

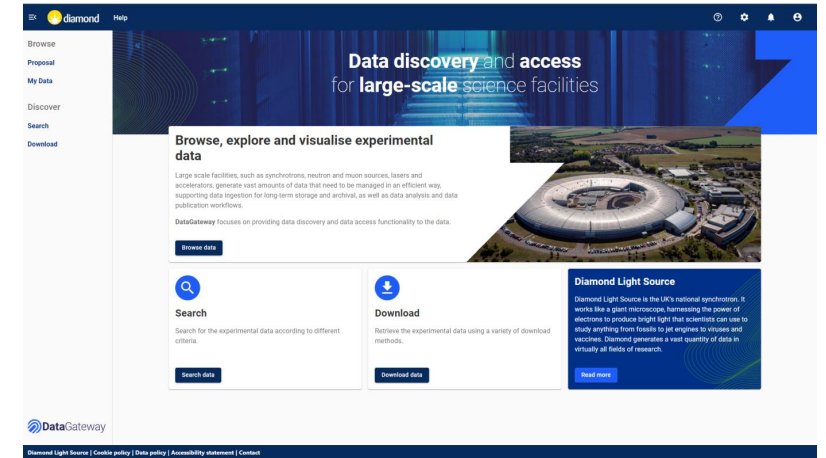
Retrieval of Data

Retrieval of Data: Summary

- Archiving Policy
- ISPyB
 - Auto-processed data retrieval
- rsync < 20 Gb Live data
 - <https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/rsync.html>
- Globus
 - Live data retrieval
- Data Gateway
 - Archived raw & processed Data retrieval

Archiving & Data Release Policy

- Raw Data (i.e. Diffraction Images) can be archived any time from **40 days** after the visit for academic users
- Archived data is stored on tape – a single non-redundant copy is kept on tape
- Retrieval after 40 days is via Data Gateway
- **3 years** after data is collected - Data can be made available on an open-access basis under a **CC-BY-4.0 licence**.
 - An extension to this period can be made by the principal investigator
- <https://www.diamond.ac.uk/Home/Legal-and-Compliance/Policies/Experimental-Data-Management-Pol.html>

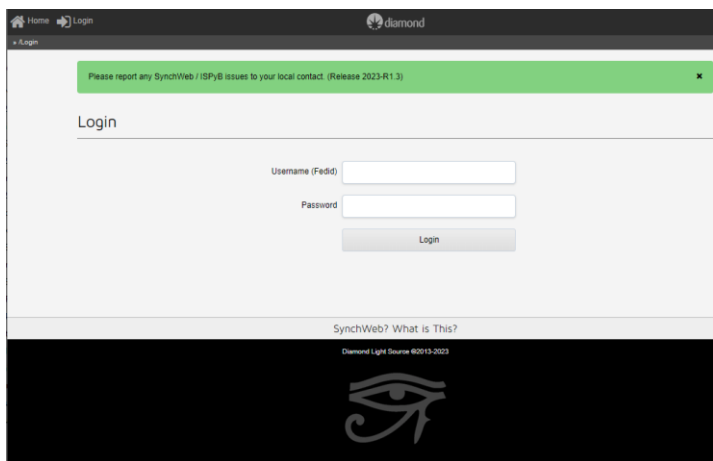


Data Gateway

<https://icat.diamond.ac.uk/>

Archived Raw Data

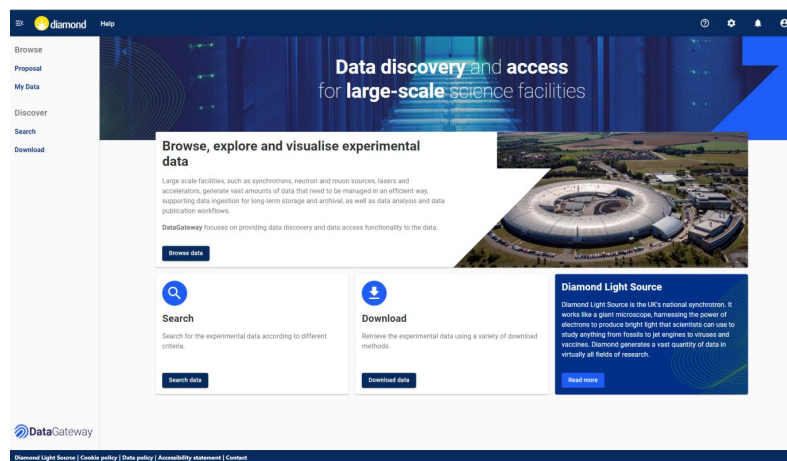
Diamond MX – Key Software Interfaces



ISPyB

<https://ispyb.diamond.ac.uk/>

Processed Data Files



Data Gateway

<https://icat.diamond.ac.uk/>

Archived Raw Data



Globus

<http://www.globus.org/>

Live Data – Raw & Processed

- Sign in with your organisation login, or personal Globus ID.

Data Retrieval

ISPyB - Logs & Files

- For each dataset, every auto-processing & downstream processing pipeline has a [log and files](#) button
- Here you can retrieve the processed mtz files, logs and other intermediate files
 - An archive of all the files can be downloaded
 - Some intermediate files are only on disk
- Raw images cannot be retrieved via ISPyB

