PROJECT DELIVERABLE REPORT

Project Title:

Advanced personalised, multi-scale computer models preventing osteoarthritis
SC1-PM-17-2017 - Personalised computer models and in-silico systems for well-being

Deliverable number	**D6.1**
Deliverable title	Data Management Infrastructure
Submission month of deliverable	M23
Issuing partner	CETRI
Contributing partners	LJMU, CERTH
Dissemination Level (PU/PP/RE/CO):	RE
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
1.1	1/10/19	CETRI	First Draft
1.2	29/10/19	UNIC	Final

Table of Contents

Abbreviations

BMI	Body mass index
CSI	Central Sensitization Inventory
DCP	Data collection protocol
FACHS	Functional Ambulation Classification of the Hospital at Sagunto
GADS	Goldberg Anxiety and Depression Inventory
HAD	Hospital Anxiety and Depression Scale
KL	Kellgren and Lawrence
KOOS	Knee Injury and Osteoarthritis Outcome Score
OA	Osteoarthritis
WOMAC	Western Ontario and McMaster Universities Osteoarthritis Index

1. Summary

This report refers to the Deliverable 6.1 which is related to the OActive's WP.6 "Hyper-modelling framework empowered by big data and deep learning", led by CERTH.

The report describes the cloud-based data management infrastructure i.e., the tools and the services that have been developed to realize a high-level data quality and accessibility for the data analytics applications.

2. Introduction

Data management infrastructure is the set of tools and services that are used to provide data management and enforce data management policies. A data management infrastructure would include resources such as a data repository and an information catalogue. The design and implementation of the data management infrastructure includes all the steps for collecting, storing and utilizing the available information. Effective data management will help the project partners locate valuable information in large sets of unstructured data and semi-structured data from the variety of sources.

The data management infrastructure in the context of a data science workflow is depicted in the following Figure:
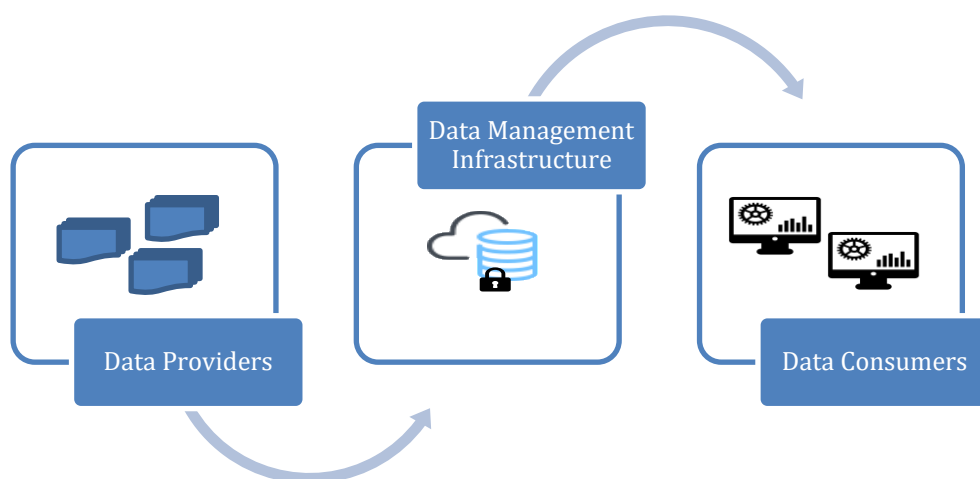

Fig.1 Data Science Workflow

Following the above workflow, the data management infrastructure should meet the following generic requirements:

1. Provide a centralized and backed up storage area of research data.
2. Provide read & write access to data amongst defined groups of researchers in a secure way.
3. Remove data redundancy, by removing the need to store the same data in multiple places.
4. Allow the structured annotation of research data to provide context to the data
5. Allow the search and dissemination of data
6. Provide a system simple enough for users to pick up and utilize without structured training.

In addition, and in the context of the OActive project, the data management infrastructure, should also follow the Data Collection Protocol described in D2.1. The data collection protocol defines the timing, content, screening, assessment, and evaluation tests and other rules relating to the ascertainment and collection of data over the life cycle of OACTIVE.

In the following sections we will describe the Data Management Infrastructure components in detail.

3. Data Management in the Cloud

The start of last decade presented that the pace and volume of data being generated is exceeding the current capacity of "institutions" data management. Cloud-based data management is in turn helping to realize the potential of large-scale data management solutions by giving effective scaling of resources. Data management is one of the most important research areas in cloud computing and cloud computing has become a major influence in data management research and plays a key role.

