MED&BS RDBFIS use cases (version 1)

Zilioli M and Carrara P

Revision/inputs 1: Martinelli M (10/6/2021)

Revision/inputs 2, 3: Bitetto I, Zupa W (19/6/2021; 28/6/2021)

Revision/inputs 4: Sabatella E (14/7/2021)

Revision/inputs 5: Angelini S (16/7/2021)

Revision/inputs 6: Bitetto I (30/7/2021, 20/08/2021; 02/09/2021)

Index

Introduction

This document represents the Version **1** of the Use Cases modelled for the development of Regional Data Base for the Mediterranean and Black Sea, planned in the framework of MARE/2020/08 grant.

It is been created in order to meet the Milestone M3.1 "Use cases and list of actions defining the interactions between a user and the system to achieve a goal" (to be reached on month 7)

Material of the use cases in Version 0

Current informative sources used:

- Technical Proposal described in Grant Agreement, Annex 4a
- Inputs from WP2 team (Milestone 2.1 - List of the RDB features needed to answer to the data collection submission and reporting obligations - Aug 2021_revIB)
- Specific meeting of the grant SC on data types and algorithms (June 16th, 2021)
- Meetings with WP2 and WP4 teams, and review inputs from Sabatella E and Angelini S (see slide 1)

For further versions the expected information (mainly from WP2) are:
- Final list of MED&BS RDBFIS input/output data/file formats
- Final list of MED&BS RDBFIS input/output/managing users

Data flow in RDBFIS

Task 4.3 - Developing data validation and quality checking tools

From the experience gained within the **STREAM** project: 2-steps process to verify the consistency of the biological data:

- <u>a priori quality checks</u> (QC), to detect possible inconsistency and inaccuracies already present in the **detailed data**
- <u>a posteriori QC</u>, designed to verify the temporal and spatial coverage, as well as that the data consistency is maintained in the **aggregated dataset**
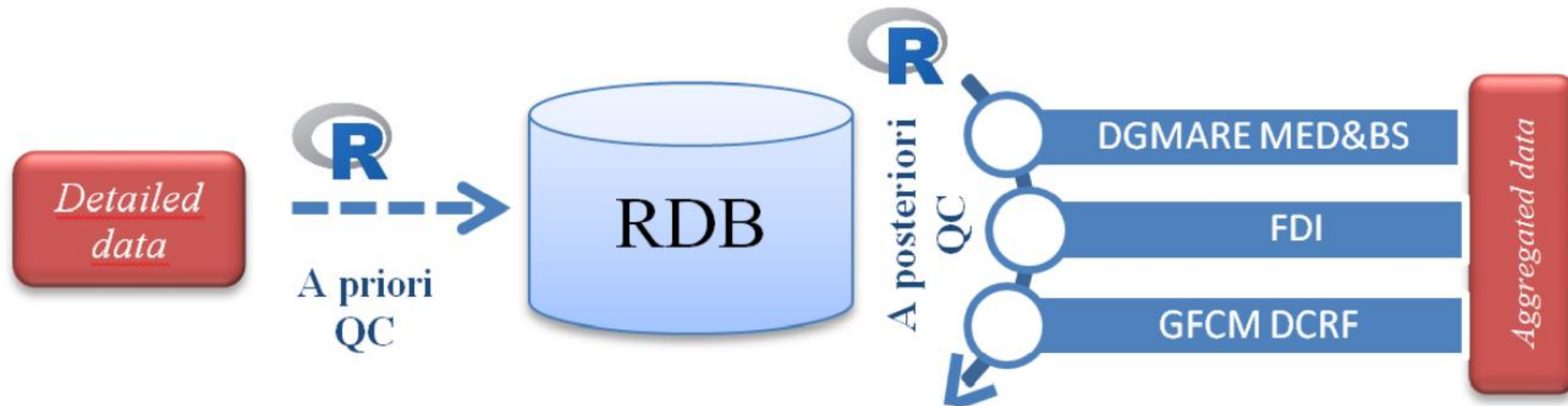
User MS, Upload Sampling data on Commercial fisheries in RCG data format (CS table)

Main Success Scenario:

1. EU Member State (MS) accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format (i.e., RCG ~~CS~~) in which the data are structured

