"Smart Big Data Platform to Offer Evidence-based Personalised Support for Healthy and Independent Living at Home"

**D4.2 (D14) - Report on SMART BEAR Cloud Enabling Components v1**

Due date of deliverable: 2021-02-28  
Actual submission date: 2021-04-26

<table>
<thead>
<tr>
<th>Grant agreement number: 857172</th>
<th>Lead contractor: CNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Start date of project: 01/09/2019</td>
<td>Duration: 60 months</td>
</tr>
</tbody>
</table>

Project funded by the European Commission within the EU Framework Programme for Research and Innovation HORIZON 2020

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PU = Public, fully open, e.g. web</td>
<td>✓</td>
</tr>
<tr>
<td>CO = Confidential, restricted under conditions set out in Model Grant Agreement</td>
<td></td>
</tr>
<tr>
<td>CI = Classified, information as referred to in Commission Decision 2001/844/EC.</td>
<td></td>
</tr>
<tr>
<td>Int = Internal Working Document</td>
<td></td>
</tr>
</tbody>
</table>
D4.2 (D14) - Report on SMART BEAR Cloud Enabling Components v1

Editor
Anca Bucur (Philips)

Contributors
Paolo Ceravolo (UMIL)
Fady Copty (IBM)
Silvio Di Iorio (ARIA)
Fanis Kalatzis (UOI)
Christos Kloukinas (CITY)
Ioannis Kouris (ICCS)
Jasper van Leeuwen (Philips)
Ioannis Leontsinis (NKUA)
Jonatan Maggesi (UMIL)
Samira Maghool (UMIL)
Mircea Marzan (ANA)
Mattia Occhipinti (UMIL)
Manuel Perez (ATOS)
Andreea Stamate (ANA)
Qiqi Su (CITY)
Stefano Siccardi (UMIL)
Mario Sicuranza (CNR)
Kostas Tsiouris (UOI)
Eleftheria Vellidou (NKUA)

Reviewers
Ioannis Basdekis (STS)
George Zissis (ATC)
Executive Summary

This document reports on the first version of the components of the SMART BEAR backend Cloud platform. The deliverable documents the backend components and the analytics, together with planned work of WP4 focused on the integration of the platform, the further development of the key components, and the implementation and evaluation of the analytics approaches adopted by the project.

The main components of the SMART BEAR cloud platform include the database and its underlying information model, the clinical repository interfaces, the Big Data Engine, the synthetic data generation component, and the analytics. The analytics and personalization components leverage the SMART BEAR FHIR-based Information Model.

The specification of the FHIR based model is detailed in this deliverable. The FHIR standard is used for the representation of the clinical data that will be collected in the project and provides well-defined semantics that use widely adopted ontologies such as LOINC and SNOMED-CT, and well-specified value sets for each resource. Based on our analysis, this approach covers well the semantic representation of data in the clinical domains of the project and for the relevant data categories: medical history, vital signs, laboratory results, medication and measurements. It also contains concepts for Hearing aid fitting and observable entities, such as the physical activity of the participant. The scales used by the clinicians to assess the patients are covered as well by SNOMED-CT, such as Epworth Sleepiness Scale, Pittsburgh sleep quality index, Euro Quality of Life, Assessment using geriatric depression scales, Montreal cognitive assessment, etc. Using the standardized representation of the data collected in the SMART BEAR platform, we will streamline the development of analytics and decision models to provide accurate personalised interventions.

The SMART BEAR database needs to persist as well non-clinical data, e.g. configurations, model parameters, etc. For this data, a custom representation has been chosen, as it is not covered by the FHIR standard.

The analytics approaches presented in this report are linked to specific clinical scenarios and relevant interventions that will be modelled. The initial data elements that will be used for the development of the models are presented as well. Additionally, the Full Homomorphic Encryption approach is included as well, which addresses the need for additional data privacy when data is stored and analysed in untrusted environments, enhancing the scalability and extensibility of our platform.

These components will be implemented and evaluated in the SMART BEAR pilots. A subsequent deliverable (D4.4) will describe the enhancements that will be implemented to address the additional requirements produced during the pilot evaluation. The first integrated system, in the form of a user guide, is presented in deliverable 4.1.
Contents

Executive Summary ............................................................................................................................................. 3
List of Figures .................................................................................................................................................. 6
List of Tables .................................................................................................................................................. 8
1 Introduction .................................................................................................................................................. 9
  1.1 Overview ............................................................................................................................................... 9
  1.2 Goals of the deliverable ....................................................................................................................... 10
2 SMART BEAR Cloud Enabling Components ..................................................................................... 11
  2.1 Database Implementation ................................................................................................................... 11
  2.2 Data Model Specification compliant with FHIR ............................................................................. 18
    2.2.1 Overview on FHIR and available tools ....................................................................................... 18
    2.2.2 Methodology used in the SMART BEAR project .................................................................. 20
    2.2.3 Data model mapping ................................................................................................................. 20
    2.2.4 Demographics ............................................................................................................................ 21
    2.2.5 Medical History .......................................................................................................................... 23
    2.2.6 Physical examination .................................................................................................................. 29
    2.2.7 Hearing Loss .............................................................................................................................. 32
    2.2.8 Mental Disorders ......................................................................................................................... 34
    2.2.9 Cognitive disorders .................................................................................................................... 34
    2.2.10 Frailty .......................................................................................................................................... 34
  2.2.11 Synthetic Data Generation ......................................................................................................... 61
    2.2.12 Questionnaire modelling .......................................................................................................... 38
      2.2.12.1 MOCA questionnaire ........................................................................................................... 38
      2.2.12.2 phq-9 questionnaire ............................................................................................................ 40
      2.2.12.3 HAM-A questionnaire ......................................................................................................... 42
      2.2.12.4 GDS-15 questionnaire ......................................................................................................... 45
      2.2.12.5 HUI questionnaire ................................................................................................................. 48
    2.2.13 FHIR artefacts produced ............................................................................................................ 50
  2.3 The Clinical Data Repository ............................................................................................................. 54
  2.4 BDA Engine ........................................................................................................................................... 56
    2.4.1 Functions and tools .................................................................................................................... 57
      2.4.1.1 Data storing and management ............................................................................................. 57
      2.4.1.2 Data analysis and tasks ........................................................................................................ 57
      2.4.1.3 Data flow management and workflow orchestration ......................................................... 57
      2.4.1.4 Data visualization ................................................................................................................ 58
      2.4.1.5 Platform Management ....................................................................................................... 58
    2.4.2 Analytics Orchestration ............................................................................................................... 58
  2.5 Synthetic Data Generation ................................................................................................................ 61
  2.6 Analytics ............................................................................................................................................... 67
    2.6.1 Analytics for Hearing Loss ....................................................................................................... 67
      2.6.1.1 Interventions .......................................................................................................................... 67
      2.6.1.2 Data ..................................................................................................................................... 67
2.6.1.3 AI Model and Approach................................................................................................................68
2.6.2 Developing AI Models with Federated and Continuous Learning for Personalization in CVD, Cognitive Disorders and Frailty ....................................................................................71
  2.6.2.1 Federated and Continuous Learning Framework ..............................................................................71
  2.6.2.2 Models and interventions in the CVD scenario ..................................................................................73
  2.6.2.3 Models and interventions in the Cognitive Diseases scenario ..........................................................75
  2.6.2.4 Models and interventions in the Frailty scenario ..............................................................................77
2.6.3 COVID-19 risk assessment ................................................................................................................78
2.6.4 Intrinsic Capacity................................................................................................................................82
  2.6.4.1 Definition ........................................................................................................................................82
  2.6.4.2 Use case ..........................................................................................................................................82
  2.6.4.3 Data ................................................................................................................................................82
  2.6.4.4 Data format .......................................................................................................................................83
  2.6.4.5 Frequency of collecting data by using SMART BEAR devices ......................................................83
  2.6.4.6 Test validity .....................................................................................................................................84
  2.6.4.7 Z-score ............................................................................................................................................84
  2.6.4.8 Performance score and parameters' weights in domain and sub-domain aggregation score ........84
  2.6.4.9 Parameters in sub-domain and domain aggregation score ...............................................................84
2.6.5 Models and interventions in the Mild Depression scenario (Mental diseases) .............................88
2.6.6 Full Homomorphic Encryption of Machine Learning Models ........................................................91
2.7 Decision Support System (DSS) ..............................................................................................................92
2.8 Dashboard.............................................................................................................................................96
  2.8.1 Registration and login .........................................................................................................................97
  2.8.2 Patients management ...........................................................................................................................98
  2.8.3 Available patients ...............................................................................................................................99
  2.8.4 Overview of a patient ...........................................................................................................................99
  2.8.5 Patient Medical History ......................................................................................................................101
  2.8.6 Patient Creation ................................................................................................................................102
  2.8.7 External ID creation .............................................................................................................................103
  2.8.8 Device allocation .................................................................................................................................104
  2.8.9 CVDs use case ...................................................................................................................................105
  2.8.10 Questionnaires ................................................................................................................................106
  2.8.11 Interventions ....................................................................................................................................107
    2.8.11.1 Intervention creation ....................................................................................................................107
    2.8.11.2 Notifications ...............................................................................................................................109
    2.8.11.3 Intervention logs ..........................................................................................................................110
  2.8.12 Dashboard requirements .................................................................................................................110
3 Concluding Remarks ...................................................................................................................................111
4 References ..................................................................................................................................................112
Appendix .......................................................................................................................................................113
List of Figures

Figure 1. BDA workflow configuration ............................................................................................................. 12
Figure 2. SUSHI results........................................................................................................................................ 20
Figure 3. SMART BEAR FHIR Implementation Guide Home page.............................................................. 50
Figure 4. Resource Profiles ............................................................................................................................. 51
Figure 5. Download page .................................................................................................................................. 52
Figure 6. Example of a JSON representation ................................................................................................. 53
Figure 7. Example of a GDS-15 questionnaire .................................................................................................. 54
Figure 8. Scheme of how data will be stored into the SMART BEAR server through the API. The HL7 FHIR compliance data is transferred in JSON format and is stored in the SB repository. .... 55
Figure 9. Scheme of the SMART BEAR HAPI-FHIR SERVER and the internal architecture.............. 55
Figure 10. The code for the POST method for patient-related data entry indicating allowed values, type of variable, meaning of the content within the whole dataset, etc. ...................................................... 56
Figure 11. BDA tools .......................................................................................................................................... 58
Figure 12. Graphic representation of a simple Synthea module ................................................................. 61
Figure 13. A simplified module to generate data for SMART BEAR .......................................................... 63
Figure 14. An example of weekly muscle mass measures generated by the simplified module for 4 patients ................................................................................................................................................. 64
Figure 15. An example of daily walked distance and number of steps generated by the simplified module for a patient.................................................................................................................................................. 65
Figure 16. A schematic view of a BN in healthcare data. The relationships among variables and their dependencies are clear. The thickness of links representing the weight of relationships.. 66
Figure 17. General FL approach ...................................................................................................................... 72
Figure 18. Data for CVD analytics .................................................................................................................. 75
Figure 19. The example of locomotion domain’s points from ICOPE handbook........................................ 83
Figure 20. Screenshot of the registration form ................................................................................................. 97
Figure 21. Screenshot of the login form ........................................................................................................... 97
Figure 22. Home page ...................................................................................................................................... 98
Figure 23. The patient page ............................................................................................................................ 99
Figure 24. Patient overview ........................................................................................................................... 100
Figure 25. Devices Management .................................................................................................................. 100
Figure 26. ID Management ............................................................................................................................ 101
Figure 27. Patient Medical History (Life Habits and Medications)............................................................. 101
Figure 28. Patient Medical History (Diet Supplements) ................................................................................. 102
Figure 29. Create Patient form ....................................................................................................................... 102
Figure 30. Create External ID form ................................................................................................................ 103
Figure 31. Assign Device form ....................................................................................................................... 104
Figure 32. Cardiovascular use case page ...................................................................................................... 105
Figure 33. The Falls Efficacy Scale ............................................................................................................... 106
Figure 34. Medical interventions page ......................................................................................................... 107
Figure 35. Creation of a Medical Intervention ............................................................................................. 108
Figure 36. Creation of a Medical Intervention. ............................................................................................. 108
Figure 37. Notifications page ............................................................................................................................... 109
Figure 38. Create Notification form .................................................................................................................. 109
Figure 39. Interventions Logs page ................................................................................................................... 110
List of Tables

Table 1. Data entities and related Smart Bear components.............................................................. 11
Table 2. Workflow configuration elements.......................................................................................... 12
Table 3. BDA results............................................................................................................................ 13
Table 4. BDA results visualization......................................................................................................... 13
Table 5. User settings.......................................................................................................................... 14
Table 6. Intermediate results............................................................................................................... 14
Table 7. Notifications.......................................................................................................................... 15
Table 8. Interventions.......................................................................................................................... 15
Table 9. Alerts....................................................................................................................................... 16
Table 10. External data......................................................................................................................... 17
Table 11. Device references................................................................................................................ 17
Table 12. Device status......................................................................................................................... 18
Table 13. Balance disorder value set.................................................................................................. 24
Table 14. Extracted Data Dictionary.................................................................................................... 68
Table 15. A Snippet of the Sampled Data and the Computation of $D_i, u_t, dt$, and HAU........... 69
Table 16. Elements of an HMM.......................................................................................................... 69
Table 17. Intrinsic Capacity data types and domains ........................................................................ 84
Table 18. Intrinsic Capacity intermediate data.................................................................................... 87
Table 19. Components and interfaces.................................................................................................. 92
Table 20. Components for alerts......................................................................................................... 93
Table 21. Components for interventions and notifications................................................................. 94
1 Introduction

1.1 Overview

Deliverable 4.2 Report on SMART BEAR Cloud Enabling Components v1 reports on the first version of the components that compose the SMART BEAR backend Cloud platform (SB@Cloud). This deliverable documents the infrastructure and analytics development activities part of WP4, which focuses on delivering all the components needed for the SB@Cloud infrastructure, including the data repository, the analytics models, the decision support system, the big data approaches, the developed interventions and their evaluation, and the visualisation.

The main components of the SB@Cloud include the database and its underlying information model, the clinical repository interfaces, the Big Data Engine, the synthetic data generation component, and the analytics.

This deliverable describes as well the SMART BEAR approach for personalization within the platform, with a focus on leveraging AI and analytics to enable effective personalized interventions. The analytics and personalization components build on the SMART BEAR Information Model. A strong semantic underpinning and a standards-based data representation, providing access to comprehensive and homogeneous datasets with full context representation, will facilitate the efficient use of the data for the implementation of high-quality, personalized interventions.

To provide the specification of the data model, we leverage the HL7 FHIR standard, and use FHIR resource profiles to define constraints and extensions to the FHIR model capturing the required information in a standard structure and with rich semantics (based on FHIR-supported ontologies and value sets).

The analytics approaches introduced in this deliverable are presented in the context of specific clinical scenarios. As previously reported, SMART BEAR develops two types of models for interventions: (a) clinical intervention models that may change patient management, and (b) interventions that do not impact the current patient management but learn from the patient-generated data to seamlessly adapt the interventions to patients’ preferences. For the first type of intervention models, the project will only undertake technical validation, while further prospective clinical validation is necessary beyond the scope of the project. Under the guidance of clinical users, model results may be shown to participating patients within regular summary reports or targeted evaluation sessions. The second type of intervention models that only improve the fit of the recommendations to the patients’ preferences will be deployed and evaluated with participating patients in the production environment.

SMART BEAR leverages an iterative approach for the development and deployment of interventions. The initial interventions will implement the current state of practice as described in the deliverable D2.1 clinical scenarios. The project collects a wide range of diverse variables that are currently used in clinical practice, that are reported as relevant in clinical research, or that the clinical experts expect could improve the accuracy of the predictive models and the outcomes of the interventions.

In each scenario and dataset, we will start with the exploratory analysis (e.g., classic statistics and small-scale AI experiments) of the data and identify relevant types of algorithms that fit the planned intervention and dataset. For each intervention, subsets of the relevant data will be explored with respect to predictive value and effect on the intervention. As the most suited type of algorithm for the problem is not always obvious, several algorithms may be developed and compared for each intervention.

Technical validation will be independently carried out for all algorithms. A best performing algorithm may be selected based on validation results, or several algorithms may be combined in line with Ensemble Learning techniques to increase the performance. Next to standard performance criteria, explainability of AI models will be considered.
The Full Homomorphic Encryption approach is introduced as well, which is applicable to a wide range of Machine Learning algorithms, as described in Section 2.6.6. This will be implemented and evaluated for several SMART BEAR models. The introduction of this privacy-preserving approach will facilitate the scalability and extensibility of the SMART BEAR infrastructure beyond the end of the project, enabling the use of external cloud resources in full compliance with the privacy requirements (e.g. for scalability, cost optimization, or increased performance), and the ingestion in our infrastructure of datasets external to the project that may have different privacy contexts.

1.2 Goals of the deliverable

This deliverable provides the technical specification of the SB@Cloud enabling components, describing their design, interactions, interfaces and functionality. A core topic of this deliverable is the solution to data representation and management, including a full specification of the FHIR-compliant information model for clinical data, the description of the full data model (including clinical and non-clinical data) and of the database implementation.

We present as well the selected approaches for the implementation of AI and analytics to support the development of effective interventions that will be deployed and evaluated in the SMART BEAR pilots. Each analytics approach is presented in the context of a specific clinical scenario, describing the targeted interventions, the process of developing and validating the model, and the used data. Next to these specific analytics approaches, the Full Homomorphic Encryption approach is introduced as well.
2 SMART BEAR Cloud Enabling Components

2.1 Database Implementation

The Data Repository component of SB@Cloud (SB@Repository) is utilising a combination of FHIR and non-FHIR databases. All the data that represent medical data entities are stored in the FHIR database (details of the data model are summarized in Section 2.2), while data related to non-medical entities are stored in the secondary database of the Cloud Backend. Such data include the entities described in the following table. Those contain elements that are not mapped to FHIR models (such as SB@Dashboard user settings), and intermediate results of the analytics models which some of them relay data back to the FHIR database. Data arriving from external vendors and/or synergy projects, will also be stored in the non-FHIR database and then relayed to the FHIR database.

Table 1. Data entities and related Smart Bear components

<table>
<thead>
<tr>
<th>Description</th>
<th>Related SMART BEAR Component(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BDA workflow configuration</td>
<td>BDA</td>
</tr>
<tr>
<td>BDA analysis results</td>
<td>BDA</td>
</tr>
<tr>
<td>BDA results visualization</td>
<td>BDA, Dashboard</td>
</tr>
<tr>
<td>User settings</td>
<td>Dashboard</td>
</tr>
<tr>
<td>Intermediate data and XAI results</td>
<td>Analytics</td>
</tr>
<tr>
<td>Notifications</td>
<td>DSS</td>
</tr>
<tr>
<td>Interventions</td>
<td>DSS</td>
</tr>
<tr>
<td>Alerts</td>
<td>DSS</td>
</tr>
<tr>
<td>Data from device vendors/synergies</td>
<td>Data repository, Dashboard</td>
</tr>
<tr>
<td>Patient devices</td>
<td>Dashboard, DSS</td>
</tr>
<tr>
<td>Patient device status</td>
<td>Dashboard, DSS</td>
</tr>
</tbody>
</table>

The following section provides the details of the data to be stored in the FHIR database. In this section we focus on the non-medical data that will not be modelled in FHIR. As mentioned above, for the non-medical data and data required for the execution of the analytics, along with metadata that relate to the interventions and the notifications, an additional database is required. In this database, the status of the patient devices will be stored, along with data arriving from non SMART BEAR sources, such as the data from the synergies and data from device vendors that do not transmit data using the FHIR data model. For each of the entities the details are provided below.
BDA workflow configuration

This data model stores the workflow templates that are used the BDA component (SB@BDA, or BDA). It contains the configuration details for the BDA Engine to create the proper template in the HDFS of the SB@BDA to use it in the execution of the analytics. This structure is used by the Dashboard in order to allow clinicians/scientists to customize the execution of a workflow. A sample of the workflow template data is available in the Appendix on this document. The configuration file is generated via the Dashboard component and stored in the database, along with the scheduled time to be executed. The image below provides a schematic representation of the flow.

![Figure 1. BDA workflow configuration](image)

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorkflowId</td>
<td>Integer</td>
<td>Identifier of the workflow</td>
</tr>
<tr>
<td>Workflow</td>
<td>JSON</td>
<td>Workflow configuration for the BDA Engine</td>
</tr>
<tr>
<td>ExecutionDatetime</td>
<td>Datetime</td>
<td>The scheduled time to execute the workflow</td>
</tr>
<tr>
<td>CreateByUserID</td>
<td>Integer</td>
<td>The identifier of the user that created the workflow via the dashboard</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
<tr>
<td>ModifiedDatetime</td>
<td>Datetime</td>
<td>The modification datetime</td>
</tr>
</tbody>
</table>
BDA analysis results

This data model stores the workflow execution results, which are produced by the BDA component. It contains the results and reference to the workflow that triggered the execution. A sample of the analysis results is available in the Appendix of this document.

*Table 3. BDA results*

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>WorkflowResultId</td>
<td>Integer</td>
<td>Identifier of the workflow result</td>
</tr>
<tr>
<td>WorkflowId</td>
<td>Integer</td>
<td>Reference to WorkflowId that triggered the execution</td>
</tr>
<tr>
<td>ExecutionStartDatetime</td>
<td>Datetime</td>
<td>The time the execution of the workflow started</td>
</tr>
<tr>
<td>ExecutionEndDatetime</td>
<td>Datetime</td>
<td>The time the execution of the workflow completed</td>
</tr>
<tr>
<td>Results</td>
<td>JSON</td>
<td>Results of the execution of the BDA Workflow</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
</tbody>
</table>

BDA results visualization

This data model stores data that are stored and consumed by the Dashboard to visualize the results of the BDA analysis. Example of the real data are available in the Appendix.

*Table 4. BDA results visualization*

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>DashboardChartId</td>
<td>Integer</td>
<td>Identifier of the dashboard chart</td>
</tr>
<tr>
<td>WorkflowId</td>
<td>Integer</td>
<td>Reference to WorkflowId that triggered the execution</td>
</tr>
<tr>
<td>ExecutionDatetime</td>
<td>Datetime</td>
<td>The scheduled time to execute the workflow</td>
</tr>
<tr>
<td>ChartData</td>
<td>JSON</td>
<td>The actual data that are used by the chart</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
<tr>
<td>ModifiedDatetime</td>
<td>Datetime</td>
<td>The modification datetime</td>
</tr>
<tr>
<td>Revision</td>
<td>Integer</td>
<td>Value to keep number of updates of the visualizations</td>
</tr>
</tbody>
</table>
User settings

This data model is used to store the user settings for the dashboard visualization. Example is available in the Appendix.

