PROJECT DELIVERABLE REPORT

Project Title:

Advanced personalized, multi-scale computer models preventing osteoarthritis

SC1-PM-17-2017 - Personalized computer models and in-silico systems for well-being

Deliverable number	D6.6
Deliverable title	Ontology-based framework for data standardisation
Submission month of deliverable	M41
Issuing partner	CERTH
Contributing partners	LEITAT, UNIC, HULAFE, ANIMUS
Dissemination Level (PU/PP/RE/CO):	PU
Project coordinator	University of Nicosia (UNIC)
Tel:	**+357 22 841 528**
Fax:	**+357 22 357481**
Email:	**felekkis.k@unic.ac.cy & giannaki.c@unic.ac.cy**
Project web site address	www.oactive.eu

Revision History

Version	Date	Responsible	Description/Remarks/Reason for changes
1.0		CERTH	First Draft
1.1		LEITAT, UNIC, HULAFE, ANIMUS	Review of First Draft
2.0		CERTH	Second Draft

Contents

Abbreviations

AR	Augmented Reality
AI	Artificial Intelligence
BMI	Body Mass Index
BML	Bone Marrow Lesions
CDASH	Clinical Data Acquisition Standards Harmonization
CDISC	Clinical Data Interchange Standards Consortium
CSI	Central Sensitisation Inventory
D	Deliverable
DoF	Degrees of Freedom
DSS	Decision Support System
EMG	ElectoMyoGraphy
EoM	Equations of Motion
EU	European Union
FACHS	Functional Ambulation Classification of the Hospital at Sagunto
GADS	Generalised Anxiety Disorder Scale
GPU	Graphic Processing Unit
HADS	Hospital Anxiety and Depression Scale
ICT	Information and Communication Technologies
ID	Invers Dynamics
IK	Invers Kinematics
IMU	Inertial Measurement Unit
ISD	Inherent Structural Differences
JRA	Joint Reaction Analysis
KOOS	Knee injury and Osteoarthritis Outcome Score
MDM	Medical Data Models
MRI	Magnetic Resonance Imaging
OA	Osteo Arthritis
OCN	OsteoCalciN
ODM	Operational Data Model
QST	Quantitative Sensory Test
SO	System Optimization
UML	Unified Modelling Language
UMLS	Unified Modelling Language System
VR	Virtual Reality
WBAN	Wireless Body Area Network
WOMAC	Western Ontario and McMaster Universities Ostreoarthritis Index
WP	Work Package

1 Summary

The purpose of the Deliverable 6.6 is sensing and gathering information about the patient's on motion information, joint kinematics, movement detection, and gait analysis. It also gathers information from many other sources. The perception layer stores this information in a data warehouse. By utilizing an ontology-based framework, the goal of which is the coherency and security of the data and the privacy of the patients. The overall procedure aims to gather the necessary information for the personalization of the treatment plan.

This report refers to Deliverable 6.6, which relates to the OACTIVE WP 6, "Hyper-modelling framework empowered by big data and deep learning" led by CERTH and specifically Task 6.6 "Ontology-based framework for data standardisation", also led by CERTH.

2 System general architecture

The architecture of the system and the relation with the WPs is shown in Figure 2.1 below. As explained in the System Description Section above, the main input into the deep learning hypermodel core of the OACTIVE system will be the data collected through clinical studies and OA registries informed by the mechanistic and phenomenological models in WPs 3, 4 &5. The bidirectional arrows indicate that the hypermodeling core will be running and updated continuously with additional data as these are generated from the clinical studies or mined from various OA registries. Likewise, the outcome of the hypermodeling component will be tested and validated in the existing databases and improvements and refinements of the prediction models will be rerun and re-tested again continuously to improve the sensitivity of the models and the predictions as indicated by the bi-directional arrows between WP 6 and WP8. Refinement of the hyper-model and identification of critical prediction factors will also be informing the development and refinement of the augmented reality and gait retraining tools as indicated by the third bi-directional arrow in the schematic system architecture description below.

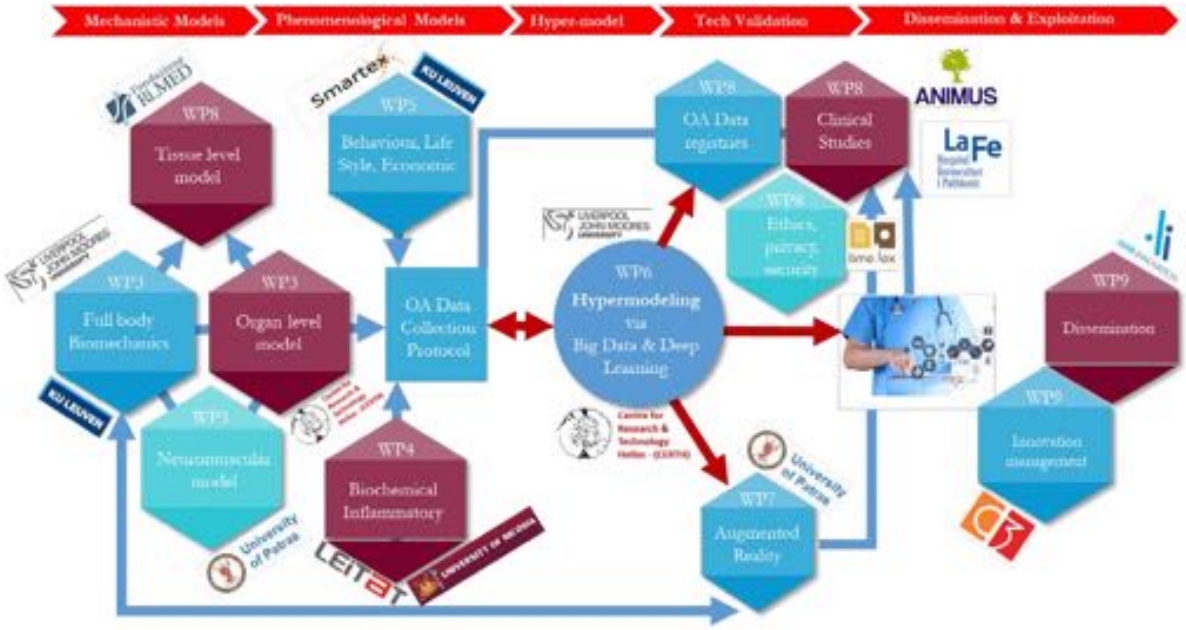


Figure 2.1: Schematic description of the OACTIVE system architecture and the relation with the

1. Mechanistic modelling framework of the musculoskeletal system

The development of in silico multiscale biomechanical models of healthy and knee joints with OA based on subject-specific joint and tissue level experimental mechanics that are capable of predicting tissue loading and responses in individuals and provide inputs for the mathematical 'hyper-models' accounting for mechanical loading of tissues in different conditions and individuals. These mechanistic models include:

- Development of personalized neuromusculoskeletal models used to predict knee OA onset and improve treatment.
- Development of novel calibration pipelines for the transformation of generic musculoskeletal models to personalized models by scaling anatomic geometry, kinematics and muscle kinetics and activation parameters.

- Development of organ and tissue level models for the incorporation of detailed bone and cartilage models capable of predicting tissue responses following estimation of forces from the rigid body musculoskeletal models.

2. Systemic health and inflammation modelling framework

The modelling of biochemical health indicators and inflammatory biomarkers that is used to assess a number of different systemic and joint condition indicators. Serum biochemical markers (3 Prognostic Biomarkers of Bone and Cartilage Degradation and Synthesis and 3 Inflammatory Biomarkers) are monitored and correlated with the clinical outcomes. These are used to examine the relationship between biochemical markers for OA and clinical diagnosis as well as the progression of disease in affected individuals or elevated inflammatory that precedes the development of the condition. In detail:

- Development of a system of prognostic biomarkers of bone and cartilage degradation and synthesis applied to OA based on serum markers.
- Development of a system of inflammatory prognostic biomarkers for OA monitoring based on biofluid samples (blood, urine and synovial fluid).

3. Behaviour, social, environmental modelling framework

To detect a user's physical, mental and social behaviours and identify higher-level physical, mental/emotional, and social states of the user and information used for providing individualised diagnosis and recommendations for patient-specific treatments. In detail:

- Assess and model behaviour of users related to physical activity using flexible platforms of wearable body sensors.
- Development and implement behaviour analysis to create a set of behaviour models and "normality patterns".
- Investigate the effect of socio-economical risk factors including (a) Compositional attributes of socioeconomic status at the individual level, (b) Social context (community-level) risk factors, (c) Personal risk factors associated with OA, including smoking, age, gender, occupation-tasks, BMI (i.e., obesity), injury, family history, race and (d) Community perceptions, will be also included with mediators and moderators and dimensions of psychological influences.

4. Hypermodelling framework empowered by big data.

The hyper-modelling framework of OACTIVE which includes:

- Data management mechanisms to ensure a high level of data quality and accessibility for the big data analytics applications.
- Development of data pre-processing algorithms to improve data quality and consequently facilitate the efficiency of the data mining task. Tools: Discretization algorithms, Instance Selection, sophisticated undersampling/oversampling techniques, filtering/denoising, data transformation
- Development of data mining techniques for knowledge discovery using interpretable rule-based models to provide insights for the understanding of OA disease development and its progression. Identification of patient-specific significant risk factors associated with the onset as well as factors related to OA progression using computational efficient Feature Selection algorithms.
- Development of the ICT deep learning infrastructure. (1) Machine learning Tools: neural networks, support vector machines, decisions trees and discriminant analysis, (2) Deep Learning Tools: fully connected neural networks, convolutional neural networks and recurrent neural networks on state-of-the-art GPU-accelerated tools.

- Design and implementation of personalized predictive Decision Support (DS) models that address specific OA stages in the disease continuum of a patient (DS-early, DS-mild, DS-mod and DS-treat).

5. Ontology-based framework for data/model reusability and sharing.
   - Employment of model and data encoding and exchange standards for multiscale modelling to ensure model reproducibility and sharing.
   - To develop modular approaches to ensure that self-contained models are developed and validated independently before being incorporated into a hierarchy of imported models.
   - Employment of Semantic web technology, to make knowledge interpretable by web agents, thereby enabling the integration and re-usability of heterogeneous resources for knowledge discovery.
   - Issuing of authentication mechanisms (via X.509 certificates) assuring the secure access to data.
   - Employment of enhanced replication mechanisms to warrant the integrity of data including the prevention of loss.
   - Insurance of a certain k-anonymity using pseudo-anonymization techniques.

6. Personalised interventions using Augmented Reality (AR)
   - Issuing personalized intervention relying on the AR gaming concept.
   - Employ assistive, real-time visual and vibrotactile feedback for OA gait retraining.
   - Calculation of biomechanical indicators for assessment and clinical decision support.
   - Implementing personalized stimuli to impact on game task completion performance.
   - The application of the model in knee OA patients to investigate what effect simulated biomechanical treatments have on the mechanical load characteristics in knee joint structures in different groups of knee OA patients.

7. Technology Validation

To test the OACTIVE system using a comprehensive validation strategy that includes:

- Clinical studies in human populations for the validation of the efficiency of the non-invasive risk factors.
  - Validation scenario A: early detection of OA (population size>100 patients, where: LAFE)
  - Validation scenario B: system evaluation in elderly people (population size>130 patients, where: NIC – hosted in Apollonion Private Hospital in Cyprus)
  - Validation scenario C: post-traumatic assessment in athletes (population >100 patients, where: AMIMUS)

i. In vitro Clinical trials: to validate in vitro the relationship between cellular responses of osteochondral tissue and (a) biomarkers and imaging data (diagnostics), and (b) the tissue level mechanical activation during AR rehabilitation (therapy).
ii. Validation in large data registries (i) Osteoarthritis Initiative (OAI) >5000 patients, >100 months follow-up.

3 Ontology - Based Framework Architecture

The purpose is the semantic annotation of data elements data integration, of matching data elements, from different data sources to be supported. By utilizing semantic web technologies web agents will be able to interpret the knowledge, as a result, the integration and re-usability of heterogeneous resources for knowledge discovery.

Ontologies were chosen as the structure for the reference data model since they provide the means and methods to integrate the various standards' data models into one unique reference model without dropping the original notation of the individual standards. In addition, using ontologies ensures that data consistency is done on a semantic level and that semantic reasoning and inference can be carried out.

The volume and complexity of patient data are steadily increasing in personalised medicine. The identification of matching data elements in different sources applied through semantic annotation of data elements. These annotations should be equable, in terms of coding, (the same annotations must be contained to matching data elements) but large terminologies (SNOMED CT / UMLS) do not provide that type of coding. To achieve it, semantic annotations are re-used for the matching data elements in the metadata repository. ISO/IEC 11179 Standard used for data elements representation.

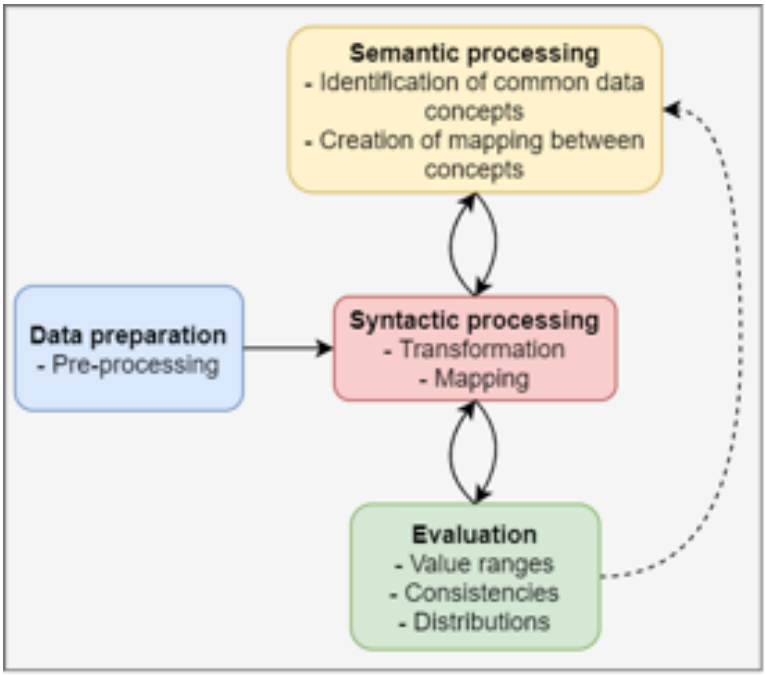


Figure 3.1: Schematic representation of the ontology-framework

The data model, with uniform semantic annotations, is implemented with the use of the ODMedit tool. ODMedit tool is a semi-automatic approach where data integration is simplified if the same code is applied for all data items with the same (or at least very similar) meaning. By re-using annotation codes, we aim to achieve equable codes. The portal of medical data models (MDM) used as an open-access repository for metadata in CDISC ODM format with ~5,800 forms and ~450,000 data elements (https://medical-datamodels.org/).

