D4.3. FAIR4Health platform user guide

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Document Information

GRANT AGREEMENT NUMBER 824666			ACRONYM: FAIR4Health	
Full title	Improving Health Research in EU through FAIR Data			
Horizon 2020 Call	SwafS-04-2018: Encouraging the re-use of research data generated by publically funded research projects			
Type of action	Research and Innovation Action			
Start Date	1st December 2018	**Duration**	36 months	
Website	www.fair4health.eu			
Project Officer	Pepa Krasteva			
Project Coordinator	Carlos Luis Parra Calderón, Andalusian Health Service (P1–SAS)			
Deliverable	D4.3 FAIR4Health platform user guide			
Work Package	WP4. FAIR4Health platform development			
Date of delivery	**Contractual**	M24	**Actual**	M24
Nature	Report	**Dissemination Level**	Public	
Lead beneficiary	P4 - ATOS SPAIN SA			
Responsible Author	P4 - ATOS (Carlos Cavero, Miriam Quintero)			
e-mail	miriam.quintero@atos.net	**Phone number**	+34 822 110 214	
Reviewers	P2 - IACS (Beatriz Poblador Plou) P9 – UNIGE (Jessica Rochat)			
Keywords	Platform, Feature Set, Data model, Algorithm, Agents, Prototype, User guide.			

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Document history

Version	Issue Date	Stage	Changes	Contributor(s)
0.1	2020-10-26	Draft	Table of contents	J. Mario Rodríguez (ATOS), Carlos Cavero (ATOS)
0.2	2020-10-02	Draft	First version	Miriam Quintero (ATOS)
0.3	2020-11-11	Draft	Merge content, table and figures	Miriam Quintero (ATOS)
0.4	2020-11-13	Draft	FAQs questions section	Miriam Quintero (ATOS)
0.5	2020-11-23	Internal Reviewed	Internal Reviewed to improve document in general terms	Jessica Rochat (UNIGE)
0.5	2020-11-23	Internal Reviewed	Internal Reviewed to improve document in general terms	Beatriz Poblador (IACS)
0.6	2020-11-25	Final review	Final version	Celia Alvarez (SAS), Miriam Quintero (ATOS)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Table of Contents

List of Figures

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

List of acronyms

API	Application Programming Interface
FAQ	Frequently asked question
GUI	Graphical User Interface
SAB	Scientific Advisory Board

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Executive summary

This document is focusing on end user, to guide them to access and move on FAIR4Health platform and is the perfect complement to D4.2 FAIR4Health platform released which describes the FAIR4Health platform and agents released.

There is a FAQ section at the end of the document in order to solve common questions raised in different focus groups but also from members of the Scientific Advisory Board (SAB).

1. Introduction

This deliverable guide end users on how to use FAIR4Health Platform, containing instructions, workflows and interfaces figures to fully explain the system, the objective at this stage is to focus on the final user, clinicians and researchers.

This document is the complement to "D4.1. FAIR4Health platform beta release" where architecture is described and "D4.2. FAIR4Health platform released" which explains the FAIR4Health platform and agents released.

FAIR4Health platform is a set of tools; (1) FAIR4Health GUI, (2) Data Curation Tool and (3) Data Privacy Tool. However, this document contains FAIR4Health GUI guidelines. This is due to the user guides of Data Curation Tool and Data Privacy Tool as well as their installation guides are included in D4.1 into annex section.

This document is also focused on FAIR4Health GUI (Graphical User Interface).

This D4.3 document is structured as follows:

- **Workflow section**: This section identifies the main activity process, from use case creation to data model identification and helps to understand the general steps to carry out.
- **FAIR4Health Platform user manual section:** This section is the core of the document. This section helps end users to go through all platform features and understanding each step. This section follows workflow definition from previous section.
- **FAQs section:** This section lists the most common problems or doubts that users may have and resolves all of them. These relevant questions become from all clinical partners.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

2. Workflow

FAIR4Health use cases have been defined as: (1) Use case one, or association type; and (2) Use case two, or prediction type. Use case one workflow is showed in following figure:

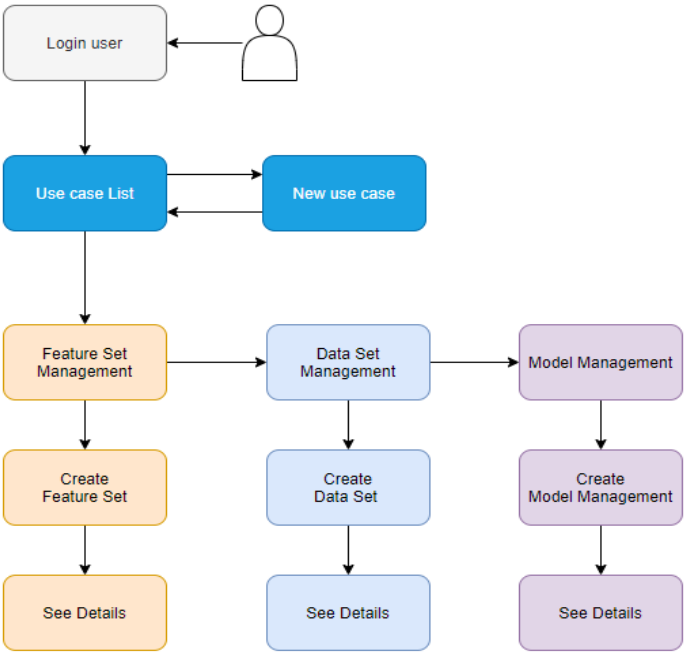


Figure 1. Use case one, or association type, workflow.

Use case two workflow, or prediction type, is showed on following image:

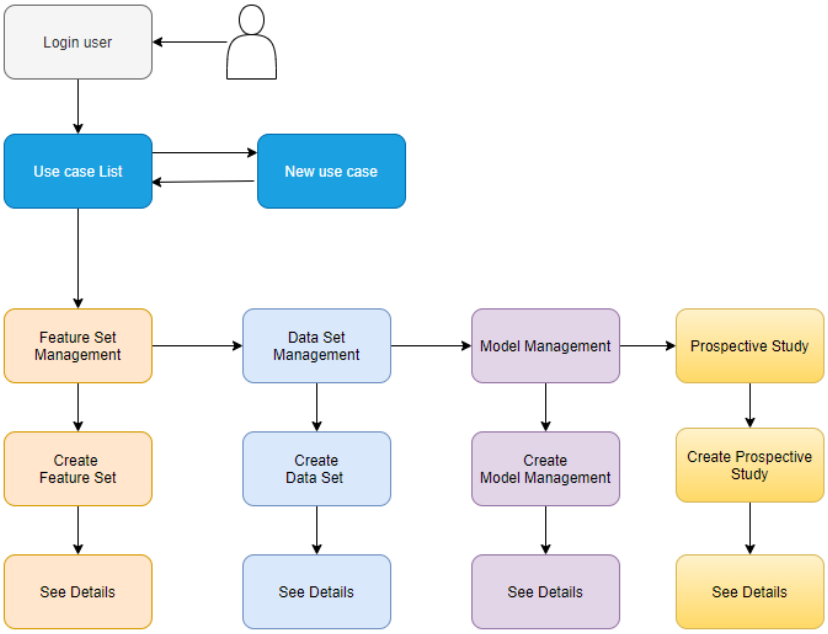


Figure 2. Use case two, or prediction type, workflow.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Use case two, or prediction type, workflow shows an additional step (from use case one) "Prospective Study" since this step is only relevant in prediction type use case.