Data management in the cloud addresses the challenges in managing large collections of data in the cloud computing environment. Huge volumes of data in cloud computing environments pose big infrastructure challenges, including data storage, massively parallel query execution, facilities for analytical processing, and online query processing. There is a high degree of complexity involved in ensuring that they can sustain consistent and reliable operations under peak loads. However, cloud-based data management systems will not replace the traditional RDBMS in the near future; however, it supplies another choice for the applications which are suitable to be deployed in the cloud.

Moving data and critical applications to the cloud though leads to security challenges. The main of these challenges is that the owner of the data may not have control of where the data is placed. Other issues that have been pointed out is that the users are unaware of cloud security and the concern about the protection of confidential data.

OActive's Data Management infrastructure follows a "cloud-based" approach as initially discussed. However, the current datasets examined in the project, in terms of volume and complexity, indicates that it cannot be considered as a Big Data case as initially thought to be. Therefore, more conventional, health-data-appropriate databases should be considered.

In particular, we have employed the Airtable, which is a cloud collaboration platform for creating, using, and sharing relational databases. It fuses the features of a database and the format of a spreadsheet so as users have always control of the process. We present in details the Airable component and infrastructure in the following sections.

4. Airtable

Airtable is a cloud-based app that combines the functionalities of spreadsheets and databases to create a powerful tool that can be used as a data management infrastructure. Airtable organizes through six components: *Bases, Tables, Fields,* and *Records.* A short description of these components is provided in the following subsections:

4.1 Airtable Base

An Airtable Base contains all of the information required from a project. Each Base should be unique, focus on one key area, and contains all the of the information is available for that key area.

We can consider Bases are simple spreadsheets that can contain any number of Tables. The advantage of Baess is that in contrast to spreadsheets are not simply flat-grids but multi-dimensional, e.g. like a calendar and thus can be more flexible.

4.2 Airtable Table

Tables are the building blocks of Bases and each Base could include any number of Tables. Tables are like worksheets in a spreadsheet. Each Table could hold information concerning one item. For example, a project Base could have a Table for data from different project's domains. These domains are relevant to the project Base, but each is unique from any other domain Table.

4.3 Airable Field

Fields are Airtable's equivalent to columns in a spreadsheet. Each Field communicates the same kind of data across a number of rows. Unlike traditional spreadsheet cells, Fields have a number of different types. These types include file attachments, checkboxes, dropdowns, and more.

4.4 Airtable Record

Records are the second part of an Airbase table. While Fields represent columns in a traditional spreadsheet, Records represent the rows. Each Record is a unique iteration of data defined by the table's rows.

5. **Airtable Security And Data Privacy**

Airtable platform considers privacy and security as core functions and thus the highest privacy and security standards. In terms of network and system security, transmission of information between the data provider's device and AirTable's servers is servers is protected using 256-bit TLS encryption. At rest, Airtable encrypts data using AES-256. Airtable servers are located in the US, in data centers that are SOC 1, SOC 2 and ISO 27001 certified. Airtable's data centers have round-the-clock security, automatic fire detection and suppression, fully redundant power systems, and strict controls for physical access.

In terms of service reliability and durability, Airtable exploits the Amazon Web Services (AWS) hosting infrastructure. Backups are geo-redundantly replicated across multiple availability zones for data durability and maintains business continuity and disaster recovery plans.

As far as the product security, collaborator permissions can be managed at the workspace level or the base level. These permissions allow you to control who you share a workspace or base with and whether they can modify the workspaces or bases that you've shared with them. Airtable also provides record-level revision history that shows a visual activity feed of the changes made to each record.

Airtable is fully compliant with the General Data Protection Regulation ("GDPR"). and can help you meet your data portability requirements for the GDPR, for example by allowing data export from Airtable to CSV files or by using the Airtable's API.