Table 5. User settings

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>UserSettingsId</td>
<td>Integer</td>
<td>Identifier of the user settings</td>
</tr>
<tr>
<td>UserId</td>
<td>Integer</td>
<td>The identifier of the user</td>
</tr>
<tr>
<td>Settings</td>
<td>JSON</td>
<td>Data related to the dashboard configuration</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
<tr>
<td>ModifiedDatetime</td>
<td>Datetime</td>
<td>The modification datetime</td>
</tr>
<tr>
<td>Revision</td>
<td>Integer</td>
<td>Value to keep number of updates of the user settings</td>
</tr>
</tbody>
</table>

Intermediate data and XAI results

This data model is required to store data that contain intermediate results of the data analytics and the XAI component.

Table 6. Intermediate results

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>IntermediateResultId</td>
<td>Integer</td>
<td>Identifier of the intermediate result</td>
</tr>
<tr>
<td>SourceComponent</td>
<td>String</td>
<td>Reference to the component that stores the data</td>
</tr>
<tr>
<td>Data</td>
<td>JSON/TEXT</td>
<td>Field that allows storing data in any textual or structured format</td>
</tr>
<tr>
<td>CreateByUserId</td>
<td>Integer</td>
<td>The identifier of the user that created the workflow via the dashboard</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
<tr>
<td>ModifiedDatetime</td>
<td>Datetime</td>
<td>The modification datetime</td>
</tr>
<tr>
<td>Revision</td>
<td>Integer</td>
<td>Value to keep number of updates of the results</td>
</tr>
</tbody>
</table>
Notifications
This data model is storing the notifications that are generated by the platform, via the DSS algorithms.

*Table 7. Notifications*

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>NotificationId</td>
<td>Integer</td>
<td>Identifier of the notification</td>
</tr>
<tr>
<td>Source</td>
<td>String</td>
<td>The component that triggered the notification</td>
</tr>
<tr>
<td>Type</td>
<td>String</td>
<td>Type of the notification</td>
</tr>
<tr>
<td>Content</td>
<td>JSON/TEXT</td>
<td>The actual content of the notification</td>
</tr>
<tr>
<td>InterventionId</td>
<td>Integer</td>
<td>Reference to the intervention that relates to the notification</td>
</tr>
<tr>
<td>TargetUserId</td>
<td>Integer</td>
<td>The identifier of the user to whom the notification is being sent</td>
</tr>
<tr>
<td>ScheduledDatetime</td>
<td>Datetime</td>
<td>Indicates the datetime the notification will be transmitted to the user</td>
</tr>
<tr>
<td>TransmissionDatetime</td>
<td>Datetime</td>
<td>Indicates the datetime the notification was sent</td>
</tr>
<tr>
<td>SendStatus</td>
<td>String</td>
<td>The status of the transmission (e.g. pending, sent, failed, etc)</td>
</tr>
<tr>
<td>AcknowledgeStatus</td>
<td>String</td>
<td>Indicates if the user acknowledged the notification</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime of the notification</td>
</tr>
</tbody>
</table>

Interventions
The intervention data model, stores the interventions generated by the DSS, which subsequently generate the notifications.

*Table 8. Interventions*

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>InterventionId</td>
<td>Integer</td>
<td>Identifier of the intervention</td>
</tr>
<tr>
<td>TargetUserId</td>
<td>Integer</td>
<td>The identifier of the user to whom the notification is being sent</td>
</tr>
</tbody>
</table>
### Condition
| String | The condition of the user for whom the intervention is created |

### ExecutionDatetime
| Datetime | The scheduled time to execute the workflow |

### CreateByUserId
| Integer | The identifier of the user that created the workflow via the dashboard |

### CreatedDatetime
| Datetime | The creation datetime |

### ModifiedDatetime
| Datetime | The modification datetime |

### Revision
| Integer | Value to keep number of updates of the workflow configuration |

### Alerts
This data model is used to store the alerts that are generated by the DSS in relation to the measurements and the personalised interventions. The alerts allow the timely notification of the clinical team and/or the significant others.

**Table 9. Alerts**

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>AlertId</td>
<td>Integer</td>
<td>Identifier of the alert</td>
</tr>
<tr>
<td>Source</td>
<td>String</td>
<td>The component that triggered the alert</td>
</tr>
<tr>
<td>Severity</td>
<td>String</td>
<td>The severity of the alert</td>
</tr>
<tr>
<td>Content</td>
<td>JSON/TEXT</td>
<td>Details on the alert</td>
</tr>
<tr>
<td>InterventionId</td>
<td>Integer</td>
<td>Reference to the intervention that relates to the alert</td>
</tr>
<tr>
<td>TargetUserId</td>
<td>Integer</td>
<td>The identifier of the user to whom the notification is being sent</td>
</tr>
<tr>
<td>ScheduledDatetime</td>
<td>Datetime</td>
<td>Indicates the datetime the notification will be transmitted to the user</td>
</tr>
<tr>
<td>TransmissionDatetime</td>
<td>Datetime</td>
<td>Indicates the datetime the notification was sent</td>
</tr>
<tr>
<td>SendStatus</td>
<td>String</td>
<td>The status of the transmission (e.g. pending, sent, failed, etc)</td>
</tr>
<tr>
<td>AcknowledgeStatus</td>
<td>String</td>
<td>Indicates if the user acknowledged the notification</td>
</tr>
</tbody>
</table>
Data from device vendors and synergies

This data model is required in order to store the data that will be transmitted by the device vendors and the synergies. Any medical data that does not comply with the FHIR data model cannot be stored in the FHIR database. This intermediate table allows to be used as a temporary storage for that data.

Table 10. External data

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ExternalDataId</td>
<td>Integer</td>
<td>Identifier of the external data</td>
</tr>
<tr>
<td>Content</td>
<td>JSON/TEXT</td>
<td>The actual data that are received</td>
</tr>
<tr>
<td>UserId</td>
<td>Integer</td>
<td>The identifier of the user who relates to the data</td>
</tr>
<tr>
<td>Source</td>
<td>String</td>
<td>The source that transmitted the data</td>
</tr>
<tr>
<td>Metadata</td>
<td>String</td>
<td>Additional data required to perform calculations or transformations</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
</tbody>
</table>

Patient devices

This data model holds a reference to the devices that have been assigned to the users of SMART BEAR.

Table 11. Device references

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeviceId</td>
<td>Integer</td>
<td>Identifier of the device</td>
</tr>
<tr>
<td>UserId</td>
<td>Integer</td>
<td>The patient identifier</td>
</tr>
<tr>
<td>Type</td>
<td>String</td>
<td>The type of the device (smartwatch, smartphone, etc)</td>
</tr>
<tr>
<td>StudyGroup</td>
<td>String</td>
<td>Indicator of the study group of the user</td>
</tr>
<tr>
<td>SerialNumber</td>
<td>String</td>
<td>Details of the device</td>
</tr>
<tr>
<td>Comments</td>
<td>TEXT</td>
<td>Additional details to be stored via the Dashboard</td>
</tr>
</tbody>
</table>
Patient device status
This data model holds a reference to the status of the devices that have been assigned to the users of SMART BEAR.

Table 12. Device status

<table>
<thead>
<tr>
<th>Property</th>
<th>Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>DeviceStatusId</td>
<td>Integer</td>
<td>Identifier of the device status</td>
</tr>
<tr>
<td>DeviceID</td>
<td>Integer</td>
<td>Reference to the device identifier</td>
</tr>
<tr>
<td>Status</td>
<td>String</td>
<td>The status of the device (online, offline, disabled, etc)</td>
</tr>
<tr>
<td>CreatedDatetime</td>
<td>Datetime</td>
<td>The creation datetime</td>
</tr>
<tr>
<td>ModifiedDatetime</td>
<td>Datetime</td>
<td>The modification datetime</td>
</tr>
</tbody>
</table>

2.2 Data Model Specification compliant with FHIR

2.2.1 Introduction

2.2.1.1 Overview on FHIR and available tools

FHIR is a standard created by HL7 International for exchanging digital health information by exploiting the best characteristics of HL7 Version 2 and Version 3, although it is not compatible with them. The standard is spreading rapidly in the international context and the most important international organizations that provide solutions to specific problems in healthcare, like Integrating the Healthcare Enterprise (IHE) and Personal Connected Health Alliance (PCHA), are updating their technical specifications to include FHIR.

FHIR defines the key health entities as resources, which are a collection of information models specifying data elements, constraints and relationships for the “business objects”. The last stable FHIR specification published is Release 4, version 4.01, which is based on the FHIR Composition Framework that includes 145 different resource types covering many clinical and administrative concepts of the healthcare sector. Each resource type, like Patient, Observation, and Device, permits to represent a number of properties related to a specific concept. In order to facilitate a homogeneous representation of the data into these resources, the FHIR specification provides a concrete formalization of them in three different data formats: eXtensible Markup
Language (XML), JavaScript Notation Object (JSON), and Turtle (TTL). In FHIR, many elements are used to memorize terminologies represented by coded values, like the element code of the Observation resource. Some of these codes are directly defined by the FHIR specification, while other codes may be defined elsewhere, like in external terminology systems.

Another important component of the FHIR specification is represented by the RESTful Application Programming Interfaces (APIs), which are a collection of well-defined interfaces for making different applications able to interoperate. Their use is useful to use shared communication protocols via Hypertext Transfer Protocol (HTTP), even if not mandatory.

The FHIR specification is a “platform specification”, as it describes a large number of base resources, frameworks and APIs. However, in order to implement a FHIR-based solution for a specific subdomain of the healthcare able to take into account different regulations, requirements, etc., the FHIR specification requires further adaptations to the particular context. These adaptations include a specification of the particular standard resource elements that will be used, possible additional elements, the APIs and terminologies to use, etc. Several FHIR artefacts may be used for adapting the specification to a specific context: i) Implementation Guide (IG), which is a set of rules well documented that specify how FHIR resources have to be used to solve a particular problem; ii) Package, which is a group of related adaptations published within an IG; iii) Conformance Resource, which is a specific FHIR resource in a package that represents the adaptations in a computable way.; iv) Profile, which is a set of constraints and extensions on a single resource.

A number of frameworks and tools exist for facilitating the implementation and adaptation of solutions in compliance with the FHIR standard.

Simplifier.net\(^1\) is a web platform for publishing, collaborating and developing the FHIR specification, offering capabilities for resource management and team collaboration. The platform provides a multitude of available FHIR resources organized in projects, consisting of profiles, extensions, value sets, dictionaries, mappings, examples and more. Forge can be used for creating FHIR profiles through a Windows desktop application, in an integrated way with Simplifier.net.

ART-DÉCOR\(^2\) is an open-source tools suite that supports the creation and maintenance of standard HL7 templates, value sets, scenarios, and datasets. It therefore represents both a methodology for the creation of standard artefacts and a set of tools to simplify the operations of managing artefacts. The tool also includes the management of the templates and value sets through cloud-based federated Building Block Repositories and therefore allows to simplify the entire phase of use and reuse of published templates and value sets. ART-DECOR supports collaboration between members of the same team or between different teams through governance operations.

FHIR Shorthand (FSH)\(^3\) is a language specified by HL7 for defining the contents of the FHIR IGs. It permits to express resource profiles, value sets, examples, etc. using any text editor, enabling thus distributed, team-based development using source code control tools such as GitHub. It is based on a formal grammar defined in the ATLR parser generator. SUSHI Unshortens ShortHand Inputs (SUSHI) is a FSH compiler, which conver scripts written according to FSH language to FHIR artefacts. As shown in the next figure, SUSHI permits to compile and produce artefacts concerning profiles, extensions, value sets, code systems and instances.

---

\(^1\) https://simplifier.net/
\(^2\) https://art-decor.org/mediawiki/index.php/Main_Page
\(^3\) https://build.fhir.org/ig/HL7/fhir-shorthand/
Figure 2. SUSHI results

IG Publisher is a tool provided by the FHIR team for translating the IG content in: i) generated resources that can be included into the published guide (XML, JSON, TTL formats); ii) HTML files; iii) zip files used by implementers. The tool is used after the compilation of the FSH scripts by the SUSHI compiler.

ART-DÉCOR and FSH are free, whereas Simplifier.net has both free and paid profiles and Forge is free only for educational purposes.

2.2.1.2 Methodology used in the SMART BEAR project

The nature of the data treated in the project, in accordance with the rules of the FHIR standard, led to the necessity of using a specific IG. For this reason, an analysis of the IGs published on the FHIR registries was carried out. Among these, particular attention was paid the Personal Health Device (PHD) and International Patient Summary (IPS) IGs.

The PHD IG adapts FHIR resources to convey measurements and supporting data from PHDs to different kind of systems, like platforms for electronic medical records, clinical decision support, etc. The interest for this IG was captured considering that it is based on the Continua Design Guidelines and upon the ISO/IEEE 11073 PHD Domain Information Model (DIM). Regardless, considering that in SMART BEAR many health data are not acquired by PHDs, but by collecting patients’ answers to specific questionnaires, this IG was not considered adequate for the SMART BEAR project.

The IPS IG defines the rules for the production of a document containing the essential healthcare information about a subject of care, designed for supporting unplanned, cross border care, although it is not limited to it. Although this IG provides an important contribution to identify a minimal, specialty-agnostic, condition-independent, clinically relevant dataset for a patient, it was not considered relevant for the SMART BEAR project.

For these reasons, in compliance with the FHIR standard and in line with the choices adopted in many European projects, the approach taken for the definition of the information model for the project is to define a dedicated SMART BEAR IG by profiling a set of identified FHIR resources and individuating the terminologies from international standard code systems as well as internal value sets. The tools chosen for modelling the FHIR information model are FHIR Shorthand, SUSHI and IG publisher, considering that they are the official tools provided by HL7.

2.2.1.3 Data model mapping

This section gives the specification of the data collected, and the associated constraints and extensions to the base FHIR model. In a subsequent iteration, this specification will be formalized in a FHIR Implementation
Guide. Where applicable, relevant profiles and definitions are mentioned. Of particular relevance is the Vital Signs FHIR profile (https://www.hl7.org/fhir/observation-vitalsigns.html) which sets minimum expectations for the Observation resource to record, search and fetch the vital signs associated with a patient that include the primary vital signs plus additional measurements such as height, weight and Body Mass Index. The International Patient Summary Implementation Guide (http://hl7.org/fhir/uv/ips/index.html) also specifies definitions relevant to the data in SMART BEAR.

The rest of the section details the SMART BEAR patient information, the mapping to and constraints on FHIR resources, relevant profiles and notes.

### 2.2.1.4 Demographics

<table>
<thead>
<tr>
<th>Date of birth</th>
<th>type</th>
<th>date format</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR requirements</td>
<td>Patient.birthDate shall have a date</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>age</th>
<th>type</th>
<th>number</th>
</tr>
</thead>
</table>
| FHIR requirements | 1. One Coding in Observation.code which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code= ‘445518008’  
  2. a fixed Observation.valueQuantity.code = ‘year’ |

<table>
<thead>
<tr>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>relevant profiles</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>biological gender</th>
<th>type</th>
<th>male / female</th>
</tr>
</thead>
</table>
| FHIR requirements | 1. One Coding in Observation.code which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code= ‘365873007’  
  2. One Coding in Observation.valueCodeableConcept which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code = ‘3365873007’ |

<table>
<thead>
<tr>
<th>relevant profiles</th>
</tr>
</thead>
<tbody>
<tr>
<td>source of referral</td>
</tr>
<tr>
<td>--------------------</td>
</tr>
<tr>
<td><strong>type</strong></td>
</tr>
<tr>
<td><strong>FHIR requirements</strong></td>
</tr>
<tr>
<td><strong>Notes</strong></td>
</tr>
<tr>
<td><strong>relevant profiles</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ethnicity</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>type</strong></td>
<td>Greek, Romanian etc</td>
</tr>
</tbody>
</table>
| **FHIR requirements** | 1. One Coding in Observation.code which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code= ‘365456003’  
  2. One Coding in Observation.valueCodeableConcept which must have  
  a. coding.system=TBD  
  b. coding.code = << TBD |
| **Notes** |  |
| **relevant profiles** |  |

<table>
<thead>
<tr>
<th>education level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>type</strong></td>
<td>primary, secondary, post-graduate</td>
</tr>
</tbody>
</table>
| **FHIR requirements** | 1. One Coding in Observation.code which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code= ‘365458002’  
  2. One Coding in Observation.valueCodeableConcept which must have  
  a. coding.system=’ http://snomed.info/sct’  
  b. coding.code = << 365458002 |
| **relevant profiles** |  |

<p>| living situation |  |</p>
<table>
<thead>
<tr>
<th>type</th>
<th>alone, couple, own / home care etc</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHIREquirements</strong></td>
<td></td>
</tr>
<tr>
<td>1. One Coding in Observation.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code= ‘365508006’</td>
<td></td>
</tr>
<tr>
<td>2. One Coding in Observation.valueCodeableConcept which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code = &lt;&lt; 365508006</td>
<td></td>
</tr>
<tr>
<td><strong>relevant profiles</strong></td>
<td></td>
</tr>
</tbody>
</table>

### 2.2.1.5 Medical History

#### Diabetes or pre-diabetes

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHIREquirements</strong></td>
<td></td>
</tr>
<tr>
<td>1. One Coding in Condition.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code= ‘73211009’</td>
<td></td>
</tr>
<tr>
<td>or</td>
<td></td>
</tr>
<tr>
<td>2. One Coding in Condition.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code = ‘714628002’</td>
<td></td>
</tr>
<tr>
<td><strong>Notes</strong></td>
<td></td>
</tr>
<tr>
<td><strong>relevant profiles</strong></td>
<td></td>
</tr>
</tbody>
</table>

#### Hearing Loss

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHIREquirements</strong></td>
<td></td>
</tr>
<tr>
<td>1. One Coding in Condition.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code= ‘15188001’</td>
<td></td>
</tr>
<tr>
<td><strong>Notes</strong></td>
<td></td>
</tr>
<tr>
<td><strong>relevant profiles</strong></td>
<td></td>
</tr>
</tbody>
</table>

#### Falls over the last 12 months

<table>
<thead>
<tr>
<th>type</th>
<th>number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>FHIREquirements</strong></td>
<td></td>
</tr>
<tr>
<td>1. One Coding in Condition.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. coding.system=’ <a href="http://snomed.info/sct%E2%80%99">http://snomed.info/sct’</a></td>
<td></td>
</tr>
<tr>
<td>b. coding.code= ‘15188001’</td>
<td></td>
</tr>
<tr>
<td><strong>Notes</strong></td>
<td></td>
</tr>
<tr>
<td><strong>relevant profiles</strong></td>
<td></td>
</tr>
</tbody>
</table>
FHIR requirements

1. One Coding in Observation.code which must have
   a. `coding.system=’http://snomed.info/sct’`
   b. `coding.code=’161898004’`
2. One value in Observation.valueInteger
3. A period in Observation.effectivePeriod specifying the last 12 months

Notes

relevant profiles

Balance disorder

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR requirements</td>
<td></td>
</tr>
<tr>
<td>1. One Coding in Condition.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system conforming the “Code System” column in Table 13</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code conforming the “Code System” column in Table 13</code></td>
<td></td>
</tr>
<tr>
<td>2. One coding in Condition.bodySite in case of a corresponding value in the “bodySite” column in Table 13 with:</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system = ‘http://snomed.info/sct’</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code conforming the “bodySite” column in Table 13</code></td>
<td></td>
</tr>
<tr>
<td>Notes</td>
<td>Extend Table 13 with an appropriate coding – possibly project managed – where currently empty</td>
</tr>
<tr>
<td>relevant profiles</td>
<td></td>
</tr>
</tbody>
</table>

Table 13. Balance disorder value set

<table>
<thead>
<tr>
<th>Balance disorder</th>
<th>Code system</th>
<th>code</th>
<th>bodySite</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPPV</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>111541001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibular migraine</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>232284007</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meniere</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>13445001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPPD</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>103293001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vestibular Neuritis</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>186738001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPA lesions</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a></td>
<td>300577008</td>
<td>21362003</td>
<td></td>
</tr>
<tr>
<td>Condition</td>
<td>Code</td>
<td>FHIR requirements</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-----------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unilateral vestibular failure</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 76797004</td>
<td>279254007 Not: unilateral</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pontine lesion</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 300577008</td>
<td>49557009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cerebellar lesion</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 300577008</td>
<td>113305005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stroke</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 230690007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple sclerosis</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 24700007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vertigo</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 399153001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unsteadiness</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 271713000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dizziness</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 404640003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Motion sickness</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 37031009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oscillopsia</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 246650003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty walking in uneven surfaces</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 715015003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty walking in darkness</td>
<td>Not available in standard ontology</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual vertigo</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 103293001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drunken feeling</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 386705008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Light headedness</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 386705008</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disorientation</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 62476001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nausea</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 422587007</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tendency to fall</td>
<td><a href="http://snomed.info/sct">http://snomed.info/sct</a> 279992003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Near fall incidents</td>
<td>Not available in standard ontology</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CVD history</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>Yes/no</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Condition.code which must have</td>
<td></td>
</tr>
</tbody>
</table>
### Cognitive issues

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. One Coding in `Condition.code` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='266995000'`

### Weight Loss

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. One Coding in `Observation.code` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='107647005'`
2. One Coding in `Observation.valueCodeableConcept` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='89362005'`

### Depression or Anxiety disorder

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. One Coding in `Condition.code` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='35489007'`
   or
2. One Coding in `Condition.code` which must have
   a. `coding.system='http://snomed.info/sct'`
<table>
<thead>
<tr>
<th>Notes</th>
<th>relevant profiles</th>
</tr>
</thead>
</table>

### Other medical history

<table>
<thead>
<tr>
<th>type</th>
<th>text or ICD10 or other similar system (Conditions, accidents, surgery)</th>
</tr>
</thead>
</table>
| FHIR requirements | Conditions and accidents:  
  1. One Coding in Condition.code which must have  
     a. `coding.system=’ http://snomed.info/sct’`  
     b. `coding.code` must have a value  
  Surgery:  
  Conditions and accidents:  
  1. One Coding in Procedure.code which must have  
     a. `coding.system=’ http://snomed.info/sct’`  
     b. `coding.code` must have a value |
| Notes | |
| relevant profiles | |

### Current medication use

<table>
<thead>
<tr>
<th>type</th>
<th>Name and dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR requirements</td>
<td></td>
</tr>
</tbody>
</table>
  1. a `MedicationStatement` which must have  
     a. One `CodeableConcept` in `MedicationStatement.medicationCodeableConcept` with:  
        i. `Coding.system = ` [http://www.whocc.no/atc](http://www.whocc.no/atc)  
        ii. `Coding.code` must have a value  
     b. A `Dosage` in `MedicationStatement.dosage` |
| Notes | |
| relevant profiles | |