Notice that, **both figures** show the main and more traditional and general workflow. Once all objects are created, the most common step would be to go straight to the last step, "Prospective study".

3. FAIR4Health Platform user manual

3.1. User access

Users can access to the FAIR4Health platform by clicking on the following link https://portal.fair4health.eu/ . Then click on login button and type username and password.

For demo purpose, ATOS has created demo user with following data:

- Username: demo
- Password: demo2020

Only validated user can access into the system by typing username and password.

Figure 3. Login user interface.

3.2. Use case list

Once the user access into the system a use case list is displayed as showed in following figure 4. In this interface the user can:

- Create a new use case.
- Select one use case to work with.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

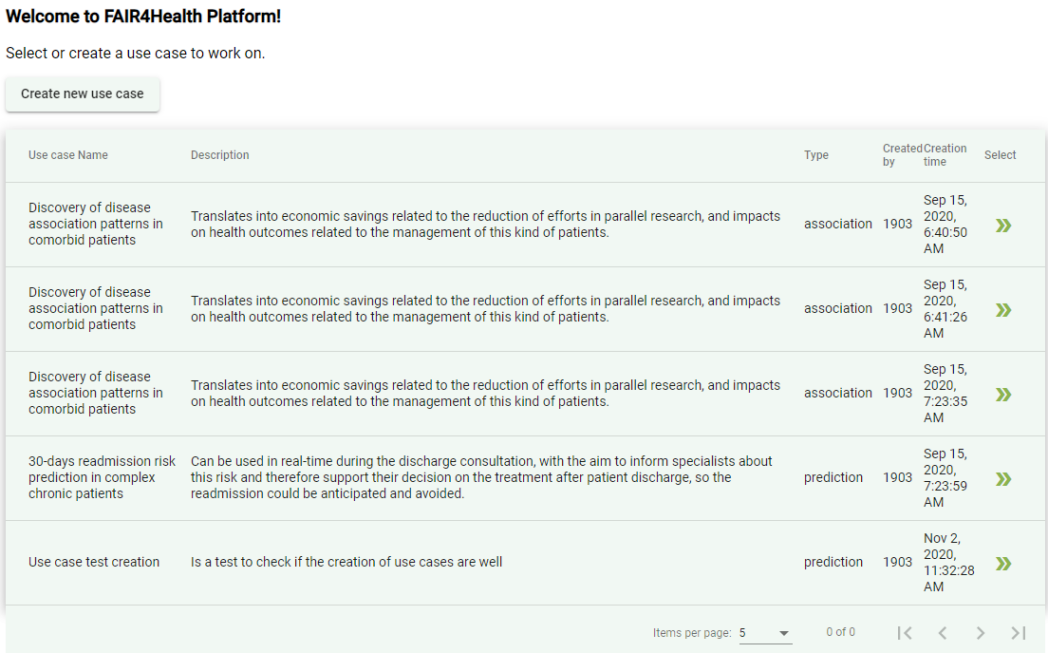


Figure 4. Use case list interface.

3.3. Use case creation

It is important to highlight that there are two types of use cases: (1) Use case one or association; and (2) Use case two or prediction. The use case one or association follows the steps indicated for the first 3 sections of the platform: Feature Set Management, Data Set Management and Model Management. In the second use case or prediction, there are one additional section, the Prospective Study (described on section 3.7) as represented on Figure 1. and Figure 2. Use case type.

Different use cases are focusing on:

(1) Use case one or association: The objective to this use case is to measure the impact of multimorbidity patterns and polypharmacy on 6-months mortality rate and cognitive impairment among elderly individuals in different health care settings.

(2) Use case two or prediction: The objective to this use case is to Develop, validate and assess the accuracy of a clinical decision support tool for predicting 30-day readmission risk in patients suffering from Chronic Obstructive Pulmonary Disease (COPD).

To create a new use case, logged user must provide required details:

- Name: Name of use case, it must be identifying.
- Description: Short description to understand the purpose of the use case.
- Model Type: Select model type: prediction or association.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

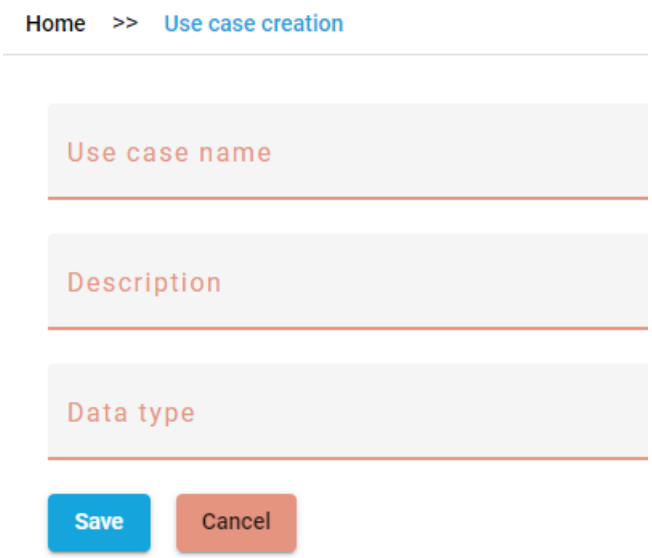


Figure 5. Use case details interface**.**

Once the use case is created, user can select it to work with and the platform display the main menu.

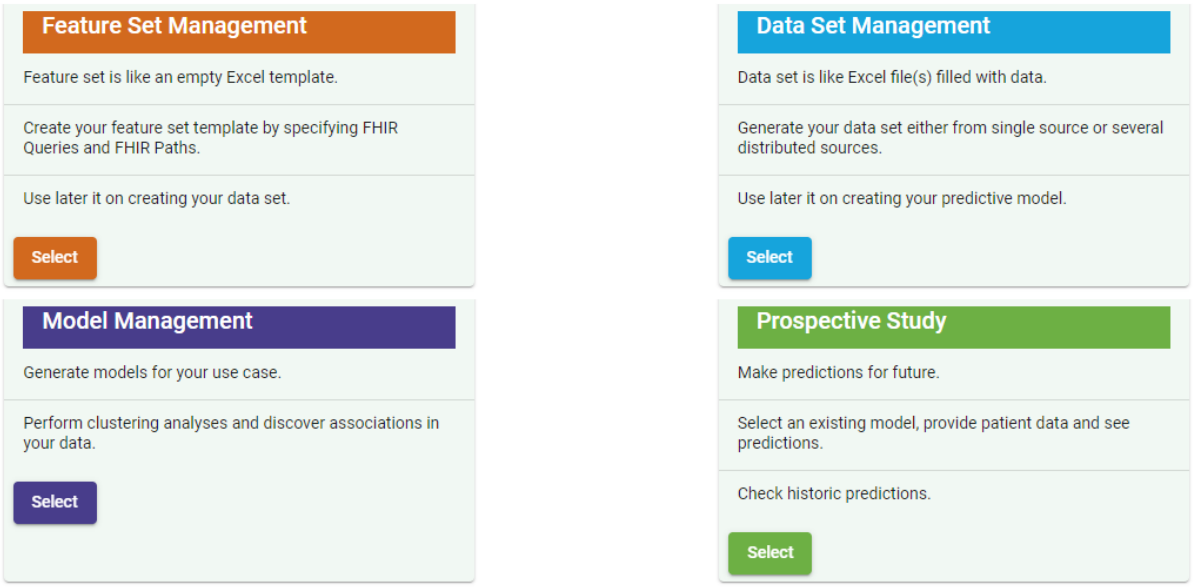


Figure 6. Main menu, after user case selection.