6. **OActive's Airtable Infrastucture**

The Data Management infrastructure has been initially built on an Airtable's Base as shown in the following Figure.
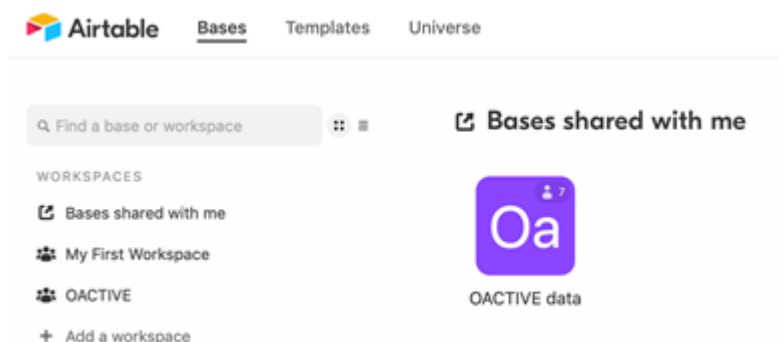
In this Base, a number of Tables following the Data Management Protocol have been created to collect the data. As described in the Protocol, the medical visits that lead to data collection from patients are divided into the following main sections: demographics, anamnesis, physical examination, body fluid tests, and functional and psychological scales. The Tables that have been are shown in the Figure and discussed in details below:

5.1 Demographics Data Table

The Table includes general demographic information about the examined individual, i.e. the institution that performed the experiments, person, such as its sex, age (in years), birth country, ethnicity and occupation.

The Table Fields with their values are given below:

Field Name	Field Type	Values
Identification Code	Text	ID of the examined individual
Data Provider	Text (Single Select)	UNIC/HULAFE/ANIMUS
Date	Date	
Sex	Text (Single Select)	Yes/No
Age	Numeric	
Birth Country	Text (Single Select)	Cyprus/Spain/Greece/Other EU Countries/Other Non EU Countries
Ethnicity	Text (Single Select)	White/Black/Hispanic American/Asian/Pacific Islander/American Indian/Other Race/Two or More Races
Occupation	Text	Free Text

5.2 Socieconomics Data Table

The Table includes extended demographic information about the examined individual. The Table Fields with their values are given below:

Field Name	Field Type	Values
Name/Identification Code	Text	ID of the examined individual
Level of Education (Individual)	Text (Single Select)	Elementary School Completed/Elementary School Not Completed/College or University
Level of Education (Parents)	Text (Single Select)	Elementary School Completed/Elementary School Not Completed/College or University
Marital Status	Text (Single Select)	Married-civil partnership/Single/Window
Residency	Text (Single Select)	Livining with Family/Living Independently
Household income	Text (Single Select)	Easily/Fairly Easily/With some difficulty/With great difficulty
Housing status	Text (Single Select)	Owing/Renting

5.3 Anamnesis Data Table

This Table includes the results obtained by clinical interviews regarding the Familial OA, defined as parents, siblings or grandparents having a diagnosis of OA, and the individual's history of OA. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
Group	Text (Single Select)	Established OA/Initial OA/Healthy/Athlete
Current Medication	Text	
High Blood Pressure	Text (Single Select)	Yes/No
Family OA History	Text (Single Select)	Yes/No
Personal History of hand OA	Text (Single Select)	Yes/No
Personal History of hip OA	Text (Single Select)	Yes/No
Knee OA diagnosis	Text (Single Select)	Yes/No
Knee OA diagnosed by a doctor	Text (Single Select)	Yes/No
Occupational risk	Text (Single Select)	Never/Always/Seldom/Once twice/week
Smoking	Text (Single Select)	Yes/No
Number of cigarettes	Number	
Alcohol	Text (Single Select)	Never/Always/Seldom/Once twice/week
Hormonal status (women)	Text (Single Select)	Premenopause/Postmenopausal
Previous Knee Injury (left)	Text (Single Select)	Yes/No
Previous Knee Injury (right)	Text (Single Select)	Yes/No
Regular Sport leisure activity	Text (Single Select)	Yes/No
Type of Sport	Text	
Sports frequency	Text (Single Select)	Daily/Once-twice week/Once-twice month
Knee Pain	Text (Single Select)	Yes/No
Pain Side	Text (Single Select)	Left/Right/Both
Time since pain start	Number	
Resting VAS	Number	
Walking VAS	Number	
Knee pain (NHANES) left	Text (Single Select)	A/C/No pain
Knee pain (NHANES) right	Text (Single Select)	A/C/No pain
Knee instability left	Text (Single Select)	Yes/No
Knee instability right	Text (Single Select)	Yes/No
Pain rhythm	Text (Single Select)	Mechanical/Inflammatory
Neuropathic Component	Text (Single Select)	Yes/No