### MOCA questionnaire

<table>
<thead>
<tr>
<th>type</th>
<th>questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR requirements</td>
<td></td>
</tr>
</tbody>
</table>
### HUI questionnaire

<table>
<thead>
<tr>
<th>type</th>
<th>questionnaire</th>
</tr>
</thead>
</table>

### Smoking

<table>
<thead>
<tr>
<th>type</th>
<th>Units/day</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
</tr>
</thead>
</table>
| 1. One Coding in Observation.code which must have  
 a. coding.system='http://loinc.org'  
 b. coding.code= ‘72166-2’  
 2. One Coding in Observation.valueCodeableConcept where  
 a. Quantity.code from valueset  
    http://hl7.org/fhir/uv/ips/ValueSet/current-smoking-status-uv-ips  

<table>
<thead>
<tr>
<th>Notes</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>relevant profiles</th>
</tr>
</thead>
</table>
| http://hl7.org/fhir/uv/ips/StructureDefinition-Observation-tobacco-use-uv-ips.html  
 http://hl7.org/fhir/uv/ips/index.html |

### Diet supplement use

<table>
<thead>
<tr>
<th>type</th>
<th>Name and dose</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
</tr>
</thead>
</table>
| 1. a MedicationStatement which must have  
 a. One CodeableConcept in MedicationStatement.medicationCodeableConcept  
 b. A Dosage in MedicationStatement.dosage |
<table>
<thead>
<tr>
<th>relevant profiles</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Questionnaire/Scale</th>
<th>Response Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Godin leisure-time exercise questionnaire</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Rapid Geriatric Assessment</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>IADL</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Social Functioning Scale (SFS)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Pittsburgh Sleep Quality Index (PSQI)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Epworth Sleepiness Scale (ESS)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Euro Quality of Life (EQ-5D)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>The System Usability Scale (SUS)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Technology Acceptance (TAM)</td>
<td>question - answer, total score</td>
</tr>
<tr>
<td>Mobile Device Proficiency Questionnaire</td>
<td>question - answer, total score</td>
</tr>
</tbody>
</table>

2.2.1.6 Physical examination

<table>
<thead>
<tr>
<th>Body height</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
</tr>
<tr>
<td>FHIR requirements</td>
</tr>
</tbody>
</table>

Notes

relevant profiles

https://www.hl7.org/fhir/bodyheight.html
https://www.hl7.org/fhir/observation-vitalsigns.html

<table>
<thead>
<tr>
<th>Body weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
</tr>
<tr>
<td>FHIR requirements</td>
</tr>
</tbody>
</table>
### Body Mass Index (BMI)

<table>
<thead>
<tr>
<th>type</th>
<th>kg</th>
</tr>
</thead>
</table>
| FHIR要求   | 1. One Coding in Observation.code which must have  
|            | a. coding.system='http://loinc.org'  
|            | b. coding.code='39156-5'  
|            | 2. One valueQuantity in Observation.valueQuantity where  
|            | a. One numeric value invalueQuantity.value  
|            | b. a fixedvalueQuantity.system="http://unitsofmeasure.org"  
|            | c. a UCUM unit code in valueQuantity.code = ‘kg/m2’ |

### Heart Rate after subject has been lying down for 3-5 minutes

<table>
<thead>
<tr>
<th>type</th>
<th>b/min</th>
</tr>
</thead>
</table>
| FHIR要求   | 1. One Coding in Observation.code which must have  
|            | a. coding.system='http://loinc.org'  
|            | b. coding.code='68999-2'  
|            | 2. One valueQuantity in Observation.valueQuantity where  
|            | a. One numeric value invalueQuantity.value  
|            | b. a fixedvalueQuantity.system="http://unitsofmeasure.org"  
|            | c. a UCUM unit code in valueQuantity.code = ‘b/min’  
|            | 3. a Period in Observation.effectivePeriod |

### Standing Blood pressure systolic/diastolic

<table>
<thead>
<tr>
<th>type</th>
<th>mmHg/mmHg</th>
</tr>
</thead>
</table>
| FHIR要求   | 1. One Coding in Observation.code which must have  
|            | a. coding.system='http://loinc.org'  
|            | b. coding.code='85354-9' |
2. One component in `Observation.component` where
   - Code.coding.system = `http://loinc.org`
   - Code.coding.code = `'8460-8`
   - valueQuantity.system = `http://unitsofmeasure.org`
   - valueQuantity.code = `'mm[Hg]`
   - valueQuantity.value must exist
   - valueQuantity.unit must exist

3. One component in `Observation.component` where
   - Code.coding.system = `http://loinc.org`
   - Code.coding.code = `'8454-1`
   - valueQuantity.system = `http://unitsofmeasure.org`
   - valueQuantity.code = `'mm[Hg]`
   - valueQuantity.value must exist
   - valueQuantity.unit must exist

Notes

relevant profiles

https://www.hl7.org/fhir/bp.html

### Supine Blood pressure systolic/diastolic

<table>
<thead>
<tr>
<th>type</th>
<th>mmHg/mmHg</th>
</tr>
</thead>
</table>

FHIR requirements

1. One Coding in `Observation.code` which must have
   - coding.system='http://loinc.org'
   - coding.code= `85354-9`

2. One component in `Observation.component` where
   - Code.coding.system = `http://loinc.org`
   - Code.coding.code = `'8461-6`
   - valueQuantity.system = `http://unitsofmeasure.org`
   - valueQuantity.code = `'mm[Hg]`
   - valueQuantity.value must exist
   - valueQuantity.unit must exist

3. One component in `Observation.component` where
   - Code.coding.system = `http://loinc.org`
   - Code.coding.code = `'8455-8`
   - valueQuantity.system = `http://unitsofmeasure.org`
   - valueQuantity.code = `'mm[Hg]`
   - valueQuantity.value must exist
   - valueQuantity.unit must exist

Notes

relevant profiles

https://www.hl7.org/fhir/bp.html
### 2.2.1.7 Hearing Loss

#### Family history of hearing loss

<table>
<thead>
<tr>
<th>type</th>
<th>Yes/no</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. One Coding in Condition.code which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='439750006'`

**Notes**

**relevant profiles**

#### Noise exposure history

<table>
<thead>
<tr>
<th>type</th>
<th>dB and years (number)</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. One Coding in Observation.code which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code='6300007'`

2. One valueQuantity in Observation.valueQuantity where
   a. One numeric value in valueQuantity.value
   b. a fixed valueQuantity.system="http://unitofmeasure.org"
   c. a UCUM unit code in valueQuantity.code = 'dB(SPL)'

3. a Period in Observation.effectivePeriod

**Notes**

**relevant profiles**
### Otoscopy

<table>
<thead>
<tr>
<th>type</th>
<th>appearance of external and middle ear (text normal/abnormal)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Observation.code which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system=' http://snomed.info/sct'</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code='247234006'</code></td>
<td></td>
</tr>
<tr>
<td>2. One Coding in Observation.valueCodeableConcept which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system=' http://snomed.info/sct'</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code='&lt;&lt;247234006'</code></td>
<td></td>
</tr>
<tr>
<td>3. One Coding in Observation.method which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system=' http://snomed.info/sct'</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code='76517002'</code></td>
<td></td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**

### Tympanometry

<table>
<thead>
<tr>
<th>type</th>
<th>Peak pressure (daPa), Middle ear volume (cm3), Compliance (cm3)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Observation.method which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system=' http://snomed.info/sct'</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code='91573000'</code></td>
<td></td>
</tr>
</tbody>
</table>

**Notes**

**TO BE DISCUSSED**

**relevant profiles**

### Pure tone audiometry

<table>
<thead>
<tr>
<th>type</th>
<th>dBs per frequency (air conduction 250, 500, 1000, 2000, 3000, 4000, 6000, 8000 Hz and bone conduction 500, 1000, 2000, 4000 Hz)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Observation.method which must have</td>
<td></td>
</tr>
<tr>
<td>a. <code>coding.system=' http://snomed.info/sct'</code></td>
<td></td>
</tr>
<tr>
<td>b. <code>coding.code='252577008'</code></td>
<td></td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**
2.2.1.8 Mental Disorders

Questionnaires for:
- Geriatric Depression Scale (GDS)
- Hamilton Rating Scale for Depression (HAM-D)

2.2.1.9 Cognitive disorders

Questionnaires for:
- Yesavage Geriatric Depression Scale (GDS)
- Montreal Cognitive Assessment (MoCA)

2.2.1.10 Frailty

### Oxygene Saturation

<table>
<thead>
<tr>
<th>FHIR requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Observation.code which must have</td>
</tr>
<tr>
<td>a. coding.system=’<a href="http://loinc.org%E2%80%99">http://loinc.org’</a></td>
</tr>
<tr>
<td>b. coding.code= ‘2708-6’</td>
</tr>
<tr>
<td>2. One valueQuantity in Observation. valueQuantity where</td>
</tr>
<tr>
<td>a. One numeric value invalueQuantity.value</td>
</tr>
<tr>
<td>b. a fixedvalueQuantity.system=&quot;<a href="http://unitsofmeasure.org">http://unitsofmeasure.org</a>&quot;</td>
</tr>
<tr>
<td>c. a UCUM unit code in valueQuantity.code = ‘%’</td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**

https://www.hl7.org/fhir/observation-vitalsigns.html
https://www.hl7.org/fhir/oxygen saturation.html

### Body Temperature

<table>
<thead>
<tr>
<th>FHIR requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. One Coding in Observation.code which must have</td>
</tr>
<tr>
<td>a. coding.system=’<a href="http://loinc.org%E2%80%99">http://loinc.org’</a></td>
</tr>
<tr>
<td>b. coding.code= ‘8310-5 ‘</td>
</tr>
<tr>
<td>2. One valueQuantity in Observation. valueQuantity where</td>
</tr>
<tr>
<td>a. One numeric value invalueQuantity.value</td>
</tr>
<tr>
<td>b. a fixedvalueQuantity.system=&quot;<a href="http://unitsofmeasure.org">http://unitsofmeasure.org</a>&quot;</td>
</tr>
<tr>
<td>c. a UCUM unit code in valueQuantity.code = &quot;Cel&quot;, or</td>
</tr>
<tr>
<td>’[degF]’</td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**

https://www.hl7.org/fhir/observation-vitalsigns.html
https://www.hl7.org/fhir/bodytemp.html
### Ambient Temperature

<table>
<thead>
<tr>
<th>type</th>
<th>FHIR requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. One Coding in Observation.code which must have</td>
</tr>
<tr>
<td></td>
<td>a. coding.system='<a href="http://snomed.info/sct">http://snomed.info/sct</a>'</td>
</tr>
<tr>
<td></td>
<td>b. coding.code= ‘422629000’</td>
</tr>
<tr>
<td></td>
<td>2. One valueQuantity in Observation. valueQuantity where</td>
</tr>
<tr>
<td></td>
<td>a. One numeric value invalueQuantity.value</td>
</tr>
<tr>
<td></td>
<td>b. a fixedvalueQuantity.system=&quot;<a href="http://unitofmeasure.org">http://unitofmeasure.org</a>&quot;</td>
</tr>
<tr>
<td></td>
<td>c. a UCUM unit code in valueQuantity.code = ‘Cel’, or ‘[degF]’</td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**

### Steps

<table>
<thead>
<tr>
<th>type</th>
<th>FHIR requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. One Coding in Observation.code which must have</td>
</tr>
<tr>
<td></td>
<td>a. coding.system='<a href="http://loinc.org%E2%80%99">http://loinc.org’</a></td>
</tr>
<tr>
<td></td>
<td>b. coding.code= ‘55423-8’</td>
</tr>
<tr>
<td></td>
<td>2. One valueQuantity in Observation. valueQuantity where</td>
</tr>
<tr>
<td></td>
<td>a. One numeric value invalueQuantity.value</td>
</tr>
<tr>
<td></td>
<td>b. a fixedvalueQuantity.system=&quot;<a href="http://unitofmeasure.org">http://unitofmeasure.org</a>&quot;</td>
</tr>
<tr>
<td></td>
<td>c. a UCUM unit code in valueQuantity.code = ‘steps’</td>
</tr>
<tr>
<td></td>
<td>3. a date in Observation. effectiveDateTime</td>
</tr>
</tbody>
</table>

**Notes**

**relevant profiles**

https://healthedata1.github.io/mFHIR/index.html

### Distance walked

<table>
<thead>
<tr>
<th>type</th>
<th>FHIR requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1. One Coding in Observation.code which must have</td>
</tr>
<tr>
<td></td>
<td>a. coding.system=’<a href="http://loinc.org%E2%80%99">http://loinc.org’</a></td>
</tr>
<tr>
<td></td>
<td>b. coding.code= ‘41953-1’</td>
</tr>
<tr>
<td></td>
<td>2. One valueQuantity in Observation. valueQuantity where</td>
</tr>
<tr>
<td></td>
<td>a. One numeric value invalueQuantity.value</td>
</tr>
</tbody>
</table>
### Body fat

**type**

**FHIR requirements**

1. One Coding in Observation.code which must have  
   a. `coding.system='http://loinc.org'`
   b. `coding.code= '41982-0'`

2. One valueQuantity in Observation. valueQuantity where  
   a. One numeric value in valueQuantity.value  
   b. a fixedvalueQuantity.system="http://unitsofmeasure.org"  
   c. a UCUM unit code in valueQuantity.code = '%'

**Notes**

**relevant profiles**

### Sleep duration

**type**

**FHIR requirements**

1. One Coding in Observation.code which must have  
   a. `coding.system='http://loinc.org'`
   b. `coding.code= '93832-4'`

2. One valueQuantity in Observation. valueQuantity where  
   a. One numeric value in valueQuantity.value  
   b. a fixedvalueQuantity.system="http://unitsofmeasure.org"  
   c. a UCUM unit code in valueQuantity.code = 'h'

**Notes**

**relevant profiles**

### HA usage time

**type**
## FHIR requirements

1. One Coding in `DeviceUseStatement.device.type` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code= '6012004'`
2. One `Period` in `DeviceUseStatement.timingPeriod`

### Notes

[https://www.hl7.org/fhir/deviceusestatement](https://www.hl7.org/fhir/deviceusestatement) (DRAFT)

### relevant profiles

## HA usage program

<table>
<thead>
<tr>
<th>type</th>
</tr>
</thead>
</table>

### FHIR requirements

1. One Coding in `DeviceUseStatement.device.type` which must have
   a. `coding.system='http://snomed.info/sct'`
   b. `coding.code= '6012004'`
2. One `CodeableConcept` in `DeviceUseStatement.device.property.type`
3. One `CodeableConcept` in `DeviceUseStatement.device.property.valueCode`
4. A `Period` in `DeviceUseStatement.timingPeriod`

### Notes

[https://www.hl7.org/fhir/deviceusestatement](https://www.hl7.org/fhir/deviceusestatement) (DRAFT)

Usage Program is a SMART BEAR specific construct, define our own system/value for 2 and 3

### relevant profiles

## Muscle mass

<table>
<thead>
<tr>
<th>type</th>
</tr>
</thead>
</table>

### FHIR requirements

1. One Coding in `Observation.code` which must have
   a. `coding.system='http://loinc.org'`
   b. `coding.code='73964-9'`
2. One `valueQuantity` in `Observation.valueQuantity` where
   a. One numeric value in `valueQuantity.value`
   b. A fixed `valueQuantity.system='http://unitsofmeasure.org'`
   c. A UCUM unit code in `valueQuantity.code = 'kg'`

### Notes

Body muscle mass Calculated

### relevant profiles
2.2.2 Questionnaire modelling

Part of the data that is collected within SMART BEAR consist of filled questionnaires. We model these questionnaires conforming the HL7 FHIR approach:

The questionnaire templates are described using the HL7 FHIR resource Questionnaire (https://www.hl7.org/fhir/questionnaire.html) and the response using the HL& FHIR resource QuestionnaireResponse (https://www.hl7.org/fhir/questionnaireresponse.html).

The generic model is defined as follows:

<table>
<thead>
<tr>
<th>Questionnaire (template)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>Questionnaire template</td>
</tr>
<tr>
<td>FHIR requirements</td>
<td></td>
</tr>
<tr>
<td>1. A Questionnaire resource where:</td>
<td></td>
</tr>
<tr>
<td>a. url shall have a value</td>
<td></td>
</tr>
<tr>
<td>b. name shall have a value</td>
<td></td>
</tr>
<tr>
<td>c. title shall have a value</td>
<td></td>
</tr>
<tr>
<td>d. version might have a value</td>
<td></td>
</tr>
<tr>
<td>e. Recursively for each entry in item:</td>
<td></td>
</tr>
<tr>
<td>i. linkId shall have a value</td>
<td></td>
</tr>
<tr>
<td>ii. type shall have a code</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>QuestionnaireResponse</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>Filled in Questionnaire</td>
</tr>
<tr>
<td>FHIR requirements</td>
<td></td>
</tr>
<tr>
<td>1. A QuestionnaireResponse resource where:</td>
<td></td>
</tr>
<tr>
<td>a. questionnaire have a value</td>
<td></td>
</tr>
<tr>
<td>b. subject shall have a value</td>
<td></td>
</tr>
<tr>
<td>c. Recursively for each entry in item:</td>
<td></td>
</tr>
<tr>
<td>i. linkID shall have a value</td>
<td></td>
</tr>
<tr>
<td>ii. answer optionally has a value</td>
<td></td>
</tr>
</tbody>
</table>

2.2.2.1 MOCA questionnaire

<table>
<thead>
<tr>
<th>Questionnaire (template)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>MOCA Questionnaire template</td>
</tr>
</tbody>
</table>
1. A Questionnaire resource where:
   a. url = “https://www.smart-bear.eu/fhir/Questionnaire/moca1”
   b. name = “MOCA”
   c. title = “Montreal Cognitive Assessment Test Form”
   d. version = “7.1”
   e. item[0] must have
      i. linkId = 0
      ii. type = “integer”
      iii. text = “Visuospatial / executive”
      iv. answerInteger must have a value
   f. item[1] must have
      i. linkId = 1
      ii. type = “integer”
      iii. text = “Naming”
      iv. answerInteger must have a value
   g. item[2] must have
      i. linkId = 2
      ii. type = “integer”
      iii. text = “ATTENTION – read list of digits”
      iv. answerInteger must have a value
   h. item[3] must have
      i. linkId = 3
      ii. type = “integer”
      iii. text = “ATTENTION – read list of letters”
      iv. answerInteger must have a value
   i. item[4] must have
      i. linkId = 4
      ii. type = “integer”
      iii. text = “ATTENTION – serial 7 substraction”
      iv. answerInteger must have a value
   j. item[5] must have
      i. linkId = 5
      ii. type = “integer”
      iii. text = “LANGUAGE - Repeat”
      iv. answerInteger must have a value
   k. item[6] must have
      i. linkId = 6
      ii. type = “integer”
      iii. text = “LANGUAGE - fluency”
      iv. answerInteger must have a value
   l. item[7] must have
      i. linkId = 7
      ii. type = “integer”
### 2.2.2.2 phq-9 questionnaire

**Questionnaire (template)**

<table>
<thead>
<tr>
<th>type</th>
<th>phq-9 Questionnaire template</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. A Questionnaire resource where:
   - a. `url = “https://www.smart-bear.eu/fhir/Questionnaire/phq-9”`
   - b. `name = “phq-9”`
   - c. `title = “Patient Health Questionnaire”`
   - d. `item[0] must have`  
     - i. `linkId = 0`
     - ii. `type = “choice”`
     - iii. `text = “Little interest in pleasure or doing things”`
     - iv. `answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-anwser-value-set`
   - e. `item[1] must have`  
     - i. `linkId = 1`
     - ii. `type = “choice”`
     - iii. `text = “Feelign down, depressed or hopeless”`
     - iv. `answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-anwser-value-set`
   - f. `item[2] must have`
i. linkId = 2
   ii. type = “choice”
   iii. text = “Trouble falling or staying asleep, or sleeping too much”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set

g. item[3] must have
   i. linkId = 3
   ii. type = “choice”
   iii. text = “Feeling tired or having little energy”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set

h. item[4] must have
   i. linkId = 4
   ii. type = “choice”
   iii. text = “Poor appetite or overeating”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set

i. item[5] must have
   i. linkId = 5
   ii. type = “choice”
   iii. text = “Feeling bad about yourself – or that you are a failure or have let yourself or your family down”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set

j. item[6] must have
   i. linkId = 6
   ii. type = “choice”
   iii. text = “Troubles concentrating on things, such as reading the newspaper or watching television”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set

k. item[7] must have
   i. linkId = 7
   ii. type = “choice”
   iii. text = “Moving or speaking so slowly that other people could have noticed – or the opposite – being so fidgety or restless that you have been moving around a lot more than usual”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answser-value-set
l. item[8] must have
   i. linkId = 8
   ii. type = “choice”
   iii. text = “Thoughts that you would be better off dead, or of hurting yourself”
   iv. answerValueSet = https://www.smart-bear.eu/fhir/ValueSet/Moca-answer-value-set

m. item[9] must have
   i. linkId = 8
   ii. type = “integer”
   iii. text = “Total”
   iv. answerInteger must have a value

2. A Valueset where:
   a. url = “https://www.smart-bear.eu/fhir/ValueSet/Moca-answer-value-set”
   b. compose.include.concepts[0] shall have
      i. code = 0
      ii. display = “Not at all”
   c. compose.include.concepts[1] shall have
      i. code = 1
      ii. display = “Several days”
   d. compose.include.concepts[2] shall have
      i. code = 2
      ii. display = “More than half the days”
   e. compose.include.concepts[3] shall have
      i. code = 3
      ii. display = “Nearly every day”

