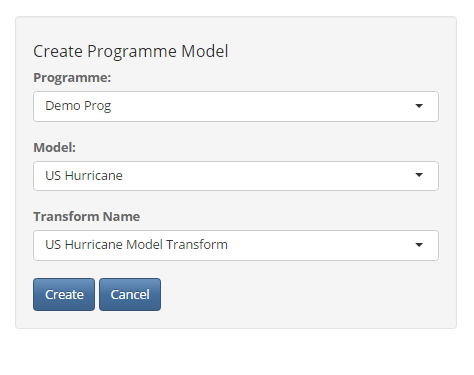
You should now see the Programme Details table is complete



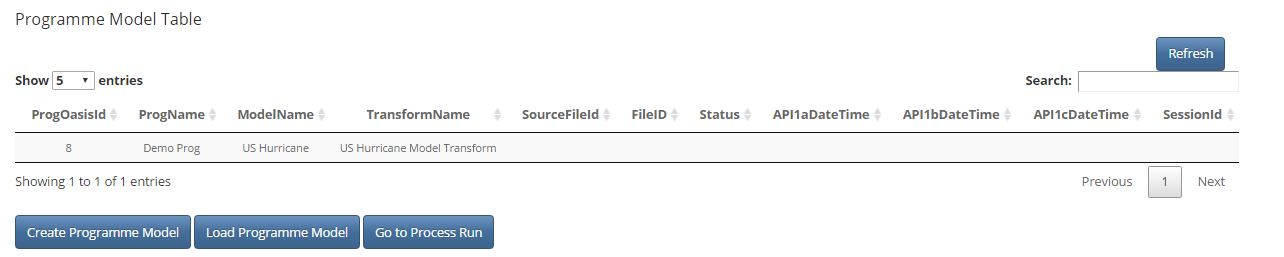
Scroll down and click on “Create Programme Model” button



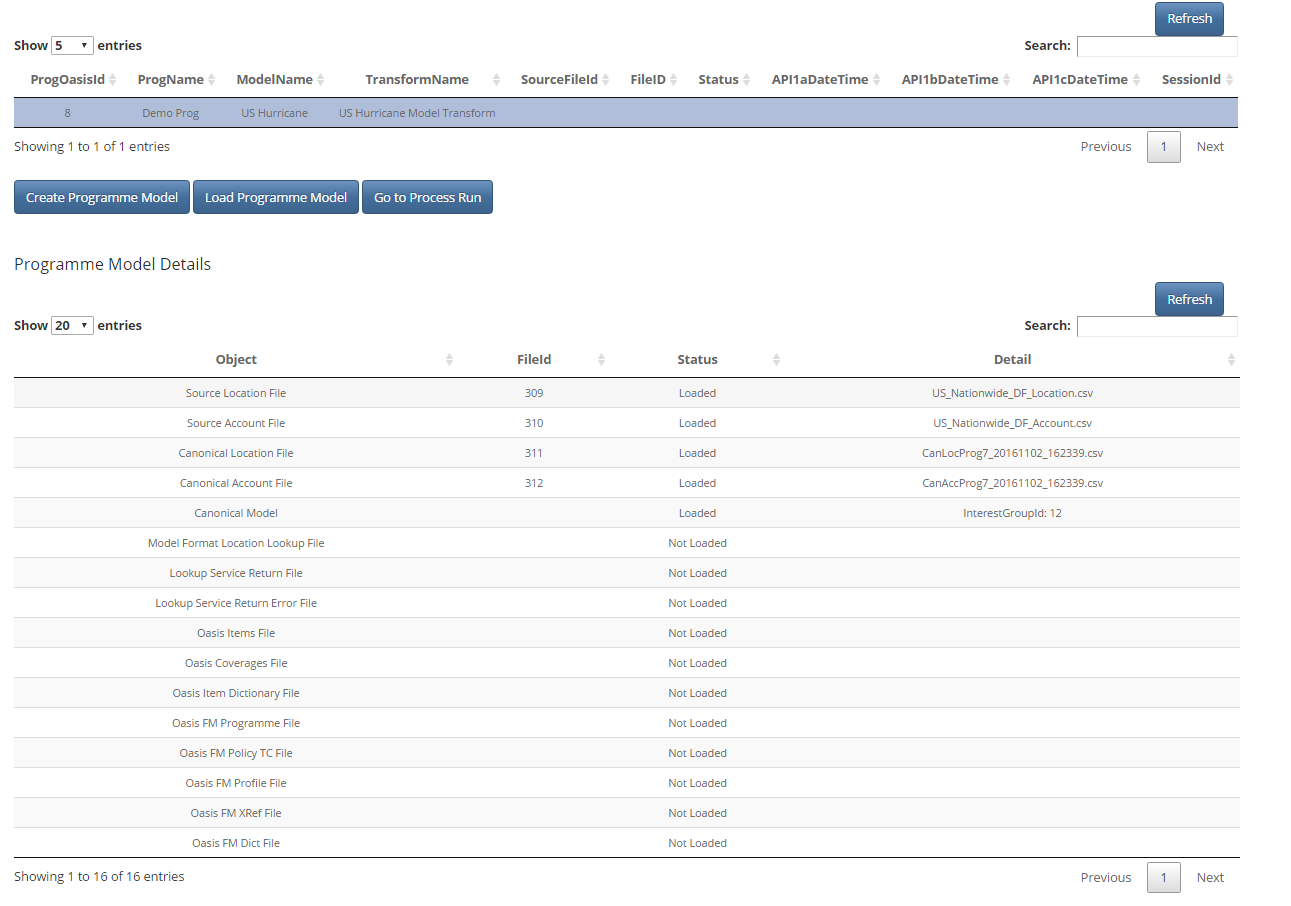
Select the Programme, Model and Transform that you want to use from the drop-down menus and click on “Create”



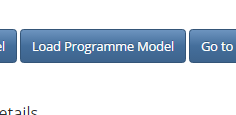
You should see a row created in the Programme Model Table



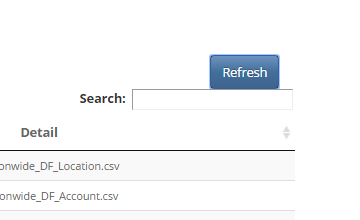
Select the row and you should see some details appear below in the Programme Model Details Table