Beam Centre	X	Y
Start	108.70	109.70
Refined	108.74	109.70
Δ	-0.04	0.00

Space Group	A	B	C	α	β	γ
P 21 2 2	45.11	110.58	135.74	90.00	90.00	90.00

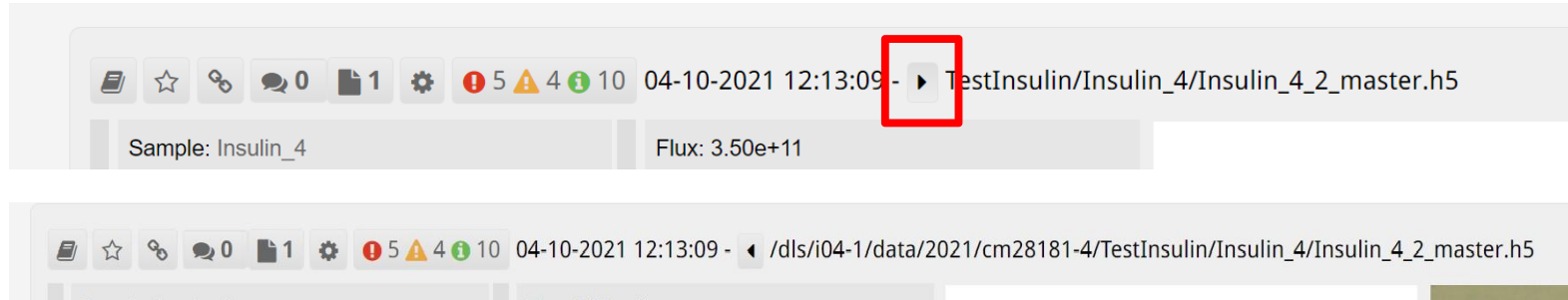
Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	CC Half	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity	CC Anom
outerShell	1800	1103	1.75 - 1.78	1.203	0.7	0.2	31.4	1.7	19.0	1.1	0.0
innerShell	2713	1074	4.75 - 49.12	0.001	29.8	1.0	28.0	2.5	21.9	1.0	0.1
overall	49700	24573	1.75 - 45.11	0.107	8.7	1.0	34.7	2.0	24.1	1.2	0.0

File	Type	Download	View
xia2.txt	Log	Download	View
xia2.html	Log	Download	View
xia2.cif	Result	Download	
AUTOMATIC_DEFAULT_SAD_SWEEP1_CORRECT.HKL	Result	Download	
AUTOMATIC_DEFAULT_free.mtz	Result	Download	
AUTOMATIC_DEFAULT_scaled_unmerged.sca	Result	Download	
AUTOMATIC_DEFAULT_SAD_SWEEP1_INTEGRATE.mtz	Result	Download	
AUTOMATIC_DEFAULT_scaled_unmerged.mtz	Result	Download	
AUTOMATIC_DEFAULT_scaled.sca	Result	Download	
AUTOMATIC_DEFAULT_SAD_SWEEP1_INTEGRATE.HKL	Result	Download	
AUTOMATIC_DEFAULT_SAD_SWEEP1.expt	Result	Download	
xia2.mmcif	Result	Download	
AUTOMATIC_DEFAULT_SAD_SWEEP1_INTEGRATE.log	Log	Download	View
Xtrriage.log	Log	Download	View
AUTOMATIC_DEFAULT_SAD_merging-statistics.txt	Log	Download	View

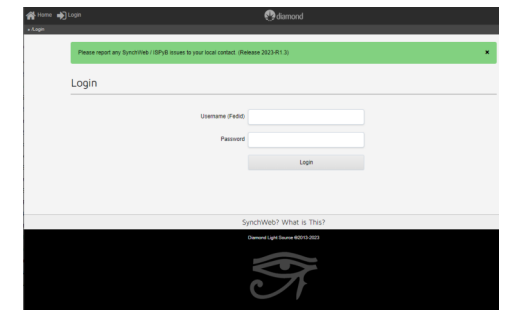
ISPyB
<https://ispyb.diamond.ac.uk/>
Processed Data Files

ISPyB - Logs & Files

- These can also be retrieved from the disk under the `processed` subdirectory for a visit
- The path for the raw data which can be seen on ISPyB by clicking the triangle icon on a dataset:



- The end of the path would need to be changed:
- `.../<visit>/processed/<protein>/<sample_name>/<dataset_name>/<processing-pipeline>/`



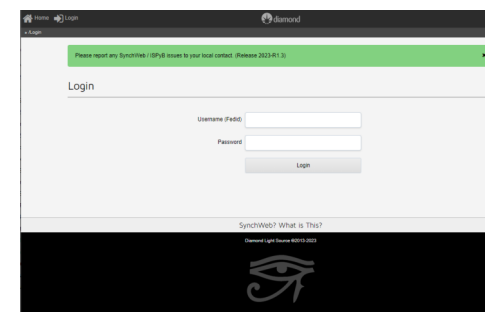
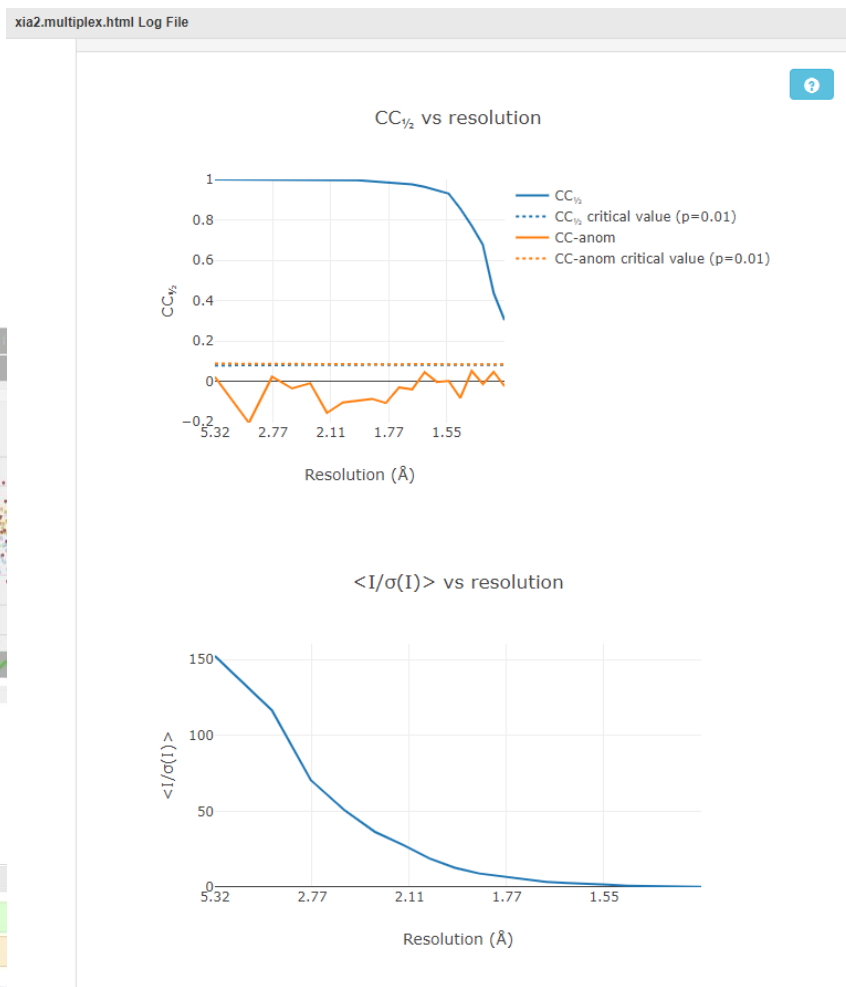
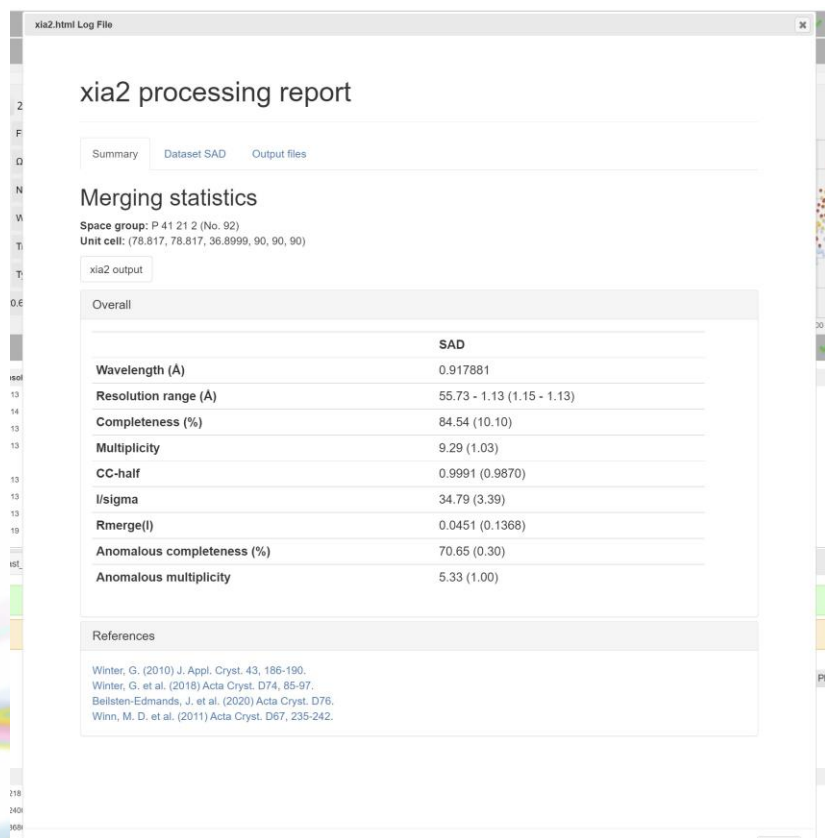
ISPyB