Figure 6. Shows on top, logged username (demo), Logout option to logout, selected use case details (name and description) and different options related to selected use case. All options are described on following sections.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

3.4. Feature Set Management

Feature Set is a template to specify FHIR Queries and Paths to identify dataset source.

This Feature Set will be used on Data Set Management section.

On Feature Set Management section, user can create a Feature Set to use it later (on Data Set Management) or see details from previously created Feature Sets.

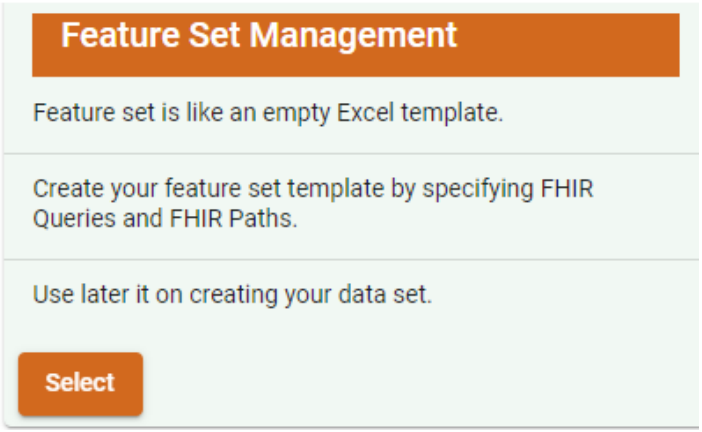


Figure 7. Feature Set menu Interface.

In this section the user can:

- See details of previously created Feature Set: Logged user can display any Feature Set details.
- Create a new Feature Set: Logged user can create a new Feature Set to work with.

Both functionalities are explained in following sections.

3.4.1. Feature Set list

To see Feature Set list, user must select related menu (Figure 7.), and the application will show the list of Feature Sets available.

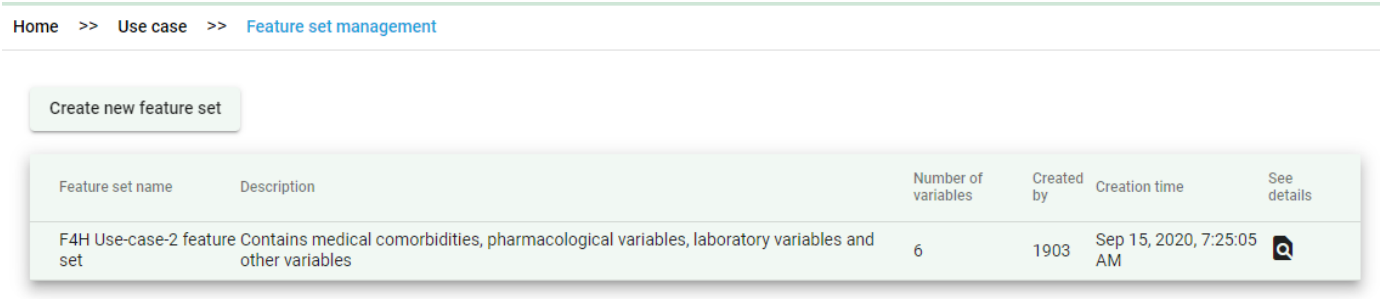


Figure 8. Feature Set list Interface.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

To see all details from any Feature Set, click on "See details" column of each related Feature Set:

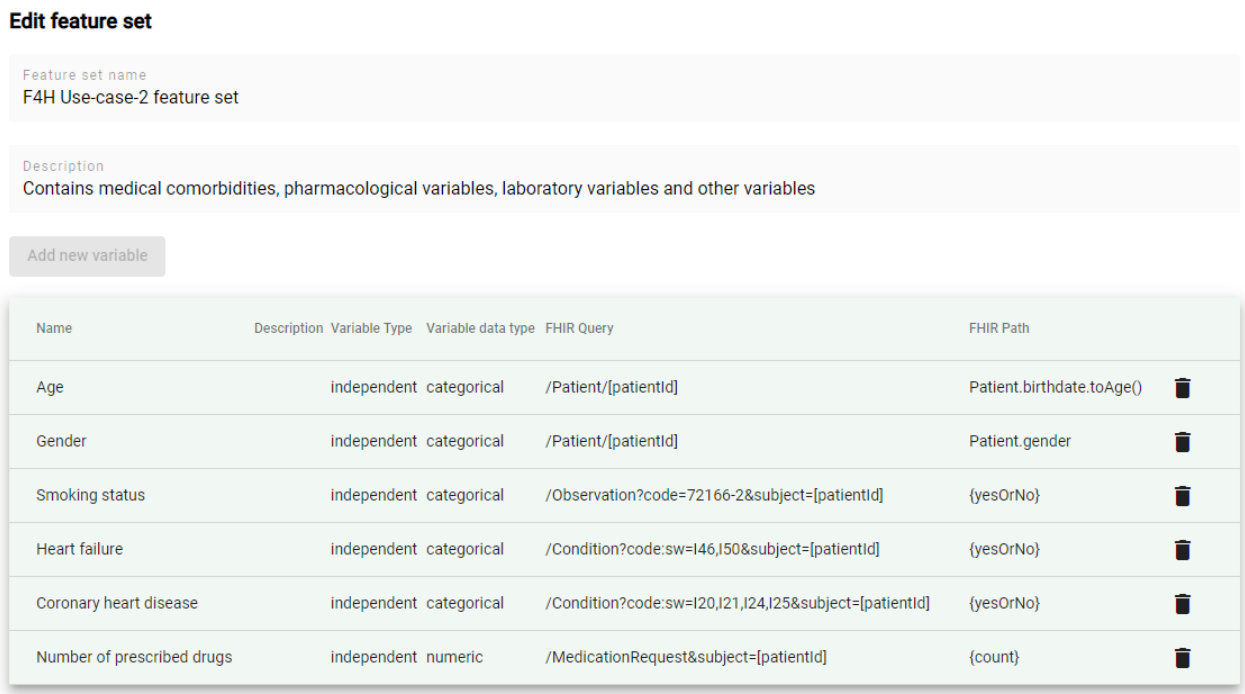


Figure 9. Feature Set details.

Notice that, once the Feature Set is created it cannot be updated or removed since the data model depends on those details. If Feature Set is removed related model analysis will be isolated.

