

## Deliverable No. 6.2

Project acronym:  
**FarFish**

Project title:  
**Responsive Results-Based Management and capacity building for EU Sustainable Fisheries Partnership Agreement- and international waters**

Grant agreement No: **727891**  
Project co-funded by the European Commission within the  
Horizon2020 Research and innovation programme

Start date of project: **1<sup>st</sup> June 2017**  
Duration: **48 months**

|                          |                    |
|--------------------------|--------------------|
| Due date of deliverable: | 31/05/2018         |
| Submission date:         | 01/06/2018         |
| File Name:               | FarFish D7.2       |
| Revision number:         | 01                 |
| Document status:         | Final <sup>1</sup> |
| Dissemination Level:     | PU <sup>2</sup>    |

### Revision Control

| Role                          | Name                           | Organisation | Date       | File suffix <sup>3</sup> |
|-------------------------------|--------------------------------|--------------|------------|--------------------------|
| Authors/<br>Task Leaders      | Jamie Lentin                   | STL          | 01/06/2018 | JL                       |
| Author                        | Margarita María Rincón Hidalgo | CSIC         | 01/06/2018 | MMRH                     |
| Author                        | Javier Ruiz                    | CSIC         | 01/06/2018 | JR                       |
| WP leader                     | Javier Ruiz                    | CSIC         | 01/06/2018 | RJ                       |
| Coordinator                   | Jónas R. Viðarsson             | MATIS        | 01/06/2018 | JRV                      |
| Administrative<br>Coordinator | Oddur M. Gunnarsson            | MATIS        | 01/06/2018 | OMG                      |

<sup>1</sup> Document will be a draft until it was approved by the coordinator

<sup>2</sup> PU: Public, PP: Restricted to other programme participants (including the Commission Services), RE: Restricted to a group specified by the consortium (including the Commission Services), CO: Confidential, only for members of the consortium (including the Commission Services)

<sup>3</sup> The initials of the revising individual in capital letters



## **Deliverable D6.2**

# **FFDB Pilot Version 1**

01/06/2018



## Executive Summary

One of many products of the FarFish project is the FarFish Data Base (FFDB), which is intended to make the data collected within the various tasks within the FarFish project available for project partners in an operational way. The FFDB will be a repository for wide range of data and will allow for users to easily upload new data and make queries e.g. for input to visualisation- and decision support tools developed in other tasks in the project. The FFDB will be open access on the FarFish webpage and towards the end of the project the relevant datasets will be uploaded from FFDB to OpenAIRE; the data repository of the H2020 Open Research Data Pilot.

A data input protocol for the FFDB (D6.1) was developed within the first four months of the FarFish project. The work on the development of the FFDB has continued since then and the core infrastructure for data collection and on-line visualisation tools is now live i.e. the FFDB pilot version 1 is available. This report presents the tools that have been built on this platform so far, which are now ready for the various work packages and case studies within the FarFish project to use. The report does as well include a discussion on the work ahead in developing the FFDB further, in particular what are the next directions for development in FFDB Pilot 2.



## Table of Contents

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>INTRODUCTION .....</b>                               | <b>5</b>  |
| <b>2</b> | <b>INFRASTRUCTURE.....</b>                              | <b>6</b>  |
| <b>3</b> | <b>WEBSITE DATA INPUT FORM / STORAGE.....</b>           | <b>7</b>  |
| <b>4</b> | <b>EXAMPLE OF USING THE FFDB DATA INPUT FORM .....</b>  | <b>9</b>  |
| <b>5</b> | <b>HOSTING / BACK END FOR VISUALISATION TOOLS .....</b> | <b>10</b> |
| <b>6</b> | <b>CONCLUSION .....</b>                                 | <b>11</b> |
|          | <b>REFERENCES.....</b>                                  | <b>12</b> |



## 1 Introduction

One of many products of the FarFish project is the FarFish Data Base (FFDB), which is intended to make the data collected within the various tasks within the FarFish project available for project partners in an operational way. The FFDB will be a repository for wide range of data and will allow users to easily upload new data and make queries e.g. for input to visualisation- and decision support tools developed in other tasks in the project. The FFDB will be open access on the FarFish webpage and towards the end of the project the relevant datasets will be uploaded from FFDB to OpenAIRE; the data repository of the H2020 Open Research Data Pilot.

A data input protocol for the FFDB (D6.1) was developed within the first four months of the FarFish project. The work on the development of the FFDB has continued since then and the core infrastructure for data collection and on-line visualisation tools is now live i.e. the FFDB pilot version 1 is available. This report presents the tools that have been built on this platform so far, which are now ready for the various work packages and case studies within the FarFish project to use. The report does as well include a discussion on the work ahead in developing the FFDB further, in particular what are the next directions for development in FFDB Pilot 2.

