The GlyGen Upload Interface
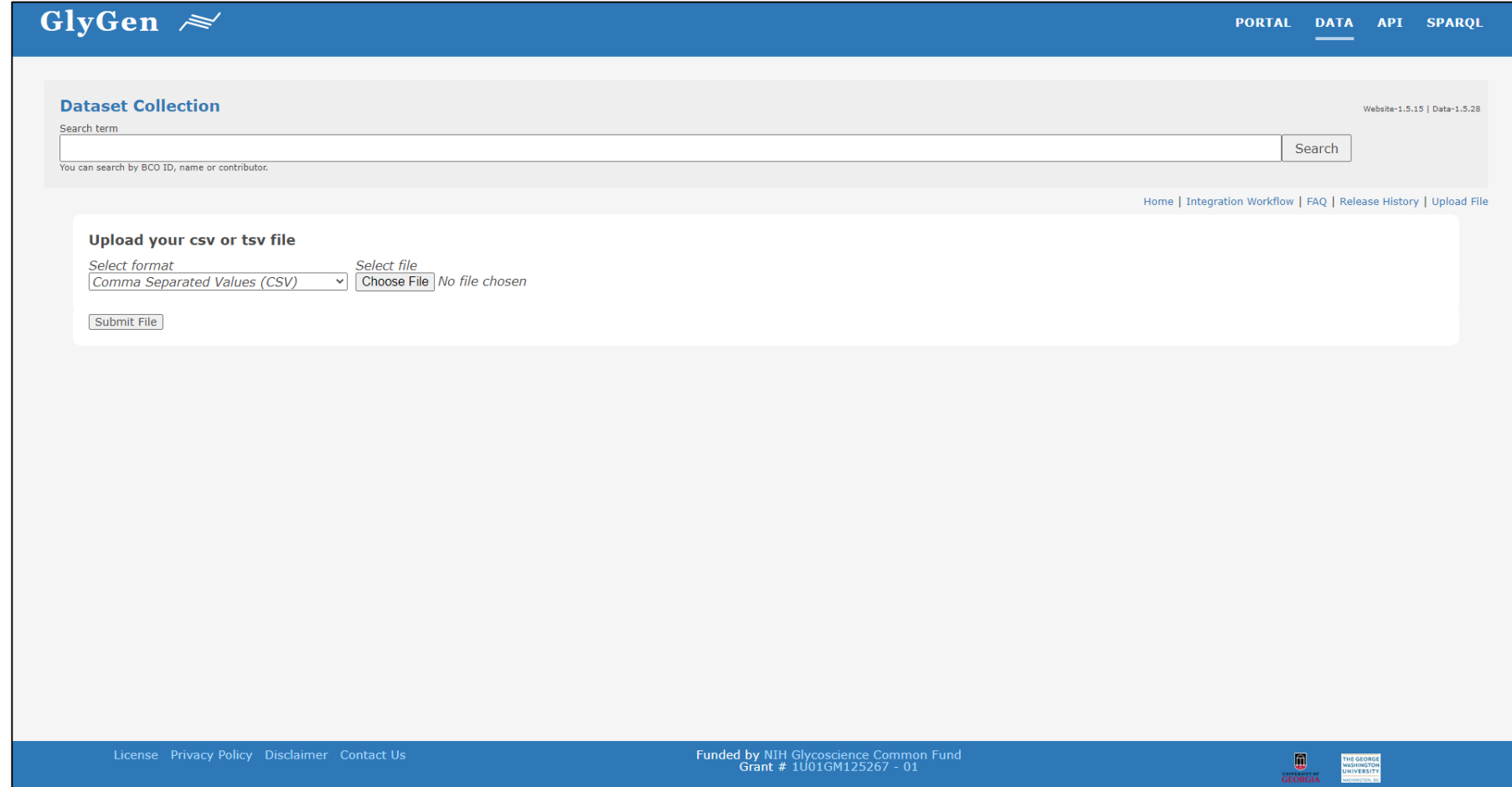
Access the Upload Interface at
https://data.glygen.org/upload

Overview

Upload CSV or TSV data files to GlyGen

Use automated QC checks to ensure data integrity

For any issues contact GlyGen help:
https://glygen.org/contact.html

Upload a Data File

Choose the file format and select the file to upload

For this example we are uploading a CSV file:
glygen_upload_example_file.csv

Review the Basic Sanity Check

The First QC step is called the **Basic Sanity Check**

The report will list the number of rows, columns, and any failed rows that need review