### relevant profiles

#### 2.2.2.3 HAM-A questionnaire

<table>
<thead>
<tr>
<th>Questionnaire (template)</th>
<th>HAM-A questionnaire template</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>type</strong></td>
<td>HAM-A questionnaire template</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>FHIR requirements</th>
<th>1. A Questionnaire resource where:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>a. url = “<a href="https://www.smart-bear.eu/fhir/Questionnaire/ham-a%E2%80%9D">https://www.smart-bear.eu/fhir/Questionnaire/ham-a”</a></td>
</tr>
<tr>
<td></td>
<td>b. name = “ham-a”</td>
</tr>
<tr>
<td></td>
<td>c. title=”Hamilton Anxiety Rating Scale (HAM-A)”</td>
</tr>
<tr>
<td></td>
<td>d. item[0] must have</td>
</tr>
<tr>
<td></td>
<td>i. linkId = 0</td>
</tr>
</tbody>
</table>
ii. type = "choice"
iii. text = "Anxious mood"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

e. item[1] must have
i. linkId = 1
ii. type = "choice"
iii. text = "Tension"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

f. item[2] must have
i. linkId = 2
ii. type = "choice"
iii. text = "Fears"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/Moca-anwser-value-set

g. item[3] must have
i. linkId = 3
ii. type = "choice"
iii. text = "Insomnia"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

h. item[4] must have
i. linkId = 4
ii. type = "choice"
iii. text = "Intellectual"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

i. item[5] must have
i. linkId = 5
ii. type = "choice"
iii. text = "Depressed mood"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

j. item[6] must have
i. linkId = 6
ii. type = "choice"
iii. text = "Somatic (muscular)"
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

k. item[7] must have
i. linkId = 7
ii. type = “choice”
iii. text = “Somatic (sensory)”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

l. item[8] must have
i. linkId = 8
ii. type = “choice”
iii. text = “Cardiovascular symptoms”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

m. item[9] must have
i. linkId = 9
ii. type = “choice”
iii. text = “Respiratory symptoms”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

n. item[10] must have
i. linkId = 10
ii. type = “choice”
iii. text = “Gastrointestinal symptoms”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

o. item[11] must have
i. linkId = 11
ii. type = “choice”
iii. text = “Genitourinary symptoms”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

p. item[12] must have
i. linkId = 12
ii. type = “choice”
iii. text = “Autonomic symptoms”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

q. item[13] must have
i. linkId = 13
ii. type = “choice”
iii. text = “Thoughts that you would be better off dead, or of hurting yourself”
iv. answerValueSet = https://www.smartbear.eu/fhir/ValueSet/ham-a-anwser-value-set

r. item[14] must have
2. A Valueset where:
   b. compose.include.concepts[0] shall have
      i. code = 0
      ii. display = “Not present”
   c. compose.include.concepts[1] shall have
      i. code = 1
      ii. display = “Mild,”
   d. compose.include.concepts[2] shall have
      i. code = 2
      ii. display = “Moderate,”
   e. compose.include.concepts[3] shall have
      i. code = 3
      ii. display = “Severe”
   f. compose.include.concepts[4] shall have
      i. code = 4
      ii. display = “Very severe”

2.2.2.4 GDS-15 questionnaire

**Questionnaire (template)**

<table>
<thead>
<tr>
<th>type</th>
<th>GDS-15 questionnaire template</th>
</tr>
</thead>
</table>

**FHIR requirements**

1. A Questionnaire resource where:
   b. name = “gds-15”
   c. title = “Geriatric Depression Scale (Short Form)”
   d. item[0] must have
      i. linkId = 0
      ii. type = “integer”
      iii. text = “Are you basically satisfied with your life?”
      iv. answerInteger must have a value
   e. item[1] must have
i. linkId = 1
ii. type = "integer"
iii. text = “Have you dropped many of your activities and interests?”
iv. answerInteger must have a value

f. item[2] must have
   i. linkId = 2
   ii. type = “integer”
   iii. text = “Do you feel that your life is empty?”
   iv. answerInteger must have a value

g. item[3] must have
   i. linkId = 3
   ii. type = “integer”
   iii. text = “Do you often get bored?”
   iv. answerInteger must have a value

h. item[4] must have
   i. linkId = 4
   ii. type = “integer”
   iii. text = “Are you in good spirits most of the time?”
   iv. answerInteger must have a value

i. item[5] must have
   i. linkId = 5
   ii. type = “integer”
   iii. text = “Are you afraid that something bad is going to happen to you?”
   iv. answerInteger must have a value

j. item[6] must have
   i. linkId = 6
   ii. type = “integer”
   iii. text = “Do you feel happy most of the time?”
   iv. answerInteger must have a value

k. item[7] must have
   i. linkId = 7
   ii. type = “integer”
   iii. text = “Do you often feel helpless?”
   iv. answerInteger must have a value

l. item[8] must have
   i. linkId = 8
   ii. type = “integer”
   iii. text = “Do you prefer to stay at home, rather than going out and doing new things?”
   iv. answerInteger must have a value

m. item[9] must have
i. linkId = 9
ii. type = “integer”
iii. text = “Do you feel you have more problems with memory than most people?”
iv. answerInteger must have a value

n. item[10] must have
   i. linkId = 10
   ii. type = “integer”
   iii. text = “Do you think it is wonderful to be alive?”
   iv. answerInteger must have a value

o. item[11] must have
   i. linkId = 11
   ii. type = “integer”
   iii. text = “Do you feel pretty worthless the way you are now?”
   iv. answerInteger must have a value

p. item[12] must have
   i. linkId = 12
   ii. type = “choice”
   iii. text = “Do you feel full of energy?”
   iv. answerInteger must have a value

q. item[13] must have
   i. linkId = 13
   ii. type = “integer”
   iii. text = “Do you feel that your situation is hopeless?”
   iv. answerInteger must have a value

r. item[14] must have
   i. linkId = 14
   ii. type = “integer”
   iii. text = “Do you think that most people are better off than you are?”
   iv. answerInteger must have a value

s. item[15] must have
   i. linkId = 15
   ii. type = “integer”
   iii. text = “Total”
   iv. answerInteger must have a value
2.2.2.5 HUI questionnaire

<table>
<thead>
<tr>
<th>Questionnaire (template)</th>
<th>HUI questionnaire template</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>type</strong></td>
<td></td>
</tr>
<tr>
<td><strong>FHIR requirements</strong></td>
<td>1. A Questionnaire resource where:</td>
</tr>
<tr>
<td></td>
<td>a. url = “<a href="https://www.smartbear.eu/fhir/Questionnaire/hui3%E2%80%9D">https://www.smartbear.eu/fhir/Questionnaire/hui3”</a></td>
</tr>
<tr>
<td></td>
<td>b. name = “hui3”</td>
</tr>
<tr>
<td></td>
<td>c. title = “Health Utilities Index Mark 3 (HUI3)”</td>
</tr>
<tr>
<td></td>
<td>d. item[0] must have</td>
</tr>
<tr>
<td></td>
<td>i. linkId = 0</td>
</tr>
<tr>
<td></td>
<td>ii. type = “choice”</td>
</tr>
<tr>
<td></td>
<td>iii. text = “DEXTERITY”</td>
</tr>
<tr>
<td></td>
<td>iv. answerOption[0] must have</td>
</tr>
<tr>
<td></td>
<td>1. valueCoding.code = 1</td>
</tr>
<tr>
<td></td>
<td>2. valueCoding.display = “Full use of two hands and ten fingers.”</td>
</tr>
<tr>
<td></td>
<td>v. answerOption[1] must have</td>
</tr>
<tr>
<td></td>
<td>1. valueCoding.code = 2</td>
</tr>
<tr>
<td></td>
<td>2. valueCoding.display = “Limitations in the use of hands or fingers, but does not require special tools or help of another person.”</td>
</tr>
<tr>
<td></td>
<td>vi. answerOption[2] must have</td>
</tr>
<tr>
<td></td>
<td>1. valueCoding.code = 3</td>
</tr>
<tr>
<td></td>
<td>2. valueCoding.display = “Limitations in the use of hands or fingers, is independent with use of special tools (does not require the help of another person).”</td>
</tr>
<tr>
<td></td>
<td>vii. answerOption[3] must have</td>
</tr>
<tr>
<td></td>
<td>1. valueCoding.code = 4</td>
</tr>
<tr>
<td></td>
<td>2. valueCoding.display = “Limitations in the use of hands or fingers, requires the help of another person for some tasks (not independent even with use of special tools).”</td>
</tr>
<tr>
<td></td>
<td>viii. answerOption[4] must have</td>
</tr>
<tr>
<td></td>
<td>1. valueCoding.code = 5</td>
</tr>
</tbody>
</table>
|                         |     2. valueCoding.display = “Limitations in use of hands or fingers, requires the help of another person for most tasks (not
| relevant profiles | independent even with use of special tools.|”
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ix. answerOption[5] must have</td>
<td></td>
</tr>
<tr>
<td>1. valueCoding.code = 6</td>
<td></td>
</tr>
<tr>
<td>2. valueCoding.display = “Limitations in use of hands or fingers, requires the help of another person for all tasks (not independent even with use of special tools).”</td>
<td></td>
</tr>
</tbody>
</table>
2.2.3 FHIR artefacts produced

The specifications described in the previous sections have been translated into a formal FHIR IG using FSH and IG Publisher. The corresponding web pages have also been created. This section shows some illustrations of the web specifications produced.

![Figure 3. SMART BEAR FHIR Implementation Guide Home page](image)

*Figure 3. SMART BEAR FHIR Implementation Guide Home page*
### 3.0.1 Structures: Resource Profiles

These define constraints on FHIR resources for systems conforming to this implementation guide.

<table>
<thead>
<tr>
<th>Resource Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ambient Temperature Observation</td>
<td>Observation of the ambient temperature</td>
</tr>
<tr>
<td>Anxiety Disorder Condition</td>
<td>Conditions on anxiety disorder</td>
</tr>
<tr>
<td>Blood Pressure Standing Observation</td>
<td>Observation of the blood pressure standing</td>
</tr>
<tr>
<td>Blood Pressure Observation</td>
<td>Observation of the blood pressure</td>
</tr>
<tr>
<td>Body Fat Observation</td>
<td>Observation of the body fat</td>
</tr>
<tr>
<td>Body Height Observation</td>
<td>Observation of the body height</td>
</tr>
<tr>
<td>Body Mass Index Observation</td>
<td>Observation of the body mass index</td>
</tr>
<tr>
<td>Body Temperature Observation</td>
<td>Observation of the body temperature</td>
</tr>
<tr>
<td>Body Weight Observation</td>
<td>Observation of the body weight</td>
</tr>
<tr>
<td>Cardiovascular Disease Condition</td>
<td>Conditions on cardiovascular disease</td>
</tr>
<tr>
<td>Cognitive Issue Condition</td>
<td>Conditions on cognitive issue</td>
</tr>
<tr>
<td>Current Medication Statement</td>
<td>Statement concerning a medication being assumed by a patient</td>
</tr>
<tr>
<td>Depression Disorder Condition</td>
<td>Conditions on depression disorder</td>
</tr>
<tr>
<td>Diet Supplement Use Medication Statement</td>
<td>Statement concerning a medication assumed by a patient for diet</td>
</tr>
<tr>
<td>Distance Walked Observation</td>
<td>Observation of the distance walked</td>
</tr>
<tr>
<td>Falls Observation</td>
<td>Observation of the falls over the last 12 months</td>
</tr>
<tr>
<td>Family Hearing Loss Condition</td>
<td>Conditions on family hearing loss</td>
</tr>
<tr>
<td>GDS-15 Questionnaire</td>
<td>Geriatric Depression Scale (Short Form)</td>
</tr>
<tr>
<td>Gender Observation</td>
<td>Observation of the gender</td>
</tr>
</tbody>
</table>

*Figure 4. Resource Profiles*
Figure 5. Download page
Figure 6. Example of a JSON representation
2.3 The Clinical Data Repository

In this section, we describe the medical standardized repository and the API for performing data CRUD\(^4\) operations.

The SMART BEAR Clinical Data Repository (CDR) is based on the Health Data Hub developed by Atos. The Health Data Hub is built around the HL7 FHIR standard, structuring and disposing clinical information using this standard as specification. Therefore, the SB CDR repository stores and serves clinical information in HL7 standardized, safe and scalable way. This allows Big Data Analytics (BDA) and Clinical Decision support (CDS) developers to focus on having the algorithms or applications that best suit the SMART BEAR pilots requirements, enable them to build a common set of solutions and products smoothly connected using standardized data. Medical terminology not fully covered by FHIR will be annotated using SNOMED CT. The interoperability with some different clinical terminologies (ICD9, LOINC) used across the healthcare industry will be reached by adapting the Atos Terminology Sever (ATS). The ATS will be customized and implemented in a second phase after the finalization of the PoP and it will be ready for being used for the pilots. A RESTful API is provided together, with the Clinical Data Repository (CDR). The RESTful API is able to safely access clinical information, allowing developers to abstract from the integration with clinical information sources, and simply focus on their applications. All the clinical information is stored and served following the definition of HL7

---

\(^4\) CRUD stands for create, read, update and delete
FHIR. Standardized data is available in a highly scalable manner. That is achieved by using relational database technologies configurable for each scenario, and their indexing through information retrieval engines.

![Diagram](image1.png)

Figure 8. Scheme of how data will be stored into the SMART BEAR server through the API. The HL7 FHIR compliance data is transferred in JSON format and is stored in the SB repository.

The internal architecture of the HAPI-FHIR showing the dependencies of their subcomponents is depicted below (Figure 7.).

![Diagram](image2.png)

Figure 9. Scheme of the SMART BEAR HAPI-FHIR SERVER and the internal architecture.

The API allows smoothly interaction with the SB CDR by interacting with the repository in a more intuitive and human-readable way. Additionally, the API provides interfaces for visualizing how data is stored in HL7 FHIR and codified with SNOMED CT terminology. An initial version of the API is provided as a part of the Pilot of Pilot with some basic functionalities. The example of the POST API below shows the code for the insertion of a patient with the required associated parameters and type of data format.
The deployment of the CDR will be done using Docker Containers and compose scripts. The API is accessed through HTTPS. The puppet scripts for VM are also provided. Linux Ubuntu is the preferred OS. The minimum required infrastructure for the installation is at least 40 GB HDD and a VM with 8 GB swap 2 GB, and 2 CPUs

### 2.4 BDA Engine

The BDA engine mainly addresses the functionalities required for processing DAWs (Data Analysis Workflows) and providing/storing execution results. BDA engine exposes a set of API to compute and to get raw data and analyses. Application programs must request the computation before being able to get the data. In terms of Machine Learning, a preliminary extraction of data analytics - that will be carried on the pre-processed datasets - are going to indicate variables or combinations of variable for the feature selection approaches. All ML methods and techniques are data-driven, and the “best” method will be decided after its application as already said in Deliverable 12.1. ML techniques will be used to make predictions for example in relation to Intrinsic Capacity as it is described later in this Deliverable. Given IC is a time series, ML techniques will be used to predict the behaviour in the next months using a ML model. The BDA Engine is based on a set of sub-components that are described in Deliverable 2.2:

1. **BDA Infrastructure**: Big Data Processing infrastructure that supports analytic workflow execution and libraries to supports statistical and data mining tasks, such as Apache Spark.
2. **Task Catalogue**: The Task Catalogue handles the available tasks for managing data ingestion, data orchestration, and platform management.
3. **Management/Catalogue Backend**: management backend for the BDA Infrastructure based on Ambari and for the Task and Workflow catalogues. It is used mainly for administrative management of the BDA Engine and Catalogues management.
4. **API Module**: RESTFUL APIs for the SMART BEAR components interacting with the BDA.
5. **Analysis Transformation Tool**: Subcomponent that transforms a DAW into an executable form.

![Figure 10](image.png)

*Figure 10. The code for the POST method for patient-related data entry indicating allowed values, type of variable, meaning of the content within the whole dataset, etc.*
These components are fundamental to provide the capability to process a DAW in the SMART BEAR framework. A DAW is an ordered sequence of Data Analytic Tasks (just Tasks in the following) of the following types:

- **Data Processing task**: focused principally on data preparation like data source selection for feature reduction, data cleaning, or data type transformation.
- **Statistical Analysis task**: focused on performing statistical analysis on a given dataset like ANOVA, Breusch-Pagan Test, etc.
- **Data Mining task**: focused on more elaborated analysis (e.g., clustering, machine learning) involving supervised or unsupervised algorithms like Random Forest, K-means, etc.

The output of one DAW could be the input of a consecutive one. For instance, one DAW can implement a PCA to select features, and the consecutive one can use the output to compute classification.

### 2.4.1 Functions and tools

The BDA tools are shown in Figure 9.

#### 2.4.1.1 Data storing and management

The SMART BEAR BDA Engine adopts the Hadoop⁵, a tool for data-intensive distributed applications, based on YARN programming model and a distributed file system called Hadoop Distributed Filesystem (HDFS). The data storage part of the BDA Engine adopts HBase, Hive and Phoenix (Kalakanti et al., 2015)⁶.

#### 2.4.1.2 Data analysis and tasks

The SMART BEAR BDA provides a Task Catalogue based on a set of computation libraries to provide analytic and processing capabilities. The selected libraries isSpark⁷.

#### 2.4.1.3 Data flow management and workflow orchestration

The SMART BEAR BDA also offers Workflows Catalogue where executable workflows of analytics Tasks are stored. As orchestrator manager, we selected Oozie that permits to organize, manage and schedule workflows based on events and actions (Chuan-kai, 2014)⁸.

---


⁷ [https://spark.apache.org/](https://spark.apache.org/)

2.4.1.4 Data visualization

The SMART BEAR BDA offers to the Dashboard a visualization tool to visualize results of analytics or directly ask for a simple aggregation or query. We adopt Zeppelin\(^9\), a web-based and multi-purpose notebook that enables interactive data analytics, and Vue.JS\(^10\), a progressive framework for building user interfaces.

2.4.1.5 Platform Management

The tool adopted by SMART BEAR BDA for platform management is OpenShift\(^11\), belonging to the Apache ecosystem.

![Figure 11. BDA tools](image)

2.4.2 Analytics Orchestration

The SMART BEAR BDA provides a Task Catalogue based on a set of computation libraries to provide analytic and processing capabilities. In this section the main supported algorithms are listed here below.

<table>
<thead>
<tr>
<th>BASIC STATISTICS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Description</strong></td>
</tr>
</tbody>
</table>
| Use of statistics to quantitatively describe or summarise features of a collection of information. The supported correlation methods are currently Pearson’s and Spearman’s correlation. | • Mean, Variance  
• Correlation  
• ANOVA |

\(^9\) [https://zeppelin.apache.org/](https://zeppelin.apache.org/)  
\(^10\) [https://vuejs.org/](https://vuejs.org/)  
\(^11\) [https://www.openshift.com/](https://www.openshift.com/)
### Related technologies | Preconditions | Responses
---|---|---
Spark MLib | Data types: Vector, Matrix  
Model: N  
Baseline: N | Prediction: N  
Evaluation: N  
Data storing: Y

## MODEL TRAINING

**Description**
Variety of statistical techniques (e.g., predictive modelling, machine learning) that analyse current and historical facts to generate models of an expected behaviour.

**Supported Algorithms**
- Binary Classification
- Multiclass Classification
- Regression

## PREDICTIVE ANALYTICS

**Description**
Variety of statistical techniques (e.g., predictive modelling, machine learning) that uses models to make predictions about future or unknown events.

**Supported Algorithms**
- Binary Classification
- Multiclass Classification
- Regression
### SEGMENTATION

<table>
<thead>
<tr>
<th>Description</th>
<th>Supported Algorithms</th>
</tr>
</thead>
</table>
| Dividing a broad population into sub-groups of consumers based on some types of shared characteristics such as common needs, interests, similar lifestyles or even similar demographic profiles. | • Clustering  
• Filtering  
• Mixture Models |

<table>
<thead>
<tr>
<th>Related technologies</th>
<th>Preconditions</th>
<th>Responses</th>
</tr>
</thead>
</table>
| Spark MLib | **Data types:** Vector, Matrix  
**Model:** N  
**Baseline:** Y | **Prediction:** Y  
**Evaluation:** Y  
**Data storing:** Y |

### EVALUATION METRICS

<table>
<thead>
<tr>
<th>Description</th>
<th>Supported Algorithms</th>
</tr>
</thead>
</table>
| Predictive analytics need to be evaluated by assessing the accuracy of the results. Specific machine learning algorithms like classification, regression, clustering, have well-established metrics for performance evaluation. | • Clustering  
• Classification  
• Regression |

<table>
<thead>
<tr>
<th>Related technologies</th>
<th>Preconditions</th>
<th>Responses</th>
</tr>
</thead>
</table>
| Spark MLib | **Data types:** Vector, Matrix  
**Model:** N  
**Baseline:** Y | **Prediction:** N  
**Evaluation:** N  
**Data storing:** Y |

### FREQUENT PATTERN MINING

<table>
<thead>
<tr>
<th>Description</th>
<th>Supported Algorithms</th>
</tr>
</thead>
</table>
| Mining frequent items, itemsets, subsequences, or other substructures is usually among the first steps to analyze a large-scale dataset, which has been an active research topic in data mining for years. | • Association Rules  
• FP-growth  
• PrefixSpan |
2.5 Synthetic Data Generation

In this section we will describe the data generated by a set of software tools that will permit us to simulate the first phase of tests. In details, these synthetic data will be generated with the objective to test the main functionalities of the system. In SMART BEAR we adopt Synthea, a synthetic patient generator that permits to model the medical history of patients.

In Synthea, clinical care maps and statistics are used to construct models of disease progression and treatment in a Generic Module Framework, that encodes these models in a Synthea module as state transition machines. In other terms, modules describe a progression of states and the transitions between them. On each Synthea generation time-step, the generic framework processes states one at a time to trigger conditions, encounters, medications, and other clinical events. In Figure 10, which is taken from the Synthea documentation, a simplified example of childhood ear infections can be found. It shows the flow of a generic module. In this instance, children get ear infections at different rates based on their age, are then diagnosed at an encounter, and then are prescribed either an antibiotic or a painkiller.

![Figure 12. Graphic representation of a simple Synthea module](image)

It is possible to activate different modules, which compute state transitions (if any) for every person at every timestep in the synthetic world. Each state transition in a module can trigger other events, like e.g., condition

---

12 [Home · synthetichealth/synthea Wiki · GitHub](https://github.com/synthetichealth/synthea/wiki)

onset, encounters with physicians, observations, prescriptions and so on. A list of the official Synthea modules can be found at Module Gallery · synthetichealth/synthea Wiki · GitHub.

The main Synthea concepts adopted in the preliminary phase of the SMART BEAR testing are: Observation and Condition Onset / End states, the synthetic patient personal data and the resulting EHRs.

Measures, like vital signs, laboratory tests, exams and so on, are recorded in the Observation states, which may only be processed during an Encounter. Observations are recorded along with their relevant codes, normally taken from the LOINC code system. The generated data may also include Conditions start and end to record the time periods when a patient suffers from specific diseases or is in a certain condition (for instance: Frailty). Conditions are normally specified through their codes in the SNOMED system.

The measures that will be generated include all the data considered in SMART BEAR information model variables list, that includes: measures coming from the devices, the App and the initial and follow up visits. This implies that we are going to write a new Synthea module to fully comply to the SMART BEAR model. Details will be tuned following the evolution of the specifications for the measures, the relevant LOINC / SNOMED codes to be used, the units of measure and value types and the frequency of measurement that can be found in the device measurement specification documents.

The data is generated according to the measurement frequency: e.g., 6 months for the measurements taken during the visits, 1 week for body weight, twice a day for blood pressure, etc.

The synthetic patient population is generated using a set of probabilities to get a mixture of conditions corresponding to the relevant scenarios. The probabilities are adjustable, so that, if needed, specific sets of synthetic data can be generated to test each scenario.

