



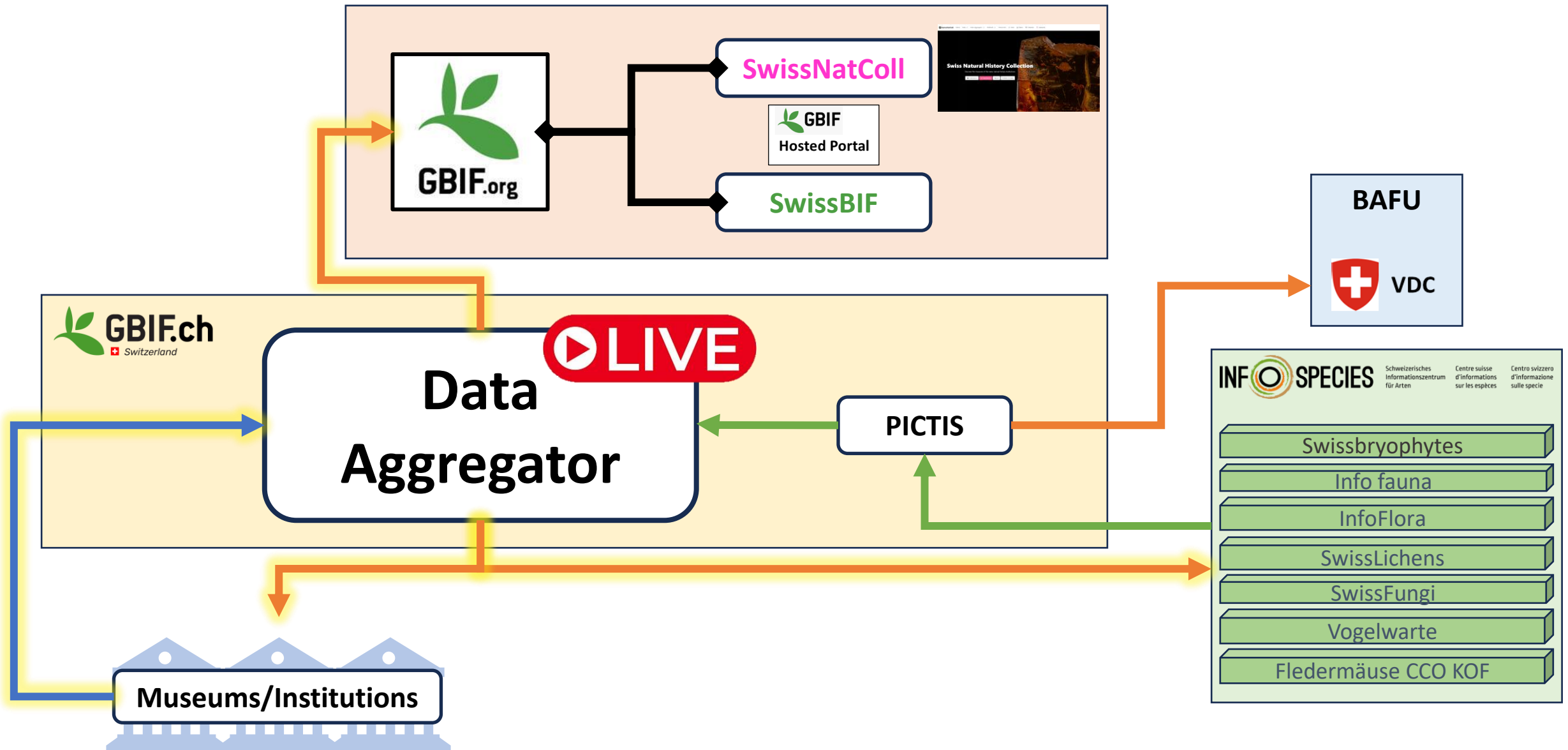
Data Aggregator – Tutorial Session 1 Lesson 1