The annotation of new data items is achieved through an initial search, of the repository, for items with the same names. The determination of the new data item's meaning, over existing data, is done by an expert. Experts can select codes according to the maximum specificity principle. On the other hand, for items that

an annotation is not available in the repository, matching annotation codes are retrieved from UMLS diagrams. In case a single matching code does not exist, post coordination is applied. In this case also, experts identify matching codes. When the next data item, with the same meaning, must be annotated, annotation codes are available. This enables equable annotation of data items, even when several UMLS codes with similar meanings are available. The decision whether two data items have the same meaning is taken semi-automatically to ensure high coding quality.

The scope of the evaluation is to demonstrate that this software tool is able to perform uniform semantic annotation for real data models from clinical studies. CDISC develops international data standards for clinical research. As a result of the Clinical Data Acquisition Standards Harmonization (CDASH) initiative, CDISC developed a set of forms frequently used data items, for instance regarding demographics data or adverse events. These items are coded with CDASH codes. It is determined how many of these data items can be annotated with UMLS codes. Correctness of UMLS codes is assessed by manual comparison with CDASH codes.

In addition, a set of data models from the MDM portal was manually processed with ODMedit to determine technical feasibility of this tool.

ODMedit is intended to foster uniform semantic annotation. A random set of data elements from an established data standard was selected to test this feature. For each of those data elements available UMLS codes were identified with the UMLS Metathesaurus Browse. Suitable codes were identified from the output of the UMLS Metathesaurus Browser by manual review. Available annotations in the MDM portal were analysed for each data element regarding uniform semantic annotation and compared to UMLS codes.

The data management plan describes plans for creating, organizing, documenting, storing and sharing data. It takes into account issues such as data protection and confidentiality, data preservation and curation and provides a framework that supports researchers and their data throughout the course of their research and beyond. It is constantly updated describing what kind of research data are generated, what policies apply to the data. Funding and legal policies and data management practices as backups, storage, access control, archiving will be determined. It is clear who owns and has access to what data and who is responsible for each aspect of the plan. Moreover, partners try to clearly describe which data will retain value after the end of the OACTIVE project and how its reuse will be enabled and how the long-term preservation ensured after the original research is completed. Partners share research data generated allowing others to replicate, validate, or correct their results, thereby improving the scientific record. This increases the research's integrity and replication as those who make use of their data and cite it in their own research will disseminate the results possibly in other disciplines, sectors, and countries. Moreover, partners can identify, retrieve, and understand the data themselves after they have lost familiarity with it, perhaps several years hence. Partners do not share specific parts of their research data if the achievement of the project's main objectives would be jeopardized by making those specific parts of the research data openly accessible. They do not also share data that they may not have the legal right to share, maybe because the parts of them belong to other authors or entities.

There are a lot of challenges in signal processing of the multi-scale data, given their current state and the non-standardized structure. But there are opportunities in each step of the process towards providing systemic improvements. Despite the need for further research in the area of data wrangling, aggregating, and harmonizing continuous and discrete medical data formats, there is also a similar need for the development of novel signal processing techniques specialized towards physiological signals. Research relevant to biomarkers and clandestine patterns within biosignals to understand and predict disease cases have ability in providing actionable information. However, there are opportunities for developing algorithms to address data filtering, interpolation, transformation, feature extraction and feature selection. Furthermore, with the notoriety and improvement of machine learning algorithms, there are opportunities

in improving and developing robust Clinical Decision Support Systems for clinical prediction, prescription, and diagnostics.

The computational modelling layer will provide subject-specific information, based on principles of physics and physiology, that will be used to define the appropriate personalized intervention strategy. The input data may also originate from other activities of the OACTIVE project, e.g., biomarkers, behaviour, hyper-modelling, etc. Appropriate abstraction barriers (data abstraction) through big data analytics and machine learning techniques, are used to decouple the implementation from the input data representation in order to provide a modular and expandable design. The sensory information, that can originate from a variety of sensor devices, such as Inertial Measurement Units (IMUs), motion capture systems, Force-plates, will be analysed, stored, processed and compared against the optimal performance dictated by the Decision Support System (DSS). Convenient interfaces should be created so as to enable the support of multiple, heterogeneous devices. The information must be synchronized using data synchronization techniques, such as network time protocol. Appropriate filtering and pre-processing will be applied to refine and remove any artifacts (e.g., noise) that can negatively affect the processing phase. Finally, the necessary information required by the personalized intervention module will be extracted. The personalized intervention module will fuse sensory and other information in order to define the subject specific strategy. Multiple rehabilitation strategies will be specified based on the expert's knowledge and results. This module will choose the most appropriate strategy of action according to the provided information in real time.

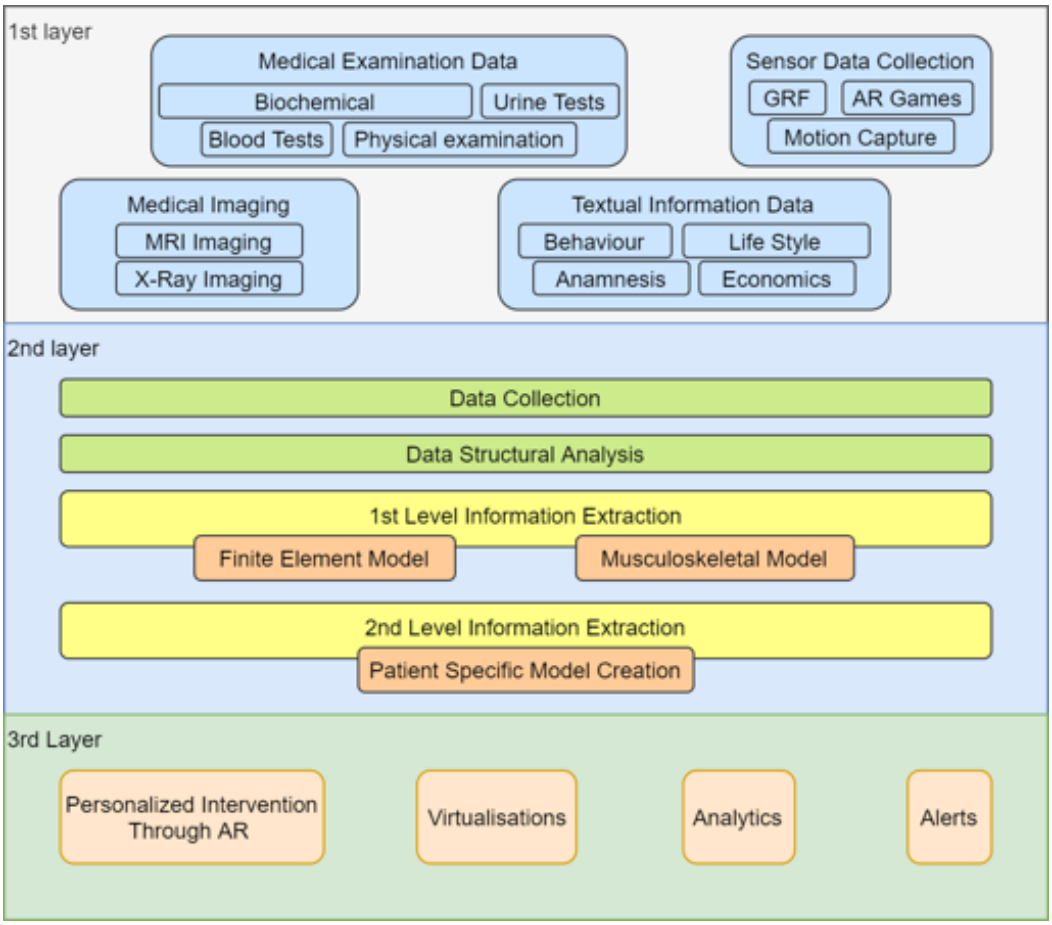


Figure 3.2: OAcitve's layered architecture overview

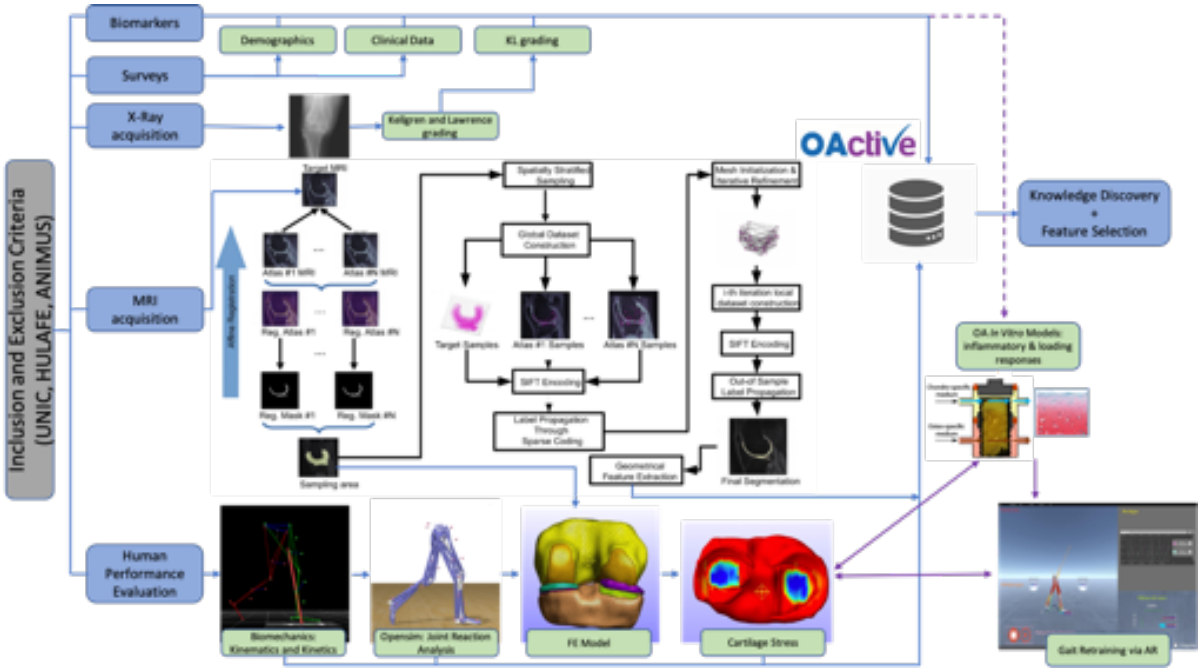


Figure 3.3: Data collection and processing structure

Ontology - Based Architecture is a three-layer architecture (Figure 3.2), namely, Data Source layer, Core layer and Utility layer.

i. *The Data Source layer* is the layer where the data from all sources is collected. It includes medical examination data, sensor data, signal data, and textual information data. All the collected data is used as input for data structural analysis and information extraction through computational models.

ii. *The Core layer,* in this layer, the collected data, are submitted in a harmonization process. This way it ensured that all the different sources of data will have the same numerical value ranges and categorical levels interpretation. Afterward, the content's structural analysis, of each variable, is applied to ensure the validity of the data collected. Finally, the final two-step procedure of the 2nd layer is storing and distributing, strictly and practically, the information extracted from the necessary OACTIVE models.

iii. *The Utility layer* is the layer where personalised interventions, alerts, analytics, and visualizations are extracted as an output of the whole procedure, with respect to the upper ontology framework.

3.1 Data Source Layer

3.1.1 Clinical Data Collection

OACTIVE uses medical imaging data from MR, X-Ray Imaging, to determine the patient's body and knee geometry. Marker-based and inertial sensors data for kinematic calibration. Ground reaction forces, and contact foot pressure data, calibrated from dynamometers, for kinetic calibration and also uses motion, load, and muscle activity, from EMG, for neurological calibration. The OACTIVE creates biomechanical models and simulations that fuse data about human movement, including joint kinematics, joint moments, and ground reaction forces, to produce estimates of muscle forces, muscle activations, and joint reaction forces.

Also, from the MRI, OACTIVE derives a full organ model. Bones are modelled, to provide all biomechanical indicators predicted by the model. Muscle and joint forces computed by the body model are applied as boundary conditions to the bone model, which predicts displacements, stresses, and strains at each point of the bone, during the entire motion. Knee computational models developed based on finite element representation. The geometry of the knee will rely on the MRI of the subject's anatomy. Image segmentation will establish the three-dimensional reconstruction of the tibiofemoral and patellofemoral joint components. Predominantly, the femur, tibia, and patella. Femoral, tibial, and patellar cartilage. Medial and lateral collateral ligaments, anterior and posterior cruciate ligaments, and patellar ligament. Menisci, and other passive components of the capsule, that may be deemed necessary for the investigation of tibiofemoral and patellofemoral joint mechanics. Finite element analysis is the preferred modelling and simulation tool for the proposed project due to its capacity to provide predictions of joint mechanics (knee kinematic-kinetic response) and bone or tissue mechanics. Based on the research question, abstractions (anatomical and material) both at the joint and bone level can be changed easily to balance the accuracy needed to address the scientific (or clinical) question and the cost required for the analysis.

Data generated from biomarkers used for the tissue level approach. The purpose of this tissue level engineering approach is to explore the potential of OA progress, by measuring specific molecular markers (biomarkers) in serum and fluid samples. Biomarkers, biomarker panels and methods for diagnosing osteoarthritis are disclosed, using measurement of the expression level of certain polypeptides in a test sample from a subject. The biomarkers include anabolic, catabolic as well as inflammatory molecules representing diverse biological pathways. Specifically, molecules that are released into biological fluids, including blood, during matrix metabolism of articular cartilage, subchondral bone, and synovial tissue are potential biochemical markers for the detection and monitoring of the process of osteoarthritis. The degree of articular inflammation will be associated with the disease progression and thus inflammation contributing to articular damage. To our knowledge, a comprehensive model correlating the secretion of OA biomarkers in the blood serum and the progress of the disease has not developed yet.

Moreover, osteochondral tissues from patients undergoing total joint replacement harvested from the post-surgery waste and cultured in vitro for up to 4 weeks within our osteochondral bioreactor that allows to separately interrogate the cartilaginous and osseous components. Cartilage cultured in serum-free, pro-chondrogenic medium and bone cultured in pro-osteogenic medium. Tissues from areas of healthy/macroscopically minimally damaged cartilage compared with tissues from areas of moderate to severe OA, assessing at different time points the following: (a) presence in the effluent media for cartilage and bone of catabolic markers, in particular of the known biomarkers in table 1.3.1, as well as bone markers such as the ratio of osteopontin (OPN)/RANKL, crosslinked collagen type I N-telopeptide, NTX1, and osteocalcin (OCN), and of metalloproteinases and other enzymes such as Cathepsin K; (b) microCT of the subchondral bone and of cartilage after the addition of an appropriate contrast agent; (c) quantitative real time PCR (qRT-PCR) for cartilage genes (collagen 2, aggrecan, collagen 10, Sox9, MMP-1, MMP-3, MMP-13) and bone genes (RUNX2, COL-I, OCN, BSP, and OPN accounting for osteoblasts, and

RANKL, TRAP, Cathepsin K, and MMP9 accounting for osteoclasts); (d) histology (Haematoxylin and Eosin, Alcian Blue, Alizarin Red, Masson's Trichrome, etc.) and immunohistochemistry (collagen II, collagen X, ALP, RANKL, TRAP, etc.). Imaging and biomarkers readings in patients matched the corresponding data obtained which will serve to relate clinical imaging and biomarkers with the biochemical and gene expression signature at the cellular level. The presence of bone is particularly important both for matching bone microCT data and for the key role bone in cartilage metabolism in health and disease. The use of the osteochondral bioreactor allows accounting for mechanisms of cartilage-bone crosstalk which could not otherwise be assessed. Human samples from cadavers serve as control to verify that osteochondral units extracted from macroscopically pristine areas of surgical wastes after knee replacement do not present a biochemical, structural, and cellular profile too different from normal. Also, porcine osteochondral plugs can be used as an alternative reference for healthy tissue. The results provide an initial cellular level input to the development of the hypermodel, thus covering the cellular scale.