Ranges for each measure are specific for each patient’s condition; values within the ranges are randomly generated by Synthea at each observation time, according to its standard behaviour.

The general flow of the Synthea module used to generate data for SMART BEAR is:

- Patients in the required age range enter the module main loop, with duration 6 months, where the visits are performed three times (initial visit, 6 months and 12 months follow up). The required measures are recorded. Furthermore, some conditions onset is performed, e.g., High Frailty (with proper SNOMED code) is set in a percentage of cases. In follow up visits some conditions can be ended in a percentage of cases and new conditions can be set

- An inner loop is then entered for measures with weekly frequency; ranges are set according to the patient’s conditions (as defined in the visits) and personal characteristics when relevant. For instance, different ranges are set for weight, depending on the Frailty condition (and other) and sex

- Other inner loops for daily measures (e.g., walked distance, etc.) and for measure taken more than daily (e.g., Blood Pressure) are provided

Figure 11 illustrates a simplified module to record the SARC-F and FRAIL scores for Frailty and Sarcopenia during the first and six months follow up visits, the muscle mass value once a week, the walked distance and steps daily.
Figure 13. A simplified module to generate data for SMART BEAR
The basic Synthea processing generates values randomly within the predefined ranges. In some cases, this might yield unrealistic data for two reasons: first, frequently repeated measures of the same variable for the same individual cannot randomly oscillate in the whole range considered, even when the patient’s condition is taken into account. For instance, each weekly measure of the body muscle mass cannot differ from the previous one by more than a certain percentage. In Figure 12 an example of the problem can be found: the lines represent weekly measured muscle mass values in kg for four individuals computed with the simplified module of Figure 11; it can clearly be seen that often the difference between consecutive measures is too high.

![Muscle Mass, healthy, green=Males, blue=Females](image)

*Figure 14. An example of weekly muscle mass measures generated by the simplified module for 4 patients*

Moreover, some values should be correlated, for instance the walked distance and the number of steps. In Figure 13 an example of the problem can be found: the lines represent daily measured walked distance in meters and number of steps for an individual, computed with the simplified module of Figure 11; it can clearly be seen that no correlation is found between the data.
The module that will be used to generate synthetic data for SMART BEAR takes into account the above points:

- For values that cannot freely oscillate in the whole range, once the first value (randomly taken from the admissible range) is recorded, the other values are obtained applying random increments from a smaller range to the first value.
- Correlated measures are derived using simple functions from a measure chosen as fundamental, when the correlation does not require specific assumptions or dedicated models, as in case of walked distance and number of steps. In the other cases measures are considered as not correlated.

Synthea can export patients’ data in several formats, namely:

- **FHIR**: Fast Healthcare Interoperability Resources (FHIR), versions 4.0.1 (R4), 3.0.1 (STU3) and 1.0.2 (DSTU2).
- **C-CDA**: uses the MDHT CDA Tools library along with templates from the health-data-standards Ruby gem to export patients as Consolidated Clinical Document Architecture (C-CDA) format. C-CDA is an XML-based standard defined by HL7, that uses templates from a standard library to represent clinical concepts.
- **Text**: This format does not adhere to any standards but is clear and easy for a person to read and understand.
- **CSV**: Unlike other formats which export a single record per patient, this format generates 9 total files, and adds lines to each based on the clinical events for each patient. These files are intended to be analogous to database tables, with the patient UUID being a foreign key. Files include: patients.csv, encounters.csv, allergies.csv, medications.csv, conditions.csv, careplans.csv, observations.csv, procedures.csv, and immunizations.csv.
Some examples of the data can be found in the Appendix. Synthetic data is an alternative assistant in the rapid development and validation of new tools motivated in healthcare studies. The high-quality data must contain all of the correct possibly non-linear and multivariate dependencies that are clear in the real data set distributions. At the same time, it should also preserve the patient privacy and diminish the risk of identification. In this sense the key issue in generating realistic data is preserving relationships, distributions, predictive capabilities and patients’ privacy. Furthermore, in order to validate the model, robust models are needed to ensure biases, overfitting and high variances.

Even though Synthea is a comprehensive tool considering high number of real data specifications’ complications, there are some deficiencies in the generated data using the current version of Synthea. Synthea uses the agent-based modelling in which the coarse-grained probabilities of configurations are taking into account, while changes in these probabilities that may happen due to some changes in lifestyle (accidental or on purpose like in the presence of wearable alarm-equipped devices), for individuals is out of scope of their study. Furthermore, non-coherent data and high variations are some of other problems which may lead to generation of unrealistic data. In this sense some modifications are needed in order to get more reliable data or adopt one of the already existing methods that are examined for generating more realistic synthetic data. One of the simplest approaches is data perturbation by cropping, rotating and injecting noise to the real data sets (Zhang et al., 2017)\(^\text{14}\). Another approach is generating models of data. In this case the correct correlations and distributions based on expert knowledge or existing real data using Bayesian Network (BNs) or neural network, are inferred (Young et al., 2009)\(^\text{15}\). According to these relationships, the correlated states probabilities change non-linearly as a conditional of function of its parents.

\[ P(\text{SBP})=P(\text{SBP}|\text{age})\times P(\text{SBP}|\text{smoking})\times P(\text{SBP}|\text{BMI})\times P(\text{SBP}|\text{HR}) \]


Tucker et al. (2020)\textsuperscript{16}, depicts the integration of probabilistic graphical models with latent variables and resampling to simultaneously capture many features of real-world complex primary care data, including missing data, non-linear relationships, and uncertainty, while focussing on the importance of transparency of the modelling and data generation process.

In accordance with the context of SMART BEAR project in data usability and integrity, a challenge still persist on importing the produced data format (mostly in the csv) to acceptable format by FHIR server (LOINC and SNOMED-CT).

2.6 Analytics

This section describes the range of approaches to analytics and AI leveraged in the various clinical scenarios of SMART BEAR.

2.6.1 Analytics for Hearing Loss

This section concerns the clinical use case of the Hearing Loss scenario in the SMART BEAR project.

2.6.1.1 Interventions

There are several parameters to be improved on in this clinical use case: 1) achieve more hours of usage, 2) higher patient satisfaction with the Hearing Aid (HA) usage, 3) lower the risk of drop out, and 4) fewer visits to the Audiologists’ office for HA fine tuning (Katrakazas, et al., 2017).

This section reports on the analysis on HA usage, which aims to improve the aforementioned parameter 1) – achieving more hours of usage. To improve this parameter, the strategies or policies could include alerts (e.g., notifying participants when the HAs are inactive for a period of time), enablers (e.g., auditory training programmes) (Katrakazas, et al., 2017), and providing clinical practice guidelines based on the identified variables that influence HA usage (Katrakazas, et al., 2019).

2.6.1.2 Data

For the analysis of HA usage, relevant data is extracted from the data repository of the Evotion project (Basdekis, et al., 2017). Extracted data contains pseudo-IDs for the patients, HA IDs to distinguish left or right ear, both temporal and geographical variables, as well as other variables relating to the HA settings (e.g., its program and volume) and the environmental factors (e.g., quiet, or noisy) when the HA is active at that particular time and space. Each variable in the extracted data with its data type and remark are shown in Table 14.

Table 14. Extracted Data Dictionary

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Data Type</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>PATIENT_ID</td>
<td>Varchar</td>
<td>Pseudo number of patients</td>
</tr>
<tr>
<td>RECORD_DATE</td>
<td>Timestamp</td>
<td>Date and time when HA are active recorded by the Evotion devices</td>
</tr>
<tr>
<td>TYPE</td>
<td>Smallint</td>
<td>Indicates if the data are linked to an episode or not (values 1 and 2 accordingly)</td>
</tr>
<tr>
<td>HEARING_AID_ID</td>
<td>Varchar</td>
<td>Distinguishes if the HA is worn on the left or right ear</td>
</tr>
<tr>
<td>ENVIRONMENT_CLASSIFICATION_UNIT</td>
<td>Smallint</td>
<td>Sound environment parameter (0=quiet, 1=noise, 2=speech, 3=speech in noise)</td>
</tr>
<tr>
<td>HA_PROG</td>
<td>Smallint</td>
<td>Hearing aid program</td>
</tr>
<tr>
<td>HA_VOL</td>
<td>Smallint</td>
<td>Hearing aid volume</td>
</tr>
<tr>
<td>SOUND_PARAMETERS</td>
<td>Integer[]</td>
<td>There are 20 acoustic parameters and split into five sound characteristics. Namely momentary sound pressure level (SPL), signal-to-noise ratio (SNR), noise floor (Nf), modulation index (MI), and modulation envelope (ME). Each characteristic is measured in four frequency bands: 0-1.3kHz, 1.3-4.1kHz, 4.1-10kHz, and 0-10kHz.</td>
</tr>
</tbody>
</table>

Due to the fact there is an enormous amount of data points in the extracted data, an initial analysis is carried out using a sampled data of one patient only. This sampled data contains 222,965 data points and is ensured to contain a sufficient number of varied temporal variables, which range from 10/26/2018 to 09/07/2019. This analysis can then be scaled up for all patients in the extracted data. Similarly, the methodology can be adapted to the SMART BEAR dataset in the future.

2.6.1.3 AI Model and Approach

The proposed model for analysing HA usage is the Time Series Hidden Markov Model (HMM), since flexible and general-purposed models for univariate and multivariate time series can be provided with HMM (Zucchini, MacDonald, & Langrock, 2017). There is an underlying and unobserved state in an HMM of which it determines the distribution of an observation at a given time (MacDonald & Zucchini, 1997).
Pre-processing:

Our goal here is to transform the data into a form that maps more closely to what will be available in SMART BEAR, i.e., the daily usage of HAs per patient, using the almost continuous measurements from the devices that were available in the Evotion project data that we have used to test our approach.

Firstly, the Usage Intervals, $u_t$, need to be computed. To formally define $u_t$, we first define the term Distance, $D$, to be the difference in seconds between two consecutive timestamps, $u_i$ and $u_{i+1}$, of when the measurements are taken, such that $D = u_{i+1} - u_i$. Furthermore, the Maximum Distance, $D_{\text{max}}$, is specifically set to equal 10 minutes (or 600 seconds) in this initial analysis. Therefore, two consecutive measurements $m_i$ taken at $u_i$ and $m_{i+1}$ taken at $u_{i+1}$ belong to the same interval, $u_t$, if and only if the Distance between the two timestamps is less than or equal to 600 seconds, i.e., $|D| \leq D_{\text{max}}$, and $m_{i+1}$ belongs to the subsequent interval, $u_{t+1}$, if the Distance is more than 600 seconds, such that $|D| > D_{\text{max}}$. Hence, to determine whether a measurement with timestamp $u_i$ belongs to an interval $u_j$ we have:

$$\text{inInterval}(t_i, u_j) = (i = 0 \land j = 0) \lor (\exists k . t_k \in u_i \land |t_i - t_k| \leq D_{\text{max}}) (1)$$

The Usage Interval Duration, $d_i$, is then computed by taking $t_e - t_s$ of each $u_t$, where $t_e$ is the last timestamp of $u_t$ and $t_s$ is the first timestamp of $u_t$. Finally, the HA usage, $h_t$, per day, week, or month, can be computed by taking $d_1 + d_2 + \cdots + d_n$ for all the intervals in the respective period. For clarity, Table 15 shows a snippet of the sampled data and how each component is computed.

### Table 15. A Snippet of the Sampled Data and the Computation of $D_i$, $u_t$, $d_i$, and HAU.

<table>
<thead>
<tr>
<th>'RECORD_DATE'</th>
<th>Time-stamp</th>
<th>Distance</th>
<th>Usage Interval</th>
<th>Duration</th>
<th>HA Usage for 26/10/2018</th>
</tr>
</thead>
<tbody>
<tr>
<td>'2018-10-26 11:18:10.580'</td>
<td>$t_1$</td>
<td>-</td>
<td>$u_1$</td>
<td>$d_1 = t_5 - t_1 = 660s$</td>
<td>$h_1 = d_1 + d_2 = 660s + 60s = 720s$</td>
</tr>
<tr>
<td>'2018-10-26 11:19:11.139'</td>
<td>$t_2$</td>
<td>$t_2 - t_1 = 60s$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'2018-10-26 11:20:10.476'</td>
<td>$t_3$</td>
<td>$t_3 - t_2 = 60s$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'2018-10-26 11:21:10.475'</td>
<td>$t_4$</td>
<td>$t_4 - t_3 = 60s$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'2018-10-26 11:29:10.474'</td>
<td>$t_5$</td>
<td>$t_5 - t_4 = 480s$</td>
<td>$u_2$</td>
<td>$d_2 = t_7 - t_6 = 60s$</td>
<td></td>
</tr>
<tr>
<td>'2018-10-26 11:44:10.481'</td>
<td>$t_6$</td>
<td>$t_6 - t_5 = 900s$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>'2018-10-26 11:45:10.518'</td>
<td>$t_7$</td>
<td>$t_7 - t_6 = 60s$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Main processing – HMMs:

The elements that characterise a Hidden Markov Model (HMM) are summarised in Table 16 (Rabiner, 1989).

### Table 16. Elements of an HMM
<table>
<thead>
<tr>
<th>Elements</th>
<th>Notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Hidden States</td>
<td>$N$</td>
</tr>
<tr>
<td>Number of Distinct Observation Symbols</td>
<td>$M$</td>
</tr>
<tr>
<td>Number of Observations in a Sequence</td>
<td>$T$</td>
</tr>
<tr>
<td>Hidden States</td>
<td>$Q = {q_t, \quad 1 \leq t \leq N}$</td>
</tr>
<tr>
<td>Observation Sequence</td>
<td>$O = {o_t, \quad 1 \leq t \leq T}$</td>
</tr>
<tr>
<td>State Transitional Probability</td>
<td>$A = {a_{ij}, a_{ij} = P[q_{t+1} = S_j</td>
</tr>
<tr>
<td>Emission Probability</td>
<td>$B = {b_i(k), b_i(k) = P[v_k at \ t</td>
</tr>
<tr>
<td>Initial State Distribution</td>
<td>$\pi = {\pi_i, \pi_i = P[q_1 = S_i], \quad 1 \leq i \leq N}$</td>
</tr>
</tbody>
</table>

Therefore, the compact notation of the HMM that indicates the complete set of parameters is (Rabiner, 1989):

$$\lambda = (A, B, \pi).$$

More specifically, the state transitional probability, $A$, represents the probability of moving from one state to another, the emission probabilities, $B$, is a sequence of observation likelihoods, and the initial state distribution, $\pi$, represents the probability that the Markov chain will start in some particular state.

There are three fundamental problems that can be solved given such an HMM (Rabiner, 1989):

**Problem 1 (Likelihood):** Given the observation sequence, $O$, and the model $\lambda = (A, B, \pi)$, determine the likelihood $P(O|\lambda)$ efficiently.

**Problem 2 (Decoding):** Given the observation sequence, $O$, and the model $\lambda = (A, B, \pi)$, discover the best hidden state sequence, $Q$.

**Problem 3 (Learning):** Given the observation sequence, $O$, adjust the parameters $\lambda = (A, B, \pi)$ to maximise $P(O|\lambda)$.

As an initial analysis, we will attempt to use an HMM to predict the future HA usage. The observation sequence in this analysis is the observed HA usage, where each observation, $o_t$, is the sum of all Usage Interval Durations in the $i^{th}$ period of time (here day), such that $o_t = \sum d_j$. Given the observation sequence, the parameters $\lambda = (A, B, \pi)$ can now be learned and adjusted by solving the aforementioned Problem 3. Although there is no known optimal algorithm of estimating the model parameters given any finite number of observation sequence (Rabiner, 1989), we can however choose the parameters that maximises $P(O|\lambda)$ locally using an iterative approach. One of the most popular of such iterative approaches is the Baum-Welch algorithm (Baum, 1972). The Baum-Welch algorithm is a special case of the Expectation-Maximisation (EM) algorithm and the Forward and Backward algorithms are applied at the Expectation step.

The number of hidden states heavily influence the performance of an HMM. Therefore, the performance of the HMM with different number of hidden states is evaluated using four common criteria (Nguyen, 2018): the
Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Hannah-Quinn Information Criterion (HQIC), and the Bozdogan Consistent Akaike Information Criterion (CAIC).

Each criterion is calculated as follows:

\[
AIC = -2 \ln(L) + 2k \quad (2)
\]
\[
BIC = -2 \ln(L) + k \ln(M) \quad (3)
\]
\[
HQIC = -2 \ln(L) + k \ln(\ln M)) \quad (4)
\]
\[
CAIC = -2 \ln(L) + k (\ln(M) + 1) \quad (5)
\]

where \(k\) is the number of estimated parameters in the model, \(M\) is the number of distinct observation symbols, and \(L\) is the likelihood function of the model which can be computed using the Forward algorithm and the trained parameters calculated by the Baum-Welch algorithm.

Many methods have been proposed over the years for predicting the observation at time \(T + 1\) with HMM. The most popular one adopted for Time Series Analysis is by employing the Forward algorithm, which is to calculate the likelihood function \(P[o_{t+1}, O = o_1, o_2, ..., o_t| \lambda]\) for each of the possible \(o_{t+1}\) from the known observation sequence \(O = o_1, o_2, ..., o_t\), and the \(o_{t+1}\) that gives the maximum likelihood can be estimated as the future observation. Therefore, we will also adapt this method for this initial analysis, and the predictions will be evaluated using out-of-sampled R-Squared following the work of (Nguyen, 2018).

2.6.2 Developing AI Models with Federated and Continuous Learning for Personalization in CVD, Cognitive Disorders and Frailty

This section focuses on the implementation of a federated and continuous learning approach to support interventions that are personalized to the needs of individual patients. We apply this approach to the CVD, cognitive disorders and frailty scenarios.

2.6.2.1 Federated and Continuous Learning Framework

Federated learning\(^{17}\) is a machine learning technique that can train an algorithm across multiple systems or devices holding local datasets without the need of exchanging data. Instead of data, models and model parameters are exchanged. With this approach, healthcare organizations can better utilize their patient information and even collaborate without exchanging data. The Figure below describes the overall architecture of a cross-site Federated Learning solution.

Federated Learning is an approach that could make it easier to preserve data privacy while enjoying the benefit of large-scale aggregation and collaboration. Setting up research collaborations may require less time and lower costs. The generation of valuable new insights will be faster and less expensive as well. Enabling cross-site training and validation may lead to models that are more robust and potentially have higher performance.

\(^{17}\)https://ai.googleblog.com/2017/04/federated-learning-collaborative.html
This approach is well suited as well for personalization of models to individual users, in a network of IoT devices and smart devices such as mobile phones or dedicated edge devices. In fact, the first use of the concept has been in the context of mobile devices\textsuperscript{18}.

\begin{figure}[h!]
\centering
\includegraphics[width=\textwidth]{general_fl_approach.png}
\caption{General FL approach}
\end{figure}

Several Federated Learning frameworks exist, such as PySyft\textsuperscript{19} that enable such a learning approach. However to enable their use in the healthcare domain, the development of specific components and improvements are required, such as replacing the bi-directional communication protocol with a push approach from the local workers to the coordinator (as hospital firewalls would not allow inbound messaging from the coordinator), implementing protocols managing parameter and model updates according to the needs of the use cases, monitoring and safeguarding that sufficient performance/quality is maintained, and ensuring that local outliers that would degrade the overall performance are detected and removed.

Even when no geographic distribution of data is necessary, this architecture enables the separation of local datasets for training and supports the implementation of increased autonomy with respect to local model updates and contribution to the global model between the coordinator and the individual worker nodes. Such an approach will enable optimizing the performance of the local models, while maintaining a high performance for the global model, and will allow as well to identify, assess, and flexibly handle outliers in the local datasets.

A Federated learning solution is therefore as well the underpinning for continuous learning and customization to individual systems or devices, and for personalization to individual users. While typically AI models are trained on data from many users and their development and validation is a one-off process, a wide range of

\textsuperscript{18} https://arxiv.org/abs/1602.05629
\textsuperscript{19} https://towardsdatascience.com/federated-learning-3097547f8ca3
use cases would require models that are re-trained with data of an individual user or of a specific clinical site, and that are continuously adapted and improved. There is a need to develop approaches that tailor AI models towards a particular site to incorporate local specificities and support active learning post-deployment, or towards individual users in order to provide optimized and truly personalized user experience through updates by continuous learning AI models. Local models can learn from new data and the federated learning system can implement policies that specify when and what updates are shared between the local models and the global model.

The data management scenario in SMART BEAR is centralized, all data being collected in the repository are hosted on the cloud. Therefore, we do not leverage the privacy and distribution benefits of the federated learning approach. The key benefit is in the ability to hyper-personalize to individuals and to their data. In SMART BEAR we will develop techniques and improve the federated learning framework to address the current limitations of the technology described above, and to tailor the federated learning solution towards individual user-level personalization. Specific extensions are required to assess the performance of the local and global models, to implement validation loops to help detect any performance degradation, to identify data outliers and make decisions on their use in the system, and to design patterns of data and parameters exchange between the workers responsible for the local models specific to individual patients and the coordinator responsible for the global model.

During the pilot we aim to train local patient-specific models and global models with the data collected in the first period of the pilot (e.g. 6 months), and evaluate the specific interventions in the rest of the pilot. We aim to dedicate a worker node to each patient and carry out additional training rounds with the newly collected data during pilot operation. We will explore as well the balance between patient-specific personalization and global model performance, and assess the value added by the federated and continuous learning solution compared to a single model, not personalized to individual users.

We will also research and evaluate approaches to assess the quality of the collected data in each training rounds and the impact on the global model of the local training, by introducing a validation round to detect and mitigate performance drift.

We aim to apply this approach to develop predictive models and achieve patient-level personalization in the CVD, the Cognitive Disease, and the Frailty scenarios.

This solution facilitates as well future scenarios in SMART BEAR, ensuring scalability when data from external systems becomes available for model development without the need to be transformed and loaded into the SMART BEAR repositories, or when additional cloud resources need to be leveraged for data storage or for computation.

2.6.2.2 Models and interventions in the CVD scenario

In the CVD scenario, the models developed will focus on interventions with the potential to improve outcomes and increase adherence. The baseline interventions will be based on current clinical management. The datasets collected in the project will be used to develop models aiming to select the best management for each patient, predict risk of deterioration and of negative outcomes, predict low adherence, learn patients’ preferences in the context of the current management in order to improve adherence (e.g. predict when individuals are most likely to agree to a suggestion to exercise). The trained models will be integrated in the baseline monitoring workflows and their accuracy will be further evaluated.