The Basic Sanity Check report

In this example, we have no failed rows that need review, so can proceed to the next step

Upload your csv or tsv file

Select format
Comma Separated Values (CSV)

Select file
Choose File glygen_upload...mple_file.csv

Submit File

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

Protein Site Check

Records Preview
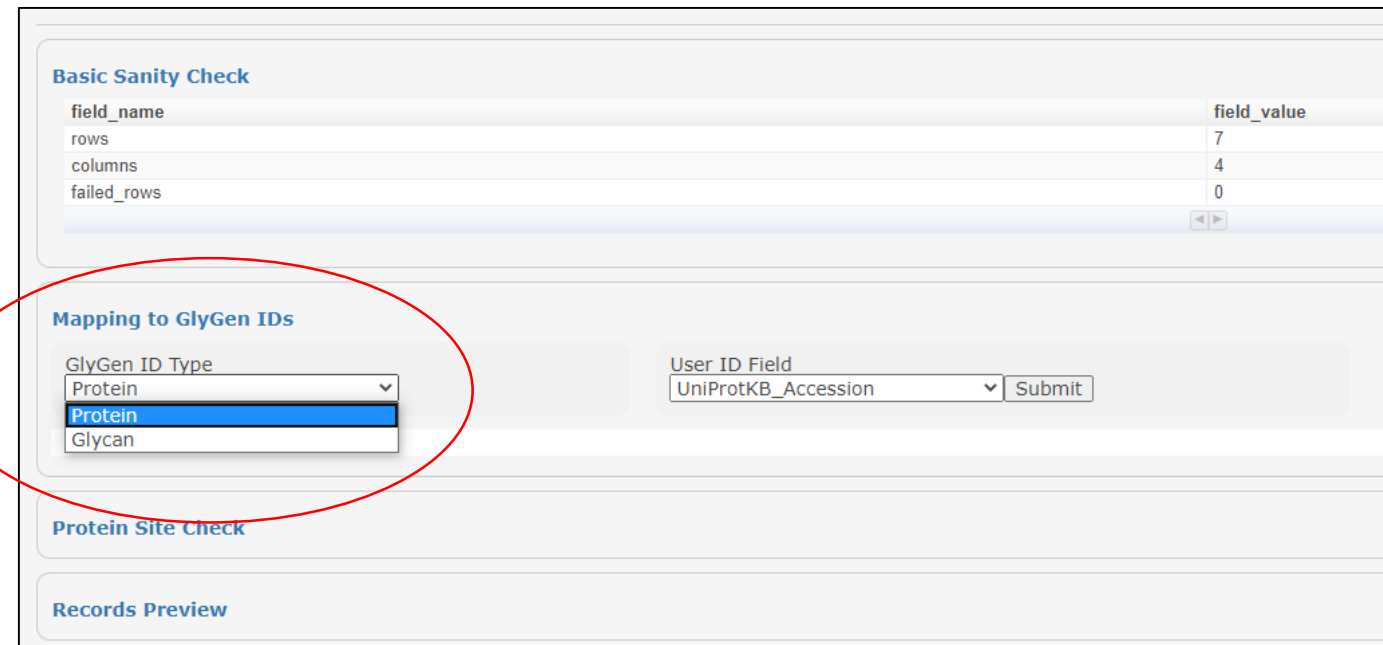
Mapping to GlyGen IDs

For **Mapping to GlyGen IDs**, a field in the upload file can be mapped to IDs already present in GlyGen