The objective is to examine the relationship between biochemical markers for OA and clinical diagnosis. These results were used for the development of advanced computer modelling and simulation tools in order to be used in early diagnosis or prognosis of the disease. In detail, this aspect includes:

- Clinical evaluation of patients.
- Determination of serum concentrations of selected biomarkers levels on patients diagnosed with OA.
- Investigation of exosome biomarkers in terms of their relationship with OA development and progression.
- Develop a method to correlate/compare concentrations of biomarkers with clinical diagnosis and OA stage.

3.1.2 Biomarkers

Osteoarthritis (OA) is a degenerative disease of the joints and the most common form of arthritis that causes pain and mobility limitation and, thus, reduces independence and overall quality of life. Osteoarthritis is a complex disease in which biochemical and biomechanical factors are involved and occurs mostly in the weight-bearing joints of the lower limbs, such as the hip and in particular the knee in addition to the hands and spine, although almost any joint can be affected. Structurally, the whole joint is usually involved including diffuse and progressive loss of articular cartilage with concomitant changes in underlying bone (osteophyte growth and increased thickening or sclerosis) and soft tissue structures in and around the joint (synovitis, meniscal degeneration, ligamentous laxity and muscle weakness). These changes affect musculoskeletal function and body movement in general, reducing general mobility and increasing disability with age. It is, therefore, of particular concern that OA is one of the most common diseases affecting old age and the single most important cause of disability in older people.

Our project establishes a concept to automatically infer system-level mechanistic models of development from collected data. The analysis and automated reasoning convert the large number of secondary phenotypes, namely the expression of tissue-specific markers in terminal cells, into the affected cells, which allows the additional inference of developmental mechanisms. Importantly, the cellular-resolution phenotype data will enable us to design novel systems biology analyses with rich biological insights. The analyses will allow us to construct an explicit model of how cell differentiation progresses and to predict the gene-gene and cell-cell signalling networks during disease progression. The aim is to examine the relationship between biochemical markers for OA and clinical diagnosis. These results are used for the development of advanced computer modelling and simulation tools in order to be used in early diagnosis or prognosis of OA. The partner NIC assessed 6 different serum biochemical markers (3 Prognostic Biomarkers of Bone and Cartilage Degradation and Synthesis and 3 Inflammatory Biomarkers) and correlated them with the clinical outcomes.

Although OA was considered as a non-inflammatory joint disease it is recently demonstrated that specific inflammatory mediators are produced by articular tissues in OA and probably implicated in the pathogenesis and progression of the disease. Specifically, it is demonstrated that cytokines and prostaglandins produced by cartilage promote cartilage degeneration. On the other hand, synovitis is associated with greater risk of cartilage loss in patients with knee OA.

Table 1.3.1.: Validation and quantification of OA biomarkers by NIC

I. Prognostic Biomarkers of Bone and Cartilage Degradation and Synthesis

Biomarker	Process	BIPEDS[1] classification	Preliminary findings	ELISA type
Serum COMP	Cartilage degradation	Knee: BPD	Elevated level in Knee OA	Competitive inhibition
Serum HA	Osteophyte burden, synovitis	Knee: BPED	Elevated level in Knee OA	Sandwich type
Serum CPII	Type II collagen degradation	Knee: D	Elevated level in Knee OA	Competitive inhibition

II. Inflammatory Prognostic Biomarkers

Biomarker	Presumed source	Inflammatory biomarker subgroup	Preliminary Findings	ELISA type
IL-1β	Cartilage, Synovium, Bone	Cytokine / chemokines, complement and lipid mediators	Associated with Knee OA	Sandwich type
TNF-a	Cartilage, Synovium, Bone	Cytokine / chemokines, complement and lipid mediators	Elevated levels in Knee OA	Sandwich type
IL-6	Peripheral blood leukocytes	Transcriptomic biomarkers	Elevated levels in Knee OA	Sandwich type

In LEITAT already known biomarkers and under clinical validation process firstly investigated. These biomarkers are followed by means of ELISA, mostly commercially available. Also, we carry out studies of a new source of biomarkers on samples coming from patients being under treatment, so the evolution of the patient has been monitored. From fluid samples (blood, urine and synovial) exosomes isolated and characterised by miRNA content. Exosomes purified from samples by means of ultracentrifugation, andmiRNA is isolated for separate analysis.. This study is completed by a metagenomics providing a profile specific for each patient type. Faecal samples taken for each patient from HULAFE. For metagenomics analysis, bacterial 16SrRNA has been sequenced and compared through a bioinformatic analysis, before and after patient treatment. It provides information on the basal gut microbiota (dysbiotic) profile and its progression after treatment. Together with the existing scientific knowledge, it leads to the identification of the most relevant phylotypes, which are interpreted in terms of microbial ecology.

[1] BIPEDS classification: B: Burden of disease; I: Investigational; P: prognostic; E: Efficacy of Intervention; D: Diagnostic and S: Safety

3.1.3 IMU

In the OACTIVE framework, novel sensing systems developed the possibility as well to adapt available existing sensors for the intended tasks investigated. Modular approaches followed in order to guarantee a flexible configuration of the wearable platform according to the operating scenarios. A set of IMU sensors, worn by the means of accessories developed ad hoc, applied to the main limbs in order to gather further information on motion.

In the modelling framework, a novel sensing system developed as well as the possibility to adapt available existing sensors for the intended tasks investigated. A modular approach followed in order to guarantee a flexible configuration of the wearable platform according to the operating scenario (indoors, outdoors, AR gaming) assuring the adequate scalability to the wireless body area network (WBAN). An electronic board developed to acquire and transmit signals via Bluetooth and/or store them on board in a micro-SD card. The electronic board runs signal processing algorithms, elaborate the acquired signals to extract several features. A 9 DOF IMU embedded in the board allows collecting information on posture and/or activity of the patient. A set of external IMU sensors, worn by the means of accessories developed ad hoc, applied to the lower limbs in order to gather further information on motion. A careful investigation conducted on state of the art to choose the solution that can fit best project requirements. Then prototypes produced to allow users to be monitored for up to 8 running hours in a comfortable way, indoor and outdoor, during standard daily activities. The patience movements collected from AWS cloud database and we take as outputs the patient's posture and activity information. At least 500 time series, data points have been used.

3.1.4 Augmented Reality (AR)

OACTIVE relies on the AR gaming technology offering both clinical assessment and rehabilitation options, usually not available with traditional rehabilitation methods. It aims at exploiting haptic and vision technologies to provide patients with assistive visual and contact feedback while performing games/rehabilitation as well as medical staff with biomechanical indicators for assessment and diagnosis support. It goes beyond the existing AR rehabilitation programs by: (i) expanding & improving the currently limited opportunities for rehabilitation scenarios, (ii) enhancing primitive spatial and temporal training scenarios, (iii) addressing staff and facility limitations as well as human factors, (iv) creating user friendly interfaces and integrating interactive environment, (v) accurately implementing crucial stimuli (force sensing, visual information) together to have a real impact on the game task completion performance.

The purpose of AR gaming is to simplify the development of OACTIVE type games and hosting patients into a single interface in order to gather data that can interact between the patient and the medical examiner. The game framework hosts all the games that are developed, for all the three different platforms, by using the decided hardware architecture in combination with the selected graphic engine. Finally, this framework fulfils two objectives: (i) simplification of the game development and (ii) host the games in a unified interface where the users are signing in, interacting, playing, logging sessions and finally exporting clinical pictures of the patient. A graphics general menu created as it facilitates the in-game menu creation to sign in and select a game. A communication system, depending on the capabilities and protocols of the tracking devices, makes the tracking data available in real-time. Filtering and synchronization into a common reference frame needed, especially in case of more than one sensor. This module, depending on the graphics engine, provides the appropriate data format. Multiple device render modules open the use to many different rendering devices, even if OACTIVE is intended to use AR devices. Because the user's performance has to be checked from therapists through more complete data, than the user's score data that is recorded through the games, the log system module records this data in real-time and it is uploaded to the user's database for each session. Finally, a server that contains a database, allows final users and professionals to access their profiles and update them with new exercises, log their gaming sessions, and leave messages as also the connection between the game system and user's profile to update the corresponding tracking data as an output.

AR glasses are responsible for the presentation of the OACTIVE games to the patient. AR glasses is a hardware device that is used for monitoring the Augmented Reality environment to the patient during the implementation of the personalized treatment plan. They inform the patient about the progress of the disease through exercises that take place in the AR games environment.

3.1.5 Exogenous/environmental risk factors

Social scientists create models to describe how individuals' function, either on their own or in complex social settings. Behaviour models identify discrete drives that researchers use these to describe and try to explain human behaviour. Mechanistic models can also attempt to explain human behaviour in terms of biochemical events in the brain and body. OA is not easy to define, predict or treat. Despite extensive research costing many billions of Euros, no drugs have been proven to modify the biological progression of OA, and only a few treatments are proven to relieve symptoms beyond the placebo effect. Given this failure to find an effective post-diagnosis treatment, attention should turn to preventing or delaying the onset of cartilage degeneration. Identification of the risk factors for developing arthritis has been limited by a lack of longitudinal data, as well as an absence of reproducible, non-invasive methods to measure changes in joint morphology and function. As a result, the disease processes governing osteoarthritis progression are still poorly understood. Although most of the existing research has focused on factors associated with the disease, the lack of longitudinal data examining the factors associated with disease onset and progression has resulted in a lack of prevention and treatment interventions that aim to target the most appropriate modifiable risk factors and, therefore, prevent or delay the onset and/or progression of the disease.

3.1.6 Medical

Medical risk factors known to influence development of the disease include advanced age, gender, hormonal status, body weight or size, usually quantified using body mass index (BMI), and a family history of disease. Additionally, there is now evidence supporting a strong genetic association. Other known risk factors for the onset and progression of OA include joint loading during occupational or physical activity and sports participation, muscle weakness, a past history of knee injury and joint operations (ACL injury and reconstruction, meniscal damage and partial meniscus removal) and depression. Although many of the above factors are fixed, other risk factors such as body weight, physical activity and occupation are modifiable. For many people occupational activities involving physically demanding jobs, such as manual handling of heavy loads or prolonged kneeling are associated with the disease.

3.1.7 Socioeconomic

Social context may interact with pathophysiological processes and individual-level variables to influence health outcomes in those living with OA. Evidence exists linking lower levels of individual socioeconomic status, life course approach and poorer health outcomes in OA. Recent data suggest that the social or socioeconomic environment of an individual may be relevant to arthritis prevalence and health outcomes as well. There are several dimensions to measure social position and social context that relate to population with OA (Figure 3.4). Compositional attributes of socioeconomic status are measured at the individual level and have commonly included variables such as occupation type (professional/managerial etc.), level of educational attainment, income (both individual and household), home ownership, and social class. The socioeconomic context of communities may affect characteristics of the environment of communities to which all residents are exposed, regardless of their social position. The complex nature of social context can be described in terms of both physical and social components, as well as their objective (i.e., actual) or subjective (i.e., perceived qualities) and their scale or immediacy to individuals and groups. The life course approach is particularly relevant in understanding the long-term effects of both social position and social context on chronic diseases. Life course measures encompass the aforementioned factors while explicitly

including the variable of time, which allows researchers to study disease outcomes across generations and at the individual level beginning from gestation and through childhood, adolescence, young adulthood, middle-age and senescence. Several proposed theoretical models that explain the influence of life course on disease risk may include critical or sensitive periods, which state that exposures during a specific window of time may have long-lasting effects that increase disease risk.

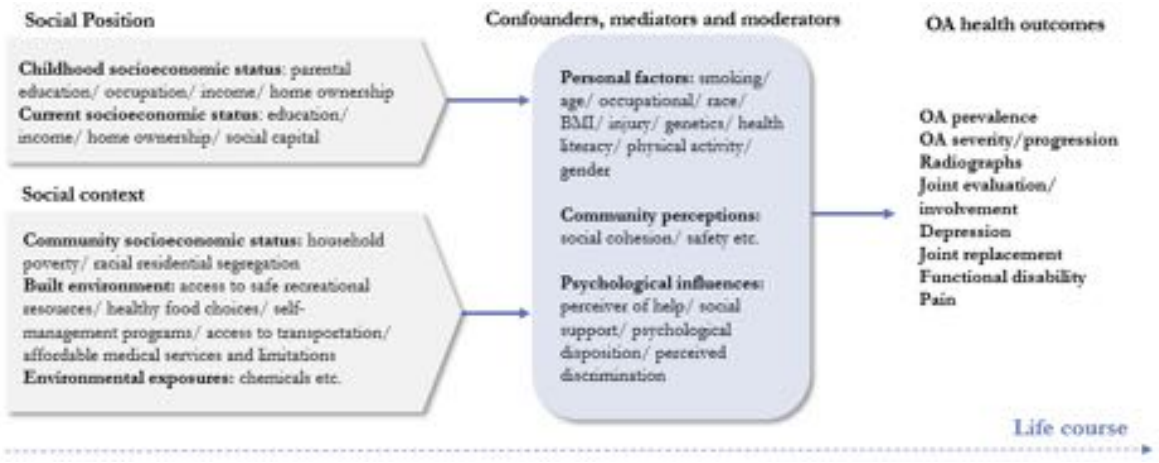


Figure 3.4: Socioeconomic risk factors considered in OACTIVE.

The life course approach is particularly relevant in understanding the long-term effects of both social position and social context on chronic diseases. Life course measures encompass the aforementioned factors while explicitly including the variable of time, which will allow us to study disease outcomes across generations and at the individual level beginning from gestation and through childhood, adolescence, young adulthood, middle-age and senescence. Our models intend to explain the influence of life course on OA risk may include critical or sensitive periods, which state that exposures during a specific window of time may have long-lasting effects that increase disease risk. The aim of this project is to analyse and validate socioeconomic status of patients with OA (Figure 3.4). More specifically in this task, data from references collected focused on social position, explored and exogenous factors related to the individual such as education, income, occupation, physical activity and behaviour analysis. There are several dimensions to measuring social position, social context personal factors that relate to populations with OA.

