# Transcription platform documentation

## Overview

The transcription platform is a stand-alone app that will have groups of specimens (projects/missions/expeditions) to be transcribed by users/volunteers using an internet browser.

These specimens should have:

1- identifier (something that allows you to identify the specimen univocally, such as a GUID, or a concatenation of Database ID+CollectionObjectID

2- at least one image (URL to an accessible image of the specimen)

For now the creation of projects/missions/expeditions will be made via a form, where an administrator can upload a csv or json file containing this basic information. Later on, automatic ways of getting information about specimens via SPECIFY API can be implemented, depending on the desired pipeline for this.

Administrators will also have the possibility to create the fields that should be filled in by the users, let's say Verbatim date, Verbatim locality, etc.

As soon as the data starts to be filled in and validated, it can be obtained incrementally by a URL request of the kind /validated/validationID, which will produce a json with the list of validated data starting with validationID value, example:

[

{

"id":200309,

"value":"1973\/06\/10",

"position":1,

"specimen\_id":106342,

"field\_id":10,

"user\_id":"system",

"status":1,

"created\_at":"2024-01-23 00:16:39",

"updated\_at":"2024-01-23 00:16:39"

},

{

"id":200310,

"value":"1958\/04\/21",

"position":1,

"specimen\_id":106368,

"field\_id":10,

"user\_id":"system",

"status":1,

"created\_at":"2024-01-23 00:16:59",

"updated\_at":"2024-01-23 00:16:59"

}

]

That's where the other stage of the process starts, which is using the validated data to populate fields in Specify installations through API. This should be programmed to insert data based on a fields mapping

## Administration

The platform has the back-office embedded in the front office, i.e. administrative functions are made from the front-office views. Administrators can perform administrative functions such as define users categories, define and manage transcription fields, create and manage groups of specimens (projects/missions/expeditions), edit languages and dialogs.

## Users

Users interacting with the platform should be logged in in order to have access to administrative functions or other specific definitions related to their profile. Though it is possible to submit values for fields unauthenticated, equivalent to the lowest user level.

There is a user category for administration tasks. Regular users of the platform never have access to administrative tasks.

### User Categories

Regular users are assigned automatically a level according to their progression on the platform, starting with the lowest level, and moving forward as they accumulate validated submissions.

User categories are defined by administrators using a specific form.:

1. Category: A standard name for the category
2. Validations required to achieve this level: number of validated submissions required to achieve each level
3. Points attributed to this level submissions: the number of points their submissions are worth
4. Name to display: what the level is called in every language
5. Text: a description for the level in every language

### User profile

Users have access to their profile: personal data, preferences, password management and account deletion

### User dashboard

Users have access to the list of their contributions with the status for each submission

## Languages

The platform is totally multilanguage. Administrators manage the list of languages displayed with all the dialogs.

New languages can be defined by submitting a form with:

1. Name: A standard name for the language
2. ISO2 Country Code: a two letter code according to ISO 3166 international standard. This assigns a flag icon to the language form an internal pool of icons.
3. Active: yes/no for the language to be displayed to users
4. Master language: one master language should be defined to be displayed as a model for translation for other languages

Dialogs for each language can be edited by administraators:

1. in a page with all the dialogs for a language (access by the languages page)
2. form the front-office by clicking any dialog. A form with all the languages will pop up to edit that dialog.

## Fields

The fields to be included in projects for users to transcribe are defined and managed by administrators. Administrators provide:

1. Name: A standard name for the field is assigned (DarwinCore term, or other recognisable term).
2. The type of field:

* Text: a text string to be typed by the user. Values form the option list are suggested as user types in. It accepts whatever the user types in.
* Textarea: a long text string to be typed by the user. Allows multiple lines. It accepts whatever the user types in.
* Select: a dropdown box with pre-defined list of options. Users should select only one of the options from the list.
* Radio: a group of radio buttons. Users should select only one of the options from the list
* Checkboxed: a list of checkboxes. Users are allowed to select multiple options.
  1. Options: a list of options available

1. Multiple values: if the field should allow single values, or an ordered list of values to be created by the user by creating new lines for the field, for example, Collector 1, collector 2, etc.
2. User minimum level: Because fields are suggested to users for transcription based on their level, administrators can assign a minimum level from which the field should be displayed. For example, if one type of field is considered too difficult for a beginner user, it could only be displayed to advanced users.
3. Name to display: The name of the field to be displayed in every language available
4. Question: a sentence in the form of a question that should help the user to understand what is required in the transcription in every language available
5. Text: a sentence that should help the user to understand what is required in the transcription in every language available
6. Helper: a few paragraphs and images with instructions/use cases that should help the user to understand what is required in the transcription in every language available

## Projects

The main section of the platform. Projects are constituted by groups of specimens to which the fields are assigned for transcription.

A project can be created/managed by accessing a group of forms, available from the projects section

### Project metadata

1. Name: A standard name for the project
2. Title: The name of the project to be displayed in every language available
3. Abstract: a few paragraphs that give some context to be displayed in every language available
4. Description: a full description of the project to be displayed in every language available
5. Visible to users: a switch to make the project available for regular users

### 1.4.2. Upload data

A form that allows the administrators to add specimens to the project:

1. File: accepts a csv or excel field with two columns:
   1. specimen\_id: a unique identifier for each specimen. This identifier is not essential for the platform to work internally, but is the one used as identifier in the output, so must be something recognisable for post-processing.
   2. Images: one or a concatenated list of URLs (pipe separated) for images of each specimen. Images should be accessible for the platform to display them.
2. Action:
   1. Replace: delete the existing list for that project and use the new one
   2. Append: keep the existing list and add the new one
3. Minimum user level: Some specimens might be considered too difficult for new users, do there is the option to assign a minimum user level, so that they are not displayed to users bellow that level (e.g. multiple specimens in one preparation). Here a value can be assigned to all specimens at once when submitting a list.

### 1.4.3. Data overview

Just a display with the list of specimens in one project. Might be used for other administrative tasks in the future (delete individual records, edit image URLs, etc.)

### 1.4.4. Edit Project Fields

The place where the fields to be transcribed are assigned to a project

## 1.5. Data export

## 

## 2. Transcription /Regular Usage

Any user with a device connected to the internet is able to access the front-office version of the platform and navigate on it as a website. Authenticated users have access to their profile and a better navigation experience based on their experience.

## 2.1. Users

### 2.1.1. Account/profile

### 2.1.2. User dashboard

## 2.2. Projects

## 2.3. Transcription

### 2.3.1. Fields