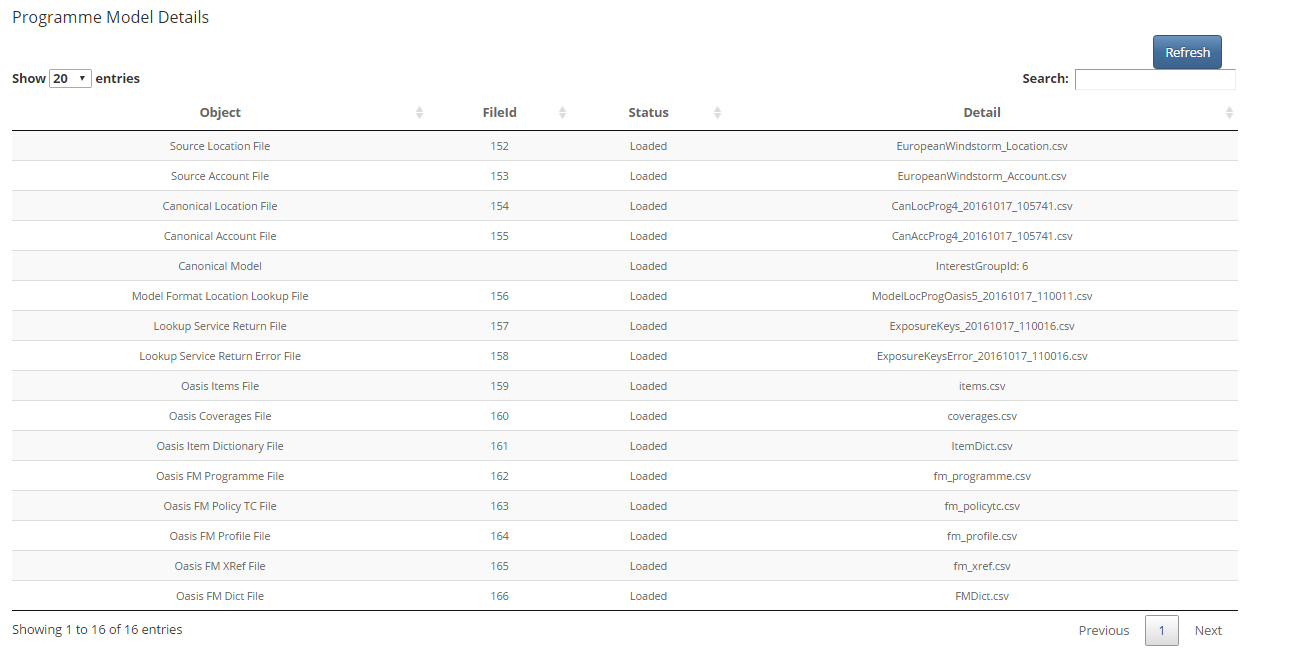
Click on the “Load Programme Model” Button. This step converts the data to the model lookup required format (say csv or upx), fires off the API lookup service, consumes the response and generates the Oasis files so can take a little while



Click the “Refresh” button above the table to monitor the progress



Once the process is complete, all rows should be populated for the table



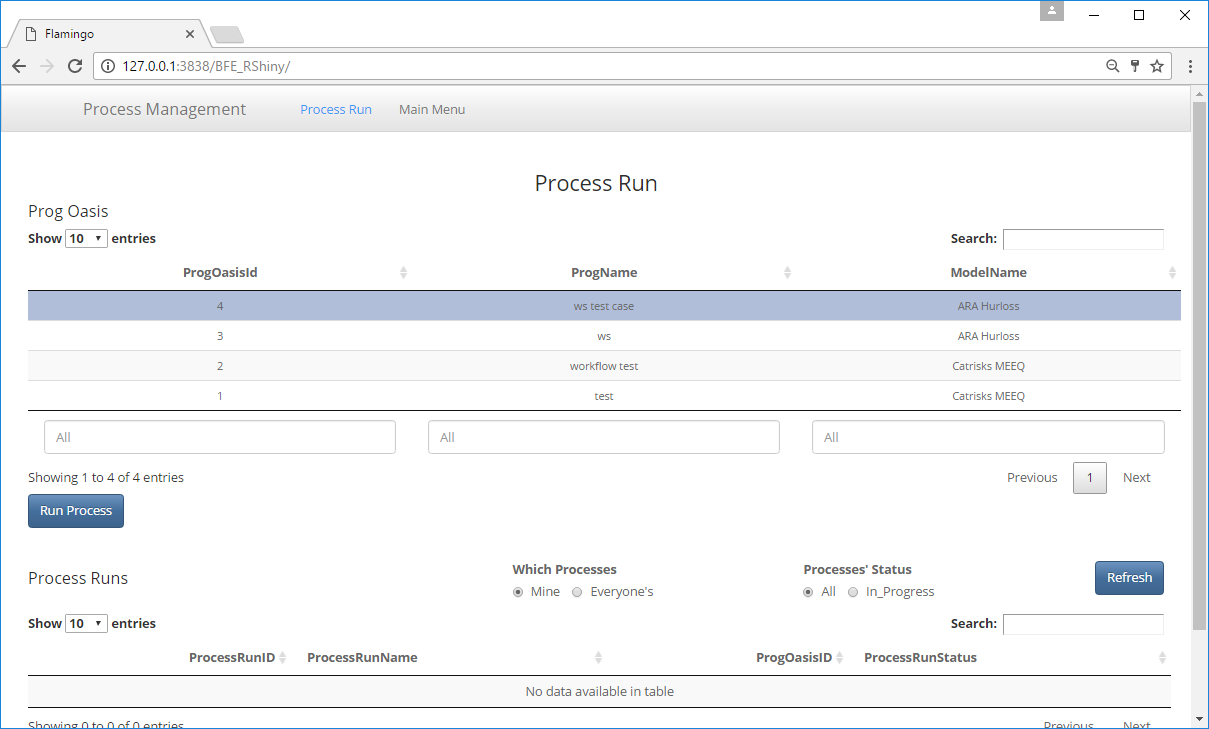
Navigate back to the main menu from the top menu bar