3.1.8 Social Position

Compositional attributes of socioeconomic status measured at the individual level including variables such as occupation type (professional/managerial etc.), level of educational (lower levels of educational attainment have frequently been associated with the increased prevalence, morbidity and mortality of many chronic diseases), attainment, income (both individual and household), home ownership, and social class. Other social position factors that are investigated are the childhood socioeconomic status: parental education; occupation; income; home ownership etc. The current socioeconomic status also examines for example the education status; income; home ownership and social capital.

3.1.9 Social Context

The socioeconomic context of communities may affect characteristics of the environment of communities to which all residents are exposed, regardless of their social position. The following social context risk factors collected and analysed in this context for OA: (i) Community socioeconomic status: this index is an

area-based measure that represents the average level of disadvantage across a geographical area (e.g., household poverty; racial residential segregation etc.) and is a composite of the income, educational attainment, levels of public sector housing, unemployment and jobs in relatively less skilled occupations. (ii)Build environment: of the neighbourhood environment examined like community mobility barriers (i.e., uneven sidewalks or other walking areas; parks and walking areas that are easy to get to and easy to use; safe parks or walking areas; places to sit and rest at bus stops, in parks, or in other places where people walk; curbs with curb cuts) and transportation facilitators (i.e., public transportation close to home; public transportation with adaptations for people who are limited in their daily activities; and adequate handicap parking, able to drive, have a car available to you at your home). (iii) Environmental exposures: physical environmental factors contributing to health outcomes.

Our model also includes commonly related personal risk factors associated with OA, including: (i) personal factors: smoking, age, gender, occupation-tasks, BMI (i.e., obesity), injury, family history, race and physical activity. Many of these personal risk factors are also associated with one's current social position. In addition (ii) Community perceptions, also included with mediator and moderators, such as community perceptions and (iii) several dimensions of psychological influences (i.e., perceived helplessness, social support/coping resources, psychological disposition, perceived discrimination and/or catastrophizing), that have been demonstrated to be associated with health outcomes. These factors provide a more transparent relationship to link social determinants to arthritic health outcomes but also, allow us to identify areas for OA behavioural interventions.

3.2 Core Layer

This section presents the OACTIVE semantic model with foundations constituted on which the OACTIVE Ontology is built. The OACTIVE semantic model consists of four specific ontologies, namely, i) Finite Element Model, ii) Musculoskeletal Model, iii) In vitro model of osteochondral tissue and iv) Patient Specific Model Creation. The first three are combined for the creation of the Personalized intervention through AR. Also, nine upper ontologies namely, MRI; X-Ray Imaging; Demographics; Physical Examination; Social Participation; Socioeconomics; Scales; Anamnesis and Biomarkers (Blood / Urine Tests) are combined for the Patient Specific Model Creation.

In order to design the taxonomies of describing OACTIVE project, the agile software development called ODMedit (ODM) has been applied to (i) define the application domain boundaries and (ii) capture elements definition. ODMedit tool is a semi-automatic approach where data integration is simplified if the same code is applied for all data items with the same, or at least very similar, meaning. By composing a top-level overview, abstract concepts facilitate system architecture planning and optimization.

ODMedit is intended to foster uniform semantic annotation. A random set of data elements from an established data standard was selected to test this feature. For each of those data elements available UMLS codes were identified with the UMLS Metathesaurus Browse. Suitable codes were identified from the output of the UMLS Metathesaurus Browser by manual review. Available annotations in the MDM portal were analysed for each data element regarding uniform semantic annotation and compared to UMLS codes.

In particular, data harmonization between populations to align data from the various OA groups before applying patient recruitment algorithms, where transformation of the data, from multiple sources, involved in this process. Because there are Inherent Structural Differences (ISD) in the datasets due to the variability in data collections protocols, the identified ISD's categorised into semantics, syntactics, and transformation differences.

Semantics category concerns those similar data elements that could be assigned to higher level common data elements, e.g., "Pain that exists in the bending of the knee" and "pain that exists in "knee extension" were mapped to "knee pain".

While syntactics is about mapping where there are differences in coding practices between datasets (e.g. {1 = positive, 2 = negative} vs {1 = no, 2 = yes}. To cope with the manipulations of representation schemes, where functions employed to map from one type to another (e.g. 'bodily pain score' = round (mean((6 - SF7) * 20, (5 - SF8) * 25)), where SF7 = 'bodily pain in the past 4 weeks' with values ranging from 1 = 'no pain' to 6 = 'very severe' and SF8 = 'pain interference in your work in the past 4 weeks' with values ranging from 1 = 'not at all' to 5='extremely').

Finally, all the above processes executed as an automated software harmonisation pipeline where data – pre-processing, semantic processing (identification of common data concepts and creation of mapping between concepts), as also syntactic processing (transformation and mapping) and quality control methods (value ranges, consistencies, distributions) combined.

3.2.1 OACTIVE Upper Ontology Framework (UOF)

Data generated includes quantitative data about the patient's anatomy and physiology, patient's health phenomenological data, movement data (being collected by clinical centres, biomechanics laboratories, wearable sensors and smartphone accelerometers), clinical data (questionnaire data for entire cohort, physical activities, exam data, bio specimen assay data and bone ancillary study data of cohort), bio-specimens (serum, plasma, and urine and DNA for entire cohort), images (replacement images, X-ray and MRI images for entire cohort), image assessments, biomarkers (serum and urine biochemical biomarkers data, mass spectrometry data of pooled serum and urine samples of subjects with and without knee OA, etc.), data on population statistical variability and healthcare value chain data. OACTIVE semantic framework will be in the form of an ontology network where each scenario specific ontology will be connected to an upper ontology through generalization. The upper ontologies are described through UML graphs below.

The graph representation of the OACTIVE upper ontology is presented in Figure 3.5. In this figure there is a description of the 1st-3rd Levels of Information Extraction, higher functions.

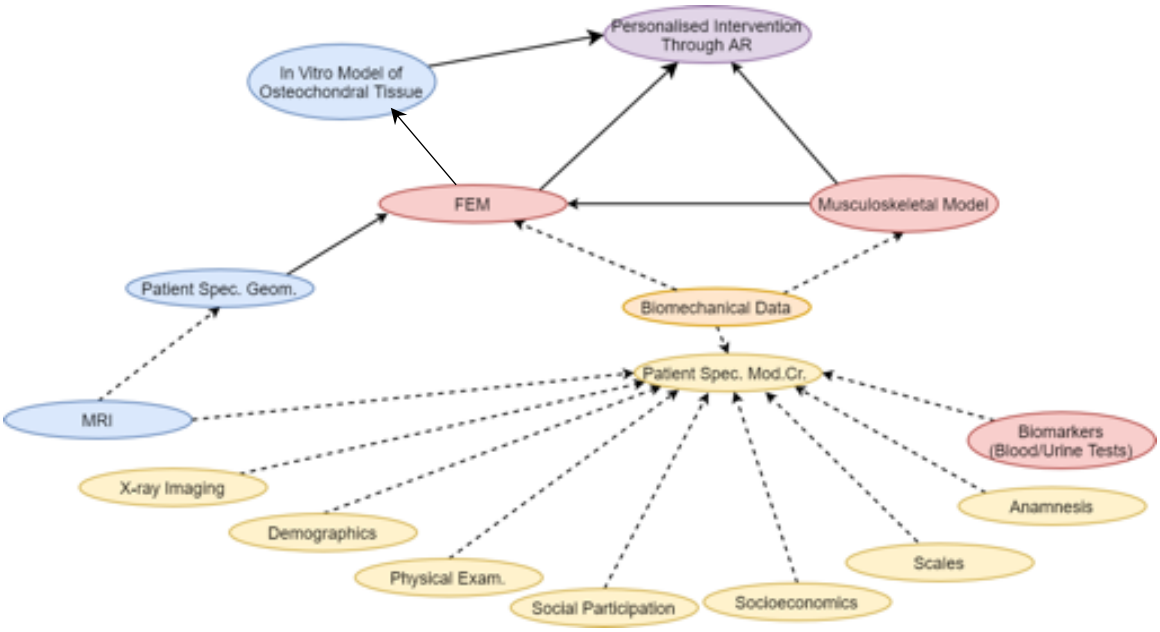


Figure 3.5: OACTIVE Ontology Framework classes visualization

The graph representation of the *Demographics* ontology is presented in Figure 3.6. In this figure feature extraction of *Demographics* data variables is represented. The variables used for feature extraction in *Demographics* data namely are: *ID code, Data provider, Date, Sex, Age (years), Birth country, Ethnicity, Occupation.*

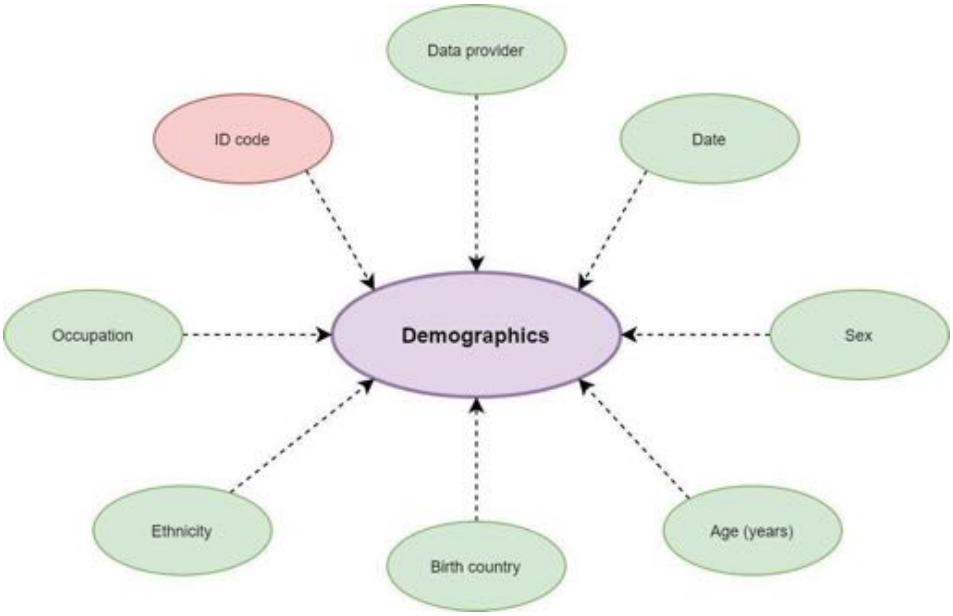


Figure 3.6: Demographics upper ontology

The graph representation of the *Socioeconomics* ontology is presented in Figure 3.7. The *socioeconomic* context of communities may affect characteristics of the environment of communities to which all residents are exposed, regardless of their social position. The following social context variable collected and analysed in this context for OA. Namely these variables are *Name, Marital status, Level of education of the patient and his/her parent*, *Residency, Household income and Housing status.*

The graph representation of the *Social Participation* ontology is presented in Figure 3.8. The Social Participation ontology consists of eight "question" variables which aim to develop the patient's social activity. Namely, the variables of Social Participation ontology are:

- *Name*
- *Have you taken part in a club, interest group or activity group, church or other similar activity?*
- *Have you been to a cultural or educational event such as the cinema, theatre, museum, talk or course?*
- *Have you eaten out?*
- *Have you been out to a pub, café or tearoom?*
- *Have you been to a public event?*
- *Have you taken part in an organised games afternoon or evening? For instance, bingo, quiz or card games*
- *Have you been on a day trip organised by a club or society?*
- *Have you carried out committee work for a club, society or other group?*
- *Have you done any organised voluntary work?*

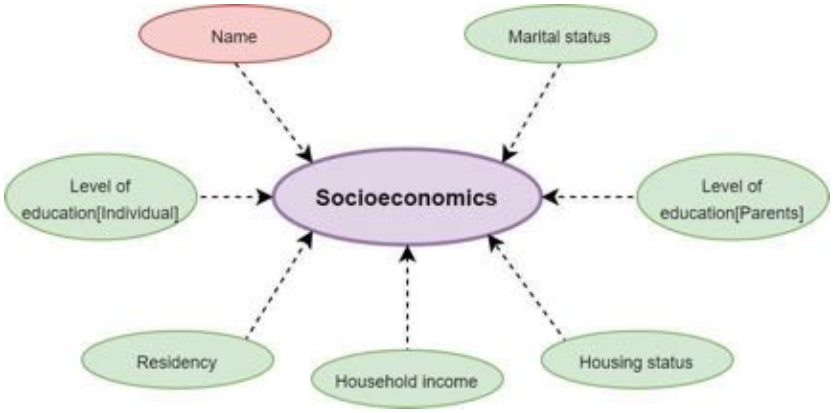


Figure 3.7: Socioeconomics upper ontology

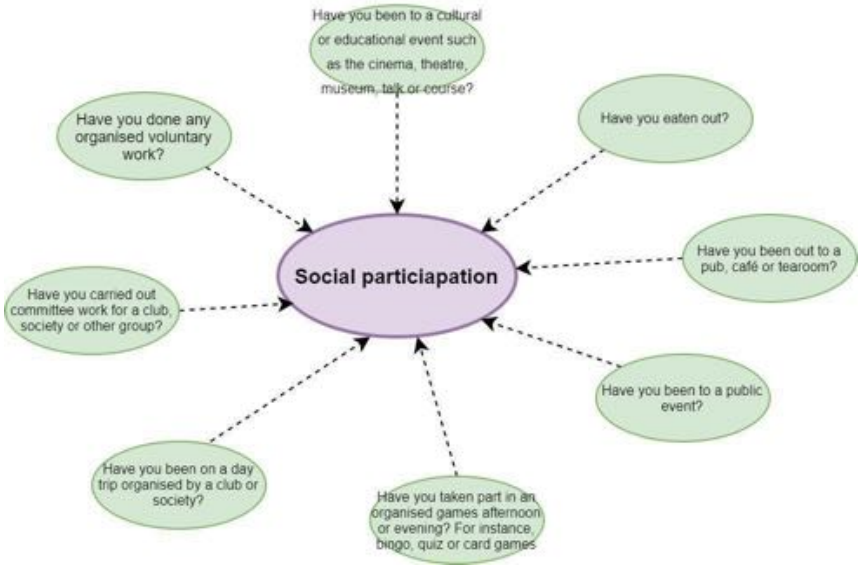


Figure 3.8: Social Participation upper ontology

As part of the patient's Medical history data *Anamnesis* data (Figure 3.9), as also *Physical examination* data (Figure 3.9), are used for features extraction. Many initial encounters with patients will include asking the patient's medical history, while subsequent visits may only require a review of the medical history and possibly an update with any changes. Obtaining a medical history can reveal the relevant chronic illnesses and other prior disease states for which the patient may not be under treatment but may have had lasting effects on the patient's health. The medical history may also direct differential diagnoses. Medical history data includes an inquiry into the patient's medical history, past surgical history, family medical history, social history, allergies, and medications the patient is taking or may have recently stopped taking. *Anamnesis* data consists of variables, namely: *ID code, Sex, Group, Data Provider, Any current medication, Birth country, Personal history of OA, Occupation, High blood pressure, Type of sport, Sports frequency, Family OA history, Meniscal damage, Do you have knee OA?, Knee instability, Have you ever been told that you have OA of your knee by a doctor?, Regular sport leisure activity, Occupational risk, Smoking, Number of cigarettes per day, Alcohol, Hormonal status (women), Knee extensor muscle weakness, Hip OA/Surgery, Resting VAS, Walking VAS, Previous knee injuries, Neuropathic component, Knee pain, Pain side, Pain Rhythm, Time since pain start, Ethnicity, Knee pain [NHANES-type questions]*.