Implementation of Federated Learning in this use case will require a worker node per each patient participating in the scenario, to learn from individual-level data. An example of personalization is predicting a drop in blood pressure with temperature increase, for patients that are sensitive to weather changes. As these predictions will be patient-specific, this illustrates as well the need to keep model parameters local to a worker node in specific use cases.
The following data will be collected at specific time points of the study, according to the defined workflows:

1. **Recruitment**
   - For all users, according to their medical conditions and the individualised cardiovascular estimated risk will be allocated to specific optimal range regarding all measured parameters (BP, HR, body weight, physical exercise etc.). i.e. optimal BP range SBP < 140mmHg and DBP > 80mmHg.
   - **Low – level processing**
     - Group users according to age, lipid profile, diabetes status, cardiovascular SB profile.

2. **First months of usage (i.e. 6)**
   - Using a decision tree algorithm, with combined multiple rule-based interventions as presented in different clinical scenarios, the platform will interact with the user, so that cardiovascular health will be promoted in a personalised fashion. Extra parameters to be utilised will be:
     i. BP averages (weekly fashion)
     ii. Resting HR (measured with BP monitors)
     iii. Average physical exercise (weekly/monthly)
     iv. Body weight (actual and changes)
     v. Arrhythmia events
     vi. Adherence to medications (% weekly/monthly)
     vii. Adherence to SB devices usage
   - In this stage the system will be able to integrate different resource to suggest interventions from different rule-based protocols. i.e. If the actual change in body weight is more than expected, the platform will identify the diet and exercise recommendation needed for the user.

3. **7th month and afterwards**
   - After more time of personalisation, the platform will be able to recognise potential extreme values (i.e. BP) or behaviours (i.e. adherence, physical exercise) and raise alerts and or recommendations.
     - i.e. weekly average physical exercise drops beyond expectations for the specific user, this could also raise suspicion for a new medical or psychological problem (i.e. depression – see on Geriatrics section). The platform will be able to suggest, based on the users previous individualised specific data (rather than the expected for the general population or a similar user), a medical review.
   - Regarding the weather conditions and temperature, the platform after a certain period of time (i.e. 8 months) will be able to predict the user’s response to any significant weather change. In a such way, it will be able to provide individualised recommendations promptly and timely.

The Figure below provides an overview of the relevant sources of data and information to be leveraged for the development of the interventions.
2.6.2.3 Models and interventions in the Cognitive Diseases scenario

In the cognitive diseases scenario, the role of analytics is to determine factors that are likely to influence the progression of cognitive disorder/impairment, predict relevant elements such as adherence, improvement and/or decline, and help design effective interventions.

We collect a wide range of variables previously presented in the literature as valid predictors and other variables for which a hypothesis exists that they may be relevant for our target interventions and outcomes.

We define several key objectives for the analytics work and for the designed models:

1. To help achieve as many possible hours of usage of the designated apps (compliance with the recommended interventions),
2. To assess if there are positive correlations between compliance (or higher hours of usage) with the interventions and disease progression (or lack thereof, as compared to the baseline cognitive assessment),
3. To assess the role of other parameters in influencing disease outcome (e.g. sleep quality)
4. To predict and reduce dropouts. We will develop models to better assess the patients’ satisfaction with the proposed interventions, detect the rate of compliance based on their satisfaction ratings, and predict the risk of drop out. The conclusions of these investigations and intervention will be provided to caregivers.
5. The expected outcome is that patients/users will have improved scores on the MOCA, GDS and sleep parameters (as measured by the sleep device - watch) as a result of the interventions.

We will explore a range of analytics approaches, such as logistic regression and classification to identify variables with higher predictive value and those that have less impact on model accuracy. The federated and continuous learning approach will be then applied with the selected set of relevant variables to develop models personalized to individual patients.

The following data elements will be collected in this scenario:
**Cognitive, Physical, social parameters**

- Type of cognitive disorder (categorical - amnestic / non-amnestic MCI),
- Cognitive and Frailty status (MOCA, EFS, numerical AND categorical),
- Self-assessment of mental status (BDI, GDS, HAM-D, STAI, numerical AND categorical),
- Self-assessment of quality of life (EQ-5D) (numerical),
- Physical activity - number of minutes walking per week (numerical),
- Weight change (numerical), Nutrition intake (numerical per category), Water intake (numerical),
- Hours of sleep (numerical), Number of wake-ups during the night (numerical),
- Quality of sleep/sleep architecture (numerical),
- Adherence to medication (%), Covid-19 incidence (bin),
- Hours of usage (numerical), Hours of usage per type of intervention cognitive games/physical activity app (numerical),
- Proportion of usage per type of intervention cognitive games/physical activity app (numerical values).

**Clinical parameters**

- History of psychiatric disease (categorical),
- History of brain injury/stroke (categorical),
- History of substance abuse (categorical),
- CSF biomarkers for Alzheimer’s disease (categorical),
- Cognitive assessment (numerical and categorical – Baseline and at 6 and 12 months),
- Heart Rate (numerical),
- Standing Blood pressure systolic/diastolic (baseline, numerical),
- Supine Blood pressure systolic/diastolic (at the baseline assessment, numerical),
- BP (through BP tracker, numerical),
- BMI (numerical),
- Body temperature (numerical),
- O2 saturation (numerical).

**Environmental parameters**

- Family history of cognitive disorders (categorical),
- Major life stressors e.g. recent loss of a loved one (categorical),
- Socioeconomic status (categorical).
2.6.2.4 Models and interventions in the Frailty scenario

The primary aim of the frailty scenario is to determine if compliance with the personalized intervention leads to better outcomes (as measured with Edmonton Frailty Scale). We selected a wide range of variables in order to evaluate several hypothesis, and determine and subsequently improve the factors that are likely to influence the disease progression.

We have defined the following objectives for this scenario:

1. To achieve as many possible hours of usage (compliance with the recommended interventions – Serious Games, physical activities, medication compliance, dietary adjustments, social interactions). We will design and evaluate interventions to increase compliance, and develop personalized models to predict compliance.
2. To analyze whether there are positive correlations between compliance (or higher hours of usage) with the interventions and disease progression (or lack thereof, as compared to the baseline cognitive assessment)
3. To evaluate the role of other parameters in influencing disease outcome (e.g. social interactions)
4. To reduce dropouts. We will develop models to better assess the patient satisfaction with the proposed interventions, detect the rate of compliance based on their satisfaction ratings, and predict the risk of drop out. The conclusions of these investigations and intervention will be provided to caregivers.

The analytics approaches in this scenario will be mapped to the desired objective, similarly to the cognitive disease scenario. The following data elements will be collected in the frailty scenario and used for analysis:

Cognitive, Physical, social parameters

- Cognitive and Frailty status (EFS, MoCA numerical AND categorical),
- Self-assessment of mental status (BDI, GDS, HAM-D, STAI, numerical AND categorical),
- Self-assessment of quality of life (EQ-5D) (numerical),
  - Physical activity (numerical), Weight-related data (numerical),
  - Nutrition intake (numerical per category), Water intake (numerical),
  - Hours of sleep (numerical), Number of wake-ups during the night (numerical), Quality of sleep/sleep architecture (numerical),
  - Adherence to medication (%), Covid-19 incidence (bin),
- Hours of usage (numerical), Hours of usage per type of intervention cognitive games/sleep app meditation/nutrition meal planning (numerical),
Proportion of usage per type of intervention cognitive games/sleep app meditation/nutrition meal planning (numerical values).

Clinical parameters
Cognitive assessment (numerical and categorical – Baseline and at 6 and 12 months),
Heart Rate (numerical),
Standing Blood pressure systolic/diastolic (baseline, numerical),
Supine Blood pressure systolic/diastolic (at the baseline assessment, numerical),
BP (through BP tracker, numerical),
BMI (numerical),
Body temperature (numerical),
O2 saturation (numerical).

Environmental parameters
Major life stressors, e.g. recent loss of a loved one (categorical), socioeconomic status (categorical).

Other relevant parameters
Gender (categorical),
Age (numerical),
Nationality (categorical),
Level of education (categorical),
Living situation (categorical),
Postal code (categorical),
Number of social interactions (categorical).

2.6.3 COVID-19 risk assessment
The goal of the COVID-19 scenario is to explore risk models for COVID-19 complications and hospitalization following a COVID-19 infection, and to develop interventions that aim to early detect patients that may be at risk of a COVID-19 infection and patients who are likely to have been infected with CoVID-19 (based among others on temperature, O2 saturation, and potential exposure to infected individuals) and should be tested. This scenario workflow comprises the 4 following tasks:
1. Acquiring data for the analytics
2. Running analytics to compute the risk value
3. Generating result
4. Visualizing the results
Workflow

Task Specifications

<table>
<thead>
<tr>
<th>Task ID</th>
<th>Task Name</th>
<th>Task Type</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>COVID-19 Risk Assessment Vector</td>
<td>Generate Local Vector</td>
<td>VPi</td>
</tr>
</tbody>
</table>

Storage

<table>
<thead>
<tr>
<th>API</th>
<th>RW</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R</td>
<td>Patient ID, timestamp</td>
</tr>
</tbody>
</table>

Data Schema

<table>
<thead>
<tr>
<th>Entry</th>
<th>Data Type</th>
<th>Evaluation</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID</td>
<td>ID</td>
<td>unique value</td>
<td>N</td>
</tr>
<tr>
<td>Physical activity</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Weight change</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Nutrition intake</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Hours of sleep</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Number of wake-ups during the night</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Adherence to medication</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Adherence to balance rehabilitation program</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Covid-19 incidence</td>
<td>binary</td>
<td>-</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Hospitalisation related to covid-19</td>
<td>binary</td>
<td>-</td>
<td>L.U.T</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>--------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Self-assessment of quality of life</td>
<td>numerical</td>
<td>EQ-5D</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Cognitive and Frailty status</td>
<td>numerical</td>
<td>MOCA-EFS</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Smoking history</td>
<td>numerical</td>
<td>years</td>
<td>L.U.T</td>
</tr>
<tr>
<td>OH history</td>
<td>binary</td>
<td>-</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Heart Rate</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T, baseline</td>
</tr>
<tr>
<td>Standing Blood pressure systolic/diastolic</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T, baseline</td>
</tr>
<tr>
<td>Supine Blood pressure systolic/diastolic</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T, baseline</td>
</tr>
<tr>
<td>BP</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T, baseline</td>
</tr>
<tr>
<td>BMP</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T, baseline</td>
</tr>
<tr>
<td>Number of social interactions</td>
<td>numerical</td>
<td>AVG 7dd</td>
<td>L.U.T</td>
</tr>
<tr>
<td>Potential exposure to infected individuals</td>
<td>numerical</td>
<td>Selfreported or based on tracker</td>
<td>L.U.T</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task ID</th>
<th>Task Name</th>
<th>Task Type</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>COVID-19 Risk Assessment Computation</td>
<td>Script</td>
<td><img src="image" alt="Risk Assessment" /></td>
</tr>
</tbody>
</table>

**Procedure**

<table>
<thead>
<tr>
<th>Function</th>
<th>Parameters</th>
<th>Return</th>
<th>Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>ComputeRisk</td>
<td>Vector Task1</td>
<td>Risk[L</td>
<td>M</td>
</tr>
<tr>
<td>Store</td>
<td>Risk[]</td>
<td>1</td>
<td>-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task ID</th>
<th>Task Name</th>
<th>Task Type</th>
<th>Icon</th>
</tr>
</thead>
</table>
### COVID-19 Risk Assessment Result Vector

#### Generate Local Vector

![VR(Pi)](image)

### Storage

<table>
<thead>
<tr>
<th>API</th>
<th>RW</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>W</td>
<td>Patient ID, timestamp</td>
</tr>
</tbody>
</table>

### Data Schema

<table>
<thead>
<tr>
<th>Entry</th>
<th>Data Type</th>
<th>Evaluation</th>
<th>Parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient ID</td>
<td>ID</td>
<td>unique value</td>
<td>N</td>
</tr>
<tr>
<td>Risk level</td>
<td>numerical</td>
<td>Task2 result</td>
<td>L.U.T</td>
</tr>
</tbody>
</table>

### Task ID

<table>
<thead>
<tr>
<th>Task ID</th>
<th>Task Name</th>
<th>Task Type</th>
<th>Icon</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>COVID-19 Risk Assessment Visualization</td>
<td>Script</td>
<td>Visualization Risk(Pi)</td>
</tr>
</tbody>
</table>

### Procedure

<table>
<thead>
<tr>
<th>Function</th>
<th>Parameters</th>
<th>Return</th>
<th>Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>VisualizeRisk</td>
<td>Vector Task, Vector Task3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.6.4 Intrinsic Capacity

2.6.4.1 Definition

Due to the fact that there is no consistency in definition and data acquisition method in assessing the Intrinsic Capacity (IC), the Integrated Care for Older People, ICOPE handbook Guidance on person-centred assessment and pathways in primary care\(^{20}\), that the World Health Organization has published, is considered as the reference. WHO defines Intrinsic Capacity as the combination of the individual’s physical and mental, including psychological capacities. Functional ability is the combination and interaction of intrinsic capacity with the environment a person inhabits. According to the abovementioned handbook, the Intrinsic Capacity score comprises six generic domains:

1. Vitality,
2. Locomotor capacity
3. Psychological capacity
4. Cognitive capacity
5. Hearing capacity

It is noteworthy that, there are different ontologies in the existing literature where the Hearing capacity and Visual capacity could be compressed in one domain named “Sensory capacity” and Psychological capacity is called the “Mood” capacity. Therefore, we will use the following names for the domains hereafter:

1. Vitality
2. Locomotion
3. Cognition
4. Sensory
5. Mood

2.6.4.2 Use case

This section concerns the measurement and analysis of the IC score based on data collected from SMART BEAR project’s devices and questionnaires. Furthermore, the main goal is to relate and validate the retrieved IC score as a holistic health state of the participants. In this sense, using continuous monitoring provided by SMART BEAR facilities, the early-stage diagnoses in the IC decline and subsequently the early interventions would be feasible. The SMART BEAR approach to IC is focused on accounting the levels achieved by a patient using both personalized and population-specific functions. These two aspects are accounted, respectively by the z-score and the performance score. Another important aspect considered in the SMART BEAR methodology is the validity range of the collected measurements that account for the reliability of the recorded data depending on the ageing of the specific measurements and the different depreciation dynamics of the IC domains.

2.6.4.3 Data

According to the ICOPE handbook, each generic capacity domain contains a set of the most reliable clinical tests and non-clinical measurements, which has been suggested and validated by clinician. Table 1 (Data format and collection frequency) illustrates the domains and sub-domains’ parameters respecting the

leveraged devices/questionnaires in implementing the measurements. Each of the measurements has a value in the IC score, an example of evaluating points adapted from ICOPE handbook is shown in Figure 15. We try to follow the proposed set of measurements as much as possible in the SMART BEAR pilots and keep the balance in the number of measurements for all domains. Some inconsistencies may happen due to resources limitation while there are also extra available measurements. Therefore, there is a possibility for extending the domains and the scope of IC using the SMART BEAR project potentials and evaluate the effectiveness of these extra measurements in validating the IC score.

Figure 19. The example of locomotion domain’s points from ICOPE handbook

2.6.4.4 Data format

The data measured and collected with SMART BEAR devices, mobile application, and questionnaires will be stored in HAPI FHIR repositories using LOINC and SNOMED-CT codes. Each measurement’s value type is mentioned in Table 5, while intermediate data are listed in Table 6 that are generated in the course of the successive steps towards drawing the IC trajectories as defined in the following subsections.

2.6.4.5 Frequency of collecting data by using SMART BEAR devices

In Table 1 the frequencies of sending/receiving data to/from SMART BEAR clouds are mentioned specifically for the measurements. According to the Intrinsic Capacity use case, the suitable time scale for the data retrieval is considered. Therefore, in the cases of repetitive measurements, an averaging process should be applied on data. The required analytical parameters such as Risk value, Expected value, Tolerance value (in case these two values are population-dependent, not personalized), and parameters’ weight are model-based values, while the z-score, performance score and personalized Expected and Tolerance values are agent-based and will be assigned to each studied agent independently.
2.6.4.6 Test validity

Due to different time expiration of each test in measuring IC, an assigned Boolean value will clarify whether the result of that specific test is still valid. In the case of any type of invalidity, the IC data point will get tagged as “invalid” in the output diagrams. Test validity is computed and stored with timestamp and patient id each time a measure is received.

2.6.4.7 Z-score

Considering the sub-domains measurements and their different valid ranges, the z-score normalizing transformation is proposed. This normalization will be defined using the “Expected” and “Tolerance” values. Considering these values as personalized or population-dependent, two different approaches in this step could proceed. "Z-score" is computed and stored with timestamp and patient id each time a measure is received.

2.6.4.8 Performance score and parameters’ weights in domain and sub-domain aggregation score

In this proposal, the performance score is an asymmetric mapping of measurements’ z-score. In assessing the performance score, the risk and expected values have crucial roles. The risk and expected values are generic parameters that should be defined specifically for each studied population. A high-performance score could be achieved in the z-score higher than expected value, while in the case that z-score is about the risk value, the minimum performance score will be assessed. Performance score is computed and stored with timestamp and patient id each time a measure is received. An approach in aggregating the sub-domains elements into domains and thereafter, mapping domains into the IC is the weighted aggregation. In this way, the importance of each domain and sub-domains are indicated by weighted values and the aggregation will take place using weighted arithmetic functions. It is noteworthy that, the weights could vary from one studied population to another one. Therefore, patient’s retrospective data in assessing the IC, following the ICOPE handbook, by clinician could provide the ground truth data in leveraging Machine Learning methods in predicting the population-specific parameters’ weights. A single value of parameters’ weight and risk value is stored for each type of measure.

2.6.4.9 Parameters in sub-domain and domain aggregation score

Performing the abovementioned data analysis, will result in assessing the time series of IC score during the monitoring process. This final value is in range [0,6] whilst the higher is the measured value the higher intrinsic capacity has the patient.

---

Table 17. Intrinsic Capacity data types and domains

---
<table>
<thead>
<tr>
<th>Domains</th>
<th>Parameters stored on HAPI FHIR repository</th>
<th>Data Type</th>
<th>Frequency of data transmission to SMART BEAR Cloud</th>
<th>Frequency of data reception from SMART BEAR Cloud (DSS/BDA)</th>
<th>Domain mapping weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vitality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>double</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>double/mmHG</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>double/mmHG</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heart rate</td>
<td>integer/BPM</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oxygen Saturation</td>
<td>%/BPM</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sleep duration</td>
<td>object/HOUR</td>
<td>per 3days</td>
<td>per 3days</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body weight (through the smart scale)</td>
<td>double/Kg</td>
<td>weekly</td>
<td>weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body fat (through the smart scale)</td>
<td>double</td>
<td>weekly</td>
<td>weekly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score on SNAQ (as part of RGA)</td>
<td>Questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sarcopenia (SARC-F)</td>
<td>Questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRAIL (as part of RGA)</td>
<td>Questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short MNA</td>
<td>Questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diet adherence (supported with biological data)</td>
<td>Questionnaire</td>
<td>monthly</td>
<td>Monthly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locomotion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>double</td>
</tr>
<tr>
<td>Steps</td>
<td>integer</td>
<td>daily</td>
<td>Daily</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance walked</td>
<td>meter/doubl</td>
<td>daily</td>
<td>Daily</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skeletal muscle (through the smart scale)</td>
<td>object</td>
<td>monthly</td>
<td>Monthly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cognition</td>
<td></td>
<td></td>
<td>double</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------</td>
<td>-----------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Cognitive screen (as part of the RGA)</strong></td>
<td>Questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Score on Montreal Cognitive Assessment</td>
<td>object</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sensory</th>
<th></th>
<th></th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Side of hearing loss (no HL, unilateral/bilateral)</td>
<td>object</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>Degree of Hearing loss (no HL, mild, moderate, severe, profound)</td>
<td>object</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>GHABP score (patients with HL)</td>
<td>questionnaire</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
<td>6-month (entry to the study, 6th month, 12th month)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mood</th>
<th></th>
<th></th>
<th>double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to read</td>
<td>object</td>
<td>weekly</td>
<td>weekly</td>
</tr>
<tr>
<td>Personal Health Questionnaire Depression Scale (PHQ-9)</td>
<td>object</td>
<td>biweekly</td>
<td>biweekly</td>
</tr>
<tr>
<td>HAM-A</td>
<td>questionnaire</td>
<td>biweekly</td>
<td>biweekly</td>
</tr>
<tr>
<td>Score on Geriatric Depression Scale</td>
<td>Questionnaire</td>
<td>monthly</td>
<td>monthly</td>
</tr>
</tbody>
</table>

| IC-SCORE | Double [0,6] | Every 6 month |
### Table 18. Intrinsic Capacity intermediate data

<table>
<thead>
<tr>
<th>PARAMETERS STORED ON HAPI FHIR REPOSITORY</th>
<th>parameters weight</th>
<th>test validity</th>
<th>Performance score</th>
<th>z-score</th>
<th>Risk value</th>
<th>Expected value</th>
<th>Frequency of data transmission to SMART BEAR Cloud</th>
<th>Frequency of data reception from SMART BEAR Cloud (DSS/BDA)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>VITALITY</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
</tr>
<tr>
<td>Heart rate</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
</tr>
<tr>
<td>Oxygen Saturation</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>after each usage - twice daily or else advised</td>
<td>after each usage - twice daily or else advised</td>
</tr>
<tr>
<td>Sleep duration</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>per 3days</td>
<td>per 3days</td>
</tr>
<tr>
<td>Body weight</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>weekly</td>
<td>weekly</td>
</tr>
<tr>
<td>Body fat</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>weekly</td>
<td>weekly</td>
</tr>
<tr>
<td>SNAQ (as part of RGA)</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>(entry to the study, 6th month, 12th month)</td>
<td>(entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>Sarcopenia (SARC-F)</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>(entry to the study, 6th month, 12th month)</td>
<td>(entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>FRAIL (as part of RGA)</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>(entry to the study, 6th month, 12th month)</td>
<td>(entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>Short MNA</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>(entry to the study, 6th month, 12th month)</td>
<td>(entry to the study, 6th month, 12th month)</td>
</tr>
<tr>
<td>Diet adherence (supported with biological data)</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>monthly</td>
<td>monthly</td>
</tr>
<tr>
<td><strong>LOCOMOTION</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steps</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>daily</td>
<td>daily</td>
</tr>
<tr>
<td>Distance walked</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>daily</td>
<td>daily</td>
</tr>
<tr>
<td>Skeletal muscle</td>
<td>double</td>
<td>boolean</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>double</td>
<td>monthly</td>
<td>monthly</td>
</tr>
</tbody>
</table>
### 2.6.5 Models and interventions in the Mild Depression scenario (Mental diseases)

The goal of the mild depression scenario is to explore risk models for the development of depression in senior patients, impact of mood deterioration on overall health, quality of life and sense of wellbeing of the patients and to develop interventions that aim to early detect seniors that may be at risk of depression or already entering depression, to prevent or modify factors that contribute to mood deterioration, to strengthen factors that prevent mood deprivation and to provide clinical recommendations based on the principles of tailor-made care due to the personalization of suggestions (based on the subjective data collected) as far as possible. The intervention models developed aiming to early detection of both classic and atypical clinical presentation of depression in seniors and early consultation and treatment before further mental health deterioration will be validated during the project.