Introduction

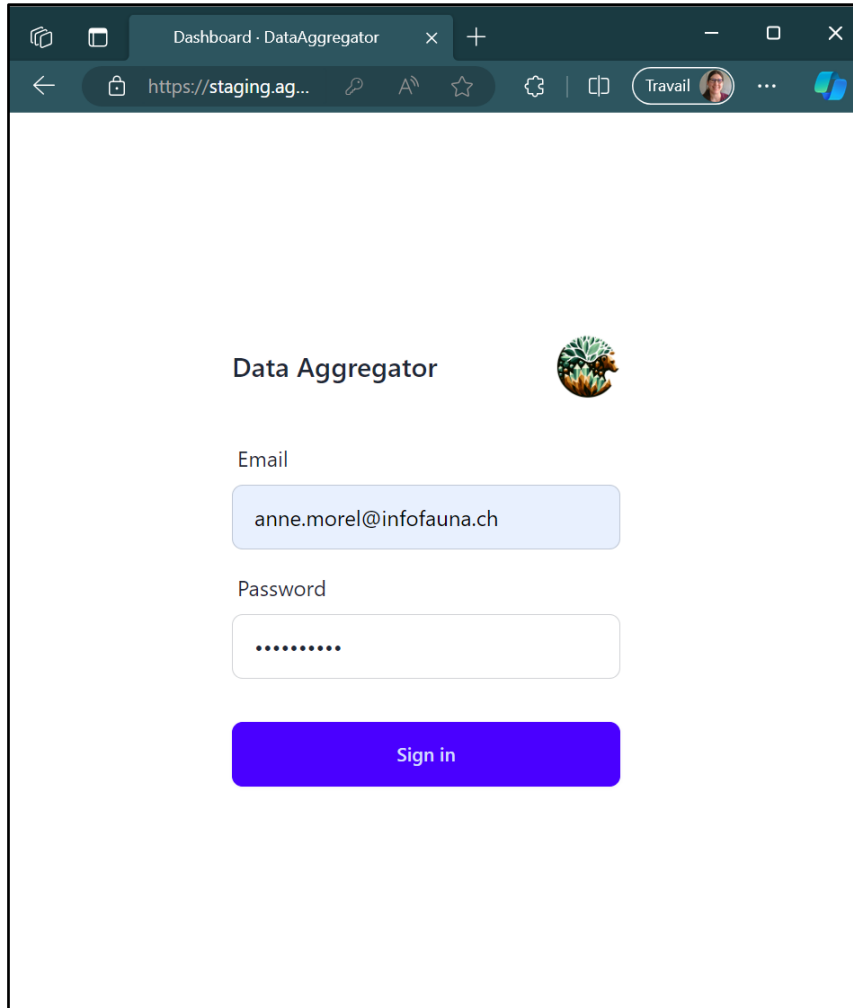
Log In

Data preparation

ZOOM sessions



Access to the Data Aggregator



The screenshot shows a web browser window with the title "Dashboard - DataAggregator". The address bar shows the URL "https://staging.ag...". The page content includes the text "Data Aggregator" next to a circular logo. Below this, there are two input fields: "Email" with the value "anne.morel@infofauna.ch" and "Password" with masked characters ".....". A blue "Sign in" button is positioned below the password field.

Current URL link to the Data Aggregator

<https://staging.aggregator.scnat.zebra.ch/sign-in>

⚠ Staging URL ≠ Production URL ⚠

To add the first user for your institution

contact-swissnatcoll@infofauna.ch

- Full name
- Email
- Institution

Data Administrator & Collection Digitizer

For all users and roles, the following rules apply:

- ♦ One **institution** can have **more than one user / role**
- ♦ One **user** is attributed to **one and only one institution**
- ♦ One **user** can have **more than one role**
- ♦ **A user profile is specific to an individual and must not be shared with others**