<https://ispyb.diamond.ac.uk/>

Processed Data Files

ISPyB - Logs & Files

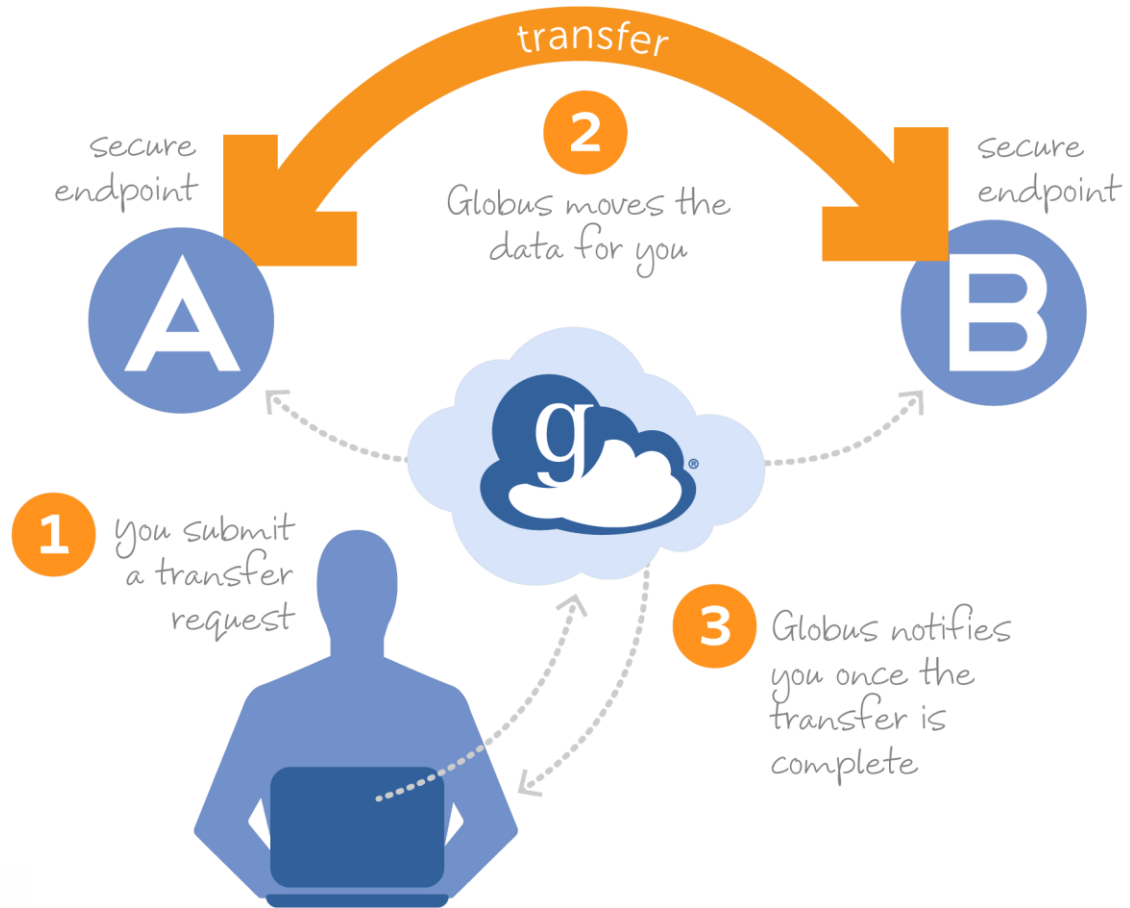
- Reports such as available from logs and files interface can have very useful graphs for assessing data quality during collection



ISPyB
<https://ispyb.diamond.ac.uk/>
Processed Data Files

Globus

- Globus is a file transfer service free at the point of service to non-profit organisations
- Many institutions have an endpoint to which data can be transferred
- Users can create a personal endpoint to transfer to a personal/local machine
 - <https://www.globus.org/globus-connect-personal>
- LS-CAT & SERCAT users
 - Some data transfers may be managed on your behalf
- Globus has Python scripting interface for automated data retrieval:
 - <https://docs.globus.org/cli/>



• Sign in with your organisation login, or personal Globus ID.