<table>
<thead>
<tr>
<th>COGNITION</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cognitive screen (as part of the RGA)</td>
<td>double boolean double double double double 6-month 6-month</td>
<td></td>
</tr>
<tr>
<td>Score on Montreal Cognitive Assessment</td>
<td>double boolean double double double double 6-month 6-month</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SENSORY</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hearing loss degree</td>
<td>double boolean double double double double (entry to the study, 6th month, 12th month) (entry to the study, 6th month, 12th month)</td>
<td></td>
</tr>
<tr>
<td>Hearing loss side</td>
<td>double boolean double double double double (entry to the study, 6th month, 12th month) (entry to the study, 6th month, 12th month)</td>
<td></td>
</tr>
<tr>
<td>GHABP score</td>
<td>double boolean double double double double (entry to the study, 6th month, 12th month) (entry to the study, 6th month, 12th month)</td>
<td></td>
</tr>
<tr>
<td>Ability to read</td>
<td>double boolean double double double double weekly weekly</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>MOOD</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Health Questionnaire Depression Scale (PHQ-9)</td>
<td>double boolean double double double double monthly monthly</td>
<td></td>
</tr>
<tr>
<td>HAM-A</td>
<td>double boolean double double double double biweekly biweekly</td>
<td></td>
</tr>
<tr>
<td>Score on Geriatric Depression Scale</td>
<td>Double boolean double double double double monthly monthly</td>
<td></td>
</tr>
</tbody>
</table>
A number of variables of interest will be collected in the mental disorder scenario as described in the following paragraphs. These variables (sleeping habits (hours of sleep, number of wake-ups during the night), appetite variations, weight changes, physical activity, leave out of home, change in living situation (e.g., lives alone, lives with spouse/partner), adherence to medication treatment, diet and the study protocol itself, increase in medical consultations) were appropriate to detect a risk of mood deterioration and to define personalized interventions that are likely to improve the person's mood or avoid its deterioration:

1. Better detect the weak signals of a senior’s mood deterioration through the submission of a regular self-assessment test (PHQ-8) and the collection and variation analysis of data on factors that can impact the senior’s mood (sleep, physical activity, diet, moving out of home, social interaction).

2. Study the possibilities of reversibility of these weak signals by sending personalized incentive recommendations to the senior as soon as a drop in morale is detected (PHQ-8 score greater than 4). The purpose of these automatic notifications will be to encourage the senior to adopt behaviours that can improve his/her health and psychological well-being. This will test the senior's adaptive capacities.

3. Reduce dropouts by detecting a decrease in compliance. We will develop models to better assess the patient satisfaction with the proposed interventions (feedbacks measure), detect the rate of compliance based on their satisfaction ratings, and predict the risk of drop out.

4. The expected outcome is that patients/users will have improved scores on the PHQ-8, HAM-A, MOCA, and healthy behaviours (sleep/diet/physical activity/leave out of home - as measured by devices) as a result of the intervention:
   a. The overall average of seniors scores on the PHQ-8 questionnaire who take part in the mental health experiment decreases by 15% between the beginning and the end of experimentation, thus showing an improvement in the senior’s morale.
   b. The incidence and prevalence of depression in the SMART BEAR program is 15% lower than these rates in the general population (taking into account differences between countries)

The variables collected in the mild depression scenario are listed below.

**Physical, psychological, social parameters**

- Physical activity - number of minutes walking per week (numerical)
- Move in-out the housing (numerical)
- Weight change (numerical)
- Water intake (numerical)
- Nutrition intake (numerical per category)
- Hours of sleep (numerical)
- Number of wake-ups during the night (numerical)
- Quality of sleep/sleep architecture (numerical)
- Adherence to medication (%)  
- Adherence to balance rehabilitation program (%)  
- Covid-19 incidence (binary)
- Self-assessment of mental status (GDS, PHQ9 HAM-A) (numerical & categorical)
Self-assessment of quality of life (EQ-5D) (numerical)
Cognitive and Frailty status (MOCA, RGA, EFS) (numerical & categorical)
Self-assessment of sleep quality (PSQI) (numerical & categorical)
Smoking history (binary and/or pack years)
Type of mental disorder (categorical - mild/moderate)
Hours of smartphone usage (numerical)
Hours of usage per type of intervention physical activity app (numerical)

Clinical parameters
Heart Rate (numerical)
Standing Blood pressure systolic/diastolic (baseline, numerical)
Supine Blood pressure systolic/diastolic (at the baseline assessment, numerical)
Blood Pressure through smart blood pressure tracker (numerical)
Body Mass Index (numerical)
Body temperature (numerical)
O₂ saturation (numerical)
History of psychiatric disease (categorical)
History of brain injury/stroke (categorical)
History of substance abuse (categorical)
Depression assessment (Baseline, 6 and 12 months, numerical & categorical)
MNA (Baseline, 6 and 12 months, numerical & categorical)

Environmental parameters
Family history of mental health problems (categorical)
Ambient temperature (numerical)
Major life stressors such as recent loss of a loved one (categorical)
Socioeconomic status (categorical)

Other relevant parameters
Number of social interactions (categorical)
Postal code (categorical)
Gender (categorical)
Living situation (categorical)
Nationality (categorical)
2.6.6 Full Homomorphic Encryption of Machine Learning Models

We will leverage our Fully Homomorphic Encryption (FHE) capabilities to transform Machine Learning (ML) models from models that operate over raw data, to models that operate over encrypted data. The FHE transformed models will process an encrypted input and produce an encrypted output without the need to decrypt the data or know the encryption key. FHE enables computation over encrypted data and so allows data scientists to create models that adhere to privacy regulations while enjoying the benefits of running on the Cloud / untrusted servers. We intend to work over structured records that include risk parameters such as pre-existing conditions, as well as current health measures to provide a warning when there is a higher probability for a health event. The individual parameters as well as the results of the inference would be completely encrypted and so adhere to GDPR and other privacy regulations.

We will leverage the FHE technology in Cloud elasticity use cases, when the SMART BEAR private cloud needs to scale and use computation capabilities provided by public cloud providers. In this case the ML model will be transformed into FHE-enabled ML model and deployed on the public cloud. The SMART BEAR platform will encrypt the data and send to the public Cloud for computation. The computation will be conducted on the Cloud while the key remains in the SMART BEAR platform. The Cloud will send back the encrypted result to the SMART BEAR platform, and the platform will decrypt the results.

Another use case for this technology is when we want to leverage the SMART BEAR private cloud to perform a computation of 3rd party data. In this case the 3rd party’s ML model will be transformed into FHE-enabled ML model and deployed on the SMART BEAR cloud. The 3rd party will encrypt the data and send to SMART BEAR for computation. The computation will be conducted on SMART BEAR private cloud while the key remains in the 3rd party’s platform. SMART BEAR will send back the encrypted result to the 3rd party’s platform, and the platform will decrypt the results.

The use of this technology is an enabler for modern cloud computing architecture that is compliant with strict regulations that govern the project. Although it is important to note that FHE is still not standardized and is not part of an international security standard. Also, it is worth mentioning that FHE comes with an overhead in terms of performance which is very much use-case dependent and so should be considered.

We will apply this technology to ML models that meet the following constraints:

- Inference models (training done in the clear in a secure location), low multiplication depth (NNs with less than 20 layers), High-throughput configurations assume data is sent in large batches. Batch sizes can be in the range 4000-16000 samples, depending on other factors.
- The batch is encrypted together, sent to be processed on server, and then a batch of predictions is sent back.
- For non-encrypted network, higher performance can be achieved: larger networks and faster inference.
- Limitations on layers: Simplest and cheapest to use are fully connect layers and square activation. Other activation functions can be costly on performance and prevent depth. Max-pooling layers may be too costly and should be avoided or replaced with alternatives. Convolutional layers are supported. Decision trees and Logistic Regression are supported as well. If possible, one should avoid hard real-time constrained use cases.

We will apply this technique on two SMART BEAR use cases. We will use the technique to approximate a given, pre-trained, ML model into an FHE-enabled ML model. The FHE-enabled ML model will have a reasonable performance overhead and a reasonable accuracy degradation. The application of FHE will be only for ML inference purposes.
2.7 Decision Support System (DSS)

Summary

The DSS component of SMART BEAR platform performed a series of tasks that target all the user roles of the platform. In that sense, DSS provides a series of features for the clinician, patients and scientists. In order the DSS to provide useful outcomes, a critical mass of input data is required. The data sources are both the FHIR database and the non-FHIR database, which provide the required information to be processed and analyzed by the DSS in order to provide useful outcomes. In the following section, each of the features that are supported by the DSS are summarized, along with the relevant interfaces to communicate with the other components of SMART BEAR platform.

Initial assessment and device selection

The DSS is designed to assist the clinicians in the initial assessment of each of the patients in terms of the optimal assessments that must be performed to assess the patient and then provide him with the optimal combination of the devices to monitor his health during the pilot study. This component is designed to evolve throughout the project, as it will continuously be trained by the data that will be digested into the platform. The initial version available for the PoP has been trained by the rules and the medical guidelines that have been provided by the clinicians (D12.1) to have a ground truth system based on the most updated medical knowledge. Starting with the PoP, the data and the analytics are combined to identify the degree of satisfaction of the patients and to what degree is the assignment of the combination of devices beneficial for the patient. After the PoP, the data that will be collected will combine the assessments of the clinicians, the preferences of the patients and the available devices, in order to provide the optimal device combination for the monitoring of the patient.

Table 19. Components and interfaces

<table>
<thead>
<tr>
<th>Related Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR Database</td>
<td>✔ Access medical history and device measurements</td>
</tr>
<tr>
<td>Non-FHIR database</td>
<td>✔ Feedback on the device usage and on the interventions</td>
</tr>
<tr>
<td>Dashboard</td>
<td>✔ Interface to the clinician</td>
</tr>
<tr>
<td>Analytics</td>
<td>✔ Perform analytics based on the input data</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Interfaces</th>
<th>Request</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>/dss/assessment</td>
<td>{</td>
<td>{</td>
</tr>
<tr>
<td></td>
<td>&quot;userId&quot;: 453,</td>
<td>&quot;userId&quot;: 453,</td>
</tr>
<tr>
<td></td>
<td>&quot;conditions&quot;: [</td>
<td>&quot;assessments&quot;: [</td>
</tr>
<tr>
<td></td>
<td>&quot;cvd&quot;,</td>
<td>&quot;EQ5D&quot;,</td>
</tr>
<tr>
<td></td>
<td>&quot;frailty&quot;</td>
<td>&quot;MOCA&quot;</td>
</tr>
<tr>
<td></td>
<td>]</td>
<td>]</td>
</tr>
<tr>
<td></td>
<td>}</td>
<td>}</td>
</tr>
</tbody>
</table>
Patient and system alerts
The DSS component implements an alerting mechanism which is continuously monitoring the device statuses and the measurements of the devices. For each user, when an outlier has been detected, the rules set by the clinicians are taken into consideration and if required an alert is being generated. The alerts are accessible via the Dashboard and are also transmitted to the significant others of the patients. The DSS is transmitting the alert to the Security Component which then notifies the relevant actors of the platform. Alerts are also generated in cases such as when a monitoring device is not transmitting data to the data repository for a period of time.

Table 20. Components for alerts

<table>
<thead>
<tr>
<th>Related Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR Database</td>
<td>✓ Access device measurements</td>
</tr>
<tr>
<td>Non-FHIR database</td>
<td>✓ Device usage and status</td>
</tr>
<tr>
<td>Dashboard</td>
<td>✓ Interface for the clinicians and technicians</td>
</tr>
<tr>
<td>Security Component</td>
<td>✓ Transmit the alert</td>
</tr>
</tbody>
</table>

Interventions and notifications
The DSS is generating the interventions that have been set by the clinicians for each disease, comorbidity and combination of them, as described in D12.1. At the end of the PoP, when enough data will be collected the DSS will provide insights for the monitored patients (based on the analytics) to propose alterations to the proposed interventions and notifications. The alterations on the interventions are based on the existing medical guidelines although they are adjusted based on the assessment of the patient actions in response to the interventions. The interventions are generated for each patient individually, combining data stored in both FHIR and non-FHIR databases. The interventions and notification that are generated are transmitted to the Security Component, so that they are routed to the patients.
### Table 21. Components for interventions and notifications

<table>
<thead>
<tr>
<th>Related Component</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>FHIR Database</td>
<td>✓ Access device measurements</td>
</tr>
<tr>
<td>Non-FHIR database</td>
<td>✓ Device usage and status</td>
</tr>
<tr>
<td>Dashboard</td>
<td>✓ Interface for the clinicians and technicians to examine the proposed intervention alterations</td>
</tr>
<tr>
<td>Security Component</td>
<td>✓ Transmit the notifications</td>
</tr>
</tbody>
</table>

**Execution of the analytics and XAI**

The DSS has implemented the mechanism to trigger the periodic execution of the analytics mentioned that generate the notifications to the patients and the BDA workflows. It is also the module that provides access to the outcomes of the trained models, to provide reasoning on the interventions proposed, in order to help the clinicians and the research understand the results of Machine Learning methods. This component performs several tasks, to allow the generation of the interventions, based on the data analytics and the guidelines provided by the clinicians for each of the conditions. As more and more data are becoming available during the pilot studies, this component will be evolving throughout the project.

The initiative for the development of the eXplainable Artificial Intelligence (XAI) component of the DSS is that Machine learning models are used to support an increasing number of important decisions, in many aspects of the everyday life, who have different needs and require different kinds of explanations. For this reason, a collection of XAI algorithms is providing diverse ways of explaining decisions generated by the machine learning models. To explore the different types of algorithmic explanations, we consider the perspective of the different platform users (patients, data scientists and clinicians), and how they require different explanations.

From the perspective of the data scientist, the scientist would like to understand the behaviour of the model as a whole, not just on specific patients. For example, it may be required to review the model before deployment to the edge devices or to compare the model to the expert’s knowledge. It may also be required to check for compliance of the current model with the medical guidelines. A clinician may take a decision for a recommendation or intervention for the patient, after he/she understands how and why the model came to that prediction, to make an informed and trusted decision. Among the algorithms that will be used in the DSS, Local Interpretable Model-Agnostic Explanations (LIME) and Fast Interpretable Prototype Selection will be evaluated. They work with existing predictive models to show how the patient compares to others who have similar profiles and had similar medical history or real-life daily records, to the model’s prediction for the current patient, which helps to evaluate and predict if the intervention will be beneficial for that patient. Based on the model’s prediction and the explanation for how it came to that recommended intervention, the clinician can make a more informed decision for the patient. On the other side, the patient may want to know how and why the decision for an intervention was made by the clinician. The explanation given to him/her will help to have an insight into what personal actions can be beneficial for himself/herself, to increase the likelihood that a recommendation/intervention is accepted.
Analytics and explainability in the DSS

The DSS will consume data generated by the Analytics & Decision Models task (T4.2), which will focus on the model creation based on a) retrospective and generated data to simulate real monitoring data, b) prospective data generated during the pilot activities. Until retrospective data are available, the initial model creation will be based on generated data as described in Section 2.1 of this document, using Synthea (the Synthetic Patient Population Simulator) to generate simulated patient records, compliant to HL7 FHIR. Those models will be combined with the BDA, which will be triggered by the DSS to execute and manage the workflows that clinicians intend to execute on the data stored in the Data Repository of the platform, to perform analysis and obtain meaningful results. DSS will provide a tool for the explainability of the trained models, to support the interpretation of the predictions of the classifiers using model-agnostic explanations. The algorithms that will be implemented in DSS will focus on the selection of the prototypical examples that capture the underlying distribution of a dataset and a model-agnostic explanation, to learn the behaviour of the underlying model is to altering the input and see how the predictions change, an approach that offers a better interpretability by humans, because input changed, makes more sense to humans when interacting with trained models. The following two approaches of XAI algorithms have been identified so far that meet the requirements of the explainability of the SMART BEAR models and will be developed as part of the execution of the workflows and will be triggered upon the completion of each execution of the BDA workflow.

The main XAI algorithms that will be integrated in DSS will be the Fast Interpretable Prototype Selection\(^\text{21}\) (FIPS) and Local Interpretable Model-Agnostic Explanations\(^\text{22}\) (LIME) algorithms. The LIME algorithm is model-agnostic where an explanation is generated by approximating the underlying model by an interpretable one (e.g. a linear model with only a few non-zero coefficients). LIME algorithm has been developed so that it can explain the predictions of any classifier or regressor in a faithful way, by approximating it locally with an interpretable model. The explanations must be interpretable by the humans into a domain that the human can make sense of. For example, the interpretable domains may include graphs and diagrams so that a human can take a look at the output and understand or interpret the content. As part of the SMART BEAR Dashboard module, the DSS will provide feedback to the clinicians in a meaningful way, to understand how the proposed interventions for a patient have been produced. For example, using data from the questionnaires the patient completes and by the measurements of the sensors, the model may predict that most probably a certain patient is about to become sick. The decision about the final diagnosis and treatment still needs to be taken by a human, although the explanation of the decision will help the clinician why the model concluded that the patient may become sick, by highlighting the symptoms specific for this particular patient that impacted most significantly the model’s prediction, along with the profile of other patients with similar diseases and/or comorbidities. With this information about the rationale behind the prediction, the domain expert will be able to assess the validity of the model’s decision.

The intuition behind LIME is that by tweaking slightly the feature values and collecting the resulting impact of each individual feature change to the prediction, focusing on how the changes affect the model locally and not globally. As part of the work that we will undertake in the development of the DSS, we will try to give answers

---


on how to choose the number of neighbors to be analyzed or how to select the vicinity limit in terms of the distance, quantity etc, to optimally select the parameters of the LIME estimator.

As an alternative approach to provide explainable and trustable indicators on the trained models, the FIPS will be examined, to discover examples in one dataset that best represent the distribution of another. This can become handy when we want to target specific patients with an intervention based upon prior experience with the outcomes for other patient groups, or retrospectively, understanding the impact of an intervention or recommendation. The goal of the algorithm is to approximate the distribution of a dataset with weighted samples from another. This is performed by selecting samples that minimize the Maximum Mean Discrepancy (MMD) (which indicates how well the distribution of a dataset is represented) by the sample of another distribution. In DSS, linear and Gaussian kernels will be evaluated, to take into consideration the mean, variance and skew, and provide a more detailed understanding of the data.

Although, as part of the research activities, we are not limited only to the abovementioned algorithm, as other algorithms may also be tested and integrated, as long as they offer better results in terms of explainability and interpretability.

Integration with the Analytics

The execution of the two algorithms will be triggered at the end of the BDA workflows, or on demand, by requesting a model to be explained. It will be executed automatically after the completion of a workflow execution, or it will be able to be called by the dashboard, providing the model to be explained and the algorithm to be used. The process will be queued and the output will be returned to the dashboard interface, while at the same time will be stored in the Data Repository, along with all the required metadata to indicate the datasets and models that are correlated to the execution of the XAI algorithm. The XAI algorithms will be developed using Python\(^\text{23}\) language, along with modules contained in Anaconda package\(^\text{24}\) and the triggering of their execution will be integrated with the Rest APIs that provide data to the Dashboard. The initial evaluation of the execution of XAI algorithms and their integration with the Cloud Backend components will be initiated upon the development of the models that will be trained with the simulated synthetic data. The results of the model explainability will be reported in the following deliverables of WP4 and will be evaluated/validated by the clinicians and the researchers are part of Task 4.7.

2.8 Dashboard

The SMART BEAR dashboard is a component aimed at providing a user-friendly graphical user interface for the clinicians of the platform. The interface allows the clinicians to register and see a patient, his/her questionnaire score and his/her medical history, in addition to the delivered interventions. This section describes the main functionalities provided by the dashboard in order to create a patient, register the results of his/her physical examination, assessment, and the delivered interventions. The functionality that are shown here support the PoP, and the implementation of interaction elements continues on the basis of the requirements elaborated in D2.1.


2.8.1 Registration and login

The clinician registration form is shown in Figure 16. The user must fill the form with his/her name and surname, email address, role, organization, a password must be set. Registration can be done after reading and agreeing upon the policy.

![Registration Form]

*Figure 20. Screenshot of the registration form*

The clinicians will be able to log in the dashboard every time after the registration is successful through the form shown in Figure 17.

![Login Form]

*Figure 21. Screenshot of the login form*
### 2.8.2 Patients management

After a clinician successfully logs in, he can see the following buttons on the left sides of the dashboard:

- **Home.**
- **Patients,** that redirects to the Patients page where the managed patients are listed.
- **Analytics,** that redirects to the Analytics page where the performed analytics and their results are shown.
- **Interventions,** that redirects to the Interventions page where the delivered interventions are shown.
- **Inbox,** that redirects to the inbox page.
- **GDPR,** containing the form to use to visualize the GDPR-related requests and create new ones.

After logging in, the Home page will appear, where general statistics on average platform use, percentage of diseases, type of devices etc. are present, in addition to notifications of delivered alerts and status of performed analytics. The Home page is shown in Figure 22.

![Figure 22. Home page](image)
2.8.3 Available patients

The Patients page is shown in Figure 23. The patient and lists the managed patients, that are identified by their IDs (real names and surnames are not reported in compliance with GDPR) and age group.

![Figure 23. The patient page](image)

2.8.4 Overview of a patient

A clinician can see the overview of a specific patient by clicking on the Show button that opens the overview window shown in Figure 24. The Overview window shows the demographics, clinical data, questionnaire scores, and evaluation outcomes of the selected patient. The conditions that will be monitored in SMART BEAR can be flagged, and data from physical examinations and questionnaire scores can be added by hand.
The devices that are associated with a patient can also be seen in the Devices Management window shown in Figure 25. Each device is indicated by a Unique identifier and a Category (e.g. Smart Home Devices or Smart Blood Pressure Monitoring), and the device status is also indicated.

The ID Management window can be seen in Figure 26 and shows the external IDs that are used to identify a device associated with a patient.
2.8.5 Patient Medical History

The user can access the Medical History of a patient from the Overview window. The Medical History is shown in Figure 27 and Figure 28 and gives information about a patient’s Life Habits (e.g. Salt Intake), taken Medications, and Diet Supplements.
2.8.6 Patient Creation

A new patient can be created through the form shown in Figure 29 where the following data have to be inserted: Pilot, Organization, email address, date of birth, date of participation consent. In addition, up to 3 Clinical Case Managers from the Pilot and up to 3 external Clinical Case Managers can be indicated to be notified in case a Yellow or Red alert is sent to the patient or an Adverse Event occurs.
### 2.8.7 External ID creation

An External ID can also be created for a patient through the form that is shown in Figure 30. The External ID is required to associate the IDs of different systems in the context of the synergies between SMART BEAR, Holobalance and Smart4Health: given that a patient is indicated by an ID for each project he/she participates to, the External IDs identifies the association of the project-related IDs assigned to a patient.