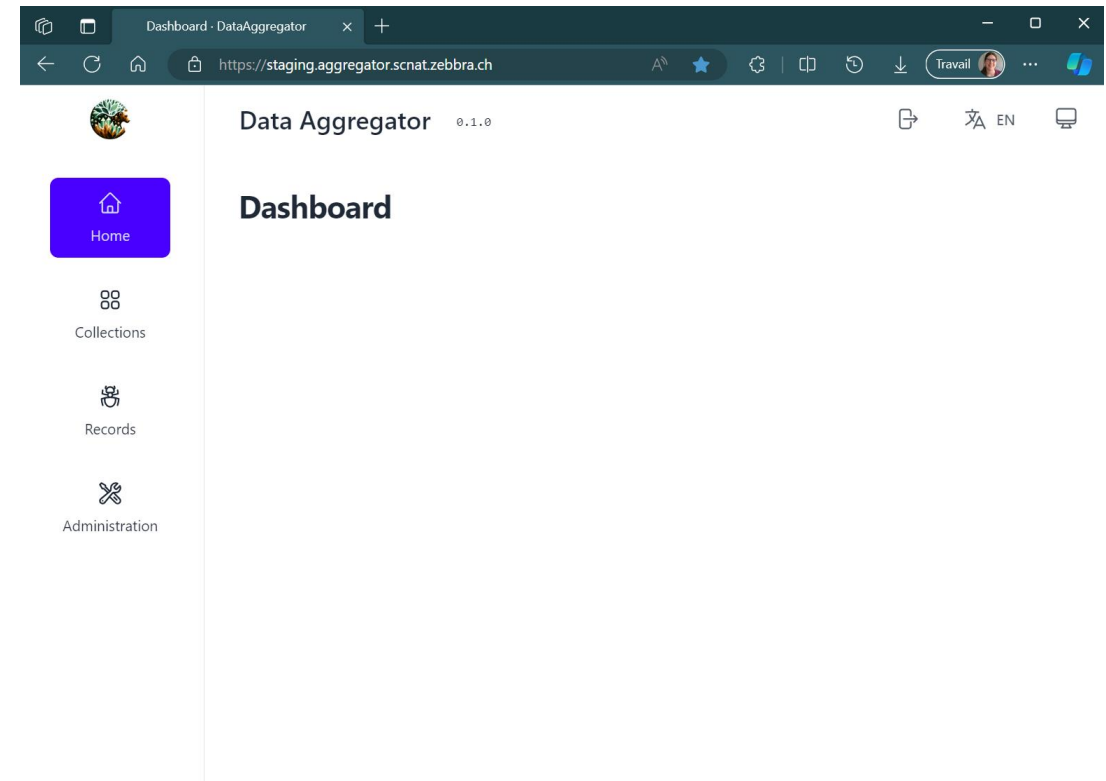
Tasks	Data Administrator	Collection Digitizer	Admin
Users -add/edit/remove Collection Digitizer	✗	✗	✓ all
Users -add/edit/remove Data Administrator	✗	✓ own institution	
Collections -add/edit	✗	✓ own institution	
Collections -view	✓ own institution	✓ own institution	
Dataset -upload/mapping/encode/export/send for validation/publish	✓	✗	

Structure of the Data Aggregator

! Page of one and only one institution !

Home (planned)

- Link to the Step-by-step guide page (<https://swissnatcoll.hp.gbif-staging.org/en/how-to-publish-data>)
- Other elements TBD