The data input protocol for the FFDB (D6.1) described in more detail the outputs to be produced as part of FFDB. The outputs identified were:

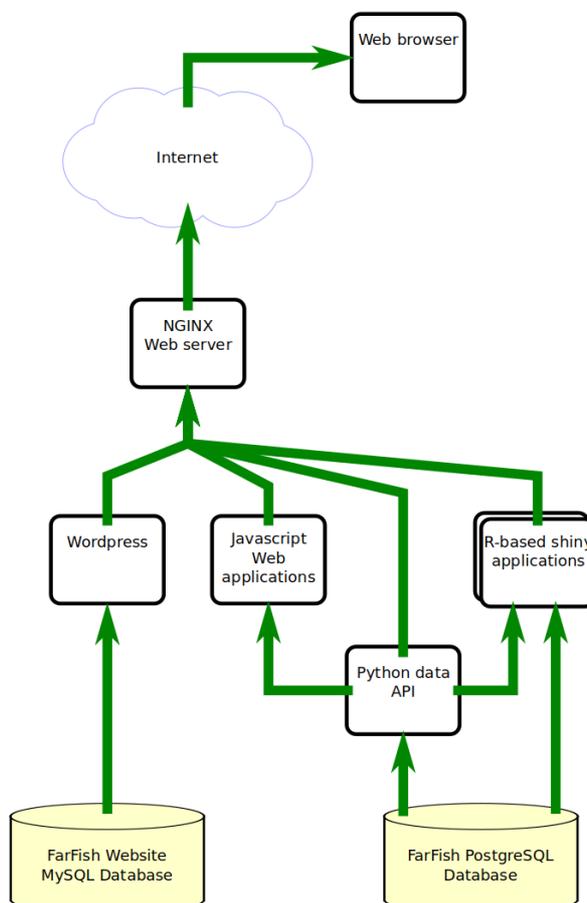
- Infrastructure
- Website data input form / storage
- Hosting / back end for visualisation tools
- Data transformation for modelling
- Website data archival to meet DMP requirements

This report details progress towards these outputs.



## 2 Infrastructure

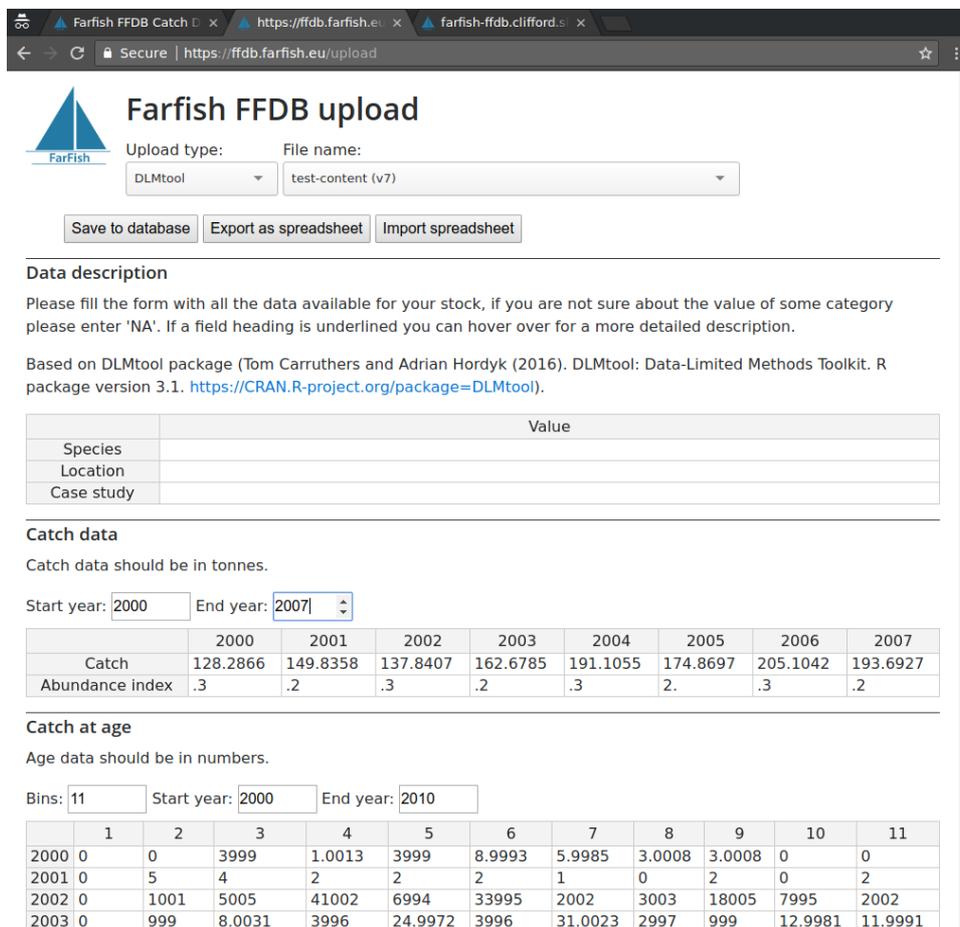
The data input protocol for the FFDB (D6.1) introduced the following architecture diagram:



The core infrastructure to support this is now live and has been used to build and host all the tools detailed in later sections of this report. All FFDB-specific parts of the FarFish website are available on the web at <https://ffdb.farfish.eu/> and source code is available at <https://github.com/farfish>.