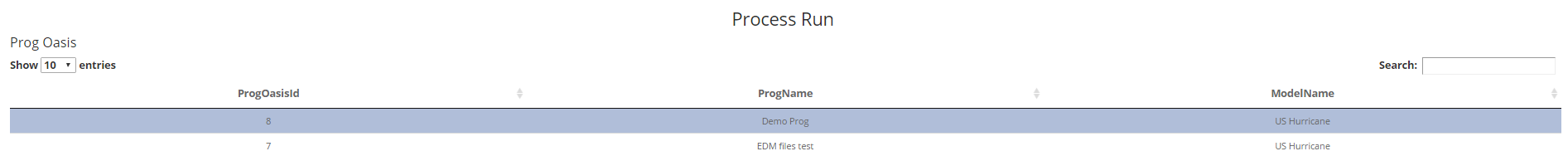
Process Management

Select “Process Management” From the menu bar on the left of the screen

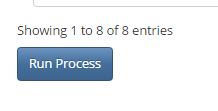




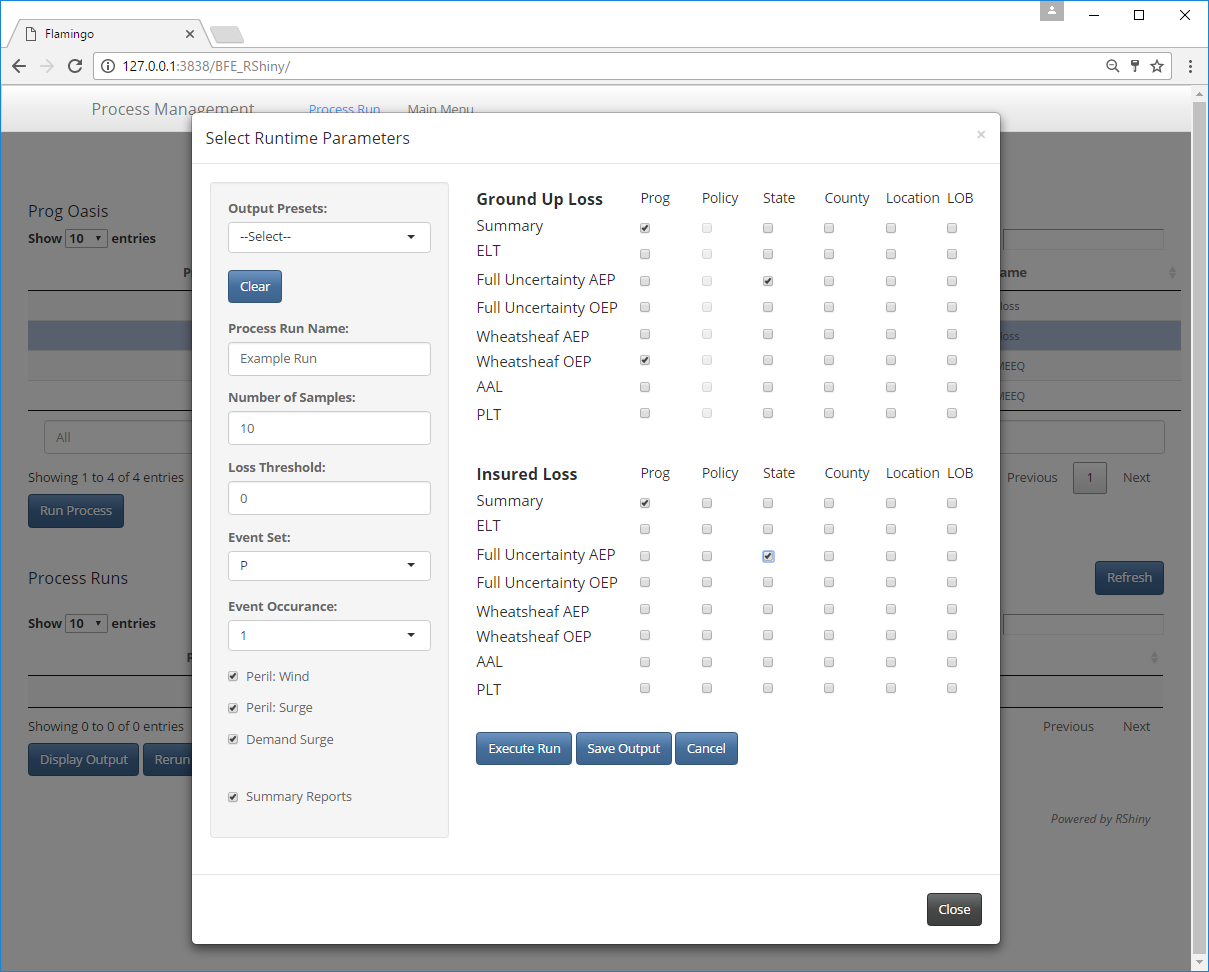
Select the Prog Oasis that you have just created form the Table on the screen



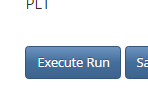
Click on “Run Process” Button



The following pop up screen should appear for you



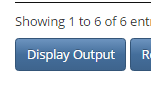
Here you can select the run time parameters that you would like to use and the output option that you would like to be returned. Once you have made your selections, click on the “Execute Run” button



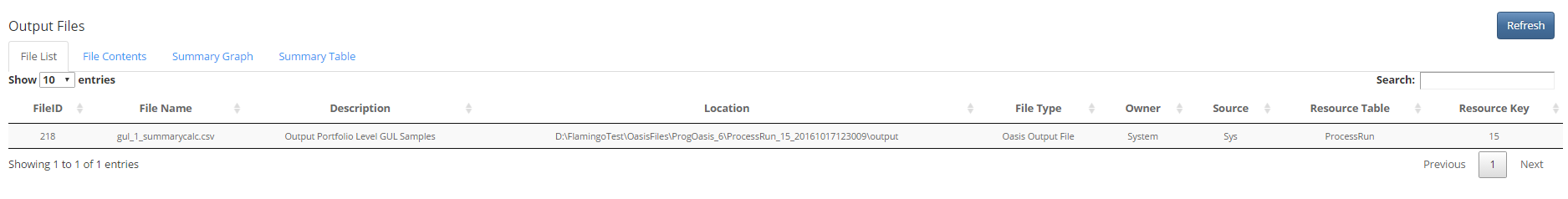
Your process will now be running. You can monitor the progress by refreshing the “Process Runs” Table for the selected ProgOasis. This will tell you if the process run is “In Progress”, “Failed” or “Completed”.

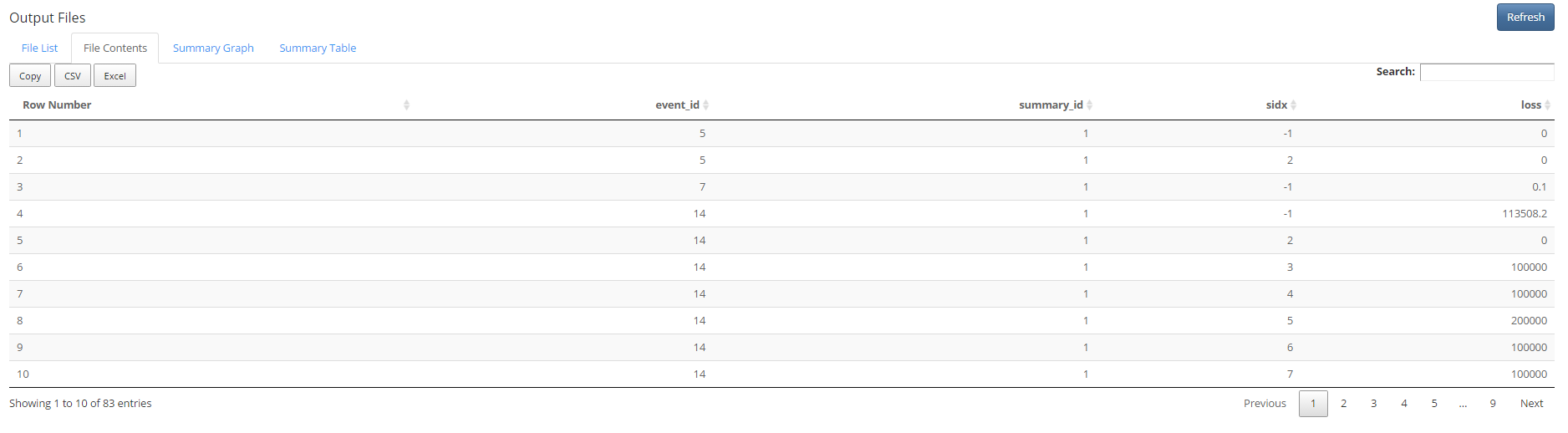


Once Completed, you can select the process run and click on the “Display Outputs” button