Globus

<http://www.globus.org/>

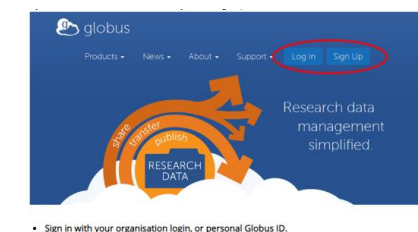
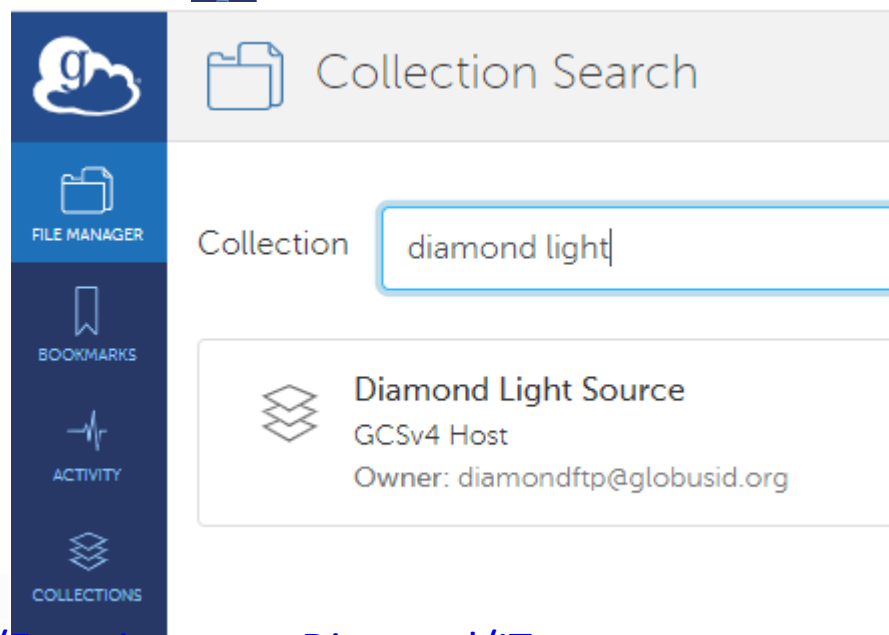
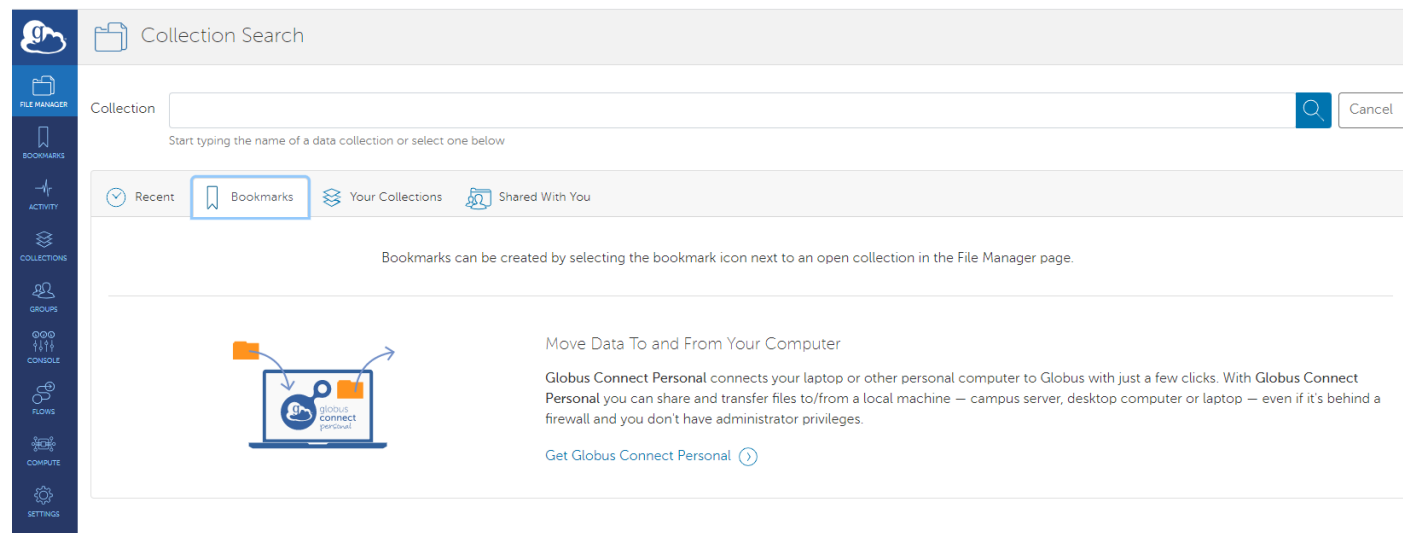
Live Data – Raw & Processed



<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>

Globus

- Globus web portal can be used once logged in to an account to setup transfers
- Search for the Diamond Light Source Endpoint



<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>

Globus
<http://www.globus.org/>
Live Data – Raw & Processed



Globus

- Log into the endpoint with diamond fedid and password

The screenshot shows the Globus File Manager interface. On the left is a dark blue sidebar with icons for FILE MANAGER, BOOKMARKS, ACTIVITY, COLLECTIONS, GROUPS, CONSOLE, FLOWS, COMPUTE, SETTINGS, and LOGOUT. The main area is titled 'File Manager' and shows the 'Collection' as 'Diamond Light Source'. Below this are fields for 'Path' and a toolbar with 'select all', 'up one folder', 'refresh list', and 'filter' options. A central message reads 'Please authenticate to access Diamond Light Source'. Below this message are fields for 'Login Server' (globus-id.diamond.ac.uk), 'Username', and 'Password', each with a red eye icon for toggling visibility. An 'Advanced' dropdown menu is visible below the password field. At the bottom are 'Authenticate' and 'Cancel' buttons.

<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>



• Sign in with your organisation login, or personal Globus ID.