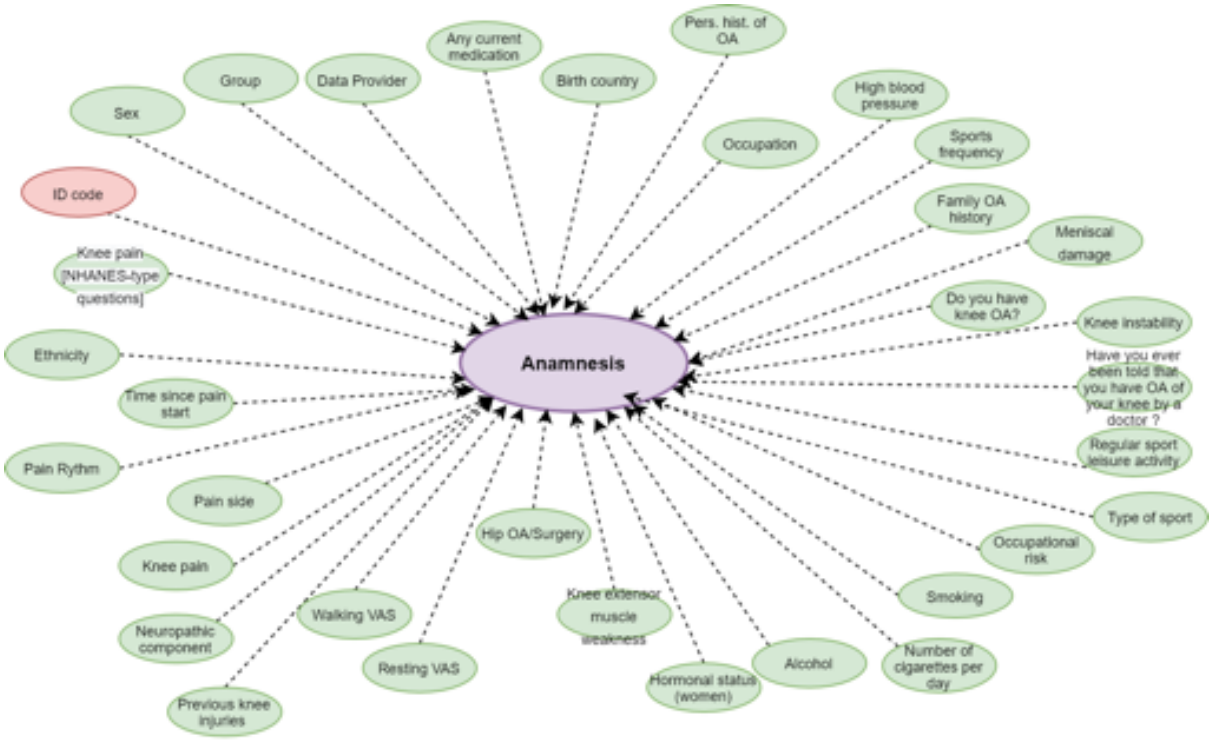


Figure 3.9: Anamnesis upper ontology

The graph representation of the *Physical examination* ontology is presented in Figure 3.10. *Physical examination* is the process of evaluating objective anatomic findings through the use of observation, palpation, percussion, and auscultation. The information obtained must be thoughtfully integrated with the patient's history and pathophysiology. Moreover, it is a unique situation in which both patient and physician understand that the interaction is intended to be diagnostic and therapeutic. The physical examination, thoughtfully performed, should yield 20% of the data necessary for patient diagnosis and management.

The *physical examination* is a key part of a continuum that extends from the history of the present illness to the therapeutic outcome. If the history and physical examination are linked properly by the physician's reasoning capabilities, laboratory tests should in large measure be confirmatory. The physical examination, however, can be the weak link in this chain if it is performed in a perfunctory and superficial manner. Understanding the pathophysiologic mechanism of a physical abnormality is essential for correct diagnosis and management.

The variables extracted from the *Physical examination* are: *ID code, Mass (Kg), Height (m), BMI , Joint line tenderness [Left], Joint line tenderness [Right], Patellofemoral pain [Left], Patellofemoral pain [Right], Crepitus [Left], Crepitus [Right], Right Flexion angle, Right Extension angle, Flexion deformity [RIGHT], Left Flexion angle, Left Extension angle, Flexion deformity [LEFT], Muscle atrophy LEFT-Specify the measurement of LEFT limb (cm), RIGHT-Specify the measurement of RIGHT limb (cm), Knee laxity [Left], Knee laxity [Right], Joint proprioception [Left], Joint proprioception [Right] Abdominal perimeter (cm), LEFT EXTENSION Dynamometric/HHD evaluation of knee strength, RIGHT EXTENSION Dynamometric/HHD evaluation of knee strength, LEFT FLEXION Dynamometric/HHD evaluation of knee strength, RIGHT FLEXION Dynamometric/HHD evaluation of knee strength, 5 Sit to stand test (sec), Walking Speed: 10 meter walk test (s), Knee_Morphology, Joint_Effusion, Increased_Local_Temperature, Local_Redness, Bakers_Cyst, Muscle strength (MRC) - LEFT [Hip flexors], Muscle strength (MRC) - LEFT [Hip abductors], Muscle strength (MRC) - LEFT [Knee extensors], Muscle strength (MRC) - LEFT [Knee flexors], Muscle strength (MRC) - LEFT [Plantar flexors], Muscle strength (MRC) - RIGHT [Hip flexors],*

Muscle strength (MRC) - RIGHT [Hip abductors], Muscle strength (MRC) - RIGHT [Knee extensors], Muscle strength (MRC) - RIGHT [Knee flexors], Muscle strength (MRC) - RIGHT [Plantar flexors], Leg length discrepancy.

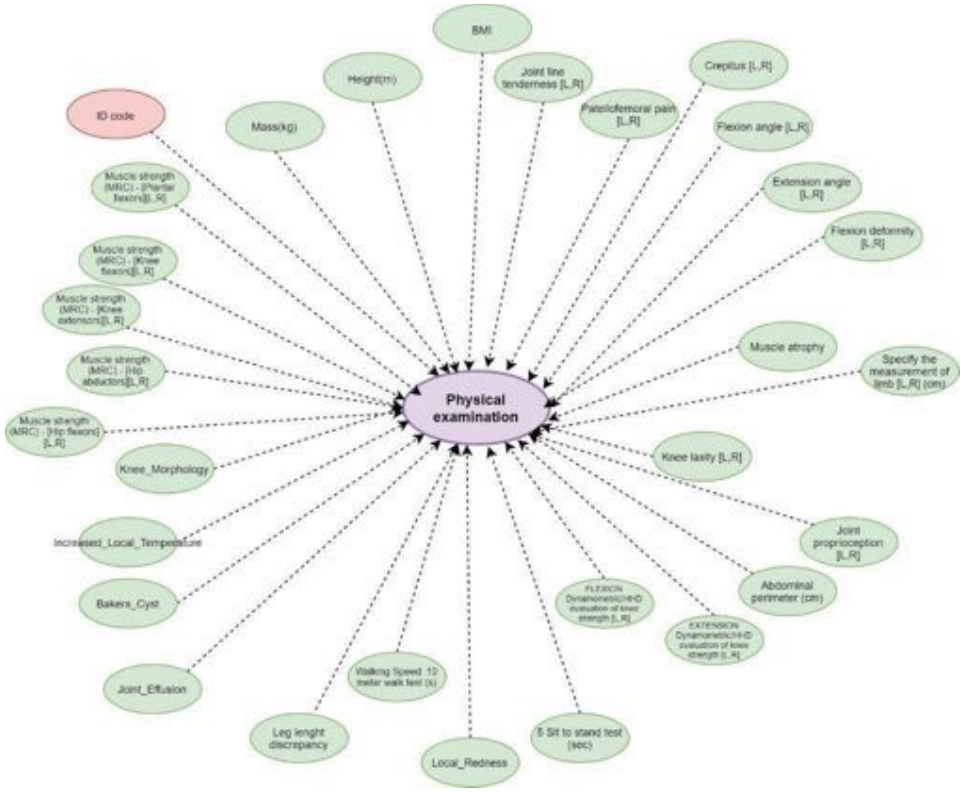


Figure 3.10: Physical examination upper ontology

The graph representation of the *Blood Tests and Urine Tests* ontology is presented in Figure 3.11 and Figure 3.12 respectively.

Already known biomarkers under clinical validation processes, will be included. These biomarkers will be followed by means of ELISA, mostly commercially available. Also, studies were carried out for a new source of biomarkers on samples coming from patients being under treatment, so the progression of patients can be monitored. From fluid samples (blood, urine and synovial) exosomes isolated and characterised by miRNA content. Exosomes purified from samples by means of ultracentrifugation. microRNA is isolated.

The variables extracted from the *Blood Tests* are namely *Name, Uric acid (mg/dL), Total cholesterol (mg/dL), HDL-cholesterol (mg/dL), LDL-cholesterol (mg/dL), Triglycerides (mg/dL), Protein C reactive (mg/L), Vitamine D (mg/L), PTH (pg/mL), Glycated haemoglobin (%), Serum COMP (ng/mL), Serum HA (ng/mL), PIICP (pg/mL), IL-1β (pg/mL), TNF-a (pg/mL), IL-6 (pg/mL) and Vitamine K.*

The variables extracted from the *Urine Tests* are namely *Name, Urine COMP (ng/ml), Urine HA (ng/mL), Urine PIICP (pg/mL), Urine IL-1β (pg/mL), Urine TNF-a (pg/mL) and Urine IL-6 (pg/mL).*

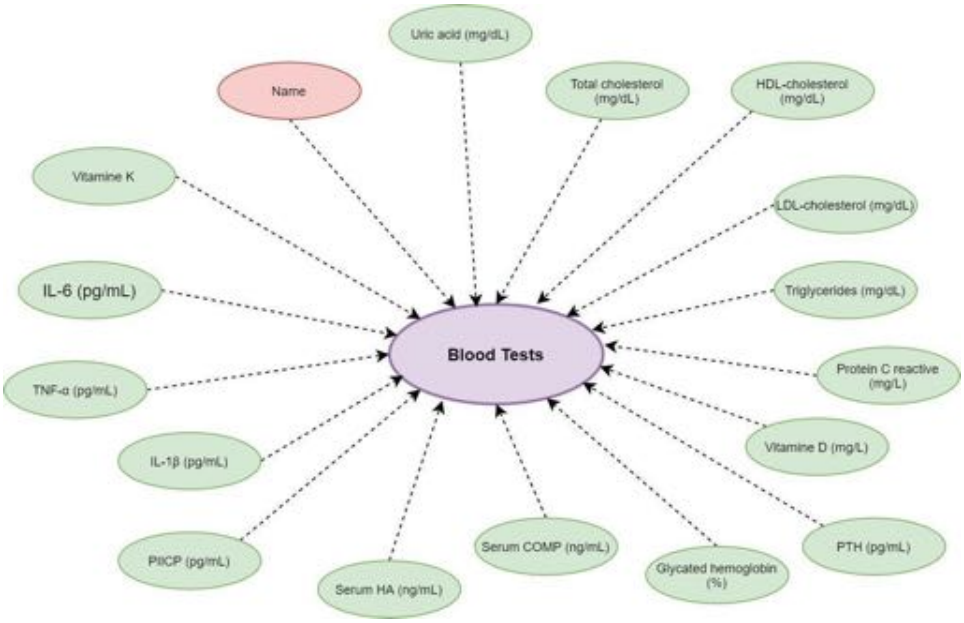


Figure 3.11: Blood Tests upper ontology

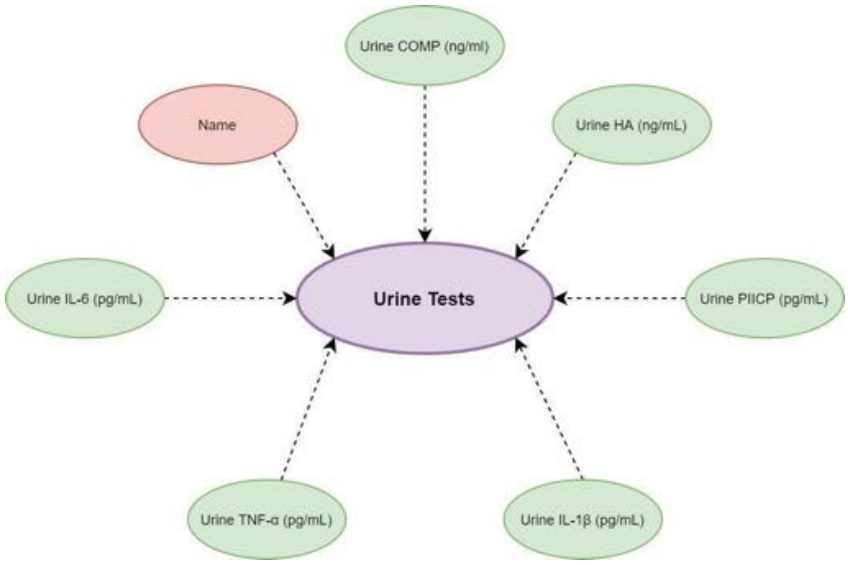


Figure 3.12: Urine Tests upper ontology

The graph representation of the *Scales* ontology is presented in Figure 3.13. In this figure feature extraction of *Scales* data variables is represented. In this upper ontology the results generated from state-of-the-art scales of questionnaires are contained. The scales used in the OACTIVE project are FACHS, WOMAC, KOOS, HADS and GADS. For each of the scales the necessary data are saved in this upper ontology.

Functional Ambulation Classification of the Hospital at Sagunto (FACHS)[2,3] is a validated scale to assess gait and categorizing patients into different walking abilities, with a simple and quick management (See Table 3.1).

Table 3.1 Functional Ambulation Classification of the Hospital at Sagunto (FACHS)[2,3]

Level 0	Non-ambulation.
Level 1	Non-functional or dependent ambulation.
Level 2	Household ambulation.
Level 3	Surroundings of the house ambulation (neighborhood).
Level 4	Community ambulation.
Level 5	Normal ambulation.

Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) is a questionnaire used to assess the health status of osteoarthritis patients introduced in 1988. It is consisted of 33 items which evaluates the health and function of the patient from various aspects including: clinical symptoms (5 questions), severity of joint stiffness (2 questions), degree of pain (9 questions), and activity of daily living (17 questions). Osteoarthritis of the knee is the most common chronic joint disease that involves middle aged and elderly persons. There are different clinical instruments to quantify the health status of patients with knee osteoarthritis and one example is the WOMAC score that has been translated and adapted into different languages. Reliability testing resulted in a Cronbach's alpha of 0.917, showing the internal consistency of the questionnaire to be a reliable tool. Different validation studies of WOMAC make this clinical instrument usable for knee OA evaluation before and at follow-up of treatment protocols including nonoperative and operative. These validation studies for WOMAC index also enable clinical investigators to assess those clinical outcome reports using this index for knee OA management from different parts of the world collectively. The WOMAC score is saved under one variable in the upper ontology.