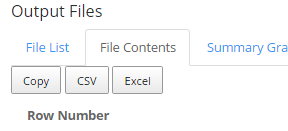


This will display the output files table(s) including a list of output files, the contents of a selected file (note, you must select a file from the list before viewing it in the file contents tab)

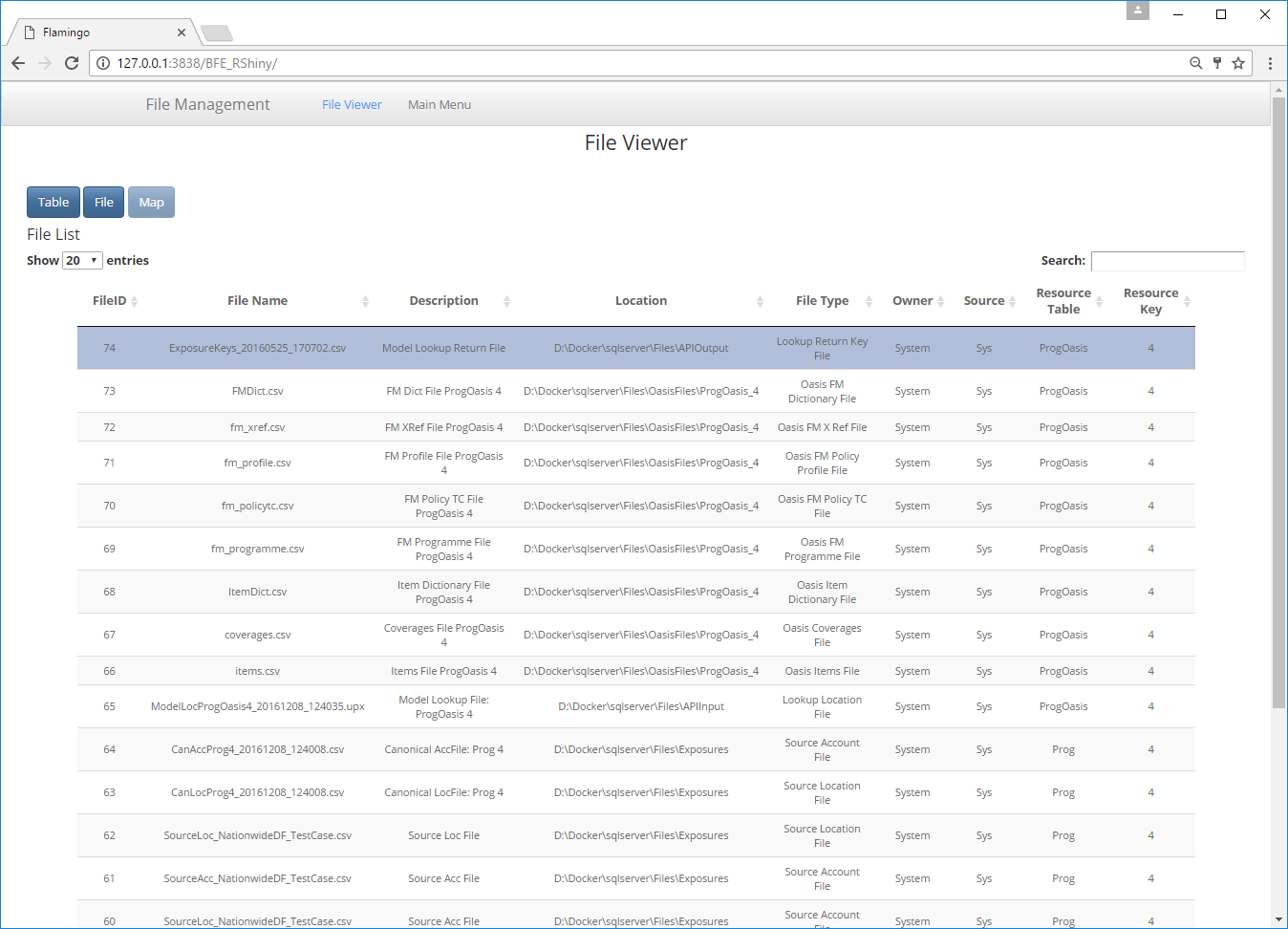


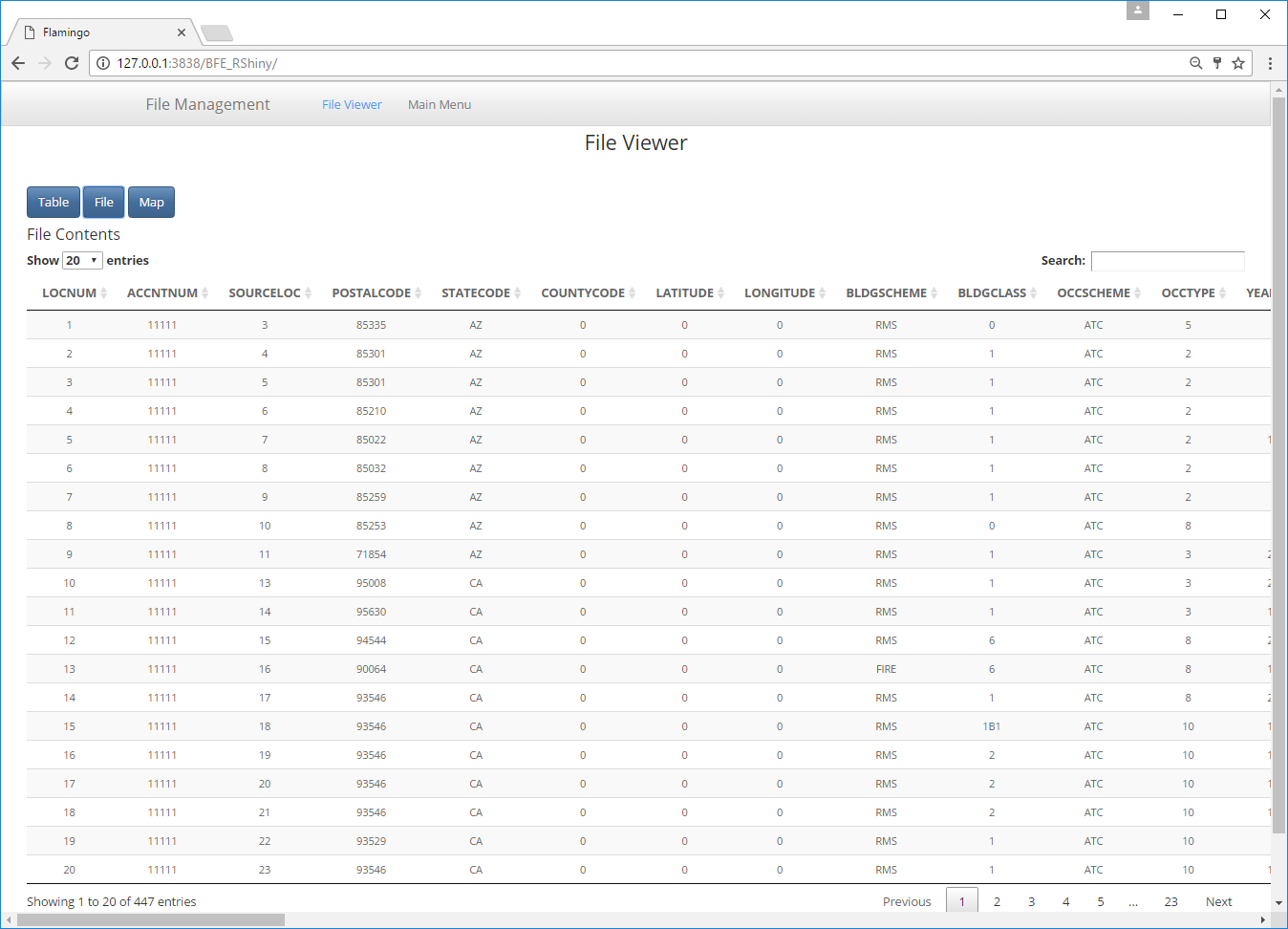


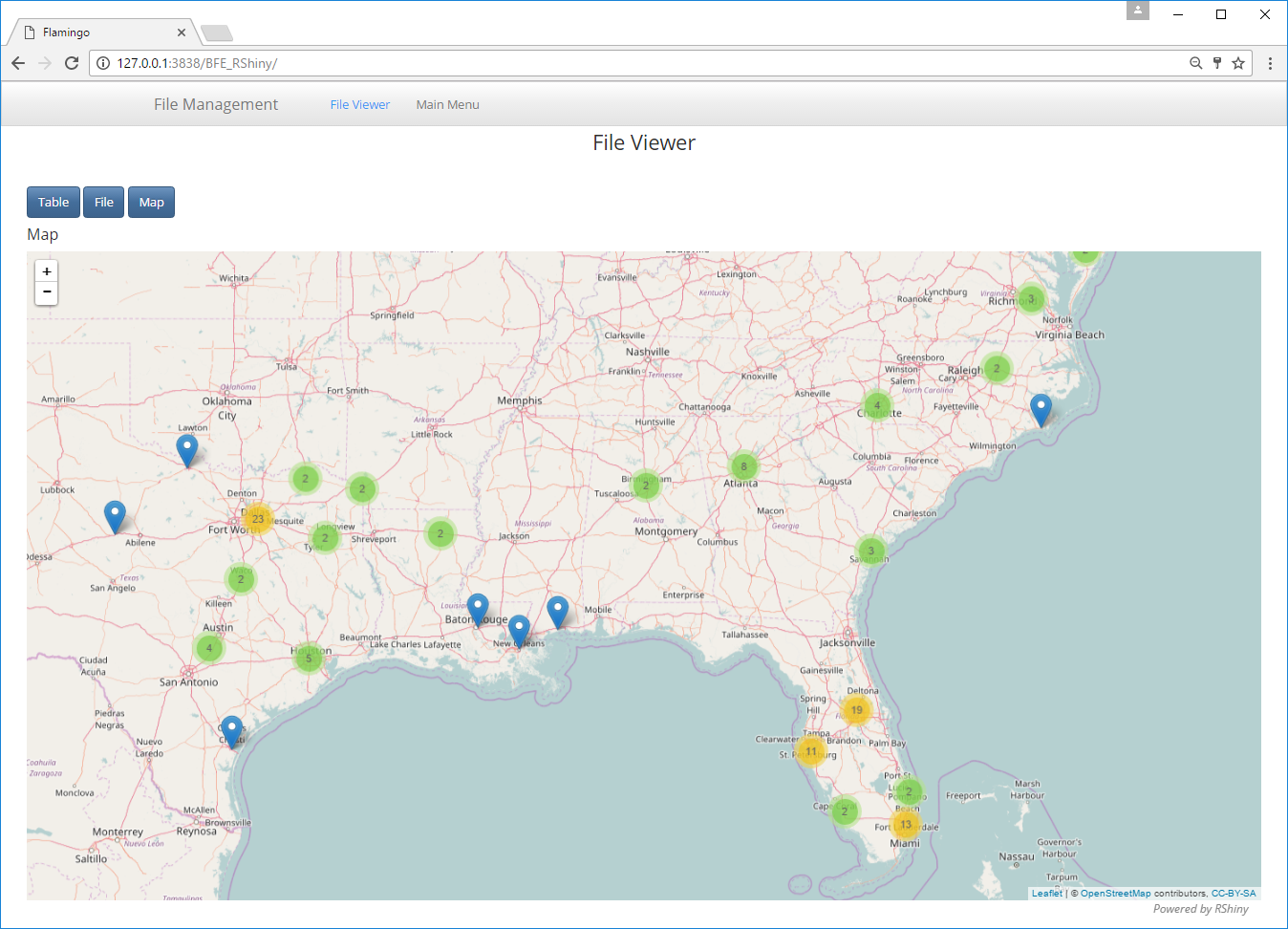
The files can be exported with the following options



File Management







# Outputs

Flamingo enables the user to generate multiple output reports for several summary levels and perspectives in a single process run. Reports are delivered as csv files which can be viewed through the user interface, download or launched in Microsoft Excel.

The Oasis kernel is a Monte-Carlo simulation engine allowing users to specify the number of samples to run. There are two types of statistical outputs, and both types will generally be delivered in the final report if more than 1 sample is run.

* Numerically integrated (type 1) – meaning the loss statistic is calculated directly from the underlying probability distribution of loss by numerical integration.
* Sample statistic (type 2) – meaning the probability distributions of loss are sampled N times and the loss statistic is calculated from the samples.

The reports are abstract, containing kernel ids to identify dimensions of the data. These are;

* **summary\_id** – A number identifying a particular group within a summary level grouping of exposures, such as an individual risk in location summary level, or a particular state code in state summary level.
* **event\_id** – A number uniquely identifying an event in the model
* **period\_no** – A number uniquely identifying a period of event occurrences in the model
* **sidx** – The sample index number.