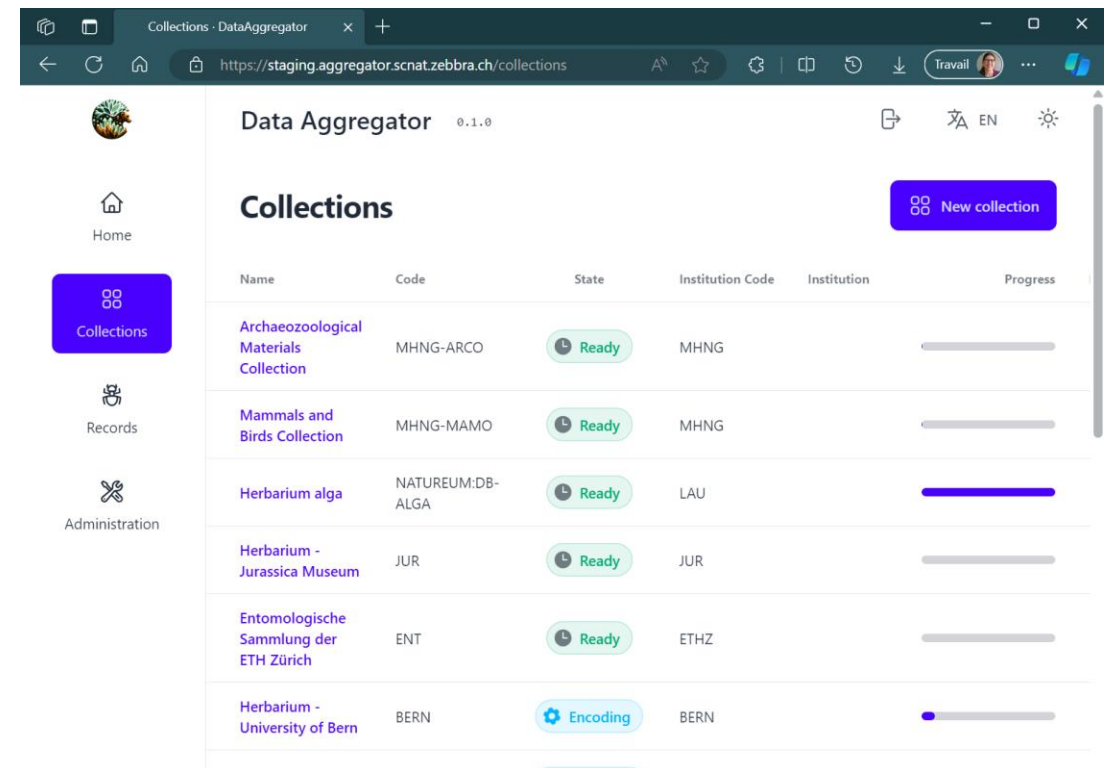
Structure of the Data Aggregator

! Page of one and only one institution !

Home

Collections

- List of all collections already added on the Aggregator
- State and general overview
- Last update timestamp
- Button to add more collections (source: [GRSciColl \(gbif.org\)](https://grscicoll.org/))



The screenshot shows the 'Data Aggregator' web interface. The browser address bar indicates the URL 'https://staging.aggregator.scnat.zebra.ch/collections'. The interface has a sidebar with navigation links: Home, Collections (highlighted), Records, and Administration. The main content area is titled 'Data Aggregator 0.1.0' and 'Collections'. It features a table with columns: Name, Code, State, Institution Code, Institution, and Progress. There are five rows of data, each representing a collection. The 'State' column shows 'Ready' for most collections and 'Encoding' for the last one. The 'Progress' column shows a progress bar for each collection.

Name	Code	State	Institution Code	Institution	Progress
Archaeozoological Materials Collection	MHNG-ARCO	Ready	MHNG		
Mammals and Birds Collection	MHNG-MAMO	Ready	MHNG		
Herbarium alga	NATUREUM:DB-ALGA	Ready	LAU		
Herbarium - Jurassica Museum	JUR	Ready	JUR		
Entomologische Sammlung der ETH Zürich	ENT	Ready	ETHZ		
Herbarium - University of Bern	BERN	Encoding	BERN		