3.4.1. Create Feature Set

Click on "Create new Feature Set" button to create a new Feature Set and specify all its details:

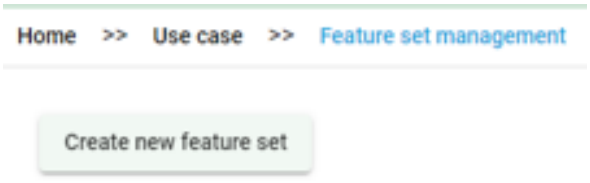


Figure 10. Feature Set details.

The system will ask for all its details and guide the user through all the process:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

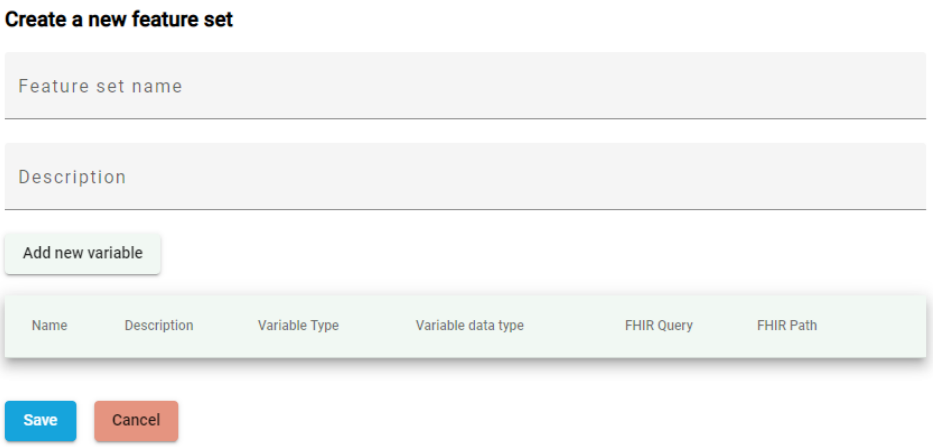


Figure 11. Create Feature Set, name and description details.

By clicking on "**Add new variable**" user can add a new variable, in the same mode, user can remove any variable by clicking on delete icon on each variable row (right). At this moment, the full Feature Set is not saved on the server, so the user can add or remove variables additional information or message confirmation (e.g. variable will be remove) are not necessary.

Variables must have following required details:

- Name: Variable name to identify it.
- Description: Variable short description to know additional details.
- Type:
  - Manual: User must type FHIR Path[1].
  - Select: User must select FHIR Path from a list.
- Variable type: Dependent or Independent.
- Data type: Categorical or numeric.

[1] https://www.hl7.org/fhir/fhirpath.html

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

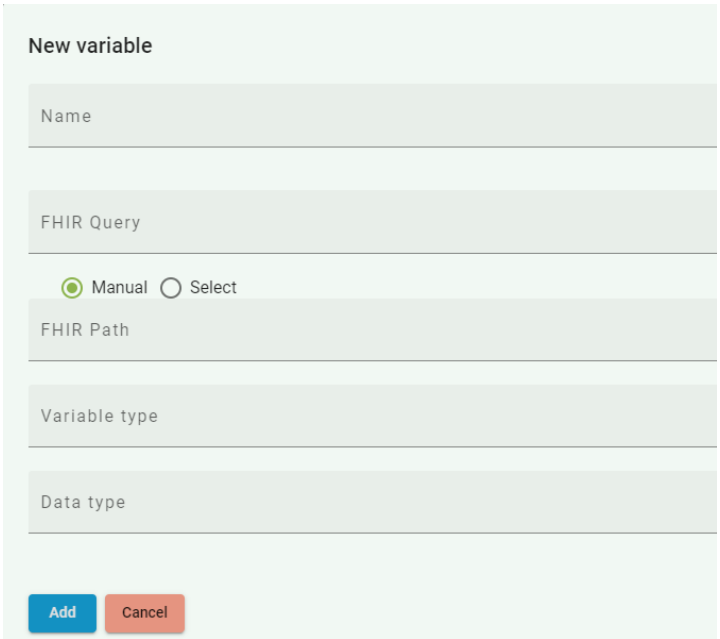


Figure 12. Add new variable interface.

The use of drop-down menus on Variable type and Data Type helps the definition of features and variables, given the fact that the number of choices is currently limited. Related to FHIR Path, the app will follow regular expression to help insertion details.

3.5. Data Set Management

On "Data Set Management" section, user can create a data set to use it later (on Model Management) or see details from previously created data sets.

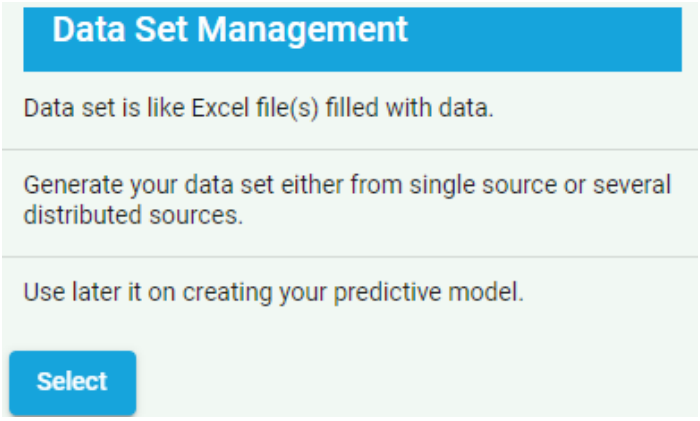


Figure 13. Data Set menu.

After getting into Data Set Management section by clicking on "**Select**" button, the system displays all Data sets and its details.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

3.5.1. Data set list

To see Data set list, user must select related menu, and the application will show the list of data sets available.