The user will be interested to know what the summary\_ids in the reports mean, for example that summary\_id 1 means Florida and summary\_id 2 means Texas for state summary level reports. Initially this will have to be looked up from the item dictionary associated with the Programme in File Management. In future, the reports will include the meaningful user and model data.

The list of available reports are as follows;

**Reports**

* Sampled losses
* Average annual loss and standard deviation
* Event loss tables
* Period loss tables
* Single loss exceedance curve (AEP/OEP)
* Multiple loss exceedance curve (AEP/OEP)

**By Summary Level**

* Location
* Line of business
* County
* State
* Programme (whole portfolio)
* Policy (insured loss only)

**By Perspective**

* Ground up
* Insured loss

**Output file description and format**

**Sampled losses**

This report contains sampled losses by summary level, event and sample index.

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| event\_id | int | Model’s unique event identifier |
| sidx | int | Kernel sample index |
| loss | float | Loss sample |

sidx has values with special meanings as follows;

-1 = numerically integrated mean

-3 = total exposed value to event

Example

| **summary\_id** | **event\_id** | **sidx** | **loss** |
| --- | --- | --- | --- |
| 1 | 5 | -1 | 3627.83 |
| 1 | 5 | -3 | 90040000.00 |
| 1 | 5 | 6 | 1001.43 |
| 1 | 5 | 8 | 682.29 |
| 1 | 5 | 10 | 8795.99 |

All other reports outlined below are calculated from the sample loss output.

**Average annual loss and standard deviation**

Average annual loss is calculated by taking the sum of the mean aggregate losses by period and then dividing by the total number of periods, for each summary level. The standard deviation measures the spread of mean aggregate losses by period. The exposure value is the maximum exposed value across all periods.

Two types of AAL and standard deviation of loss are calculated; numerically integrated (type 1) and sample (type 2). If the analysis is run with zero samples, then only type 1 statistics are returned.

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| type | int | 1 for numerically integrated statistics, 2 for sample statistics |
| mean | float | Average annual loss |
| standard\_deviation | float | Standard deviation of loss |
| exposure\_value | float | Maximum exposure value in any period |

Example

| **summary\_id** | **type** | **mean** | **standard\_deviation** | **exposure\_value** |
| --- | --- | --- | --- | --- |
| 1 | 1 | 6927.54 | 7268.98 | 591599.63 |
| 1 | 2 | 6852.44 | 7165.04 | 591599.63 |
| 2 | 1 | 9852.43 | 4995.02 | 67536.58 |
| 2 | 2 | 9354.68 | 5112.30 | 67536.58 |
| 3 | 1 | 4331.58 | 3574.95 | 185136.03 |

**Event Loss Tables**

For each event and summary\_id, representing a summary level grouping, the sample mean and standard deviation is calculated from the sampled losses. The exposure\_value is the total exposure value to the event of the subset of exposures represented by each summary\_id.

Both type 1 and type 2 loss statistics are reported, although type 1 standard deviation is not calculated and is set to zero.

When zero samples are run, only type 1 losses are output, and both type 1 and 2 are output when more than one sample is run.

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| type | int | 1 for analytical mean, 2 for sample mean |
| event\_id | int | Model’s unique event identifier |
| mean | float | Mean loss by event |
| standard\_deviation | float | Sample standard deviation, or 0 for type 1 |
| exposure\_value | float | Exposure value for summary\_id affected by the event |

Example

| **summary\_id** | **type** | **event\_id** | **mean** | **standard\_deviation** | **exposure\_value** |
| --- | --- | --- | --- | --- | --- |
| 1 | 1 | 1 | 8812.15 | 0 | 215971.65 |
| 1 | 2 | 1 | 9289.46 | 4590.79 | 215971.65 |
| 1 | 1 | 2 | 2967.23 | 0 | 773622.84 |
| 2 | 2 | 2 | 3417.72 | 2265.98 | 773622.84 |
| 2 | 1 | 1 | 3343.97 | 0 | 202025.43 |

**Period Loss Tables**

These reports are similar to Event Loss Tables, except the sampled losses are first summed by period before the loss statistics are calculated. It also contains an event occurrence date in one of two formats.

A period can be any length of time and there can be 0, 1 or more event occurrences in each period, as specified in the model files. Typically, a period is one year.

The exact format depends on how the occurrence date is specified in the model which is either an occurrence date\_id, or separate fields for occurrence year, month and day. The loss statistics will be the same regardless of which date format is used.

Only type 2 (sampled) loss statistics are provided in this report at present.

Format 1

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| event\_id | int | Model’s unique event identifier |
| period\_no | int | Identifies an abstract period of time, such as a year |
| mean | float | Mean loss |
| standard\_deviation | float | Standard deviation of loss |
| exposure\_value | float | Represents a summary level grouping of exposures |
| occ\_date\_id | int | The date\_id of the event occurrence |

Format 2

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| event\_id | int | Model’s unique event identifier |
| period\_no | int | Identifies an abstract period of time, such as a year |
| mean | float | Mean loss |
| standard\_deviation | float | Standard deviation of loss |
| exposure\_value | float | Exposure value for summary\_id affected by during the period |
| occ\_year | int | The year number of the event occurrence |
| occ\_month | int | The month of the event occurrence |
| occ\_day | int | The day of the event occurrence |

Example

| **summary\_id** | **event\_**  **id** | **period\_no** | **mean** | **standard\_**  **deviation** | **exposure\_**  **value** | **occ\_date\_id** |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 1 | 9932.43 | 1604.47 | 65809.55 | 693 |
| 1 | 1 | 102 | 7552.92 | 6876.1 | 654795.8 | 37459 |
| 1 | 2 | 56 | 1668.34 | 7151.88 | 217547.78 | 20636 |
| 1 | 2 | 89 | 6966.3 | 6074.69 | 404692.86 | 32801 |
| 1 | 2 | 104 | 9472.54 | 2361.71 | 96472.52 | 37982 |

| **summary\_id** | **event\_**  **id** | **period\_no** | **mean** | **standard\_**  **deviation** | **exposure\_**  **value** | **occ\_**  **year** | **occ\_**  **month** | **occ\_**  **day** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 1 | 9932.43 | 1604.47 | 65809.55 | 1901 | 11 | 23 |
| 1 | 1 | 102 | 7552.92 | 6876.1 | 654795.8 | 2002 | 7 | 22 |
| 1 | 2 | 56 | 1668.34 | 7151.88 | 217547.78 | 1956 | 6 | 30 |
| 1 | 2 | 89 | 6966.3 | 6074.69 | 404692.86 | 1989 | 10 | 20 |
| 1 | 2 | 104 | 9472.54 | 2361.71 | 96472.52 | 2003 | 12 | 27 |

**Single loss exceedance curve (OEP/AEP)**

Loss exceedance curves, also known as exceedance probability curves, are computed by a rank ordering a set of losses by period and computing the return period interval for each level of loss in any given period based on relative frequency. The reciprocal of the return period interval is termed the loss exceedance probability.