### 3 Website data input form / storage

The website data input form software introduced in the data input protocol for the FFDB (D6.1) has now been developed. The form is accessible at <https://ffdb.farfish.eu/upload> as seen below.



**Farfish FFDB upload**

Upload type:  File name:

**Data description**

Please fill the form with all the data available for your stock, if you are not sure about the value of some category please enter 'NA'. If a field heading is underlined you can hover over for a more detailed description.

Based on DLMtool package (Tom Carruthers and Adrian Hordyk (2016). DLMtool: Data-Limited Methods Toolkit. R package version 3.1. <https://CRAN.R-project.org/package=DLMtool>).

|            | Value |  |  |  |  |  |  |  |
|------------|-------|--|--|--|--|--|--|--|
| Species    |       |  |  |  |  |  |  |  |
| Location   |       |  |  |  |  |  |  |  |
| Case study |       |  |  |  |  |  |  |  |

**Catch data**

Catch data should be in tonnes.

Start year:  End year:

|                 | 2000     | 2001     | 2002     | 2003     | 2004     | 2005     | 2006     | 2007     |
|-----------------|----------|----------|----------|----------|----------|----------|----------|----------|
| Catch           | 128.2866 | 149.8358 | 137.8407 | 162.6785 | 191.1055 | 174.8697 | 205.1042 | 193.6927 |
| Abundance index | .3       | .2       | .3       | .2       | .3       | 2.       | .3       | .2       |

**Catch at age**

Age data should be in numbers.

Bins:  Start year:  End year:

|      | 1 | 2    | 3      | 4      | 5       | 6      | 7       | 8      | 9      | 10      | 11      |
|------|---|------|--------|--------|---------|--------|---------|--------|--------|---------|---------|
| 2000 | 0 | 0    | 3999   | 1.0013 | 3999    | 8.9993 | 5.9985  | 3.0008 | 3.0008 | 0       | 0       |
| 2001 | 0 | 5    | 4      | 2      | 2       | 2      | 1       | 0      | 2      | 0       | 2       |
| 2002 | 0 | 1001 | 5005   | 41002  | 6994    | 33995  | 2002    | 3003   | 18005  | 7995    | 2002    |
| 2003 | 0 | 999  | 8.0031 | 3996   | 24.9972 | 3996   | 31.0023 | 2997   | 999    | 12.9981 | 11.9991 |

The upload software provides a templated upload form for documents made up of any number of table objects. Documents are stored as JSONB within a PostgreSQL database, which can then be used to construct traditional relational views on the data efficiently without having to have to construct a schema to store each of the table types. All old versions of documents are saved and can be easily restored.

There is a flexible schema to define each table within the document, and a table can have either a fixed list of columns / rows or a flexible number provided by the user.

Documents can also be exported / imported from excel spreadsheets, where each table will be on a separate sheet in the document. The import is intelligent enough to notice row / column re-orderings and reassign data as appropriate.

The template system allows an administrator to easily customise what data is collected by the tool and keeps the tool generic---it could easily be used outside FarFish for any similar data-gathering exercise.

At this point data is being gathered for use with the Data-Limited Methods Toolkit [\[DLMTool\]](#). This is an open-source modelling package designed to work in data-poor situations. Data gathered here will be fed directly into the tool. As such the FarFish team has defined a "dlmtool" template that gathers all data required by the toolkit, so it can be applied in each situation as required.

Currently the page is open for others to experiment with, however as FarFish partners start gathering data from case studies access will be restricted to selected users of the FarFish website. However, there is a demo version available at <https://farfish.github.io/ffdb-demo/upload> that cannot save entries but otherwise has all the features available.

The source code for the software is open source and available on GitHub at <https://github.com/farfish/ffdb>.



## 4 Example of using the FFDB data input form

There is a full user guide for case studies on how to enter their data available on the internal FarFish website. An abbreviated version is as below.

The process for creating a new entry in the FFDB input system is as follows:

1. Create a new document by typing in a name into the "Document name" box and press enter.
2. Tables to fill in will appear on the screen
3. For each table, fill in the date / size ranges you have data for first, for example the start year and end year. The table will automatically grow to accommodate any extra rows.
4. Fill in fields. If you do not have value for a field, enter "NA".
5. When finished, press "Save" at the top of the page.