The Knee injury and Osteoarthritis Outcome Score (KOOS) was developed as an extension of the WOMAC Osteoarthritis Index with the purpose of evaluating short-term and long-term symptoms and function in subjects with knee injury and osteoarthritis. The KOOS holds five separately scored subscales: Pain, other Symptoms, Function in daily living (ADL), Function in Sport and Recreation (Sport/Rec), and knee-related Quality of Life (QOL). The KOOS has been validated for several orthopaedic interventions such as anterior cruciate ligament reconstruction, meniscectomy and total knee replacement. In addition, the instrument has been used to evaluate physical therapy, nutritional supplementation and glucosamine supplementation. The effect size is generally largest for the subscale QOL followed by the subscale Pain. The KOOS is a valid, reliable and responsive self-administered instrument that can be used for short-term and long-term follow-up of several types of knee injury including osteoarthritis. The measure is relatively new and further use of the instrument will add knowledge and suggest areas that need to be further explored and improved.

The main reason for developing a single instrument with the purpose of covering several types of knee injury and including osteoarthritis (OA), was that traumatic knee injuries often causes concomitant damage to multiple structures (ligaments, menisci, cartilage, etc.) and frequently lead to the later development of OA. To be able to follow patients after a trauma and to gain insight into the change of symptoms, function etc. over time, a questionnaire which covers both the short-term and long-term consequences is needed.

[2] Viosca E., et al. Proposal and Validation of a New Functional Ambulation Classification Scale for Clinical Use. Arch Phys Med Rehabil 2005;86:1234-1238.
[3] Viosca E., et al. Walking Recovery After an Acute Stroke: Assessment With a New Functional Classification and the Barthel Index. Arch Phys Med Rehabil 2005;86:1239-1244.

Prior instruments such as the Lysholm knee scoring scale have focused only on the short-term consequences and instruments such as the WOMAC Osteoarthritis Index only on the long-term consequences. An instrument intended for follow-up of these patients needs to adequately monitor both the acute injury consequences in the physically active and younger patients, and the chronic outcome in the older. The KOOS index data is saved in the upper ontology using five different variables for the five different metrics KOOS index evaluates. Namely, KOOS PAIN (%), KOOS SYMPTOMS (%), KOOS ADL (%), KOOS QOL (%) and KOOS SPORT/REC (%) are the variables of the KOOS index.

There is a need to assess the contribution of mood disorder, especially anxiety and depression, in order to understand the experience of suffering in the setting of medical practice. Most physicians are aware of this aspect of the illness of their patients, but many feel incompetent to provide the patient with reliable information. The Hospital Anxiety and Depression Scale, or HADS, was designed to provide a simple yet reliable tool for use in medical practice. The term 'hospital' in its title suggests that it is only valid in such a setting, but many studies conducted throughout the world have confirmed that it is valid when used in community settings and primary care medical practice. The HADS data can be saved under one variable in the Scales upper ontology.

The GAD-7 is commonly used as a measure of general anxiety symptoms across various settings and populations. However, there has been disagreement regarding the factor structure of the GAD-7, and there is a need for larger studies investigating the psychometric properties of the measure. Patients undergoing treatment, both inpatient and outpatient patients, completed the GAD-7 at pre- and post-treatment. Measures of depression, well-being, and other anxiety measures were also completed, making it possible to investigate convergent and divergent validity. Internal consistency and convergent validity were excellent for the total sample, and there was acceptable variation related to treatment groups. The GAD-7 has excellent internal consistency, and the one-factor structure in a heterogeneous clinical population was supported. The needed data from the GADS scale is saved under one variable.

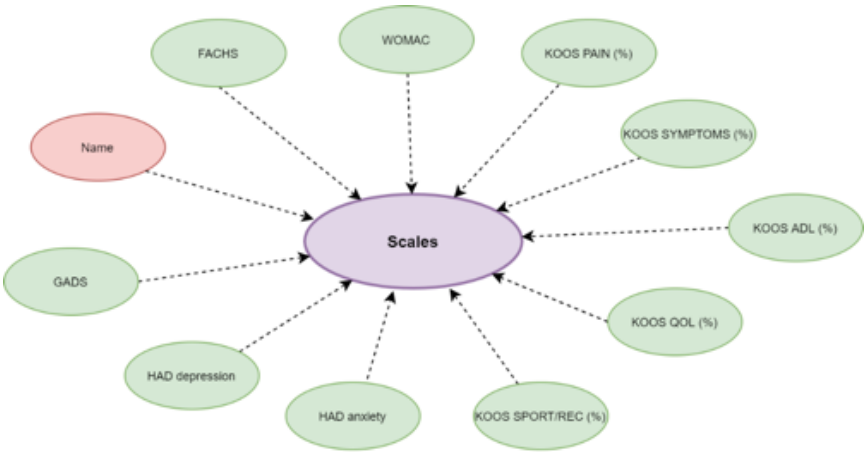


Figure 3.13: Scales upper ontology

The graph representation of the *RX* ontology is presented in Figure 3.14. RX ontology is a repository for X-Ray Kinematics to pose estimation. The kinematics model is used to calibrate joint positions and orientations in the body segments of a skeletal model. For the OACTIVE the necessary data extracted from an X-Ray, for the RX upper ontology, are *Name, Leg-length inequality, Leg-length inequality measure (mm), Right Knee alignment, Right radiographic angle (Knee alignment), Left Knee alignment, Left radiographic angle (Knee alignment), Right Kallgren and Lawrence, Left Kallgren and Lawrence, RIGHT patellofemoral lateral angle, LEFT patellofemoral*

lateral angle, RIGHT Lateral deviation patella (mm), LEFT Lateral deviation patella (mm), RIGHT Congruence angle and LEFT Congruence angle.

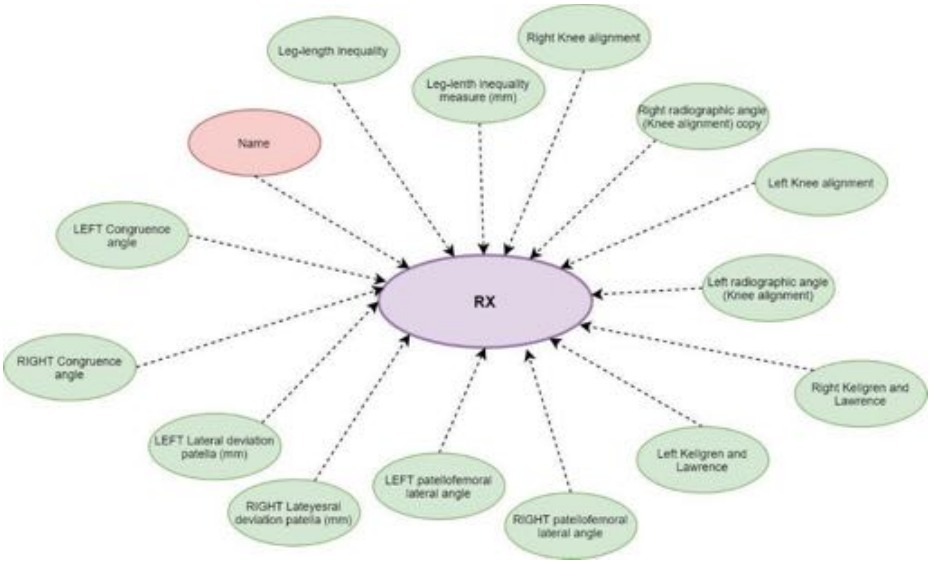


Figure 3.14: RX upper ontology

MRI upper ontology (Figure 3.15) consists of necessary data for the extraction of information needed for the patient's specific geometry, for each knee. In particular, image segmentation establishes the three-dimensional reconstruction of the tibiofemoral and patellofemoral joint components; predominantly the femur, tibia, and patella; femoral, tibial, and patellar cartilage; medial and lateral collateral ligaments, anterior and posterior cruciate ligaments, and patellar ligament; menisci; and other passive components of the capsule that may be deemed necessary for the investigation of tibiofemoral and patellofemoral joint mechanics. Each tissue assigns appropriate density to allow accurate nonlinear dynamics simulations. The bones assumed to be rigid based on their relatively high stiffness when compared to other soft tissue structures. Constitutive relationships representative of typical nonlinear stress-strain behaviour of the other underlying tissue structures relies on well-established literature available for ligaments, cartilage, and meniscus. Ligament representation incorporates in situ ligament strains. Ligament-to-bone wrapping, ligament-to-ligament wrapping, and contact interactions in between cartilage and in between cartilage and menisci are defined. Contact between components modelled as frictionless based on the low friction in synovial joints. Loading and boundary condition specifications allow prescription of the tibiofemoral and patellofemoral loads (forces and moments) or kinematics (rotations and translations) or a combination of those. In return, the model outputs unspecified kinetic variables or degrees of freedom. Complete stress-strain state for the tissues also is calculated as part of the solution process.

The MRI upper ontology is evaluated through the MOAKS score, which consists of seven upper classes which are: *Bone marrow lesions (BMLs) and cyst, Articular cartilage, Osteophytes: score, Ligaments and tendons: score, Meniscal morphology, Meniscal extrusion: score, Hoffa's fat synovitis, Synovitis /effusion and Periarticular features*. Each upper class consists of variables with the exception of Meniscal morphology which contains two sub-classes namely, Medial and Lateral. The variables for each upper class and the two subclasses are presented below.

- *Name*
- *Bone marrow lesions (BMLs) and cyst*
  - Trochlea medial containing the values for BML size, BML number and BML%V.Cyst
  - Trochlea lateral containing the values for BML size, BML number and BML%V.Cyst

- Femur central medial containing the values for BML size, BML number and BML%V.Cyst
- Femur central lateral containing the values for BML size, BML number and BML%V.Cyst
- Femur posterior medial containing the values for BML size, BML number and BML%V.Cyst
- Femur posterior lateral containing the values for BML size, BML number and BML%V.Cyst
- Patella lateral containing the values for BML size, BML number and BML%V.Cyst
- Patella medial containing the values for BML size, BML number and BML%V.Cyst
- Tibia anterior medial containing the values for BML size, BML number and BML%V.Cyst
- Tibia anterior lateral containing the values for BML size, BML number and BML%V.Cyst
- Tibia central medial containing the values for BML size, BML number and BML%V.Cyst
- Tibia central lateral containing the values for BML size, BML number and BML%V.Cyst
- Tibia posterior medial containing the values for BML size, BML number and BML%V.Cyst
- Tibia posterior lateral containing the values for BML size, BML number and BML%V.Cyst
- Tibia subspinous subregion containing the values for BML size, BML number and BML%V.Cyst
- Other subregions

- *Articular cartilage*
  - Femur: central medial: Cartilage loss % (full + partial)
  - Femur: central medial: Cartilage loss % (full)
  - Femur: posterior medial: Cartilage loss % (full + partial)
  - Femur: posterior medial: Cartilage loss % (full)
  - Femur: posterior lateral: Cartilage loss % (full + partial)
  - Femur: posterior lateral: Cartilage loss % (full)
  - Femur: central lateral: Cartilage loss % (full + partial)
  - Femur: central lateral: Cartilage loss % (full)
  - Tibia: central lateral: Cartilage loss % (full + partial)
  - Tibia: central medial: Cartilage loss % (full + partial)
  - Tibia: central lateral: Cartilage loss % (full)
  - Tibia: central medial: Cartilage loss % (full)
  - Trochlea medial: Cartilage loss % (full + partial)
  - Trochlea medial: Cartilage loss % (full)
  - Trochlea lateral: Cartilage loss % (full + partial)
  - Trochlea lateral: Cartilage loss % (full)
  - Patella lateral: Cartilage loss % (full + partial)
  - Patella lateral: Cartilage loss % (full)
  - Patella medial: Cartilage loss % (full + partial)
  - Patella medial: Cartilage loss % (full)
  - Other areas

- *Osteophytes: score*
  - Superior patella
  - Inferior patella
  - Lateral patella
  - Medial patella
  - Medial trochlea
  - Lateral trochlea
  - Central medial femur
  - Central lateral femur
  - Posterior lateral femur
  - Posterior medial femur

    - Lateral tibia
    - Medial tibia
    - Subspinous tibia
    - Other subregions
- *Meniscal extrusion: score*
    - Medial meniscus: medial extrusion
    - Medial meniscus: anterior extrusion
    - Lateral meniscus
- *Meniscal morphology: Lateral*
    - Anterior: Signal
    - Anterior Horn
    - Anterior: Meniscal cyst
    - Anterior: Other menisci morphology
    - Body: Signal
    - Body
    - Body: Meniscal cyst
    - Meniscal hypertrophy
    - Body: Other menisci morphology
    - Posterior
- *Meniscal morphology: Medial*
    - Body: Signal
    - Body: Tear
    - Body: Partial maceration
    - Body: Meniscal cyst
    - Body: Meniscal hypertrophy
    - Posterior horn: Signal
    - Posterior horn: Vertical tear
    - Posterior: Horizontal tear
    - Posterior: Radial tear
    - Posterior horn: Root tear
    - Posterior horn: Meniscal cyst
    - Posterior horn: Meniscal hypertrophy
    - Other meniscal morphology
    - Anterior horn
- *Ligaments and tendons: score*
    - ACL and PCL: score
    - BML/cyst
    - Repair
    - Patellar tendon
    - Infrapatellar bursa signal
    - Popliteal cyst
    - Other periarticular features
- *Periarticular features*
- *Hoffa's fat synovitis*
- *Synovitis / effusion*