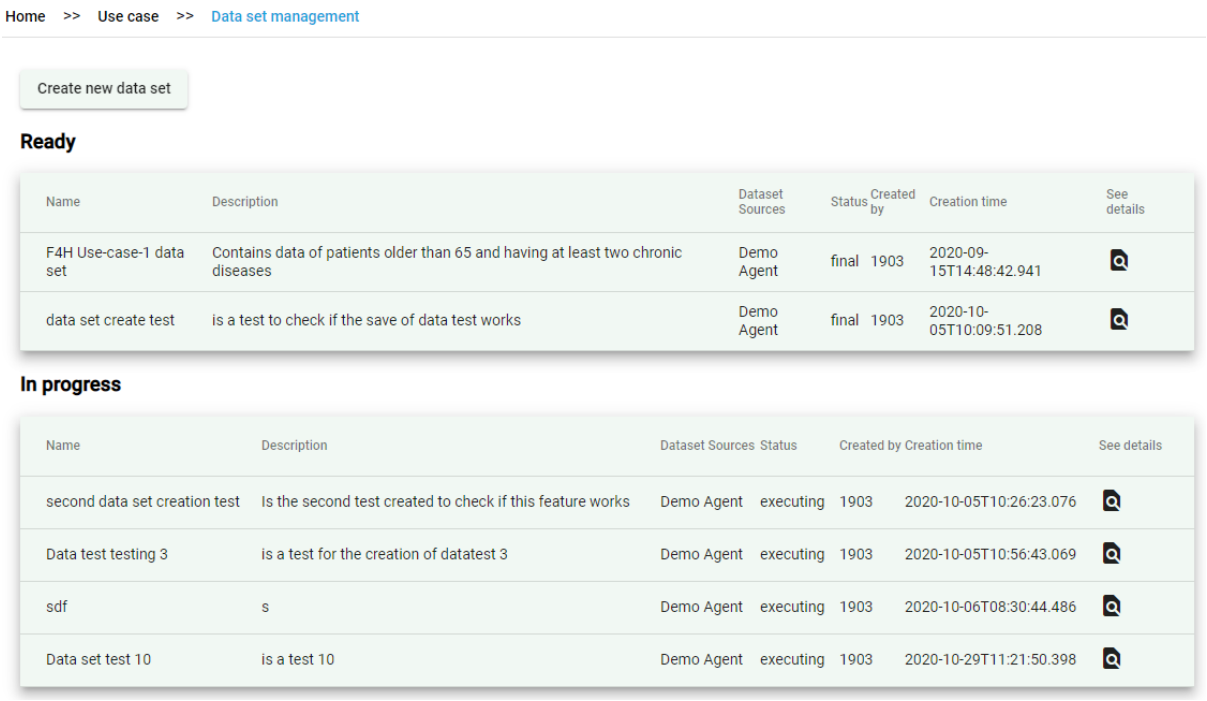


Figure 14. Data set list interface.

In this interface the user can:

- Create a new data set.
- See details of previously created data set.

3.5.2. Data set creation

To create a new Data set, logged user must follow these five identified steps:

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

1. Name and Description

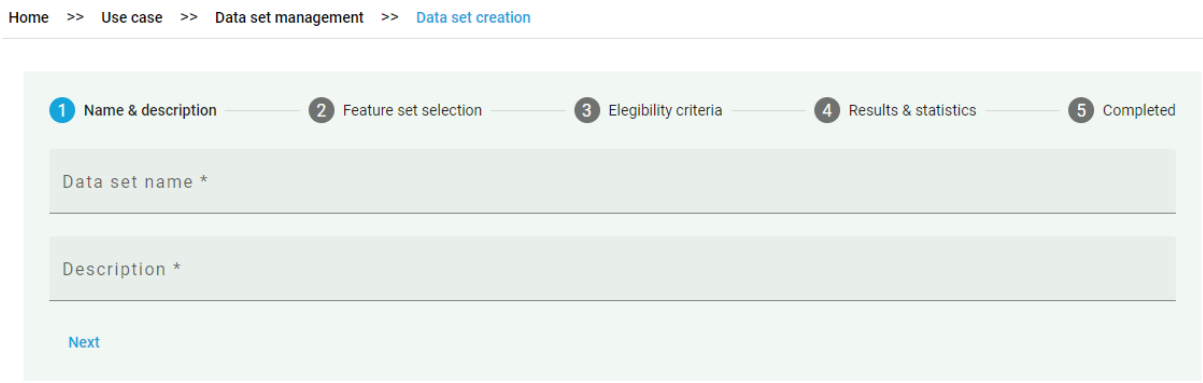


Figure 15. Data set creation workflow**.**

Each data set has following details:

- Name: Data set name to identify it.
- Description: Short description to know additional details.
- Feature Set: Feature Set selected and previously created (see section 3.3).
- Eligibility criteria: Search query and FHIR Path to identify data source and defined on section 2.

2. Feature Set selection

Logged user must select related Feature Set previously created

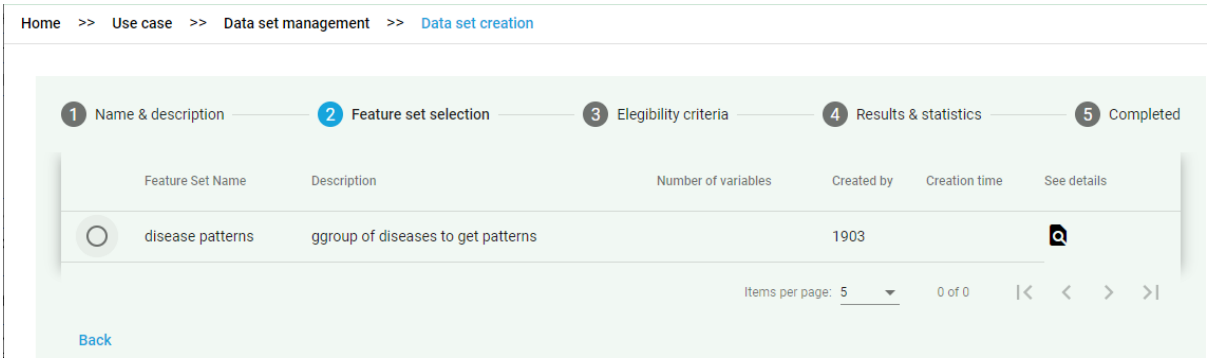


Figure 16. Data Set Management creation workflow, Feature Set selection.

3. Eligibility criteria

These two parameters; (1) FHIR query and (2) FHIR Path, are fundamentals to identify and select data source based on HL7 concepts.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

- FHIR Query[2]: Allows the platform to identify resources and conditions into them. For example, to get patients with a certain age, patient is the resource and age the condition. GUI will facilitate insertion providing format information and using regular expressions. Its format is: '/' + ResourceType + '?' + QueryParameter + Operation + Value.

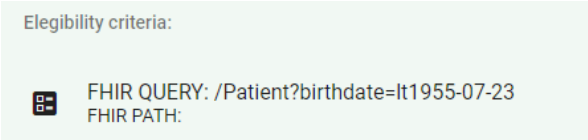


Figure 17. FHIR Query example to get Patients with a certain age.

- FHIR Path[3]: A path-based navigation and extraction language used in search parameter paths to define what contents the parameter refers to (e.g. Observation.dataAbsentReason).

4. Result & Statistics

This section shows details considering selected use case and Feature Set criteria.

5. Completed

Final section where data are stored into the system. After the data are saved those details can-not be changed or removed and Feature Set list table will be updated including the new element.

3.6. Model Management

On Model Management section, user can create a new model or see details from previously created models.

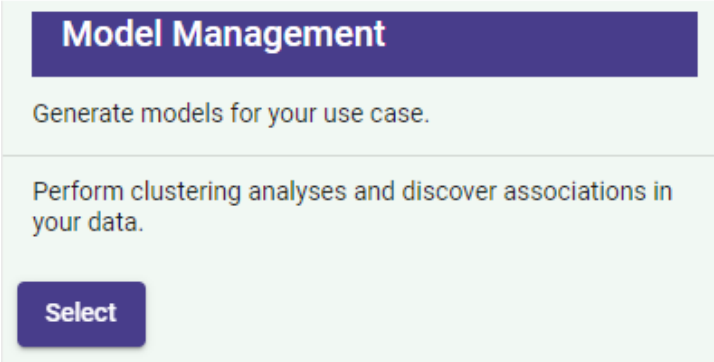


Figure 18. Model Management menu to get into Model management.

[2] https://www.hl7.org/fhir/search.html
[3] https://www.hl7.org/fhir/fhirpath.html

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

After getting into Model Management section by clicking on "**Select**" button, the system displays all Data models and its details.

In this section the user can:

- Create a new Model or.
- See details of previously created Model.