5.4 Physical Examination Table

This Table includes the results from the general physical examination, and the knee specific exploration. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
Mass	Number	
Height	Number	
BMI	Number	
Joint Line Tenderness (Left)	Text (Single Select)	Yes/No
Joint Line Tenderness (Right)	Text (Single Select)	Yes/No
Patellofemoral Pain (Left)	Text (Single Select)	Yes/No
Patellofemoral Pain (Right)	Text (Single Select)	Yes/No
Crepitus (Left)	Text (Single Select)	Yes/No
Crepitus (Right)	Text (Single Select)	Yes/No
Right Flexion Angle	Number	
Right Extension Angle	Number	
Flexion Deformity (Right)	Text (Single Select)	Yes/No
Left Flexion Angle	Number	
Left Extension Angle	Number	
Flexion Deformity (Left)	Text (Single Select)	Yes/No
Muscle Atrophy	Text (Single Select)	Yes/No
Measurement of Left Limp	Number	
Measurement of Right Limp	Number	
Knee Laxity (Left)	Text (Single Select)	Yes/No
Knee Laxity (Right)	Text (Single Select)	Yes/No
Joint Proprioception (Left)	Text (Single Select)	Yes/No
Joint Proprioception (Right)	Text (Single Select)	Yes/No
Abdominal Perimeter (in cm)	Number	
Left Extension Dynamometric	Number	

Right Extension Dynamometric	Number	
5 Sit to Stand Test	Number	
Walking Speed (10m walk)	Number	
Knee Morphology	Text (Single Select)	Normal/Altered
Joint Effusion	Text (Single Select)	Yes/No
Increased Local Temperature	Text (Single Select)	Yes/No
Local Redness	Text (Single Select)	Yes/No
Bakers Cyst	Text (Single Select)	Yes/No
Muscle Strength MRC (Left Hip Flexors)	Number	
Muscle Strength MRC (Left Hip Abductors)	Number	
Muscle Strength MRC (Left Knee Extensors)	Number	
Muscle Strength MRC (Left Knee Flexors)	Number	
Muscle Strength MRC (Left Plantar Flexors)	Number	
Muscle Strength MRC (Right Hip Flexors)	Number	
Muscle Strength MRC (Right Hip Abductors)	Number	
Muscle Strength MRC (Right Knee Extensors)	Number	
Muscle Strength MRC (Right Knee Flexors)	Number	
Muscle Strength MRC (Right Plantar Flexors)	Number	

Identification code	Mass (Kg)	Height (m)	BMI	[Left] Joint line ten...	[Right] Joint line te...	Patellofemoral pain ...
SPH001	67.5	1.640	25.097	No	No	
SPH002	78.5	1.580	31.445	No	No	
SPH003	78.0	1.735	25.912	No	No	
SPH004	81.0	1.580	32.447	No	Yes	
SPH005	75.0	1.600	29.297	No	No	
SPH006	84.0	1.740	27.745	No	No	
SPH007	73.5	1.710	25.136	No	No	
SPH008	81.5	1.735	27.074	No	No	
SPH009	76.0	1.720	25.690	No	No	
SPH010	60.0	1.650	22.039	No	No	No
SPH011	92.0	1.790	28.713	No	No	No
SPH012	54.0	1.580	21.631	No	No	No
SPH013	77.5	1.710	26.504	No	No	No
SPH014	70.5	1.720	23.830	No	No	No
SPH015	62.0	1.780	19.568	No	No	No
SPH016	78.0	1.650	28.650	No	No	No
SPH017	61.0	1.610	23.533	No	No	No
SPH018	73.0	1.650	26.814	No	No	No
SPH019	66.5	1.700	23.010	No	No	No
SPH020	89.5	1.790	27.933	No	No	No
SPH021	83.0	1.800	25.617	No	No	No

5.5 Blood Tests Table

The Blood Tests Table includes the results from the patients' blood samples needed for the search of biomarkers. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
Uric Acid (mg/dL)	Number	
Total Cholesterol (mg/dL)	Number	
HDL-cholesterol (mg/dL)	Number	
LDL-cholesterol (mg/dL)	Number	
Triglycerides (mg/dL)	Number	
Protein C reactive (mg/L)	Number	
Vitamine D (mg/dL)	Number	
PTH (pg/mL)	Number	
Glycated hemoglobin (%)	Number	
Serum COMP	Number	
Serum HA	Number	
PIICP	Number	
IL-1β	Number	
TNF-α	Number	
IL-6	Number	