Select the `GlyGen ID Type` and `User ID Field` from the drop-down menus, then click `Submit`
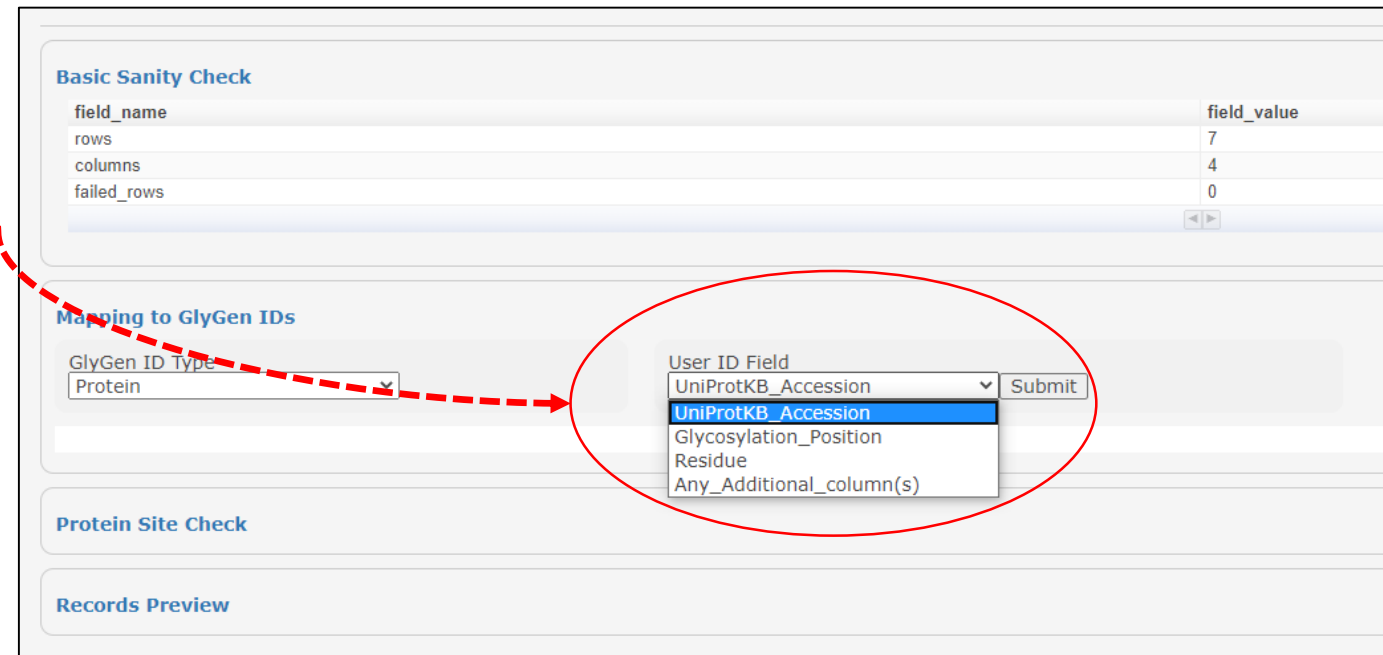
For this example we have chosen:

GlyGen ID Type: `Protein`

User ID Field: `UniProtKB_Accession`

Protein Site Check

For **Protein Site Check**, evaluate the uploaded residue information for specific positions in the protein

For the Protein Site Check, select from the drop-down menus:
`User Protein ID Field`
`User Protein Position Field`
`User Protein Residue Field`
then click `Submit`

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

GlyGen ID Type
Protein

User ID Field
UniProtKB_Accession Submit

unmapped_id   comment

Protein Site Check

User Protein ID Field
UniProtKB_Accession
- UniProtKB_Accession
- Glycosylation_Position
- Residue
- Any_Additional_column(s)

User Protein Position Field
Glycosylation_Position

User Protein Residue Field
Residue Submit

position   residue   comment

Records Preview

For this example we have chosen:

User Protein ID Field: UniProtKB_Accession

User Protein Position Field: Glycosylation_Position

User Protein Residue Field: Residue

Records Preview

With Records Preview, check that the uploaded data is previewed correctly. The report will list all the fields and data from the uploaded file

You can also download the file from the download option in top right corner.

The downloaded file will also provide "qc_comment" for Qc checks done.

The Records Preview report is displayed in the final section

Basic Sanity Check

field_name	field_value
rows	7
columns	4
failed_rows	0

Mapping to GlyGen IDs

GlyGen ID Type
Protein

User ID Field
UniProtKB_Accession   Submit

unmapped_id	comment

Protein Site Check

User Protein ID Field
UniProtKB_Accession

User Protein Position Field
Glycosylation_Position

User Protein Residue Field
Residue   Submit

protein_id	position	residue	comment

Records Preview

UniProtKB_Accession	Glycosylation_Position	Residue	Any_Additional_column(s)
P07498	143	Thr	-
Q8WWV6	423	Thr	-
Q9NQ40	182	Thr	-
P00747	308	Asn	-
P07911	76	Asn	-
P16112	1599	Ser	-

Save File

When the file has been uploaded and passed all QC checks, the uploaded file can be saved to GlyGen

Click `Save File`

Then fill out the information fields:
`First Name`
`Last Name`
`Email Address`
`Institution`
and click `Save`