To calculate a loss exceedance curve, event losses are first assigned to periods by reference to the model’s occurrence file which contains the event occurrences in each period. Event losses are summed within each period for an aggregate loss exceedance curve, or the maximum of the event losses in each period is taken for an occurrence loss exceedance curve.

A single loss exceedance curve can be calculated from sampled results when every sample is treated as an additional set of period losses in an extended timeline. For example, if the model covers a 1000 year period and it is sampled 10 times, then 10,000 periods of loss are generated, which are then rank ordered as described above. This is also termed a ‘full uncertainty’ loss exceedance curve.

If the report is run with zero samples, then the curve generated is implicitly a type 1 curve: a mean-only loss exceedance curve based on the numerically integrated mean period loss.

The list of return periods for which losses are calculated is a setting in Flamingo, typically specified by the model provider.

Format

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| return\_period | float | Return period interval |
| loss | float | Loss exceedance threshold for return period |

Example

| **summary\_id** | **return\_period** | **loss** |
| --- | --- | --- |
| 1 | 1000 | 62657.65 |
| 1 | 500 | 9305.84 |
| 1 | 250 | 1664.71 |
| 1 | 200 | 778.26 |
| 1 | 100 | 40.4 |

**Multiple loss exceedance curves (OEP/AEP)**

This report is also generated from the sampled results by relative frequency as described above, but this time each sample is used to generate a separate curve which when put together can be viewed as a fan or spaghetti chart. This type of exceedance probability curve is also termed a “Wheatsheaf”.

Unlike the single exceedance probability, which does not illustrate the amount of variability in sampled losses, this report can illustrate the spread in return period losses obtained by repeated sampling of the event set across the period timeline.

If the report is run with zero samples, then a curve is generated for sample index -1 only: a mean-only loss exceedance curve based on the numerically integrated mean period loss. This is an identical curve to the ‘type 1’ single loss exceedance curve report run with zero samples.

Format

| **Name** | **Type** | **Description** |
| --- | --- | --- |
| summary\_id | int | Represents a summary level grouping of exposures |
| sidx | int | Oasis sample index |
| return\_period | float | Return period interval |
| loss | float | Loss exceedance threshold for return period |

Example

| **summary\_id** | **sidx** | **return\_period** | **loss** |
| --- | --- | --- | --- |
| 1 | 9 | 1000 | 52657.65 |
| 1 | 9 | 500 | 10305.84 |
| 1 | 9 | 250 | 1965.71 |
| 1 | 9 | 200 | 836.26 |
| 1 | 9 | 100 | 35.4 |

# User Administration

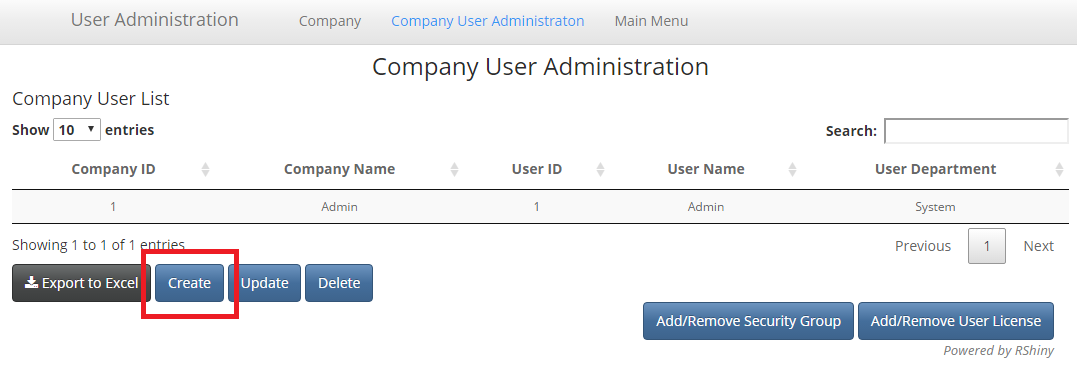
Navigate to the Landing Page and then into the “User Administration” area



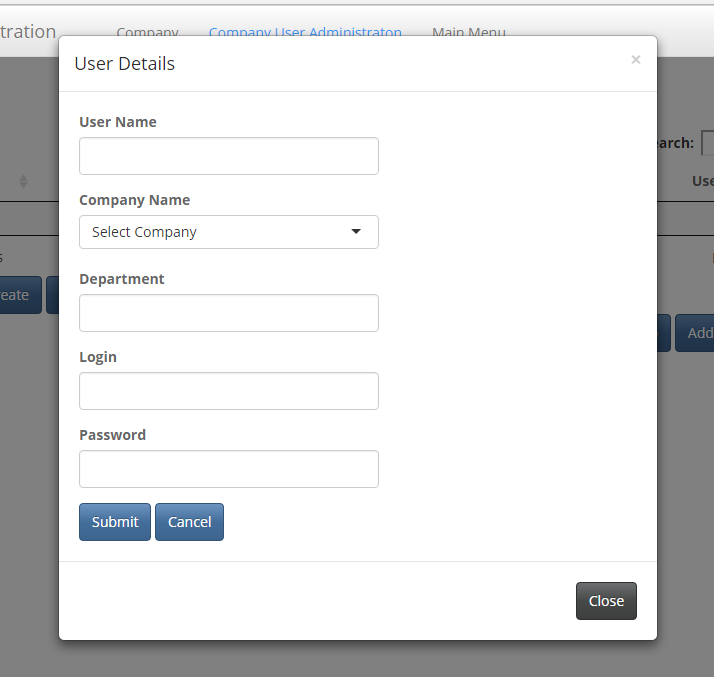
Now navigate to the “Company User Administration” area from the top menu bar



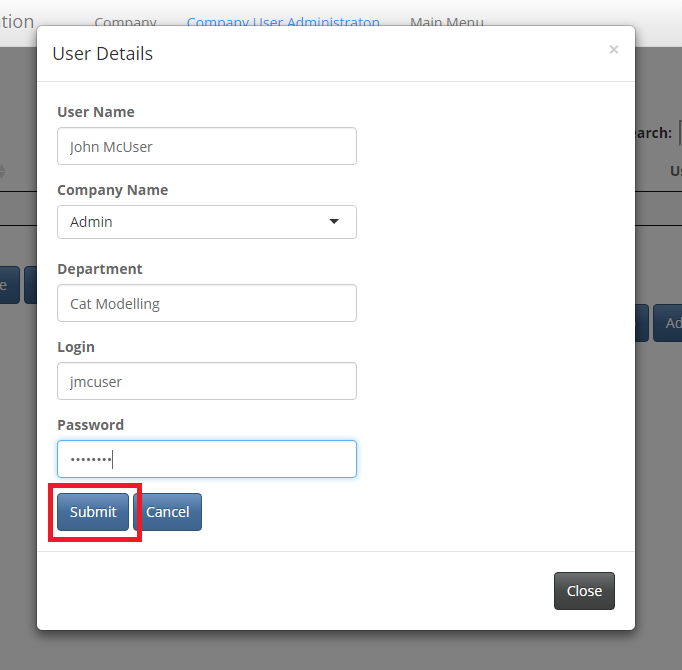
To create a new user, click on the “Create” Button under the Company User List