To edit, just select an existing document name rather than starting a new one. Alternatively, you can use Excel to edit the data as follows:

6. Select your document if not already shown
7. Press "Export as spreadsheet", you will download a copy of the document as an Excel spreadsheet.
8. Open up in Excel. Each table will be on a separate sheet, and there will be row/column headings for each.
9. Add your data. If you need extra year headings, e.g. edit the table structure appropriately. Otherwise try to keep headings as they are.
10. Once done, press "Import spreadsheet" and select your spreadsheet file. Your new data will replace the old data on-screen.
11. Press "Save" at the top of the page to upload to the database.

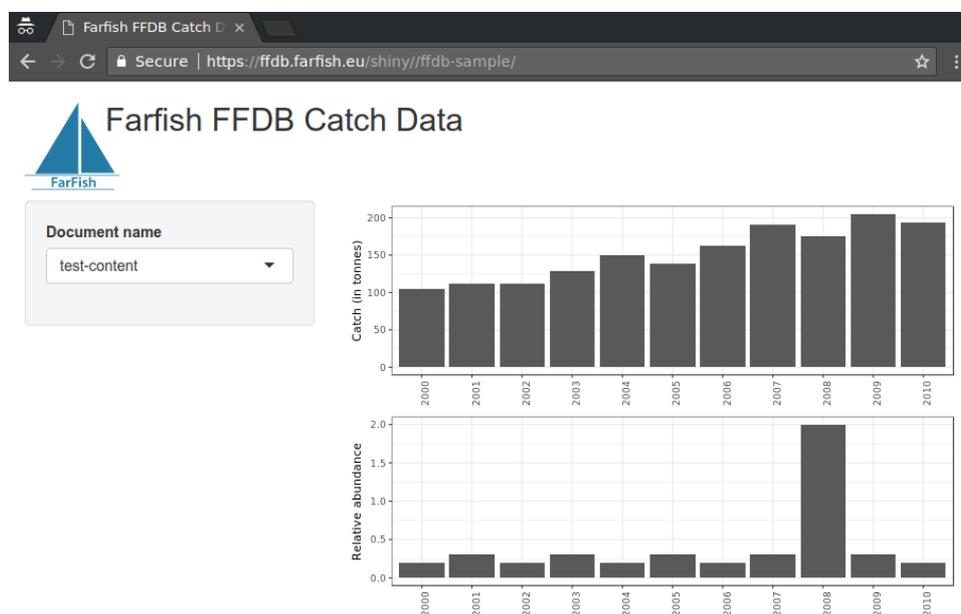


## 5 Hosting / back end for visualisation tools

An Rstudio [\[Shiny\]](#) instance has been installed for use in FarFish. Shiny is a framework for making web-based interactive visualisations using R, and will be used in other parts of FarFish to generate tools for case studies to use.

Shiny applications have direct access to the documents added to the PostgreSQL database above, as a result we can create visualisations directly from the data as users import it, and the result will be "live". Whenever data is updated, the visualisation will change.

As a proof-of concept, there is an application to graph the catch and abundance data provided in any document, as seen below:



Whenever new documents are added, or documents are updated, this will update without any administrator interaction.

The source code for the sample visualisation is open source and available on GitHub at <https://github.com/farfish/visualisation-sample>.

## 6 Conclusion

At this point the FFDB has the infrastructure required for a data-collection exercise and the ability to use this data to produce visualisations. From now onwards this will be used to collect data from the FarFish case studies. Any refinements to the template for data collection, and/or extra features will be added as real data collection is underway.

The Shiny server will also have more applications added, in particular the Data-Limited Toolkit will be set-up so it can be used for any data provided using the data input form.

Finally, the infrastructure will be used as a platform for visualisations of other data sources as they come available.

There are two parts in the introduction not previously discussed:

- Data transformation for modelling
- Website data archival to meet DMP requirements

These outputs are less urgent at this stage of the project, and as a result have not been the focus in Pilot version 1. However, in Pilot version 2 it is anticipated that there will be improvements to the R package to assist in any other modelling activity in FarFish. Finally, data archival will be considered in the final stages of FFDB.



## References

|                           |   |
|---------------------------|---|
| <a href="#">[DLMTool]</a> | Carruthers, T.R., and Hordyk, A. 2017. DLMtool: Data-Limited Methods Toolkit. <a href="http://cran.r-project.org/web/packages/DLMtool/index.html">http://cran.r-project.org/web/packages/DLMtool/index.html</a> . |
| <a href="#">[Shiny]</a>   | <a href="https://shiny.rstudio.com/">https://shiny.rstudio.com/</a>   |
| <a href="#">[MFDB]</a>    | <a href="https://github.com/mareframe/mfdb">https://github.com/mareframe/mfdb</a>   |