	Uric acid (mg/dL)	Total cholesterol (m...	HDL-cholesterol (m...	LDL-cholesterol (m...	Triglycerides (mg/dL)	Protein C reactive (...	Vitamine D (mg/L)
		196	58	125	60		30.6
		177	34		420		14.2
		113	47	54	61		30.4
	4.7	159	61	82	80	1.8	23.1
		207	54	136	84	0.7	35.9
		124	61	53	54	0.7	24.0
	5.7	152	48	83	105		24.0
		213	114	85	69	0.9	24.0
	7.0	208	46	135	135	6.2	35.6
	3.0	188	78	92	89	1.0	8.8
	5.1	220	70	136	69	3.5	21.0
	3.6	216	73	92	255	0.6	23.1
	4.0	260	139	108	69	0.3	26.0
	4.6	239	57	159	119	3.1	33.1
	4.4	189	53	93	319	0.5	6.9

5.5 Scales Data Table

Scales Data Table includes the results from the assessment of limitations in functioning among patients with OA of the lower extremities., such as Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) or the Knee Injury and Osteoarthritis Outcome Score (KOOS). The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
FACHS	Number	
WOMAC	Number	
KOOS Pain (%)	Number	
KOOS Symptoms (%)	Number	
KOOS ADL (%)	Number	
KOOS QOL (%)	Number	
KOOS Sport/Rec (%)	Number	
HAD Anxiety	Text	Normal/Abnormal/Borderline Abnormal
HAD Depression	Text	Normal/Abnormal/Borderline Abnormal
GADS	Text	Anxiety/Depression/No anxiety or depression
CSI A	Text	
CSI B	Text	

WOMAC	KOOS PAIN (%)	KOOS SYMPTOMS ...	KOOS ADL (%)	KOOS QOL (%)	KOOS SPORT/REC (...
27.00	56%	89%	60%	44%	50%
73.00	67%	68%	37%	25%	0%
0.00	100%	100%	100%	100%	100%
39.00	47%	46%	59%	6%	0%
9.00	83%	96%	94%	63%	95%
17.00	83%	89%	84%	6%	5%
1.00	100%	100%	100%	92%	100%
15.00	64%	79%	81%	50%	55%
4.00	97%	89%	97%	88%	80%

5.6 Social Participation Table

Social Participation Table includes the replies regarding the social activities of the patients. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
Social Membership Participation	Text	Not/Less than 1-week/1-2 a week
Social Activity Participation	Text	Not/Less than 1-week/1-2 a week
Eating out	Text	Not/Less than 1-week/1-2 a week
Drinking out	Text	Not/Less than 1-week/1-2 a week
Public events	Text	Not/Less than 1-week/1-2 a week
Social Games (cards etc)	Text	Not/Less than 1-week/1-2 a week

Social trips	Text	Not/Less than 1-week/1-2 a week
Social work	Text	Not/Less than 1-week/1-2 a week
Voluntary work	Text	Not/Less than 1-week/1-2 a week

5.6 RX Data Table

RX Data Table includes the results from the Radiographic examinations performed in patients. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
Leg-length inequality	Text	Yes/No
Leg-length inequality measure	Text	4 R>L/2 R<I
Right Knee alignment	Text	Varus/Valgus/Neutral
Right radiographic angle (Knee alignment)	Number	
Left Knee alignment	Text	Varus/Valgus/Neutral
Left radiographic angle (Knee alignment)	Number	
Right Kellgren and Lawrence (KL) (in points)	Number	
Left Kellgren and Lawrence (in points)	Number	

Right patellofemoral angle	Number	
Left patellofemoral angle	Number	
Right lateral deviation patella (mm)	Number	
Left lateral deviation patella (mm)	Number	
Right Congruence angle	Number	
Left Congruence angle	Number	

5.7 MRI Data Table

MRI Data Table includes the results from the MRI examinations performed in patients. The Table Fields with their values are given below:

Field Name	Field Type	Value
Identification Code	Text	ID of the examined individual
R- Bone marrow lesions (BMLs) and cyst	Number	
R- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML size	Number	

R- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML number	Number	
R- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML%V.Cyst	Number	
R- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML size	Number	
R- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML number	Number	
R- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML%V.Cyst	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML size	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML number	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML %V.Cyst	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella medial: BML size	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella medial: BML number	Number	
R- Bone marrow lesions (BMLs) and cyst: Patella medial: BML %V.Cyst	Number	
R- Bone marrow lesions (BMLs) and cyst: Other subregions	Number	
R- Articular cartilage	Number	
R- Articular cartilage: Femur: central medial: Cartilage loss % (full + partial)	Number	
R- Articular cartilage: Femur: central medial: Cartilage loss % (full)	Number	
R- Articular cartilage: Trochlea medial: Cartilage loss % (full + partial)	Number	
R- Articular cartilage: Trochlea medial: Cartilage loss % (full)	Number	
R- Articular cartilage: Patella lateral: Cartilage loss % (full + partial)	Number	
R- Articular cartilage: Patella lateral: Cartilage loss % (full)	Number	
R- Articular cartilage: Patella medial: Cartilage loss % (full + partial)	Number	