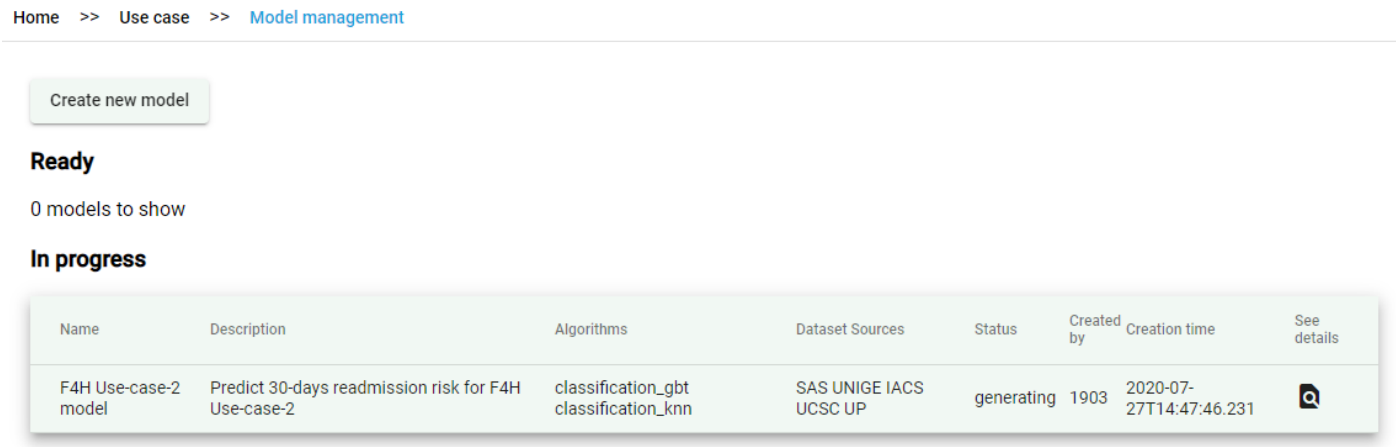


Figure 19. Data Model Management Interface**.**

3.6.1. Data Model creation

Each model has following details:

- Name: Model name to identify it.
- Description: Short description to know additional details and purpose.
- Data set: Each model has a dataset object selected from the list of data sets of the project (previously created).
- Categorical variables: The list of variables created in the Feature Set (previously created).
- Missing values: It is important for analysis purpose to identify the treatment of missing values (if any) in each variable.

To create a Model, logged user must follow these identified steps:

1. **Name and Description**

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

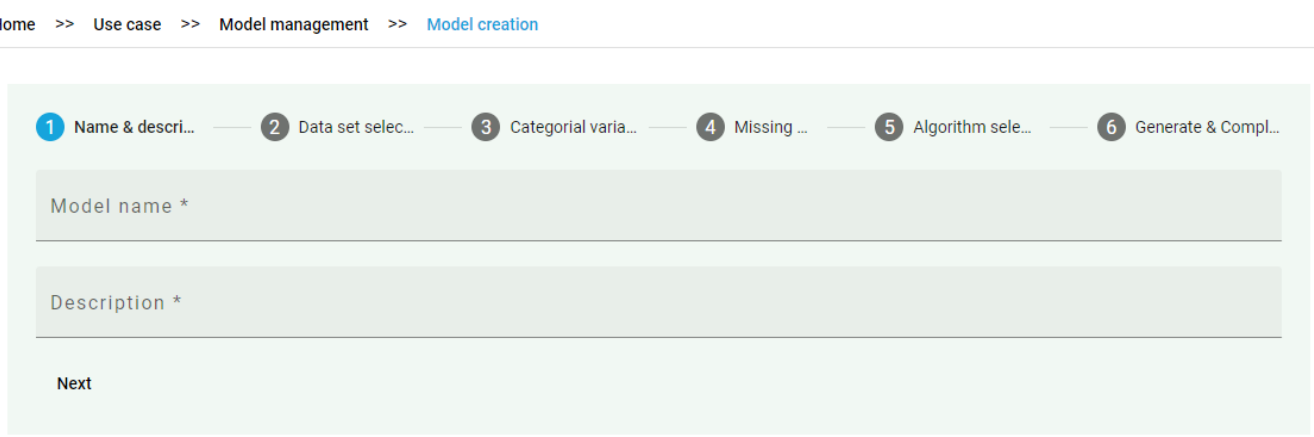


Figure 20. Model creation workflow**.**

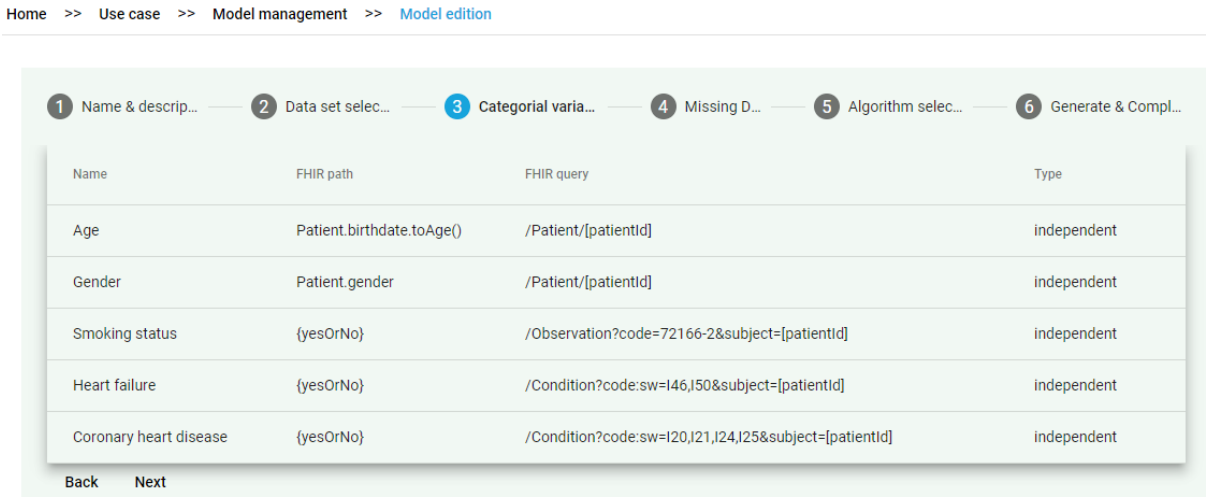
2. Data set selection

Logged user must select the appropriate Data set previously created.

Figure 21. Data set selection workflow**.**

3. Categorical variables

Each Data model has attached a Feature Set by which it is possible to recognize its variables and specify a treatment for each one of them.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Figure 22. Model creation workflow, Categorical variables.

4. Missing values data

Datasets may contain missing values for different reasons (typing errors, unknown information, etc.) those values may change analysis results, so, it is important for analysis purpose to identify the treatment of missing values (if any) in each variable.