Globus
<http://www.globus.org/>
Live Data – Raw & Processed



Globus

- You can then navigate the live files on disk



The screenshot shows the Globus File Manager interface. The top bar displays 'File Manager' and the current collection is 'Diamond Light Source'. The path is '/dls/'. Below the path, there are controls for 'select all', 'up one folder', 'refresh list', and 'filter'. A table lists the contents of the folder:

NAME	LAST MODIFIED	SIZE
a01	4/6/2023, 01:28 PM	—
b01	4/6/2023, 01:28 PM	—
b07	4/6/2023, 01:28 PM	—
b07-1	4/6/2023, 01:28 PM	—
b16	4/6/2023, 01:28 PM	—
b18	4/6/2023, 01:28 PM	—
b21	4/6/2023, 01:28 PM	—
b22	4/6/2023, 01:28 PM	—



• Sign in with your organisation login, or personal Globus ID.

Globus

<http://www.globus.org/>

Live Data – Raw & Processed



<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>

Globus

- You can then navigate the live files on disk

The screenshot shows the Globus File Manager interface. The top navigation bar includes icons for File Manager, Bookmarks, Activity, Collections, Groups, Console, Flows, Compute, Settings, Logout, and Help & Sitemap. The main content area displays a 'File Manager' window for the 'Diamond Light Source' collection. The path is '/dls/i03/2023/mx34066-6/'. The file list includes:

NAME	LAST MODIFIED	SIZE
GDAUserOptions.xml	2/23/2023, 11:49 AM	95 B
jpeg	2/23/2023, 12:09 PM	-
processed	2/23/2023, 12:17 PM	-
processing	2/23/2023, 11:08 AM	-
Test_ProteinaseK	2/23/2023, 12:28 PM	-
Test_Thermolysin	2/23/2023, 12:08 PM	-
xml	2/23/2023, 12:08 PM	-
xraycentring	2/23/2023, 12:08 PM	-
xta_snapshots	2/23/2023, 12:28 PM	-



• Sign in with your organisation login, or personal Globus ID.

Globus

<http://www.globus.org/>

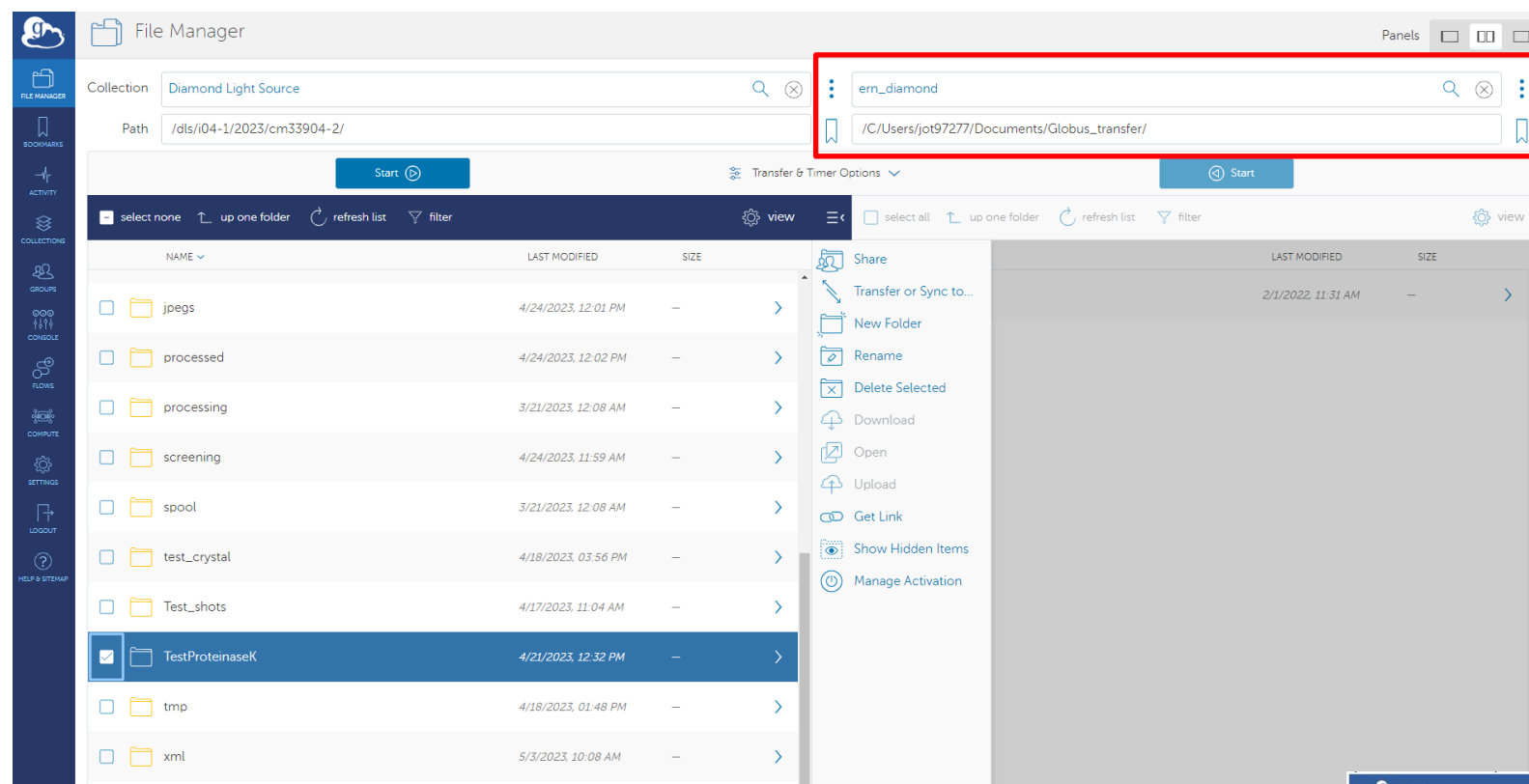
Live Data – Raw & Processed

<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>



Globus

- Select a personal or institution endpoint to transfer to



• Sign in with your organisation login, or personal Globus ID.