Structure of the Data Aggregator

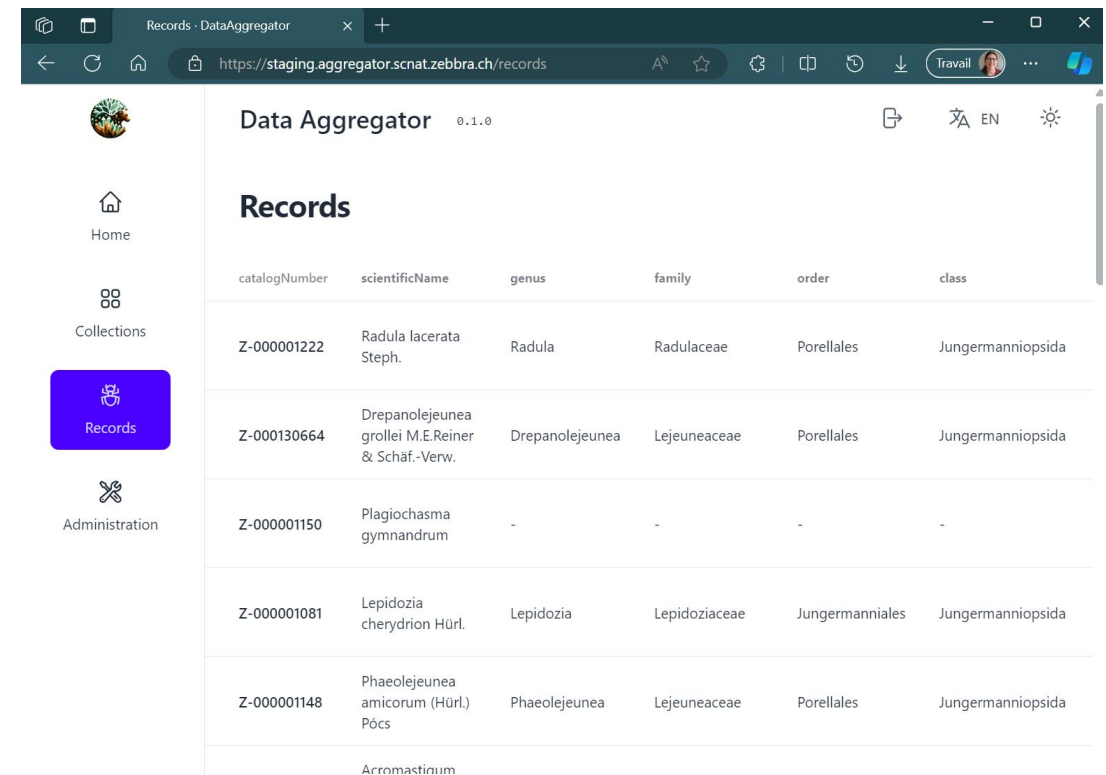
! Page of one and only one institution !

Home

Collections

Records

- List of all imported records for all collections
- State and general overview
- Last update timestamp



The screenshot shows a web browser window with the URL <https://staging.aggregator.scnat.zebra.ch/records>. The page title is "Data Aggregator 0.1.0". The left sidebar contains navigation links: Home, Collections, Records (highlighted in blue), and Administration. The main content area is titled "Records" and displays a table with the following columns: catalogNumber, scientificName, genus, family, order, and class. The table contains five rows of data.

catalogNumber	scientificName	genus	family	order	class
Z-000001222	Radula lacerata Steph.	Radula	Radulaceae	Porellales	Jungermanniopsida
Z-000130664	Drepanolejeunea grollei M.E.Reiner & Schäf.-Verw.	Drepanolejeunea	Lejeuneaceae	Porellales	Jungermanniopsida
Z-000001150	Plagiochasma gymnandrum	-	-	-	-
Z-000001081	Lepidozia cherydron Hürl.	Lepidozia	Lepidoziaceae	Jungermanniales	Jungermanniopsida
Z-000001148	Phaeolejeunea amicorum (Hürl.) Pöcs	Phaeolejeunea	Lejeuneaceae	Porellales	Jungermanniopsida

Below the table, the text "Acromastium" is visible.

Structure of the Data Aggregator

! Page of one and only one institution !

