Lepidoptera Sampling

Use Case 2 - exercise sheet

# Contact information

|  |  |
| --- | --- |
| **Full Name** |  |
| **Email** |  |

# 

# Exercise 1: Planning

## 1.1. Analyze the financial component of their new digitization plan

|  |
| --- |
| **OPTIONS** |
| **Option 1**: Buy three more computers. |
| **Option 2**: Offer financial support to the national offices to buy flatbed scanners. |
| **Option 3**: Offer financial compensation to the digitizers. |
| **Option 4**: Purchase existing biodiversity digitization software in English. |
| **Option 5**: Contract a software development company to develop customized digitization software. |
| **Option 6**: Organize a course for the digitizers. |

|  |
| --- |
| **First option selected** |
|  |
| **Second option selected** |
|  |
| **Rationale for your choice** |
|  |

## 

## 1.2. Assign roles

**Roles considered**: Collection manager, Conservator, Librarian, Registrar, Archivist, Project manager, Principal investigator, Imagers, Transcribers, Quality checker, Data specialist, Human resources, Finance, Designer, Web developer, Content developer, Programmer, Network administrator, Helpdesk, Database administrator, Researcher, Volunteer/Intern, Curator, Geographer.

|  |  |
| --- | --- |
| **Person** | **Roles** |
| **Administrative assistant** |  |
| **Volunteer 1** |  |
| **Volunteer 2** |  |
| **Volunteer 3** |  |
| **Volunteer 4** |  |
| **Volunteer 5** |  |
| **Taxonomic expert** |  |
| **Key roles missing** | |
|  | |
| **Rationale for your assignment** | |
|  | |

# 

# Exercise 2: Data capture

|  |
| --- |
| **Why is it important to share this data for science/research? Consider the dataset types that GBIF supports and list how you envision publishing this data.** |
|  |
| **Describe data that you would extract from the paper logs. Indicate how you would structure this data in a spreadsheet.** |
|  |
| **Indicate any Darwin Core fields and data that can be derived from other data not specifically included in the paper logs.** |
|  |
| **What data fields will uniquely identify the site and observations in this dataset?** |
|  |
| **Describe any data that can be standardized/transformed in preparation for publishing during the data capture process?** |
|  |
| **Link to the result spreadsheet** |
|  |

# 

# Exercise 3: Data management

|  |
| --- |
| **Provide your assessment of the quality of this dataset prior to cleaning/standardization. Describe the specific types (technical or consistency) of errors you identified.** |
|  |
| **List the tools (by name) you chose to clean/standardize the dataset and explain why you chose them. Include any specific references to nomenclatural or geographical lists you used. Indicate if you used any of the advanced features of a tool (e.g. Macros or APIs).** |
|  |
| **Describe the procedures you used to clean/standardize the dataset.** |
|  |
| **What information did you choose to document and where did you document it?** |
|  |
| **Link to the result spreadsheet** |
|  |

# 

# Exercise 4: Data publishing

|  |
| --- |
| **What is a Darwin Core Archive (DwC-A)?** |
|  |
| **List the “Core” and any extensions you used to publish the data? Why did you choose them?** |
|  |
| **When you are mapping the source fields to the Darwin Core terms, where can you find information on the definitions and examples of the terms to ensure you are mapping to the correct terms?** |
|  |
| **Describe the procedure you used to publish and subsequently register the dataset. Indicate whether the dataset was successfully registered on the GBIF UAT website (**[**https://www.gbif-uat.org/**](https://www.gbif-uat.org/)**). Did you encounter any errors in the publishing process and how did you resolve them?** |
|  |
| **Link to the public view of the published dataset on the IPT** |
|  |