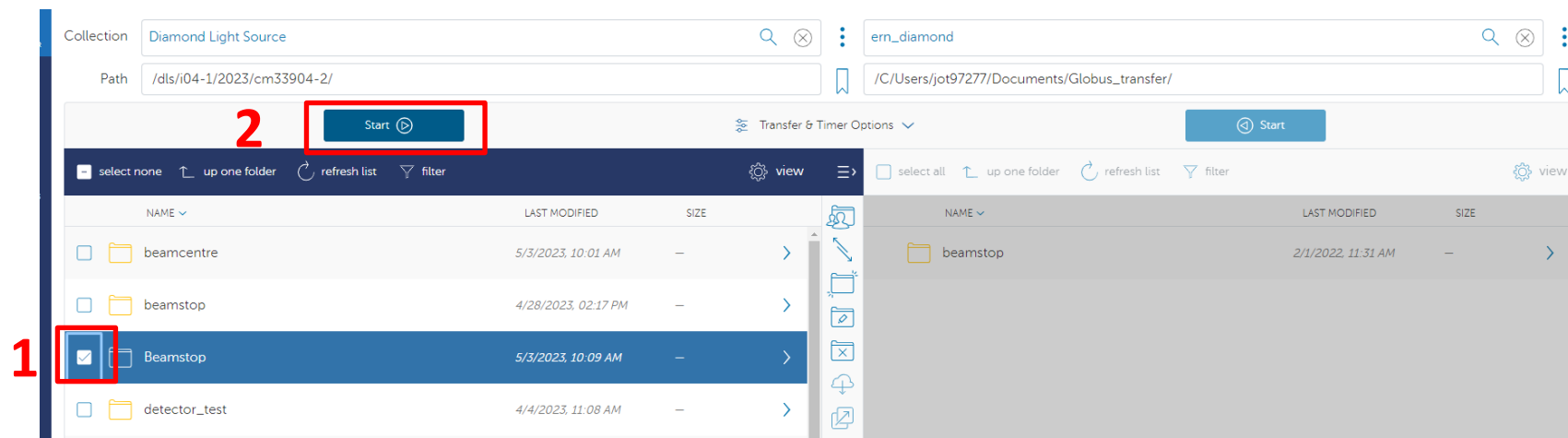
<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>

Globus
<http://www.globus.org/>
Live Data – Raw & Processed



Globus

- Once data is selected initiate transfer with start command
- You can review progress of a transfer

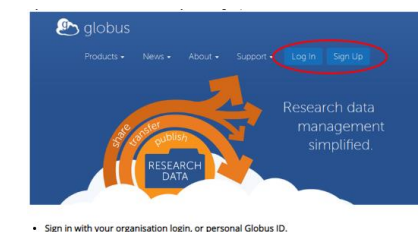
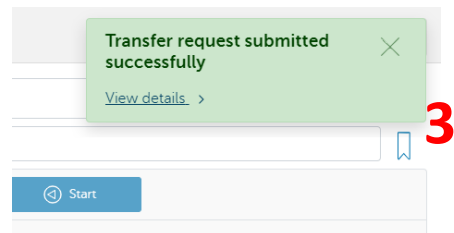


4

Diamond Light Source to ern_diamond
transfer started

Overview | Event Log

Task Label	Diamond Light Source to ern_diamond
Source	▶ Diamond Light Source
Destination	▶ ern_diamond
Task ID	2b499802-e99f-11ed-ba34-09d6a6f08166
Owner	Elliot Nelson (nelse003@gmail.com)
Condition	ACTIVE
Requested	5/3/2023, 11:42 AM
Deadline	5/4/2023, 11:42 AM
Duration	17 seconds
Transfer Settings	<ul style="list-style-type: none">• verify file integrity after transfer• transfer is not encrypted• overwriting all files on destination



<https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data/in40days/Globus.html>

Globus
<http://www.globus.org/>
Live Data – Raw & Processed



Data Gateway

- Data gateway allows access to archived academic data
 - Image data > 40 days old
- Sign in is via Fedid & password

The screenshot shows the Diamond Data Gateway website. The header includes the Diamond logo and a 'Help' link. A navigation menu on the left lists 'Browse', 'Proposal', 'My Data', 'Discover', 'Search', and 'Download'. The main content area features a large banner with the text 'Data discovery and access for large-scale science facilities'. Below this, there are three main sections: 'Browse, explore and visualise experimental data' with a 'Browse data' button; 'Search' with a 'Search data' button; and 'Download' with a 'Download data' button. A 'Diamond Light Source' section is also present with a 'Read more' button. The footer contains the 'DataGateway' logo and a link to 'Diamond Light Source | Cookie policy | Data policy | Accessibility statement | Contact'.

Data Gateway

<https://icat.diamond.ac.uk/>

Archived Raw Data

Data Gateway

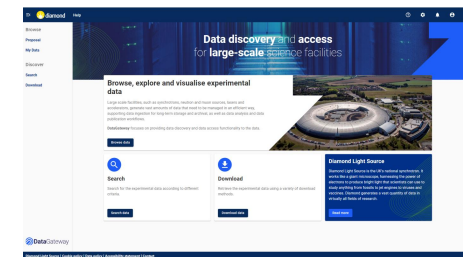
- Select the proposal of interest
- Then visit(s)

1

2

3

Visit ID	Dataset Count	Beamline	Start Date	End Date
MX34066-8	0	I04	2023-05-05 15:00:00+01:00	2023-05-05 18:00:00+01:00
MX34066-7	1	I03	2023-03-15 10:00:00+00:00	2023-03-15 11:00:00+00:00
MX34066-6	40	I03	2023-02-23 11:00:00+00:00	2023-02-23 19:00:00+00:00
MX34066-4	8	I04	2023-02-23 09:00:00+00:00	2023-02-23 17:00:00+00:00
MX34066-1	0	I03	2023-02-22 10:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-5	8	I24	2023-02-22 10:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-2	5	I23	2023-02-22 09:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-3	126	I02-2	2023-02-21 15:42:14+00:00	2023-02-25 19:42:14+00:00



Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data



Data Gateway

- A summary of the visit size and data is available

Home > Proposals > Diamond MX BAG training 2023 > Visits

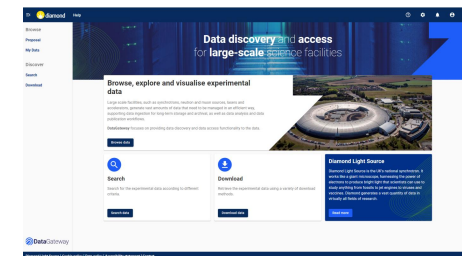
Results: 8

Display as cards Clear filters 1 item has been added to selection. Go to selection