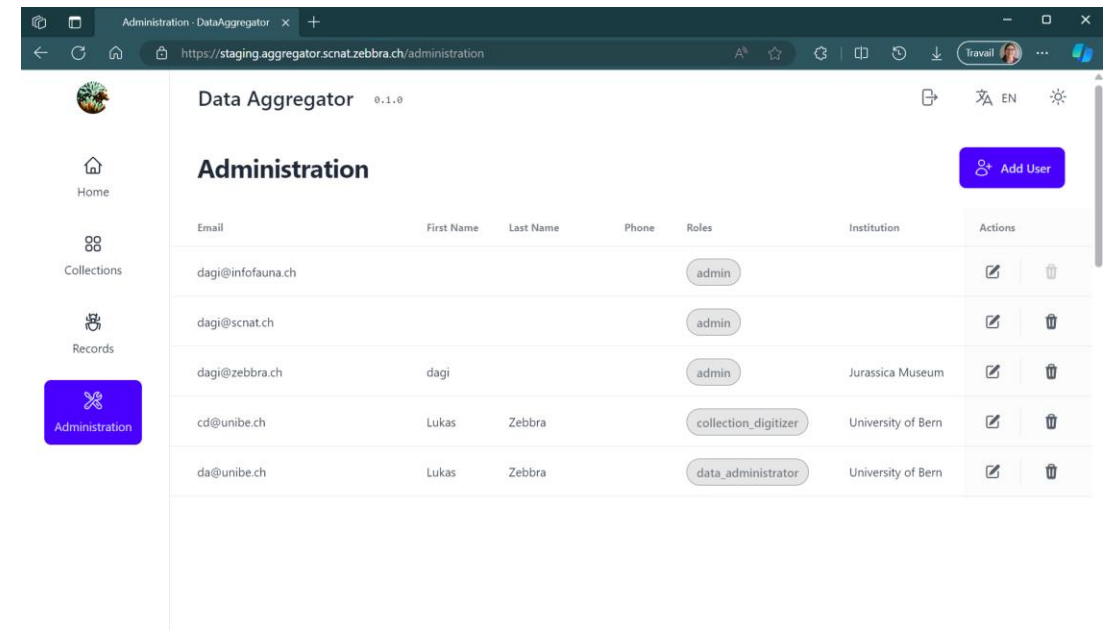
Home

Collections

Records

Administration (only for **Collection Digitizer**)

- List of all users of your institution
- Button to add more Data Administrator users