3. MS **selects the name of the table to be uploaded (CS) and** fires the upload of the CSV file

4. The system verifies if the file is already uploaded

5. **The system performs the routines to assess the quality of the data**

6. The system sends a message to the MS which reports the inconsistencies and inaccuracies detected in the dataset

7. MS validates the data as they are, allowing the import

8. The system sends a message to the MS of successful ingestion of the data in the database

9. The system updates the log of interactions of the MS with the system

10. MS may visualize the data and checks the log

Extensions:

2a: MS data are structured in ICES RDBES format

      - MS selects the input facility tool (to be developed) to translate data from ICES RDBES format to the RCG format, by removing additional information

4a: If yes, the system does not upload the file and sends a message to MS

7a: MS operates corrections on the file and re-submits the dataset

User MS, Upload Sampling data on Commercial fisheries in ICES RDBES data format

Main Success Scenario:

1. MS accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format (i.e., **ICES** RDBES) in which the data are structured

3. MS **selects the name of the table to be uploaded (i.e. rdbes_bv, rdbes_ce, rdbes_cl, rdbes_de, rdbes_fm, rdbes_fo, rdbes_ft, rdbes_le, rdbes_lo, rdbes_os, rdbes_sa, rdbes_sd, rdbes_sl, rdbes_ss, rdbes_te, rdbes_vd, rdbes_vs) and** fires the upload of the CSV file

4. The system verifies if the file is already uploaded

5. **The system performs the checks (to be defined) to assess the quality of the data**

6. The system sends a message to the MS which reports the inconsistencies and inaccuracies detected in the dataset

7. MS validates the data as they are, allowing the import

8. The system sends a message to the MS of successful ingestion of the data in the database

9. The system updates the log of interactions of MS with the system

10. MS may visualize the data and checks the log

Extensions:

2a: MS data are structured in RCG format

        - MS selects the input facility tool (to be developed) to translate data from RCG format to the new ICES RDBES format, by enriching it with additional information

4a: If yes, the system does not upload the file and sends a message to MS

7a: MS operates corrections on the file and re-submits the dataset

User MS, Upload Landings data on Commercial fisheries in RCG data format (CL table)

Main Success Scenario:

1.   MS accesses the RDBFIS through the browser and uses its credentials for authentication

2.   MS selects the format (i.e. RCG) in which the data are structured

3.   MS **selects the name of the table to be uploaded (i.e., CL) and** fires the upload of the CSV file

4.   <u>The system verifies if the file is already uploaded</u>

5.   **The system performs the checks (to be defined) to assess the quality of the data**

6.   The system sends a message to the MS which reports the inconsistencies and inaccuracies detected in the dataset

7.   <u>MS validates the data as they are, allowing the import</u>

8.   The system sends a message to the MS of successful ingestion of the data in the database

9.   The system updates the log of interactions of MS with the system

10.   MS may visualize the data and checks the log

Extensions:

4a: If yes, the system does not upload the file and sends a message to MS

7a: MS operates corrections on the file and re-submits the dataset

User MS, Upload Survey data MEDITS format (TA, TB, TC)

Main Success Scenario:

1. MS accesses the RDBFIS through the browser and uses its credentials for authentication
2. MS selects the format in which the data are structured
3. MS **selects the name of the table to be uploaded (i.e., TA, TB, TC) and** fires to upload of the CSV file
4. The system verifies if the file is already uploaded
5. **The system performs the check routines (RoME) to assess the quality of the data**
6. The system sends a message to the MS which reports warnings (>=0) detected in the dataset
7. MS validates the data as they are, allowing the import
8. The system sends a message to the MS of successful ingestion of the data in the database
9. The system updates the log of interactions of MS with the system
10. MS may visualize the data and checks the log

Extensions:

4a: If yes, the system does not upload the file and sends a message to MS

6a: The system sends a message to the MS which reports errors

   - MS operates corrections on the file and re-submits the dataset

User MS, Upload Survey detailed data MEDIAS format

Main Success Scenario:

1. MS accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format in which the data are structured **(i.e., MEDIAS)**

3. MS **selects the name of the table to be uploaded (i.e., medias_echosounder_param, medias_processed_acoustic, medias_survey_design, medias_survey_identity, medias_trawl_biodata, medias_trawl_descr, medias_trawl_haul) and** fires to upload of the CSV file

4. <u>The system verifies if the file is already uploaded</u>

5. The system performs the check routines (available?) to assess the quality of the data

6. **<u>The system sends a message to the MS which reports warnings (available?) detected in the dataset</u>**

7. MS validates the data as they are, allowing the import

8. The system sends a message to the MS of successful ingestion of the data in the database

9. The system updates the log of interactions of MS with the system

10. MS may visualize the data and checks the log

Extensions:

4a: If yes, the system does not upload the file and sends a message to MS

6a: The system sends a message to the MS which reports errors <u>(available?)</u>

       - MS operates corrections on the file and re-submits the dataset

User MS, **Export aggregated** data in Data Calls formats

Main Success Scenario:

1. MS accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format (DG MARE MED&BS, GFCM/CDRF, DG MARE FDI) in which aggregate the data

3. The system shows to MS a pre-compiled query form so as to **filter the aggregated data**

4. MS specifies in query fields the filtering values and confirms the query

5. The system performs the query and prepares an intermediate aggregated result

6. System performs a posteriori checks to verify the temporal and spatial coverage as well as the data consistency of the aggregated dataset, and evaluates if the quality is over a threshold

7. The system generates the output file to be downloaded and exported and sends a message to the MS reporting the quality details of the aggregated dataset

Extensions:

6a: The system sends a message to the MS which states that the quality of aggregated dataset is under the threshold ammitted for submission to end-users

-- how to perform correction?

User (?), Export (?) aggregated data in Data Call formats

Main Success Scenario:

1. MS accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format (DG MARE MED&BS, GFCM/CDRF, DG MARE FDI) in which aggregate the data

3. The system shows to MS a link to the R environment where appropriate routines can be launched to aggregate data starting from the detailed data

4. *Intermediate phase to produce the aggregated data need to be defined*

5. System performs a posteriori checks to verify the temporal and spatial coverage as well as the data consistency of the aggregated dataset, and evaluates if the quality is over a threshold

6. The system generates the output file to be downloaded and exported and sends a message to the MS reporting the quality details of the aggregated dataset

Extensions:

5a: The system sends a message to the MS which states that the quality of aggregated dataset is under the threshold ammitted for submission to end-users:

-- *how to perform correction?*

Methodology: Definitions according to the Unified Modeling Language (UML)

- Use cases describe the typical interactions between the users of a system and the system itself
- The interaction is represented as a **sequence of steps** (i.e. **scenario**)
- In different scenarios the user doesn't always succeed to achieve a goal, but the goal remains
- **A use case is a set of scenarios tied together by a common user goal**
- Scenarios present different outcomes (i.e. successes, failures, alternative pathways)
- In use case-speak, the users are referred to actors. An actor (e.g. humans, machines) is a role that a user plays with respect to the system
- A single actor may perform many use cases
- A use case may have several actors performing it
- One person may act as more than one actor

Fowler, M. (2004). *UML distilled: A brief guide to the standard object modeling language*. Boston: Addison-Wesley

Methodology: Structure of the use case (example)

Goal of the use case: data upload

Kinds of goal levels:
Sea-level: discrete interaction between a primary actor and the system
Fish-level: use cases included in sea-level use case (e.g. data validation in upload data use case)
Kite-level: describe wider business interaction (e.g. reporting obligation to end-user)

Goal level: Sea-level

Main Success Scenario

The Main Success Scenario is described by the sequence of steps which allow user to achieve the goal defined by the use case

1. The user (e.g. the MS) accesses the RDBFIS through the browser and uses its credentials for authentication

2. MS selects the format in which the data are structured

3. MS uploads the CSV file

4. …

5. …

6. …

7. …

An extension represents a variation on the Main Success Scenario. Extensions are referred to the steps numbered in the Main Success Scenario. Extensions illustrate alternative interactions between user and system, and can be successes or failures

Extensions

3a: The system verifies the file is already uploaded. The system does not upload the file and send a message to MS

…

Glossary

- Data format: constraint placed upon the interpretation of data in a type system

- File format: a format for encoding data for storage in a computer file

Wikipedia

Questions

- Which will the input data formats (**ICES RDBES/RCG/SDEF format**) be?

- Which will the input file formats (CSV, XML, value by value) be?

Next use cases (to be evaluated and discussed)

- Upload historical aggregated data by MS
- Export of detailed data by MS/end users
- Perform spatial analysis (VMS & MCDA) (input from S. Kavadas)
- Add a final step in the Main Success Scenario of the 'Upload data' use cases: the MS changes the status of submission from 'provisional' to 'final' according to the feedbacks received from the end users
- Updating database schema and GUIs following data calls modifications; user: system administrator
- Extension of "Export of aggregated data" for end users STECF and GFCM: request to the MSs of clarifications/corrections of exported data