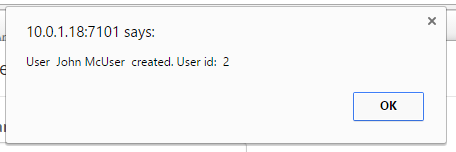
And you will be presented with the following pop up screen



Enter the Details for your new user in the boxes and click on “Submit”



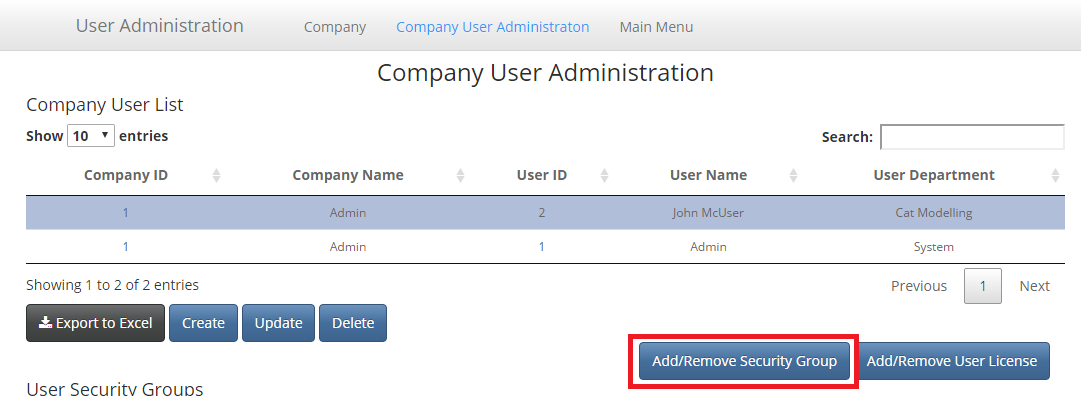
You should be presented with the following confirmation message



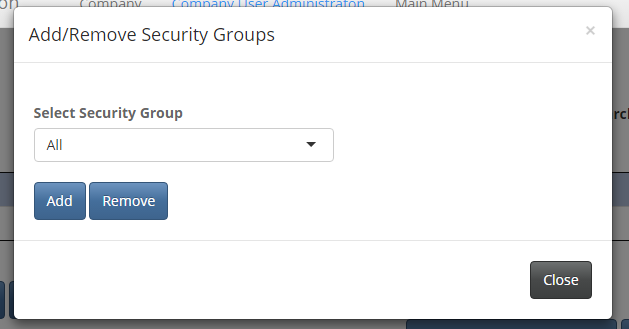
Click “ok” and you should be able to see the user that you have just set up in the Company User list



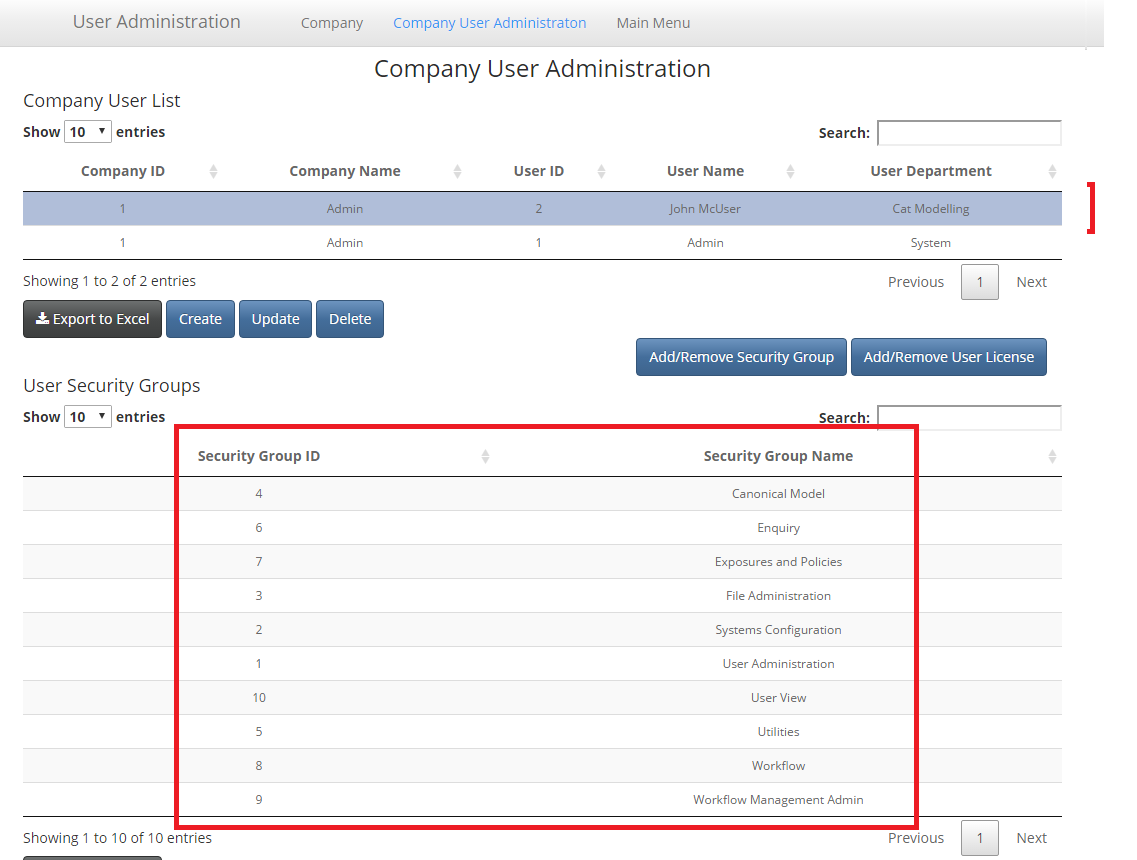
Select the user that you have just created from the list and click on the “Add/Remove Security Group” button



You should see the following pop up screen for you to add security privileges



You should choose the particular security groups that your new user and add them one at a time – or alternatively give access to “All” and then remove groups as appropriate – and then click “Add” to apply the changes. You should then see the security groups have been added to the user



You can now navigate back to the landing page by selecting “Main Menu” from the top menu and log out as the current user to test the new user credentials.