Visit ID	Dataset Count	Beamline	Start Date	End Date
MX34066-8	0	i04	2023-05-05 15:00:00+01:00	2023-05-05 18:00:00+01:00
MX34066-7	1	i03	2023-03-15 10:00:00+00:00	2023-03-15 11:00:00+00:00
MX34066-6	40	i03	2023-02-23 11:00:00+00:00	2023-02-23 19:00:00+00:00
MX34066-4	8	i04	2023-02-23 09:00:00+00:00	2023-02-23 17:00:00+00:00
MX34066-1	0	i03	2023-02-22 10:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-5	8	i24	2023-02-22 10:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-2	5	i23	2023-02-22 09:00:00+00:00	2023-02-22 18:00:00+00:00
MX34066-3	126	i02-2	2023-02-21 15:42:14+00:00	2023-02-25 19:42:14+00:00

MX34066-6

Visit Details	Visit Users	Visit Samples	Publications
MX34066			
TITLE			
Diamond MX BAG training 2023			
VISIT ID			
MX34066-6			
SUMMARY			
BAG Training to support Users in making best use of the beamlines at Diamond			
START DATE			
2023-02-23 11:00:00+00:00			
END DATE			
2023-02-23 19:00:00+00:00			
SIZE			
395 GB			



Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data



Data Gateway

- Within a visit the data is structured as it was on disk
- Hence for a single sample you may want to get centring information, processed data and raw images

The screenshot displays the Diamond Data Gateway interface. The top navigation bar includes 'Home', 'Proposals', 'Diamond MX BAG training 2023', 'MX34066-6', and 'Datasets'. A search bar on the right shows 'Results: 40'. A shopping cart icon in the top right corner is highlighted with a red box and contains the number '2'. A yellow notification bar at the top of the table area says '1 item has been added to selection. Go to selection.' The main table lists datasets with columns for Name, Datafile Count, Create Time, and Modified Time. Two rows are highlighted with red boxes: 'xraycentring/Test_ProteinaseK/10' and 'Test_ProteinaseK/10'. The 'Data Gateway' logo is visible in the bottom left corner. An inset image in the bottom right corner shows a promotional banner for 'Data discovery and access for large-scale science facilities'.

Name	Datafile Count	Create Time	Modified Time
<input type="checkbox"/> xraycentring/Test_ProteinaseK/7	8	2023-02-23 19:24:23.158000+00:00	2023-02-23 19:24:23.158000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/6	8	2023-02-23 19:24:23.029000+00:00	2023-02-23 19:24:23.029000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/9	8	2023-02-23 19:24:22.905000+00:00	2023-02-23 19:24:22.905000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/8	8	2023-02-23 19:24:22.783000+00:00	2023-02-23 19:24:22.783000+00:00
<input type="checkbox"/> Test_Thermolysin/16	15	2023-02-23 19:24:22.525000+00:00	2023-02-23 19:24:22.525000+00:00
<input checked="" type="checkbox"/> xraycentring/Test_ProteinaseK/10	8	2023-02-23 19:24:22.397000+00:00	2023-02-23 19:24:22.397000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/11	8	2023-02-23 19:24:22.266000+00:00	2023-02-23 19:24:22.266000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/12	8	2023-02-23 19:24:22.144000+00:00	2023-02-23 19:24:22.144000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/13	8	2023-02-23 19:24:22.020000+00:00	2023-02-23 19:24:22.020000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/14	8	2023-02-23 19:24:21.894000+00:00	2023-02-23 19:24:21.894000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/15	9	2023-02-23 19:24:21.754000+00:00	2023-02-23 19:24:21.754000+00:00
<input type="checkbox"/> xraycentring/Test_ProteinaseK/16	9	2023-02-23 19:24:21.611000+00:00	2023-02-23 19:24:21.611000+00:00
<input type="checkbox"/> Test_ProteinaseK/16	6	2023-02-23 19:24:21.497000+00:00	2023-02-23 19:24:21.497000+00:00
<input type="checkbox"/> Test_ProteinaseK/14	6	2023-02-23 19:24:21.387000+00:00	2023-02-23 19:24:21.387000+00:00
<input type="checkbox"/> xraycentring/Test_Thermolysin/16	8	2023-02-23 19:24:21.263000+00:00	2023-02-23 19:24:21.263000+00:00
<input type="checkbox"/> Test_ProteinaseK/15	6	2023-02-23 19:24:21.150000+00:00	2023-02-23 19:24:21.150000+00:00
<input type="checkbox"/> Test_ProteinaseK/12	6	2023-02-23 19:24:21.044000+00:00	2023-02-23 19:24:21.044000+00:00
<input type="checkbox"/> Test_ProteinaseK/13	6	2023-02-23 19:24:20.935000+00:00	2023-02-23 19:24:20.935000+00:00
<input checked="" type="checkbox"/> Test_ProteinaseK/10	6	2023-02-23 19:24:20.823000+00:00	2023-02-23 19:24:20.823000+00:00
<input type="checkbox"/> Test_ProteinaseK/11	6	2023-02-23 19:24:20.718000+00:00	2023-02-23 19:24:20.718000+00:00
<input type="checkbox"/> Test_ProteinaseK/9	6	2023-02-23 19:24:20.613000+00:00	2023-02-23 19:24:20.613000+00:00
<input type="checkbox"/> Test_ProteinaseK/8	6	2023-02-23 19:24:20.503000+00:00	2023-02-23 19:24:20.503000+00:00
<input type="checkbox"/> Test_ProteinaseK/7	6	2023-02-23 19:24:20.396000+00:00	2023-02-23 19:24:20.396000+00:00

Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data



Data Gateway

- Within a visit the data is structured as it was on disk
- Hence for a single sample you may want to get centring information, processed data and raw images
- SHIFT can be used to select multiple rows
- Filtering in this case by `/10/` can give just the processed files for the test data

Home > Proposals > Diamond MX BAG training 2023 > MX34066-6 > processed > Datafiles

Results: 343

345 items have been added to selection. [Go to selection.](#)