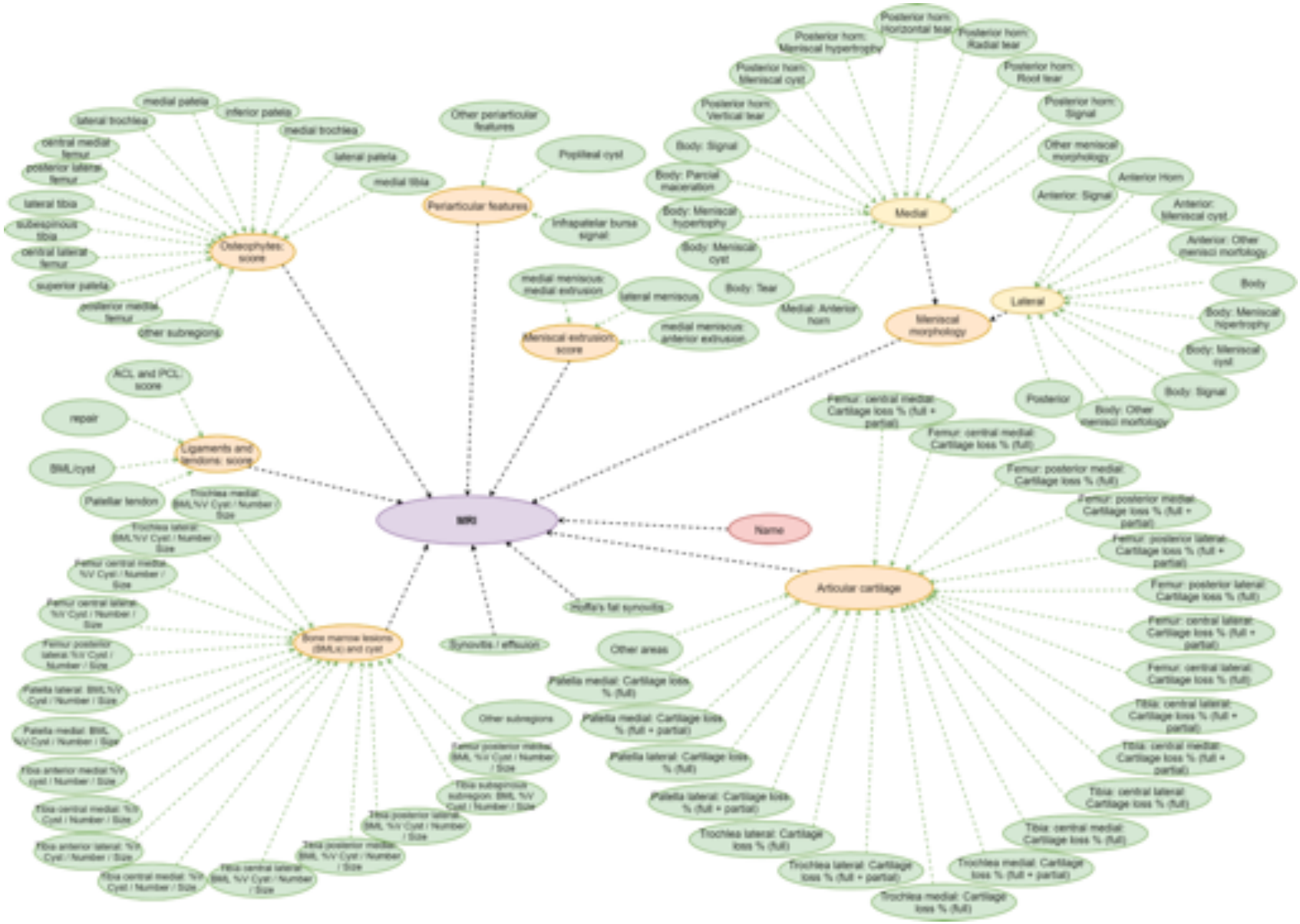


Figure 3.15: MRI upper ontology Levels of information extraction

Personalized gait retraining intervention requires an estimation of the internal state of the musculoskeletal system. Current approaches rely mostly on kinematic measurements to evaluate the performance of a subject. However, a pure kinematic analysis is not enough to quantify the stress of the knee during movement, because one must also account for the forces that act on the system. In terms of dynamics, most studies are limited only to inverse dynamics analysis due to the increased computational burden required for calculating the muscle forces and joint reaction loads in real-time. The system developed in the OACTIVE project moves one step further by estimating the muscle forces and joint reaction loads in real-time. This section presents the architecture and implementation details of the underlying system. The term "internal state of the musculoskeletal system" refers to the quantities that cannot be measured non-invasively, which however can be determined from other experimental measurements, and are important for quantifying the performance of the subject. Examples include (i) the kinematics of the skeletal system (e.g., anatomical joint angles), (ii) joint loads, (iii) muscle forces, and (iv) joint reaction loads. This information can be used to determine whether a subject is walking in an optimal manner by examining his/her reaction loads at the diseased knee and provide real-time feedback that can aim in reducing the contact pressures.

Figure 3.16 presents a diagram of the different stages that are being solved in real-time in order to calculate the internal state of the musculoskeletal system in real-time from kinematic and dynamic measurements. The experimental measured positions (markers or orientations) are fed to the Inverse Kinematics (IK) module, which determines the generalized model coordinates that best match the experimentally recorded motion. The kinematic analysis is very important for the next stages. Following Inverse Dynamics (ID) is performed to calculate the generalized forces that satisfy the Equations of Motion (EoMs), provided any externally applied force. Static Optimization (SO) can be employed to estimate the required muscle activation and forces that satisfy both the motion and the physiological muscle constraints. This step can be further improved by using information from the EMG recordings if available. The muscle forces are essential for the calculation of the joint reaction loads that is the final stage of the system.

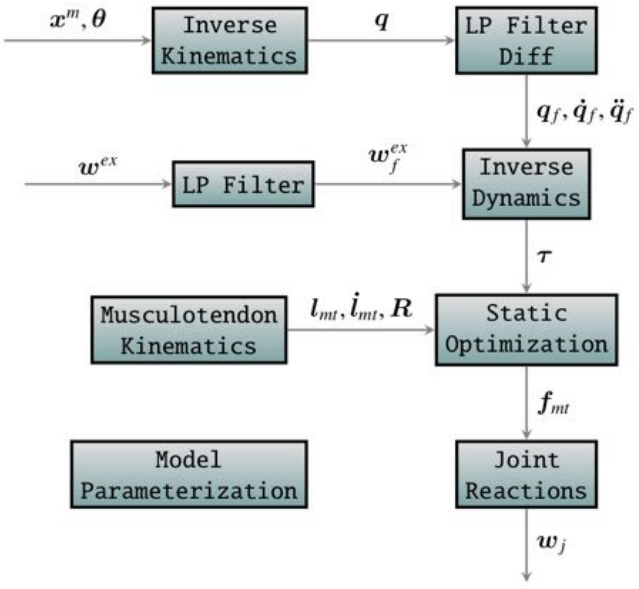


Figure 3.16: Architecture diagram of the different stages that are solved in order to determine the internal state of the musculoskeletal system in real-time.

In order to perform any kind of IK analysis one must obtain the motion of the body segments from the recorded kinematics. The motion can be obtained either by recording the position of the attached markers

or by measuring the 6D motion of IMUs. The next step would be to determine the evolution of the generalized model coordinates that best match the experimentally recorded motion. The IK method goes through each time step of the recorded motion and computes the generalized coordinates which positions the model in a pose that best matches the experimental measurements. More formally, this is expressed as a weighted least squares problem, whose solution aims to minimize the following error.

$$\sum_{i}^{measure} w_i \left|\left| x_i^{exp} - x(q) \right|\right|^2$$

where w_i represents the weight for measure i, x_i^{exp} the experimental measure and $x_i(q)$ the model measure that depends on the pose. Each measure has a weight associated with it, specifying the influence of the particular term on the overall error. The least squares problem can be solved using a general quadratic programming solver. In terms of performance, our implementation is able to solve the IK for a model with 23 Degrees of Freedom (DoFs) in less than 0.5ms without compromising the accuracy of the solution ($< 2cm$ Root Mean Square Error (RMSE)).

The next stage (Figure 3.16 second level) is to provide a solution to the ID problem. In order to solve for the generalized forces that satisfy the Equations of Motion (EoMs) one must first calculate the first and second derivative of the kinematics. Numerical differentiation is a challenging problem, especially when we deal with real-time application, where one does not know the value of the kinematics in future time instances (casual systems). The kinematics obtained from IK are noisy, resulting in bad approximations of higher order derivatives. Therefore, appropriate filtering must be implemented in order to reduce the noise artefacts without introducing lags and errors in the filtered signals. However, even if one filters the signal there will be always jitter in the higher order derivatives. To this end, we adopted the following strategy making use of the high sampling rates of the kinematics measurements in order to implement higher order IIR filters without introducing lags in the calculated quantities:

1. IK is solved in a separate thread (T1) configuring the motion capture system to operate with high sample frequencies and the solution of IK is stored in a thread safe buffer.
2. In a separate thread (T2) we solve the rest of the stages (ID, SO, and JRA).
3. Each time we proceed into T2 we read the buffer containing the solutions obtain by IK.
4. A low pass Butterworth filter of order > 50 is applying on the most recent kinematics data.
5. Then smooth splines are fitted in order to compute higher order derivatives.
6. Finally, we choose the sample delay in order to evaluate the kinematics and their derivatives. Low delay (2-3 samples) will result in low latency (lag) in the estimated quantities. Higher delay (noncausal filter) results in larger latencies but permits better approximation of the higher order derivatives as compared to offline analyses.

Figure 3.17 depicts the results obtained by applying the aforementioned procedure to filter the kinematics and obtain higher order derivatives. For a comparison, a conventional 4th order Butterworth filter is applied to the signal (right column). The latter is the filtering adopted in the literature, which as evident can lead to bad approximations of the higher order derivatives. The increased complexity of the proposed filter satisfies the requirement for accuracy that can propagate to errors in the estimated quantities of the next stages.

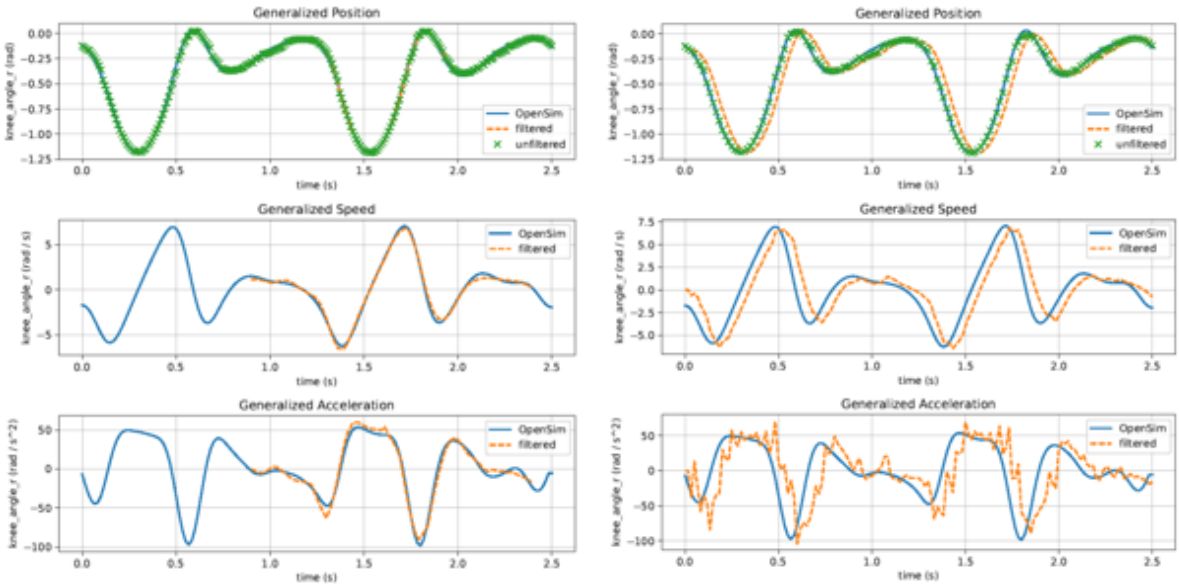


Figure 3.17: Calculated kinematics of the knee joint. Unfiltered kinematics from IK are filtered by the process outlined in this section. The filtered kinematics are compared against results obtained from OpenSim when performing the analysis in an offline mode, meaning that OpenSim can apply non-casual filters that can better remove the noisy artefacts in the signal. The figure on the right depicts what will be the actual result of the filtered signal if one used conventional 4th order low pass Butterworth filter.

For the given kinematics describing the movement of a model and any externally applied force, the ID method determines the generalized forces (e.g., net forces and torques) at each joint that satisfy the movement. More formally (see equation bellow), are solved for the unknown τ provided q, $\dot{q}$ and $\ddot{q}$. Since q is calculated from IK, $\dot{q}$ and $\ddot{q}$ must be obtained using numerical differentiation. Discontinuities in the generalized coordinates can lead to numerical singularities during the evaluation of higher order derivatives (as discussed previously) and it is thus advised to apply filters to remove any undesirable artefacts. The following notation is used for describing the EoMs.

$$M(q)\ddot{q} + f(q,\dot{q}) = \tau$$

where $M \in R^{nxn}$ denotes the symmetric, positive definite joint space inertia mass matrix, n the number of models DoFs and $q, \dot{q}, \ddot{q}$ the joint space generalized coordinates and their derivatives with respect to time. The term f models all internal and external applied forces (e.g., gravity, Coriolis, GRF, etc.), whereas τ the vector of applied generalized forces that actuate the model. Most of the quantities in the equations are a function of the generalized coordinates and derivatives, thus this dependency will be omitted for simplicity.

When ID is solved, it is preferred to ignore model constraints since a set of $q, \dot{q}$ and $\ddot{q}$ may not necessarily satisfy the constraint algebraic equations. As a result, if a model contains kinematic constraints ID may give rise to unreliable estimates of the generalized forces. In this case, ID-based methods may provide a suitable solution. Another important misconception is that ID does not necessarily depend on externally applied forces, however, it is a common sense that they should be accounted for since their omission will significantly alter the result.

Moving to the solution of ID and as indicated in Figure 3.16, the external forces w^{ex} which are measured are low passed filtered. The filtered kinematics as well as filtered external forces are fed to the ID system which is also a very fast operation ($< 0.5ms$). As indicated in Figure 3.17 where we compare the result

obtained from the real-time analysis to the one obtained by OpenSim in an offline mode, good agreement is observed.

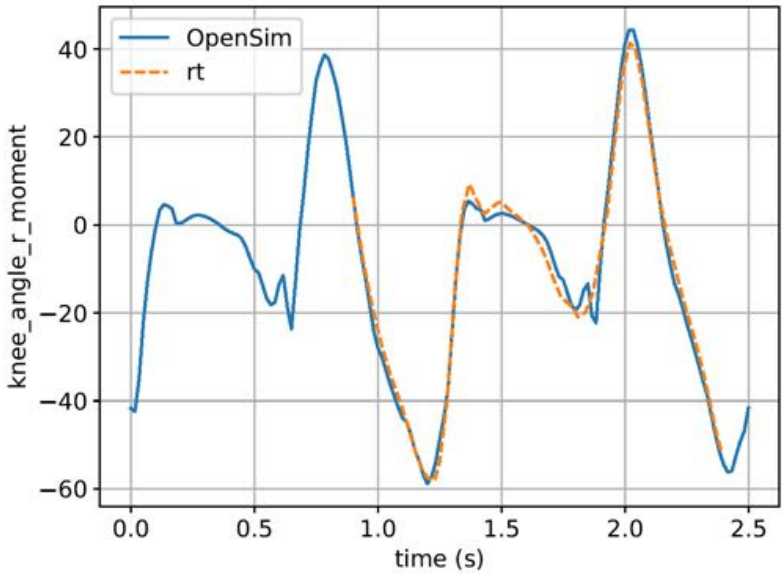


Figure 3.18: Estimated knee moments from ID. Comparison is made by performing an ID analysis through OpenSim in an offline mode, where the whole measured kinematics and external forces are available.