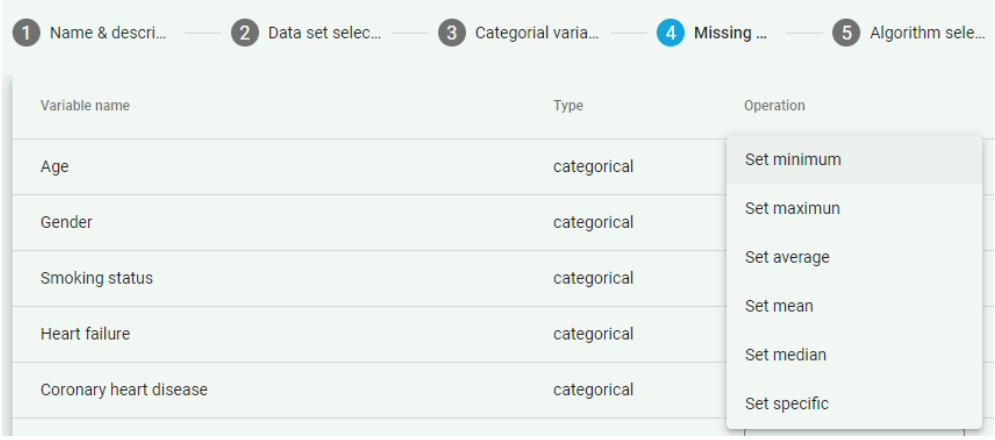


Figure 23. Missing values treatment selection.

For each variable user must identify missing values treatment by selecting one of those options:

- **Set minimum**: Missing values are replaced based on the minimum value. It can only be used with numeric data.
- **Set maximum:** Missing values are replaced based on the maximum value. It can only be used with numeric data.
- **Set average:** Missing values are replaced based on the average of no missing values. It can only be used with numeric data.
- **Set mean**: Missing values are replaced based on the mean of no missing values. It can only be used with numeric data.
- **Set median**: Missing values are replaced based on the median of no missing values. It can only be used with numeric data.
- **Set specific:** Missing values are replaced with specific values specified by the user.

5. Algorithm selection and parameters

This interface allows logged user to specify the algorithm and its parameters to use and its parameters. For example K-means algorithm will ask for the number of clusters, maximum number of iterations or number of time the k-means algorithm will be run with different centroid seeds and Decision tree regressor algorithm will ask for the function to measure the quality of a split, the maximum depth of the tree or The minimum number of samples required to split an internal node.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

Data sets will be trained using this selected algorithm and parameters running in the background since it is an asynchronous process.

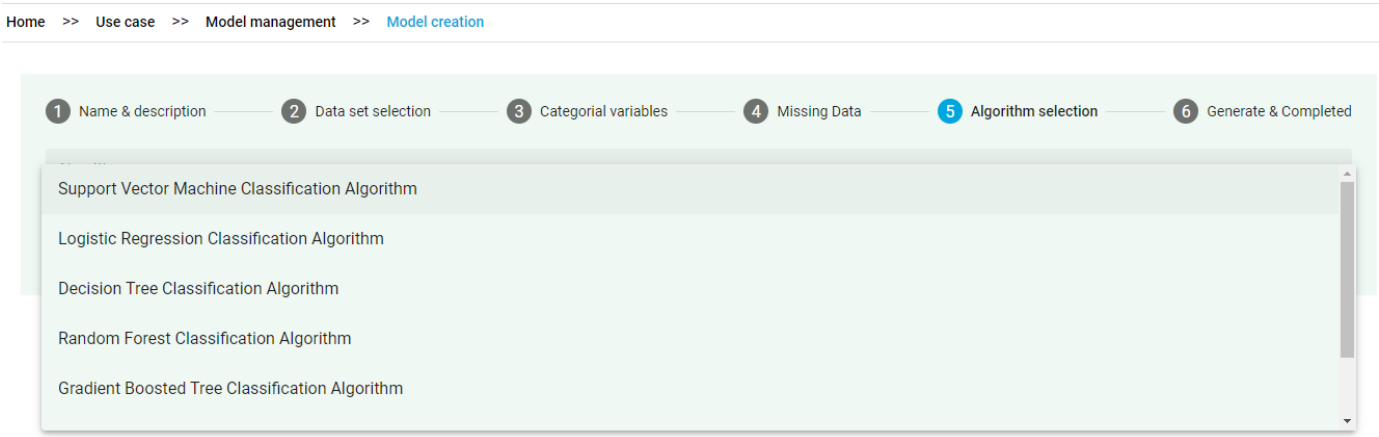


Figure 24. Algorithm selection.

For each algorithm user must fill up its corresponding parameters. Each time users select an algorithm; parameters will be updated according to selected one.

Notice that, use case one or association will show association algorithm and use case two or prediction will show only prediction algorithms.

6. Generate and completed

This is the final section where data model is stored into the system. After the data are saved those details can-not be changed or removed and data model list table will be updated with the new element.

3.7. Prospective Study

It is important to remember that this section is related to use case two or prediction. In this section users can make predictions for future by selecting an existing model, provide patient data and see predictions of the FAIR4Health platform.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

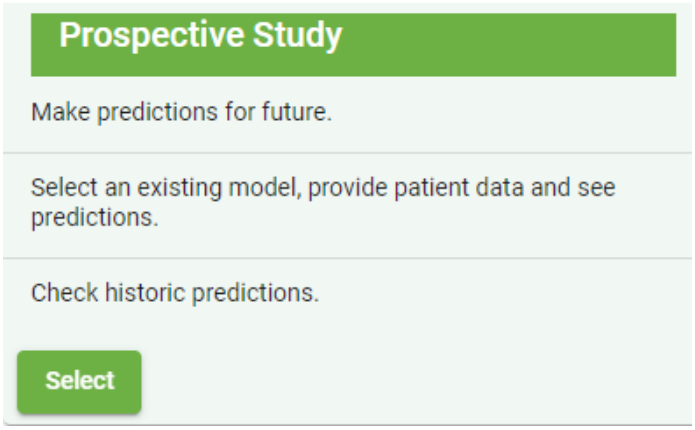


Figure 25. Prospective Study menu to get into Prospective section.

After getting into Prospective Study by clicking on "Select" button (Figure 20.), the system displays all studies and its details.

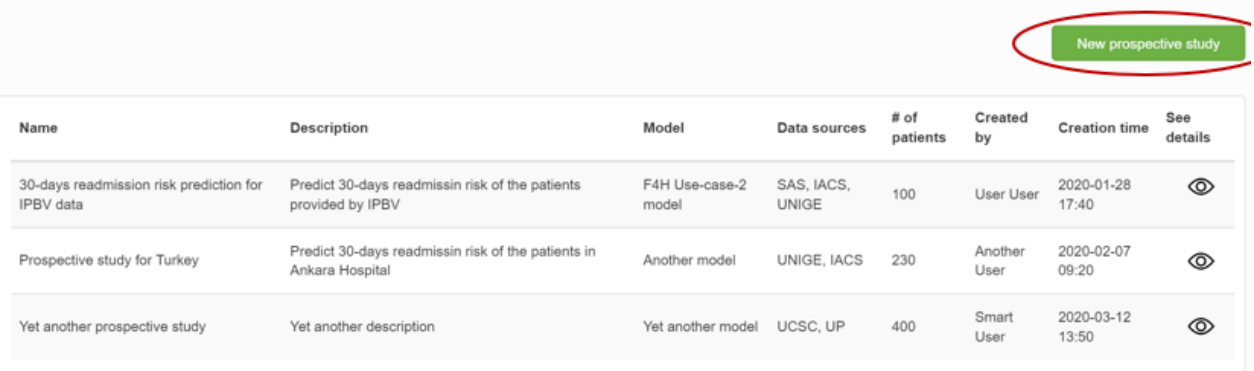


Figure 26. Prospective studies list Interface.