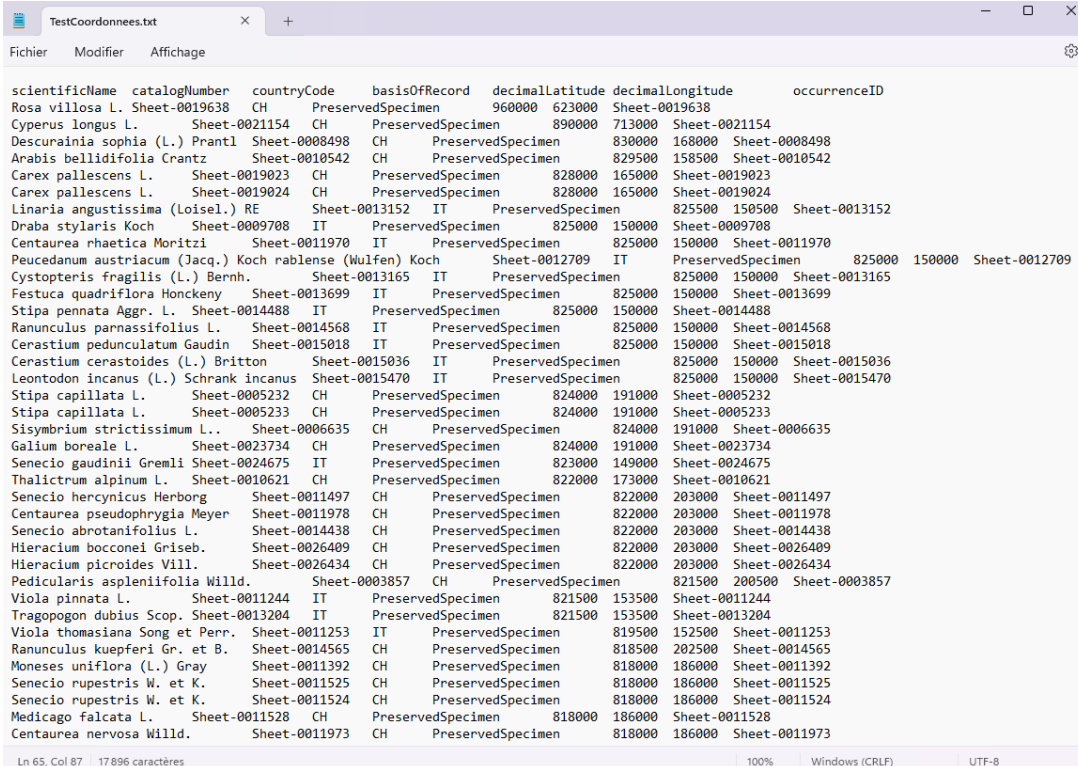


Upload file characteristics

Step 1 of 3

Import records

Please provide your collection file holding your records.



scientificName	catalogNumber	countryCode	basisOfRecord	decimalLatitude	decimalLongitude	occurrenceID
Rosa villosa L.	Sheet-0019638	CH	PreservedSpecimen	960000	623000	Sheet-0019638
Cyperus longus L.	Sheet-0021154	CH	PreservedSpecimen	890000	713000	Sheet-0021154
Descurainia sophia (L.) Prantl	Sheet-0008498	CH	PreservedSpecimen	830000	168000	Sheet-0008498
Arabis bellidifolia Crantz	Sheet-0010542	CH	PreservedSpecimen	829500	158500	Sheet-0010542
Carex pallescens L.	Sheet-0019023	CH	PreservedSpecimen	828000	165000	Sheet-0019023
Carex pallescens L.	Sheet-0019024	CH	PreservedSpecimen	828000	165000	Sheet-0019024
Linaria angustissima (Loisel.) RE	Sheet-0013152	IT	PreservedSpecimen	825500	150500	Sheet-0013152
Draba stylaris Koch	Sheet-0009708	IT	PreservedSpecimen	825000	150000	Sheet-0009708
Centaurea rhaetica Moritzi	Sheet-0011970	IT	PreservedSpecimen	825000	150000	Sheet-0011970
Peucedanum austriacum (Jacq.) Koch rablense (Wulfen) Koch	Sheet-0012709	IT	PreservedSpecimen	825000	150000	Sheet-0012709
Cystopteris fragilis (L.) Bernh.	Sheet-0013165	IT	PreservedSpecimen	825000	150000	Sheet-0013165
Festuca quadriflora Honckeny	Sheet-0013699	IT	PreservedSpecimen	825000	150000	Sheet-0013699
Stipa pennata Aggr. L.	Sheet-0014488	IT	PreservedSpecimen	825000	150000	Sheet-0014488
Ranunculus parnassifolius L.	Sheet-0014568	IT	PreservedSpecimen	825000	150000	Sheet-0014568
Cerastium pedunculatum Gaudin	Sheet-0015018	IT	PreservedSpecimen	825000	150000	Sheet-0015018
Cerastium cerastoides (L.) Britton	Sheet-0015036	IT	PreservedSpecimen	825000	150000	Sheet-0015036
Leontodon incanus (L.) Schrank incanus	Sheet-0015470	IT	PreservedSpecimen	825000	150000	Sheet-0015470
Stipa capillata L.	Sheet-0005232	CH	PreservedSpecimen	824000	191000	Sheet-0005232
Stipa capillata L.	Sheet-0005233	CH	PreservedSpecimen	824000	191000	Sheet-0005233
Sisymbrium strictissimum L.	Sheet-0006635	CH	PreservedSpecimen	824000	191000	Sheet-0006635
Galium boreale L.	Sheet-0023734	CH	PreservedSpecimen	824000	191000	Sheet-0023734
Senecio gaudinii Gremli	Sheet-0024675	IT	PreservedSpecimen	823000	149000	Sheet-0024675
Thalictrum alpinum L.	Sheet-0010621	CH	PreservedSpecimen	822000	173000	Sheet-0010621
Senecio hercynicus Herborg	Sheet-0011497	CH	PreservedSpecimen	822000	203000	Sheet-0011497
Centaurea pseudophrygia Meyer	Sheet-0011978	CH	PreservedSpecimen	822000	203000	Sheet-0011978
Senecio abrotanifolius L.	Sheet-0014438	CH	PreservedSpecimen	822000	203000	Sheet-0014438
Hieracium bocconeii Griseb.	Sheet-0026409	CH	PreservedSpecimen	822000	203000	Sheet-0026409
Hieracium picroides Vill.	Sheet-0026434	CH	PreservedSpecimen	822000	203000	Sheet-0026434
Pedicularis asplenifolia Willd.	Sheet-0003857	CH	PreservedSpecimen	821500	200500	Sheet-0003857
Viola pinnata L.	Sheet-0011244	IT	PreservedSpecimen	821500	153500	Sheet-0011244
Tragopogon dubius Scop.	Sheet-0013204	IT	PreservedSpecimen	821500	153500	Sheet-0013204
Viola thomasiana Song et Perr.	Sheet-0011253	IT	PreservedSpecimen	819500	152500	Sheet-0011253
Ranunculus kuepferi Gr. et B.	Sheet-0014565	CH	PreservedSpecimen	818500	202500	Sheet-0014565
Moneses uniflora (L.) Gray	Sheet-0011392	CH	PreservedSpecimen	818000	186000	Sheet-0011392
Senecio rupestris W. et K.	Sheet-0011525	CH	PreservedSpecimen	818000	186000	Sheet-0011525
Senecio rupestris W. et K.	Sheet-0011524	CH	PreservedSpecimen	818000	186000	Sheet-0011524
Medicago falcata L.	Sheet-0011528	CH	PreservedSpecimen	818000	186000	Sheet-0011528
Centaurea nervosa Willd.	Sheet-0011973	CH	PreservedSpecimen	818000	186000	Sheet-0011973