Static Optimization (SO) is a method for determining the muscle forces required to satisfy the motion and forces of the model. The OpenSim internal implementation requires 1s/iteration to solve the SO problem, which is unacceptable for real-time applications. There are many ways to implement and solve the SO problem. Here we will outline an approach which can result in a solution that can operate with 10ms delay for a model containing 92 muscles. The following optimization problem is being formulated.

$$\sum_{m=1}^{muscles} f_m^2$$

$$s.t. R(q) f_m = \tau and f_m \geq 0$$

where, $\boldsymbol{f_m}$ is muscle force m, $\boldsymbol{R}$ is the moment arm matrix and $\boldsymbol{\tau}$ the generalized forces obtained from ID. The objective of the optimization is to compute the minimal muscle force $\boldsymbol{f_m}$ needed so that the contribution of the muscle forces to the joint torques to be equal to the torques computed by ID for each time instance. It is called "static" optimization because the performance criterion (i.e., the cost index) is confined to quantities that can be computed at any instant in time during a simulation. The bottleneck of this formalization is (i) that one must calculate the moment arm matrix $\boldsymbol{R(q)}$ which is a time-consuming operation provided that the muscles have complex routing mechanisms and (ii) the optimization must find a solution with less iterations to achieve real-time performance (preferably <7).

To address the first problem, we derived and precomputed a symbolic representation of the moment arm matrix as a function of the generalized coordinates. In order to derive a symbolic representation, multivariate polynomial fitting was performed on samples of the muscle moment arm at different configurations. To reduce the complexity and improve the robustness of the fit, we determined the coordinates affecting each element in the moment arm matrix, by identifying the DoFs spanned by each

muscle. The sampled and symbolically obtained moment arm of the vastus intermedius (a mono-articular muscle) at the knee joint as a function of the knee flexion angle and the moment arm of the hamstring muscle at the knee joint as a function of the hip and knee flexion angles are compared. The second problem is addressed by providing the previously obtained solution as an initial guess for the next optimization. Furthermore, one could adjust the convergence tolerance of the optimization in order to reduce the iterations needed to obtain a solution within a desired accuracy of digits. The close match between the muscle forces obtained by the proposed real-time optimization and the one obtained from OpenSim can be graphically depicted.

The visualization back-end was developed with the use of the Unity 3D game engine. Unity 3D was selected because it offers multiple tools for visualization and gamification and there is already support for multiple AR and VR devices, such as the Meta 2 AR headset. A fast communication between the real-time musculoskeletal analysis and the visualization process was necessary in order to develop a game with real-time interventions and feedback. To achieve that kind of demanding communication, especially in latency, a shared memory interface was developed. The shared memory interface utilizes the OS capabilities of mapping a memory block by one process and then sharing the same block with another process. Providing access to the same memory block, the data is available on both processes with almost no delay compared to communication techniques utilizing the network stack (i.e., sockets). The shared memory interface limitation is that the processes must run on the same computer. The implementation followed a triple buffer design in order to ensure that the interface is thread-safe, meaning there are no race conditions or memory inconsistencies.

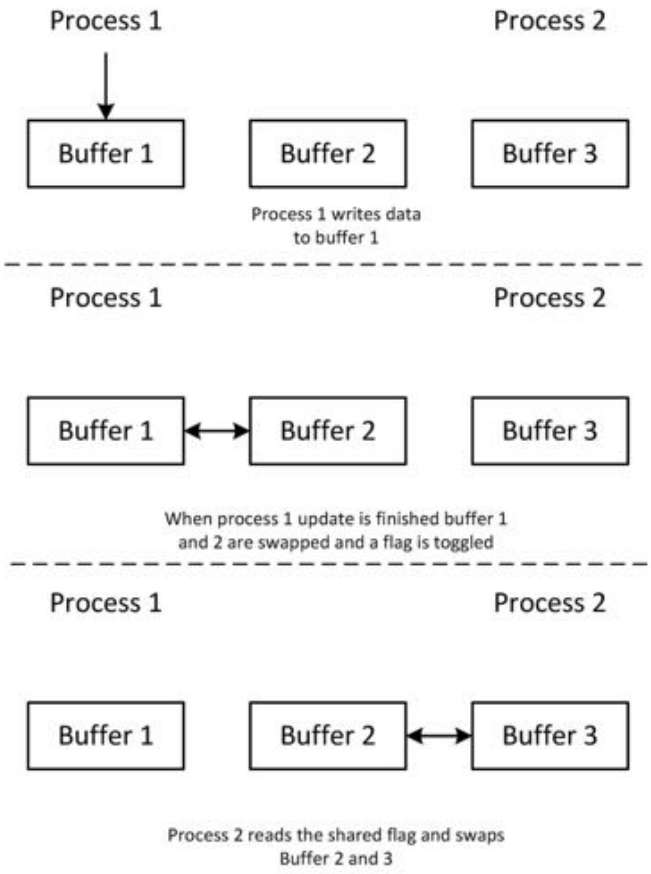


Figure 3.19: Triple buffer implementation flow diagram

The program flow follows the steps bellow:

1. The user starts the visualization process.
2. It bounds the memory block in order to share with the simulation process.
3. It starts the simulation process via an OS command with the map of the shared memory block as a parameter.
4. The simulation process performs some initialization steps and serializes the model structure (joints, bones and muscles) in a json format.
5. The simulation process provides the json to the visualization process.
6. The visualization process constructs the skeleton model along with the muscles by deserializing the json.
7. Then asynchronously the simulation process sends all the necessary data to the visualization process through the shared memory interface on each frame and the visualization process updates the model until the user terminates the application.

3.3 Utility Layer

Utility layer is the final layer of the Ontology framework, where personalised treatment prediction, visualisations, analytics and alerts are contained. Through the game framework, the walking ability of the individual can be optimised, the pressure acting on the knee joint is reduced and thus not only the progression of OA is delayed, but also the pain is relieved. The user interacts with the system that provides the necessary information for a patient's specific gait. Performance statistics can be stored into the database system and processed accordingly by clinicians and Artificial Intelligence (AI) algorithms to adapt and refine the patient intervention strategy. Furthermore, this information is presented to the user in a form of a progress report or to the clinician for evaluation (more details in section Nonetheless, an AR system can be employed to substitute the screen where the game is viewed. The purpose of using an AR system is to improve the interaction of the user with the application/game and enhance the user experience. By displaying virtual objects in the real environment, not only the user can better comprehend how to walk effectively, but also creates an environment that excites the user, stimulates his senses and thus motivating him to seek again for this "unique" experience.

Joint Reaction Analysis (JRA) is used for calculating resultant forces and moments at joint. Specifically, it calculates the joint forces and moments transferred between consecutive bodies as a result of all loads acting on the model. These forces and moments correspond to the internal loads carried by the joint structure. These loads represent the contributions of all unmodeled joint structures that would produce the desired joint kinematics, such as cartilage contact and any omitted ligaments. The reaction load acts at the joint center (mobilizer frame) of both the parent and child bodies. The loads can be reported and expressed in either the child, parent, or ground frames. The default behaviour is to express the force on the child in the ground frame. JRA from the data and statistics from the wearable devices as well as the AI DSS must operate in real time, so that the participant has instantaneous feedback in order to closely follow his "virtual trainer" or to alarm him immediately whether he is executing an irregular – dangerous activity.

Some of the secondary goals achieved from the design of the games are:

4 The games are easily understandable, and they must not require any long technical manual or specialized tutorials. Users must quickly start trying them, and they must feel confidence and security while they play.

5 Each game separately becomes a "habit" for the user. This means that users must come back on a regular or daily basis to commit mostly repetitive playing.

6 Continuously give meaning, interest, and excitement to the users so they never become boring to play the games.

7 The user must not feel anxious about the technology and the games must seem simple and without lags. For example, a simple progress bar is a great feature to help the users get through the waiting process. This makes it very clear what the site wants patients to do, how doing it, and emotionally rewards them for doing it.
8 Achievements should be awarded for sticking to the desired behaviour.
9 Badges should be awarded for long-term performance after hitting some cumulative target.
10 Levels as a measure of progress, with gradual "unlocking" of game features (e.g., advanced analytics, comparative evaluations), designed to engage the player in the early stages of the game.

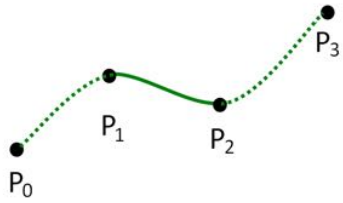
Gamification elements, such as rewards, achievements, scores, difficulty level, etc., also, have been implemented in order to improve the interaction and engage users to improve their physical activities. The effectiveness of this method was further evaluated in the test campaign. It is desirable that the system involves sensory feedback, such as IMU measurements, marker-based motion-capture cameras and ground reaction forces. In this case, this information can be transmitted to a station or to the smart device (smartphone) for processing in order to reconstruct the kinematics and dynamics in real-time. Provided this information, the system will be able to compare the performance of the user with respect to a desired profile and provide detailed guidelines for the correct execution of the gait. This information can also be used by the AI and simulation pipelines, in an offline mode, to provide detailed information on the contact pressure and other quantities of interest that can be used to refine the personal intervention strategy. Finally, instead of displaying the game and the visual information on a screen the system can incorporate an element of AR feedback, in order to enhance the user experience and effectiveness of the game. In this case, the system provides visual cues and information statistics in real-time. This results in correct execution of the gait and will motivate and engage the user to increase its daily physical activities.

The gait retraining activity is ideally adapted to the games' playing script. The user's physical performance and tolerance is accounted for by the game, in order to adjust the positioning of the patient and duration of the training. The features that the game must have in order to motivate the user, are the following:

- Enactive mastery experiences (e.g., goal setting, discussion of performance and progress) • Vicarious experiences (e.g., role modelling, storytelling)
- Verbal persuasion (e.g., education, support, encouragement)
- Physiological and effective feedback (e.g., monitoring the emotional and physical burden, managing discomfort).

The visualization of the skeleton, thus of each bone is achieved using the same 3D models of simulation. Unity 3D does not support the format VTP natively, so a special importer was developed in order to load the 3D geometry in the Unity 3D environment. The initial json file contains all the information about the position, rotation and scale for each 3D model of all bones, as some bones consist of multiple 3D models. Each update cycle again contains information about the position, rotation and scale of each bone in world coordinates.

The geometry of muscles is generated procedurally based on a couple of points provided by the simulation process. In order to achieve a smooth and close to reality result for the muscle geometry, a centripetal Catmull-rom spline is calculated from the waypoints. Catmull-rom spline was preferred as it goes through its control points.



Based on the spline S, elliptical rings (R_t) of vectors are created and the consecutive vector rings form the triangles (T_t) of the 3D mesh. To create the elliptical rings a local orthogonal system was calculated for each interpolation step. Based on the up and right vector of that system, two radius parameters a_t, b_t and the parametric representation of ellipse based on sine and cosine functions, the position of each vector was calculated

$$v_t = c_t + up_t \cdot a_t \cdot cos\ cos\ \theta\ + right_t \cdot b_t \cdot sin\ sin\ \theta\ , 0 \leq \theta \leq 2\pi, 0 \leq t \leq 1$$

A special shader was also developed in order to create the texture of the muscle based on the geometry. In detail, at the two endpoints all triangles are coloured with a white colour to look like tendons and gradually turn to a dark red colour with darker stripes based on the θ angle described above to look like myofibrils. In addition to the structural colours of the muscle (tendons, myofibrils), there is also a highlight colour that indicates the calculated activation of the muscle. The final colour of the muscle is the result of blending the structural colours and the highlight colour. Activation range is $[0,1]$, so the blending is based on this number with linear interpolation.

$C_{result} = a \cdot C_{highlight} + (1 - a) \cdot C_{structural}$, where a is the activation of the muscle.

The update of all muscles is performed in each frame according to the new waypoints provided by the simulation process, respecting the restriction surfaces and their highlight colours are updated based on the calculated muscle activations.

The reaction forces are visualized via a 3D line, where the two endpoints are provided by the equations below:

$$start\_point = force\_application\_point,$$

$$end\_point = start\_point + force\_vector$$

The point, that the force is applied, is the centre of the joint or the contact point with the ground depending on if the related visualization is a joint reaction force or a ground reaction force.

Even though the users can easily comprehend the reaction forces visualization, they do not provide an intuitive way to keep track of them over time and grasp a better overview of complex movements like walking. The most simple and intuitive method for time-varying data is the utilization of 2D plots, where the user can easily comprehend the related data through the whole movement, identify local extrema where there is possible violation of thresholds. Plots also provide an intuitive way for the user to detect patterns like the movement cycles and easily find the correspondences between the displayed data points and the points in time relevant to the movement cycle.

Apart from the reaction forces visualization and the plots, a minimal implementation of the posture score as part of the initial gamification. Scores, as gamification elements, provide an extra layer of abstraction and ease the learning curve for the users. In that context, a new user can easily track only the score and by intuitively finding correspondences between the score values and patterns in the plots, the user learns how to interpret the plots and classify the "bad" and "good" movement patterns.

Plenty of visualization elements developed in order to accomplish the gamification approach. Minimal visualizations take place during the real-time exercise in order to motivate the user's engagement and motivation, but not disorient the user's focus from the main objective.

More visualizations included with the objective to help the user identify the movement deviations from the ideal movement such as red colour highlighting of the failing part of the body (joint or muscle), along with a relevant highlighting in the corresponding plot indicating the local extrema which signifies the flaw. A semi-transparent model also is visible above the opaque model, which reflects the user's movement and acts as a guide for the user to learn the exercise and has a visual reference of the ideal movement. Basically, the ghost model shows the ideal position for each part of the body for the next step of the exercise, so the user can train in a speed of their choice and not chase a pre-recorded animation, with the exception of the exercises that the speed is important. Other supportive tools developed like:

- Auto rotation of the model, so that the failing part of the body is visible from the user.
- Auto plot selection related to objective of the user and of the exercise, since not all the measurements can be presented at once.
- Audio feedback synchronized with the visuals to improve user's identification and memory stimulation.

Apart from the real-time visualizations, some graphs and overview plots of stats and progression metrics are part of the developments. These visualizations act both as gamification elements and analysis components. Such visualizations are:

- Spider graph of scores categorized by exercise.
- User's progression graph based on scores.
- Calendar view with collected medals, daily challenges completed.
- Display measured or simulated data that are most relevant to the task goal.

4 Conclusions

Deliverable D6.6 Ontology-based framework for data standardisation constitutes a vital part of the OACTIVE project. It is the link that makes the inter-process communication of the project possible. In this deliverable the presentation of a unified way to interact with the vast amount of the collected/generated data, the interaction between the data, the important variables and the information extraction through an upper framework is implemented. All these procedures are possible via an intricate architecture that takes into consideration all the possible needs of the project's data interaction (self-generated or imported data). Through the layers of this architecture processes take place, that export information from a huge amount of clinical data collection collected or generated from biomarkers, sensors, and augmented reality. Also, this architecture manages exogenous/environmental risk factors as also, medical, and socioeconomic data. All the extracted information is used for personalised treatment prediction, visualizations, analytics, and alerts that creates a friendly, helpful, and motivational environment for the patient.

5 Acknowledgment

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 777159.

Disclaimer