In this interface the user can:

- Create a new Prospective Study.
- See details of previously created Prospective Study.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

3.7.1. Prospective Study creation

This is a complex process for which the app uses a wizard approach to facilitate this process to end user.

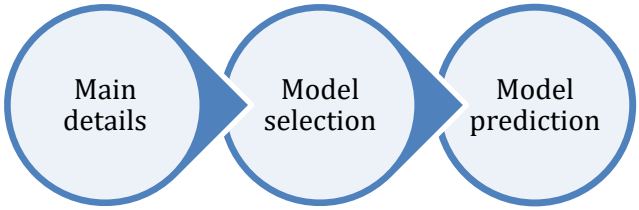


Figure 27. Prospective Study creation workflow.

1. **To create a new Prospective Study**

The user must identify its main details:

- Name of the study: Prospective name to identify its purpose.
- Description: Additional details to enforce understanding of study and purpose.

Figure 28. Prospective studies main details.

2. **Machine Learning model selection**

User must select the appropriate model from a list of them previously created.

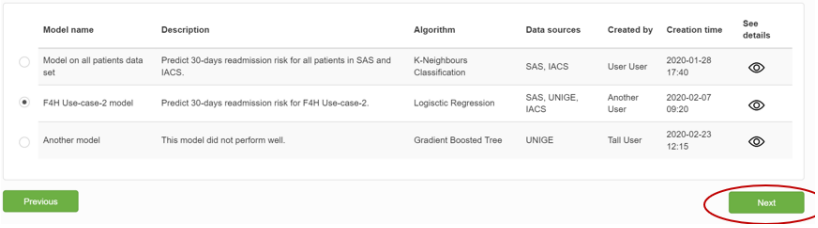


Figure 29. Machine Learning model selection.

3. **Model prediction**

- Model prediction for a user.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

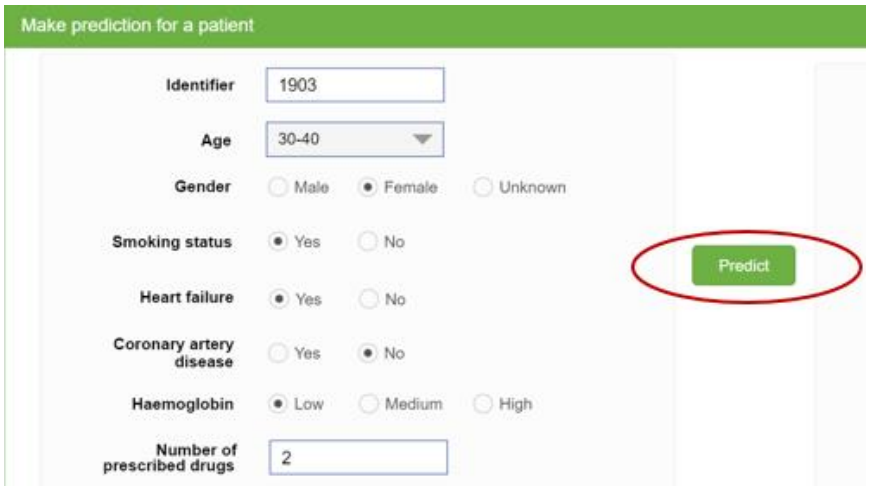


Figure 30. Machine Learning patient prediction.

- Model prediction for a set of patients: Logged user can upload a set of patient details to make predictions on them.

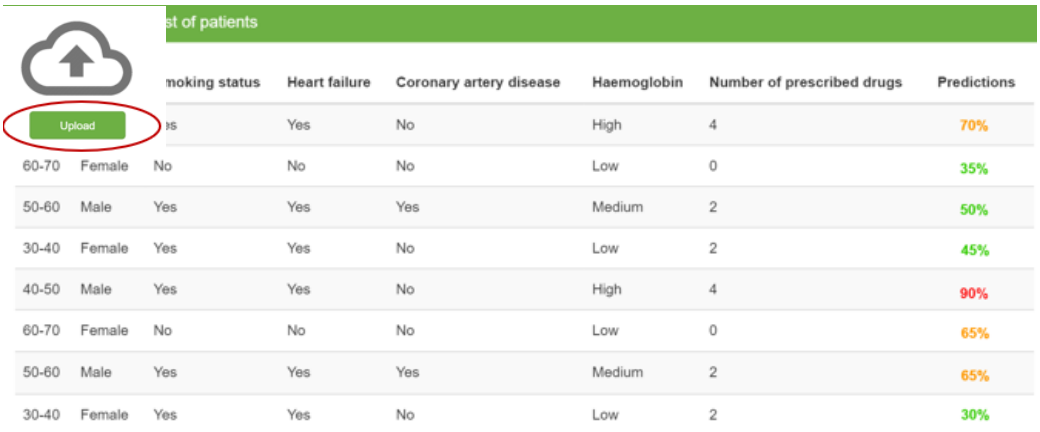


Figure 31. Machine Learning set of patient prediction.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

4. FAQs

- **How could I get my own username and password to access to FAIR4Health platform?**

At this moment, users can access using testing user demo. Once this testing phase gone each user must ask for its own username and password to the administrator.

- **Can I update a use case detail?**

No, once users save the data, they cannot be updated or changed. This happen with use case and Feature Set and data models. This happens since, the data analysis depends on those details (use case, Feature Set and data model) if we remove any of them the data analysis will get isolated and unreachable.

- **Can I edit the parameters of a model that has already been created?**

Yes, once the model is saved, is possible to edit its details by clicking on "**see details**" button. Logged user can see all its details, but can-not update them, since the data analysis related to removed data model will become isolated.

- **How can I calculate the prediction of a set of patients already created?**

User can upload an Excel file with a set of patients to calculate its predictions, This file must ensure same variables than those defined in related Feature Set, see section 3.7.1 Prospective Study creation.

- **Can I download the predictions made on the patient data in the Prospective Study?**

No at this moment, the platform could be improved to add this functionality.

- **How can I edit Feature Set?**

Log into the system and select one user case, the system will display the main menu, click on "Feature Set Management". Feature Set Management section shows a list of all Feature Set created, logged user can see its details by selecting one of them and click on "**see details"** button, the app will display all its details.

- **Can I edit the parameters of an algorithm that has already been generated?**

Yes, Log into the system and select one user case, the system will display the main menu, click on "Data Model Management". Data Model Management section shows a list of all models created, logged user can see its details by clicking on "**See details**" button, the app will display all its details.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666

- **How do you facilitate to find suitable use cases?**

After users gets into the system, the system displays a list of available use cases. Those uses cases are fully identified by name, description and model type. This information is displays since the very beginning so, logged user does not need to make additional clicks to identify the proper use case to work with. Nevertheless, if the list of use cases significantly increases, it will be possible to include a smart searcher.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 824666