Name	Location	Size	Create Time
processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	5.65 MB	2023-04-07 04:15:52.186000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	1.18 MB	2023-04-07 04:15:52.146000+01:00
processed/Test_ProteinaseK/10/10_1/ca0d3385-ff1f-4b...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	540.33 KB	2023-04-07 04:15:52.126000+01:00
processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	549 B	2023-04-07 04:15:52.078000+01:00
processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	7.04 KB	2023-04-07 04:15:52.013000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	14.14 MB	2023-04-07 04:15:51.909000+01:00
processed/Test_ProteinaseK/10/10_1/ca0d3385-ff1f-4b...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	985 B	2023-04-07 04:15:51.809000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	14.44 KB	2023-04-07 04:15:51.791000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	80.37 KB	2023-04-07 04:15:51.773000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	2.94 KB	2023-04-07 04:15:51.683000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	80.37 KB	2023-04-07 04:15:51.647000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	2.21 KB	2023-04-07 04:15:51.625000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	6.69 KB	2023-04-07 04:15:51.576000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	7.23 KB	2023-04-07 04:15:51.534000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	16.35 KB	2023-04-07 04:15:51.499000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	660.49 KB	2023-04-07 04:15:51.380000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	9.44 KB	2023-04-07 04:15:51.363000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	49.04 KB	2023-04-07 04:15:51.326000+01:00
processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	15.95 MB	2023-04-07 04:15:51.298000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	19.82 MB	2023-04-07 04:15:51.263000+01:00
processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	11.57 MB	2023-04-07 04:15:51.263000+01:00
processed/Test_ProteinaseK/10/10_1/ca0d3385-ff1f-4b...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	2.07 KB	2023-04-07 04:15:51.263000+01:00
processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-47...	/dls/i03/data/2023/mx34066-6/processed/Test_Prote...	675.68 KB	2023-04-07 04:15:51.263000+01:00

Tooltip: /dls/i03/data/2023/mx34066-6/processed/Test_ProteinaseK/10/10_1/8ecd501f-2c68-472a-ac40-7121df313f28/autoPROC/aimless.mrfana_run1.cif

Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data



Data Gateway

- Within a visit the data is structured as it was on disk
- Hence for a single sample you may want to get centring information, processed data and raw images
- SHIFT can be used to select multiple rows
 - Select top then Scroll to bottom and shift click to select all
- Filtering in this case by /10/ can give just the processed files for the test data
- The checkout icon 4 brings you to a summary of items to download

The screenshot shows the Data Gateway interface with the following annotations:

- 1**: A red box highlights the left sidebar menu containing 'Browse', 'Proposal', 'My Data', 'Discover', 'Search', and 'Download'.
- 2**: A red box highlights the 'Location' filter field in the table header, which contains the text '/10/'.
- 3**: A red box highlights a yellow notification banner at the top of the table area that reads '345 items have been added to selection. Go to selection'.
- 4**: A red box highlights the checkout icon (a shopping cart) in the top right corner of the page, next to a 'Results: 343' indicator.

The main table displays a list of data files with columns for Name, Location, Size, and Create Time. The 'Name' column contains paths like 'processed/Test_ProteinaseK/10/10_1/2de7ba42-3efe-4d...'. The 'Location' column shows paths starting with '/dls/i03/data/2023/mx34066-6/processed/Test_Prote...'. The 'Size' column shows various file sizes such as 5.65 MB, 1.18 MB, and 540.33 KB. The 'Create Time' column shows timestamps like '2023-04-07 04:15:52.186000+01:00'.

Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data



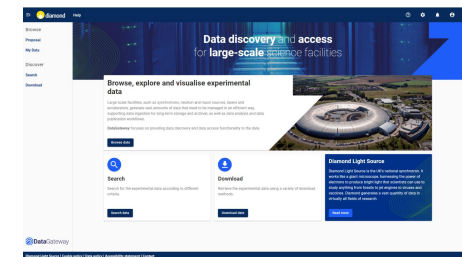
Data Gateway

- A summary of the dataset
 - 8.72GB
 - Excluding the processed files by using the gear icon and typing processed
- Up to 10TB can be retrieved at once
- Data can be downloaded, restored to Diamond disk, sent to globus

The screenshot shows the Data Gateway interface. On the left is a navigation menu with options: Browse, Proposal, My Data, Discover, Search, and Download. The main area is titled 'Selection' and contains a table with columns: Name, Type, Size, File Count, and Actions. The table lists two datasets: 'Test_ProteinaseK/10' (7.26 GB, 6 files) and 'xraycentring/Test_ProteinaseK/10' (725.22 MB, 8 files). A search filter 'pr' is applied to the Name column, and 'Include' is selected in the Type column. Below the table, it shows 'Number of files: 357 / 250000' and 'Total size: 8.72 GB / 10 TB'. At the bottom right of the table area are buttons for 'Remove All' and 'Download Selection' (highlighted with a red box).

Overlaid on the interface is a 'Confirm Your Download' dialog box. It contains the following fields and information:

- Download Name (optional):** A text input field with the placeholder text 'Enter a custom file name or leave as the default format (facility_date_time).'.
- Access method:** A dropdown menu currently set to 'Globus'.
- Access Method Information:** A section stating 'Download the data to Globus Online' and 'Download Size: 8.72 GB'.
- Email Address (optional):** A text input field with the placeholder text 'Send me download status messages via email.'
- A blue 'Download' button at the bottom right.

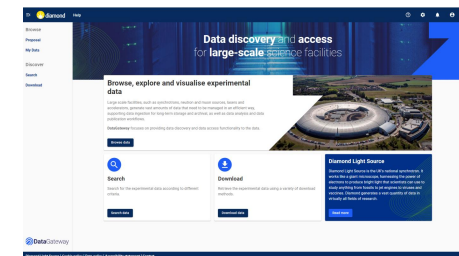


Data Gateway
<https://icat.diamond.ac.uk/>
Archived Raw Data

Data Gateway

When restored to Diamond disk:

- Users will receive an email when data is available in the DLS file system. Data will be available in one of the following directory structures:
 - **/dls/staging**/dls/\$instrument/data/\$year/\$visit
 - **/dls/staging**/dls/\$village/data/\$proposal/\$visit
- The user, beamline staff and visit users will have full write/read access to the directory that has been created at DLS and contains the data requested in DataGateway.
- No further archiving will take place.
- Users will have 15 days to process and transfer their data.
 - After that period, data will be deleted from DLS filesystem.
 - Information on when data will be deleted can be found in README/DATA_EXPIRATION_DATE file in the restored visit directory.
 - It is the responsibility of the user to ensure that data has been transferred before deletion, processed data will not be archived.



Data Gateway

<https://icat.diamond.ac.uk/>

Archived Raw Data