Ln 65, Col 87 | 17 896 caractères | 100% | Windows (CRLF) | UTF-8

Mandatory

- UTF-8 encoded file
- .csv, .tab, or .txt file
- 1 line = 1 record
- Unique catalog number
- Formatted fields

Facilitating

- Darwin Core headers
- Prioritization
- Tab separation

Fields mandatory and recommended

Mandatory	+ minimal recommended (MIDS 0-1-2)	+ minimal expected (MIDS 2-3)
<ul style="list-style-type: none"> • catalogNumber • scientificName 	<ul style="list-style-type: none"> • institutionCode • partOfOrganism* • preparations • yearCollectionEntrance* • collectionCode • collectionID • eventDate (OR day, month, year, endOfPeriodday/m/y*) • recordedBy • typeStatus • locality • decimalLatitude** • decimalLongitude** 	<ul style="list-style-type: none"> • verbatimElevation/verbatimDepth • continent • higherGeography • country • stateProvince • county • associatedMedia • recordNumber • basisOfRecord • geodeticDatum • coordinateUncertaintyInMeters • swiss_coordinates_x (LV95)* • swiss_coordinates_y (LV95)*

*borrowed terms (MIDS, GBIFch), **only for WGS84 coordinates

Useful links

- Step-by-step tutorial : <https://swissnatcoll.hp.gbif-staging.org/en/how-to-publish-data>
- Data Aggregator overview : <https://swissnatcoll.hp.gbif-staging.org/en/data-aggregator-about/>
- Data Aggregator FAQs : <https://swissnatcoll.hp.gbif-staging.org/en/data-aggregator-faqs> (is planned to be enriched by your questions too)
- Darwin Core summary : <https://swissnatcoll.hp.gbif-staging.org/en/data-aggregator-faqs#what-is-darwin-core>
- Examples with real museum specimens (more coming soon)
 - Mollusca : <https://swissnatcoll.hp.gbif-staging.org/en/label-mollusca>

Take-Home Message



Start small

- Check the **mandatory fields** in your database to make sure they are **standardised**
- Then move on to the **next level of minimum fields**
- Test first with a **small file** and then proceed with a bigger amount of records



Error messages are a good point

- During upload, error messages show inconsistencies in your database
- They are opportunities to correct them



Expected changes in the interface and functionalities

- Development is still ongoing, bugs are currently being solved
- Your contributions are essential for this user-oriented tool