![Create External ID form](image)

*Figure 30. Create External ID form*
2.8.8 Device allocation

The allocation of a device to a patient can be done through the form shown in Figure 31. The form allows the user to insert a Unique ID, the device status (e.g. active).

Figure 31. Assign Device form
2.8.9 CVDs use case

In the case of a patient that will be monitored for CVDs, the specific page opens where the user can insert data regarding the number of scheduled and unscheduled visits, the assessment outcomes (e.g. ECG), and clinical data. The Cardiovascular page is shown in Figure 32.

![Cardiovascular use case page](image-url)
2.8.10 Questionnaires

During the assessment, questionnaires are administered and their results and scores can be reported as illustrated in Figure 33, where Fall Efficacy Scale is shown as an example.

![Falls Efficacy Scale International](image)

*Figure 33. The Falls Efficacy Scale.*
2.8.11 Interventions

The dashboard allows the clinician to visualize and create interventions for all patients or just for specified targeted conditions or values of medical parameters. Please note that the Interventions that are shown here are hardcoded ones and support the PoP functionalities because the work on DSS component is currently in progress. The Medical interventions page is shown in Figure 34 and visualizes a list of Interventions each having an ID, the targeted Medical Conditions (all or a specified one), a Priority level (Low, Medium, High), a Status (Active, Inactive), a Notification text (e.g. “Repeat the measurement” in the case of CVDs), the name of the creator and the last update date. In the upper right corner, the user can find a button to create a new Intervention.

![Figure 34. Medical interventions page.](image)

2.8.11.1 Intervention creation

The creation of a Medical Intervention is shown in Figure 35 and Figure 36: after the user clicks on the button for creating an Intervention, a form will appear, where the Conditions to target, the Status, the Priority, the Criteria (i.e. specifications of ranges). In case the Intervention requires a Notification, the clinician can create it or select it from the existing ones.
Figure 35. Creation of a Medical Intervention

Figure 36. Creation of a Medical Intervention.
2.8.11.2 Notifications

The list of Notifications is shown in Figure 37. For each Notification, the following features are indicated: ID, Application (e.g. MyHeart, MyDiet, MyMemory, etc.), Available languages, Creator, Last Update.

![Figure 37. Notifications page](image)

In case a clinician needs to create a new Notification, she must click on the button in the upper right corner of the page and the form will appear that is shown in Figure 38.

![Figure 38. Create Notification form](image)
2.8.11.3 Intervention logs

In case a clinician needs to check whether an Intervention was successfully received, she can visualize the Intervention Logs that are listed on a page shown in Figure 39. The Status is indicated as Read or Unread.

![Intervention Logs page](image)

Figure 39. Interventions Logs page

2.8.12 Dashboard requirements

An example of a workflow template saved in the repository is shown in the Appendix, which describes the requirements for the dashboard. The object is used by the API of the BDA Engine to create the proper template in the HDFS of the BDA to use it in the execution of the analytic.
3 Concluding Remarks

This deliverable provides the specification for the first versions of the relevant SMART BEAR infrastructure components that compose the central cloud platform, including the database, the clinical repository, the Big Data Engine, the analytics components, and the synthetic data generator. These components will be integrated and deployed to support the SMART BEAR pilots. The user guide for this initial system is described in deliverable D4.1 SMART BEAR Cloud Enabling Components v1 (Demo), which depicts the demonstrator of the first version of the SMART BEAR backend Cloud platform enabling components.

Following the deployment and evaluation of this first version of the backend cloud platform, we will report on the updated version in D4.4 Report on SMART BEAR Cloud Enabling Components v2.
4 References


Appendix

BDA workflow configuration

Example of the BDA configuration to be stored in the non-FHIR database

```json
{
    "Oozie": {
        "WorkflowName": "awspl",
        "StartNodeName": "data-input",
        "ActionNode": [{
            "SparkNodeDetails": {
                "NodeName": "data-input",
                "JobTracker": "${jobTracker}\",
                "NameNode": "${nameNode}\",
                "Master": "${master}\",
                "ApplicationName": "Input",
                "Jar": "${taskPathInput}\",
                "Args": [{
                    "Arg": "mode=reader"
                },
                {
                    "Arg": "tables=HA_ENVIRONMENT_DATA"
                },
                {
                    "Arg": "queryInputData=${queryInputData} WHERE PATIENT_ID=${patient}"
                },
                {
                    "Arg": "resultPath=${urlcsv}/${wf:id()}"
                },
                {
                    "Arg": "zkUrl=172.20.28.14:2181:/hbase-unsecure"
                }
            ],
            "NextAction": "data-preparation",
            "ErrorToName": "fail"
        }
    }
}
```
"ClassName": "it.unimi.SMART BEAR.tasks.Runner"
}
},
{
"SparkNodeDetails": {
"NodeName": "data-preparation",
"JobTracker": "${jobTracker}",
"NameNode": "${nameNode}",
"Master": "${master}",
"ApplicationName": "data-preparation",
"Jar": "${taskPathPrepare}",
"Args": [{
   "Arg": "csvData=${urlcsv}/${wf:id()}/part-*"
 },
   {
      "Arg": "resultPath=${resultPath}/${wf:id()}/prepared"
   }
],
"NextAction": "data-cleaning",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART BEAR.tasks.DataPreparation"
}
},
{
"SparkNodeDetails": {
"NodeName": "data-cleaning",
"JobTracker": "${jobTracker}",
"NameNode": "${nameNode}",
"Master": "${master}",
"ApplicationName": "data-cleaning",
"Jar": "${taskPathClean}",
"Args": [{
   "Arg": "csvData=${resultPath}/${wf:id()}/prepared"
 },
   {
      "Arg": "labels=${labels}"
   },
   {
      "Arg": "resultPath=${resultPath}/${wf:id()}/cleaned"
   }
]
"NextAction": "awspl",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART BEAR.tasks.DataCleaning"
},
{
  "SparkNodeDetails": {
    "NodeName": "awspl",
    "JobTracker": "${jobTracker}",
    "NameNode": "${nameNode}",
    "Master": "${master}",
    "ApplicationName": "awspl",
    "Jar": "${taskPathawspl}",
    "Args": [{
      "Arg": "csvData=${resultPath}/${wf:id()}/cleaned"
    },
    {
      "Arg": "patient=${patient}"
    },
    {
      "Arg": "resultPath=${resultPath}/${wf:id()}/output"
    }],
    "NextAction": "data-save",
    "ErrorToName": "fail",
    "ClassName": "it.unimi.SMART BEAR.tasks.awspl"
  }},
{
  "SparkNodeDetails": {
    "NodeName": "data-save",
    "JobTracker": "${jobTracker}",
    "NameNode": "${nameNode}",
    "Master": "${master}",
    "ApplicationName": "Save",
    "Jar": "${taskPathInput}",

"Args": [{
  "Arg": "mode=writer"
},
{
  "Arg": "csvData=${resultPath}/${wf:id()}/output/part-*"
},
{
  "Arg": "primaryKey=AUTO"
},
{
  "Arg": "table=T${wf:id()}"
},
{
  "Arg": "urlDb=${urlDb}" 
},
{
  "Arg": "zkUrl=172.20.28.14:2181:/hbase-unsecure"
},
"NextAction": "end",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART BEAR.tasks.Runner"
],

"EndNodeName": "end",
"KillNodeDetails": {
  "KillNodeMessage": "Workflow failed, error message ${wf:errorMessage(wf:lastErrorNode())}",
  "KillNodeName": "fail"
}

"Workflow": {
  "Name": "awspl",
  "DawId": "awspl",
  "Tasks": "",
BEAR/new_tasks/awSPL.jar;taskPathDataSave=hdfs://172.20.28.14:8020/user/bda/SMART
BEAR/new_tasks/csv2hbase.jar;patient=IN161228;queryInputData=SELECT
PATIENT_ID,RECORD_DATE,SOUND_PARAMETERS FROM HA_ENVIRONMENT_DATA;labels=RECORD_DATE,SOUND_PARAMETERS;phoenixDriver=hdfs://172.20.28.14:8020/
user/bda/phoenix-4.7.0.2.6.4.0-91-spark2.jar;“

"ExecutionType": "",

"Description": "Calculates the distribution of A-weighted SPL outside of HA and gained by HA, distribution of A-weighted continuous SPL and distribution of Noise exposure level related to 16 hour per day outside of HA and gained by HA",

"DynamicParameters": [
  {
    "name": "Avg",
    "description": "Choose if we want to use Average",
    "mandatory": true,
    "valueType": "boolean",
    "default": "true"
  },
  {
    "name": "Patient",
    "description": "Patient to analyze",
    "mandatory": false,
    "valueType": "string",
    "default": "Patient1"
  },
  {
    "name": "Noise Exposure",
    "description": "The noise exposure to use as a barrier",
    "mandatory": true,
    "valueType": "numeric",
    "default": "20"
  }
]

BDA analysis results

{
  "id":23,
  "name": "awspl",
  "dawId": "awspl",
  "tasks": "",

Page 117
BEAR/new_tasks/awSPL.jar;taskPathDataSave=hdfs://172.20.28.14:8020/user/bda/SMART
BEAR/new_tasks/csv2hbase.jar;patient=IN161228;queryInputData=SELECT
PATIENT_ID,RECORD_DATE,Sound_PARAMETERS FROM
HA_ENVIRONMENT_DATA;labels=RECORD_DATE,SOUND_PARAMETERS;phoenixDriver=hdfs://172.20.28.14:8020/
user/bda/phoenix-4.7.0.2.6.4.0-91-spark2.jar;";

"executionType": "",
"description": "Calculates the distribution of A-weighted SPL outside of HA and gained
by HA, distribution of A-weighted continuous SPL and distribution of Noise exposure level related
to 16 hour per day outside of HA and gained by HA",
"dynamicParameters": [{
   "name": "avg",
   "text": "Avg",
   "description": "Choose if we want to use Average",
   "mandatory": true,
   "valueType": "boolean",
   "default": "true"
}, {
   "name": "patient",
   "text": "Patient",
   "description": "Patient to analyze",
   "mandatory": false,
   "valueType": "string",
   "default": "Patient1"
}, {
   "name": "noise-exposure",
   "text": "Noise Exposure",
   "description": "The noise exposure to use as a barrier",
   "mandatory": true,
   "valueType": "numeric",
   "default": "20"
}]
}

BDA results visualization

{
   "userID": 123,
   "jobId": "0000022-200211034642235-oozie-oozi-W",
   "results": [
      ["score", "patientID", "BMI"],
      [89.3, 58212, 25],
   ]
}
[57.1, 78254, 23],
[74.4, 41032, 24],
[50.1, 12755, 17],
[89.7, 20145, 12],
[68.1, 79146, 45],
[19.6, 91852, 14],
[10.6, 101852, 34],
[32.7, 20112, 34]
],
"conf": {
"homeChart": true,
"idChart": "avg-BMI"
}
}

User settings

{
"userID": 123,
"charts": [
{
"id": "avh-heartbeat",
"name": "Avg Heartbeat",
"show": true
},
{
"id": "avg-BMI",
"name": "Avg. BMI",
"show": false
}
],
"language": "EN"
}

Dashboard requirements

{
"Oozie": {
"WorkflowName": "awspl",
"StartNodeName": "data-input",

"ActionNode": [{
    "SparkNodeDetails": {
        "NodeName": "data-input",
        "JobTracker": "${jobTracker}",
        "NameNode": "${nameNode}",
        "Master": "${master}",
        "ApplicationName": "Input",
        "Jar": "${taskPathInput}",
        "SparkOpts": "--conf
        spark.driver.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
        spark2.jar:file:///usr/hdp/current/phoenix-client/phoenix-client.jar
        --conf
        spark.executor.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
        spark2.jar:hfile:///usr/hdp/current/phoenix-client/phoenix-client.jar
        --jars
        file:///usr/hdp/current/phoenix-client/phoenix-spark2.jar,file:///usr/hdp/current/phoenix-
        client/phoenix-client.jar",
        "Args": [
            
            
            "Arg": "mode=reader"
        ],
        
        
        "Arg": "tables=HA_ENVIRONMENT_DATA"
    ],
    "NextAction": "data-preparation",
    "ErrorToName": "fail",
    "ClassName": "it.unimi.SMART BEAR.tasks.Runner"
}]
},

{ "SparkNodeDetails": {
    "NodeName": "data-preparation",
    "JobTracker": "${jobTracker}",
    "NameNode": "${nameNode}",
    "Master": "${master}",
    "SparkOpts": "--conf
    spark.driver.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
    spark2.jar:file:///usr/hdp/current/phoenix-client/phoenix-client.jar
    --conf
    spark.executor.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
    spark2.jar:hfile:///usr/hdp/current/phoenix-client/phoenix-client.jar
    --jars
    file:///usr/hdp/current/phoenix-client/phoenix-spark2.jar,file:///usr/hdp/current/phoenix-
    client/phoenix-client.jar",
    "Args": [
        "Arg": "mode=reader"
    ],
    "NextAction": "data-preparation",
    "ErrorToName": "fail",
    "ClassName": "it.unimi.SMART BEAR.tasks.Runner"
}]
},

{ "SparkNodeDetails": {
    "NodeName": "data-preparation",
    "JobTracker": "${jobTracker}",
    "NameNode": "${nameNode}",
    "Master": "${master}",
    "SparkOpts": "--conf
    spark.driver.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
    spark2.jar:file:///usr/hdp/current/phoenix-client/phoenix-client.jar
    --conf
    spark.executor.extraClassPath=file:///usr/hdp/current/phoenix-client/phoenix-
    spark2.jar:hfile:///usr/hdp/current/phoenix-client/phoenix-client.jar
    --jars
    file:///usr/hdp/current/phoenix-client/phoenix-spark2.jar,file:///usr/hdp/current/phoenix-
    client/phoenix-client.jar",
    "Args": [
        "Arg": "mode=reader"
    ],
    "NextAction": "data-preparation",
    "ErrorToName": "fail",
    "ClassName": "it.unimi.SMART BEAR.tasks.Runner"
}]
}
"ApplicationName": "data-preparation",
"Jar": "${taskPathPrepare}",
"Args": [{
    "Arg": "csvData=${urlcsv}/$(wf:id())/part-*"
    },
    {
    "Arg": "resultPath=${resultPath}/$(wf:id())/prepared"
    }
},
"NextAction": "data-cleaning",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART BEAR.tasks.DataPreparation"
}
},
{
"SparkNodeDetails": {
    "NodeName": "data-cleaning",
    "JobTracker": "${jobTracker}"
    "NameNode": "${nameNode}"
    "Master": "${master}"
    "ApplicationName": "data-cleaning",
    "Jar": "${taskPathClean}"
    "Args": [{
        "Arg": "csvData=${resultPath}/$(wf:id())/prepared"
    },
    {
    "Arg": "labels=${labels}"
    },
    {
    "Arg": "resultPath=${resultPath}/$(wf:id())/cleaned"
    }
    ],
    "NextAction": "awspl",
    "ErrorToName": "fail",
    "ClassName": "it.unimi.SMART BEAR.tasks.DataCleaning"
}
}
{"SparkNodeDetails": {
    "NodeName": "awspl"}
"JobTracker": "${jobTracker}",
"NameNode": "${nameNode}",
"Master": "${master}",
"ApplicationName": "awspl",
"Jar": "${taskPathawspl}",
"Args": [{
    "Arg": "csvData=${resultPath}/${wf:id()}/cleaned"
},
{
    "Arg": "patient=${patient}"
},
{
    "Arg": "resultPath=${resultPath}/${wf:id()}/output"
}
],
"NextAction": "data-save",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART.BEAR.tasks.awspl"
}],
{
  "SparkNodeDetails": {
    "NodeName": "data-save",
    "JobTracker": "${jobTracker}",
    "NameNode": "${nameNode}",
    "Master": "${master}",
    "ApplicationName": "Save",
    "Jar": "${taskPathInput}",
    "SparkOpts": "--conf
    "Args": [{
        "Arg": "mode=writer"
    },
    {
        "Arg": "csvData=${resultPath}/${wf:id()}/output/part-*"
    },
    {
        "Arg": "primarykey=AUTO"
"Arg": "table=${wf:id()}",
"Arg": "urlDb=${urlDb}",
"Arg": "zkUrl=172.20.28.14:2181:/hbase-unsecure"},
"NextAction": "end",
"ErrorToName": "fail",
"ClassName": "it.unimi.SMART BEAR.tasks.Runner"}
]
,"EndNodeName": "end",
"KillNodeDetails": {
  "KillNodeMessage": "Workflow failed, error message: ${wf:errorMessage(${wf:lastErrorNode()})}",
  "KillNodeName": "fail"
}
,"Workflow": {
  "Name": "awspl",
  "DawId": "awspl",
  "Tasks": "",
  "ExecutionType": "",
  "Description": "Calculates the distribution of A-weighted SPL outside of HA and gained by HA, distribution of A-weighted continuous SPL and distribution of Noise exposure level related to 16 hour per day outside of HA and gained by HA",
  "DynamicParameters": [{
    "name": "Avg",
...."}]}
After the API of the BDA Engine receives this JSON, an object like the one below should be saved in the backend

```
```json
{    
  "id":23,
  "name": "awspl",
  "dawId": "awspl",
  "tasks": ",",
  "executionType": "",
  "description": "Calculates the distribution of A-weighted SPL outside of HA and gained by HA, distribution of A-weighted continuous SPL and distribution of Noise exposure level related to 16 hour per day outside of HA and gained by HA",
}
```
"dynamicParameters": [{
  "name": "avg",
  "text": "Avg",
  "description": "Choose if we want to use Average",
  "mandatory": true,
  "valueType": "boolean",
  "default": "true"
}, {
  "name": "patient",
  "text": "Patient",
  "description": "Patient to analyze",
  "mandatory": false,
  "valueType": "string",
  "default": "Patient1"
}, {
  "name": "noise-exposure",
  "text": "Noise Exposure",
  "description": "The noise exposure to use as a barrier",
  "mandatory": true,
  "valueType": "numeric",
  "default": "20"
}]
}
```

This structure is used by the Dashboard in order to allow the clinicians to customize the execution of a workflow

## Workflow Execution

When a user runs a workflow the Dashboard sends an object like the one here below to the BDA using the id of the workflow in the path argument:

```
POST /api/workflows/run/{edawId}
```
```
```
```
```json

```json

```

The object that the BDA should store in the backend should be an object like this one:

```
```json

```

```json

```
The object contains also contains the visualization settings that will be updated directly from the Dashboard using the Apache Echarts definition language to store the settings of the visualization chosen by the user.

## Analytic Results

To store the results of an analytic a convention could be defined like the one below, using this convention the visualization could be easier for the visualization component in the UI:

```json
{
    "userID": 123,
    "jobId": "0000022-200211050542235-oozie-oozi-W",
    "results": [
        ["score", "patientID", "predictionCVD"],
        [89.3, 58212, true],
        [57.1, 78254, true],
        [74.4, 41032, true],
        [50.1, 12755, false],
        [89.7, 20145, true],
        [68.1, 79146, true],
        [19.6, 91852, false],
        [10.6, 101852, false],
        [32.7, 20112, true]
    ],
    "conf": {
        "homeChart": false,
        "idChart": null
    }
}
```

Another solution could be to add directly in the JSON object that also contains also the workflow execution but maybe the data are bigger than expected.

## Visualization

...
The visualization settings of the analytics will be stored inside the object that will contain the reference to the workflow execution like said before.

## Inbox

Inbox should will contain the messages that the platform sends to a Clinical Case Manager end-users (i.e., clinicians), like an alert or a message, provided they have given their consent to receive those.

## Interventions

The Dashboard should create also the Interventions

![Intervention](intervention.png)

## Home Dashboard Charts Data

The Home Dashboard data should contain also some values that we should display using charts.

```json
{
    "userID": 123,
    "jobId": "0000022-200211034642235-oozie-oozi-W",
    "results": [
        ["score", "patientID", "BMI"],
        [89.3, 58212, 25],
        [57.1, 78254, 23],
        [74.4, 41032, 24],
        [50.1, 12755, 17],
        [89.7, 20145, 12],
        [68.1, 79146, 45],
        [19.6, 91852, 14],
        [10.6, 101852, 34],
        [32.7, 20112, 34]
    ],
    "conf": {
        "homeChart": true,
        "idChart": "avg-BMI"
    }
}
```
## User Settings

This object is used to store the settings that the user wants to apply.

```json
{
  "userID": 123,
  "charts": [
    {
      "id": "avh-heartbeat",
      "name": "Avg Heartbeat",
      "show": true
    },
    {
      "id": "avg-BMI",
      "name": "Avg. BMI",
      "show": false
    }
  ],
  "language": "EN"
}
```

Example of Synthetic Data: muscle mass value

```json
{
  "fullUrl": "urn:uuid:31849b47-367a-08ff-b71c-47e5f9f2c3d0",
  "resource": {
    "resourceType": "Observation",
    "id": "31849b47-367a-08ff-b71c-47e5f9f2c3d0",
    "status": "final",
    "category": [ {
      "coding": [ {
        "system": "http://terminology.hl7.org/CodeSystem/observation-category",
        "code": "vital-signs",
        "display": "vital-signs"
      } ]
    } ]
  }
}
```
Example of Synthetic Data: walked distance

```json
{
  "fullUrl": "urn:uuid:d7915f58-903f-2fac-062f-4ff644348f67",
  "resource": {
    "resourceType": "Observation",
    "id": "d7915f58-903f-2fac-062f-4ff644348f67",
    "status": "final",
    "category": [ { 
      "coding": [ { 
        "system": "http://terminology.hl7.org/CodeSystem/observation-category",
        "code": "exam",
        "display": "exam"
      } ]
    } ]
  }
}
```
"code": { "coding": [ { "system": "http://loinc.org", "code": "41953-1", "display": "distance walked" } ], "text": "distance walked" }, "subject": { "reference": "urn:uuid:95366b5e-3a32-6dfe-e91b-8025462f8a4f" }, "encounter": { "reference": "urn:uuid:e7491a01-88c7-855a-2e46-877b443c674f" }, "effectiveDateTime": "2013-01-18T04:26:39+01:00", "issued": "2013-01-18T04:26:39.209+01:00", "valueQuantity": { "value": 583.55, "unit": "m", "system": "http://unitsofmeasure.org", "code": "m" } }, "request": { "method": "POST", "url": "Observation" } }