R- Articular cartilage: Patella medial: Cartilage loss % (full)	Number	
R- Articular cartilage: other areas	Number	
R- Osteophytes: score	Number	
R- Osteophytes: superior patela	Number	
R- Osteophytes: inferior patela	Number	
R- Osteophytes: other subregions	Number	
R- Meniscal extrusion: score	Number	
R- Meniscal extrusion: medial meniscus: medial extrusion	Number	
R- Meniscal extrusion: medial meniscus: anterior extrusion	Number	
R- Meniscal extrusion: lateral meniscus	Number	
R- Meniscal morphology	Number	
R- Meniscal morphology: Lateral	Number	
R- Meniscal morphology: Lateral: Anterior: Signal	Text	Yes/No
R- Meniscal morphology: Lateral: Anterior: Meniscal cyst	Text	Yes/No
R- Meniscal morphology: Lateral: Anterior: Other menisci morfology	Text	Yes/No
R- Meniscal morphology: Lateral: Body: Signal	Text	Yes/No
R- Meniscal morphology: Lateral: Body: Meniscal cyst	Text	Yes/No
R- Meniscal morphology: Lateral: Body: Meniscal hipertrophy	Text	Yes/No
R- Meniscal morphology: Lateral: Body: Other menisci morfology	Text	Yes/No
R- Meniscal morphology: Lateral: Posterior	Text	Yes/No
R- Meniscal morphology: Medial	Text	Yes/No
R- Meniscal morphology: Medial: Body: Signal	Text	Yes/No
R- Meniscal morphology: Medial: Body: Tear	Text	Yes/No
R- Meniscal morphology: Medial: Body: Parcial maceration	Text	Yes/No
R- Meniscal morphology: Medial: Body: Meniscal cyst	Text	Yes/No
R- Meniscal morphology: Medial: Body: Meniscal hypertophy	Text	Yes/No
R- Meniscal morphology: Medial: Posterior horn: Signal	Text	Yes/No

R- Meniscal morphology: Medial: Posterior horn: Vertical tear	Text	Yes/No
R- Meniscal morphology: Medial: Posterior: Horizontal tear	Text	Yes/No
R- Meniscal morphology: Medial: Posterior: Radial tear	Text	Yes/No
R- Meniscal morphology: Medial: Posterior horn: Root tear	Text	Yes/No
R- Meniscal morphology: Medial: Posterior horn: Meniscal cyst	Text	Yes/No
R- Meniscal morphology: Medial: Posterior horn: Meniscal hypertrophy	Text	Yes/No
R- Meniscal morphology: Medial: Posterior: Other meniscal morphology	Text	Yes/No
R- Meniscal morphology: Medial: Anterior horn	Text	Yes/No
R- Ligaments and tendons: score	Number	
R- Ligaments and tendons: ACL and PCL: score	Number	
R- Ligaments and tendons: BML/cyst	Text	Yes/No
R- Ligaments and tendons: repair	Text	Yes/No
R- Ligaments and tendons: Patellar tendon	Number	
R- Periarticular features:	Number	
R- Periarticular features: Infrapatelar bursa signal:	Text	Present/Absent
R- Periarticular features: Popliteal cyst	Text	Present/Absent
R- Periarticular features: Other periarticular features	Text	Present/Absent
R- Hoffa's fat synovitis	Number	
R- Synovitis / effsuion	Number	
L- Bone marrow lesions (BMLs) and cyst	Number	
L- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML size	Number	
L- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML number	Number	
L- Bone marrow lesions (BMLs) and cyst: Trochlea medial. BML%V.Cyst	Number	
L- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML size	Number	

L- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML number	Number	
L- Bone marrow lesions (BMLs) and cyst: Trochlea lateral. BML%V.Cyst	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML size	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML number	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella lateral: BML %V.Cyst	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella medial: BML size	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella medial: BML number	Number	
L- Bone marrow lesions (BMLs) and cyst: Patella medial: BML %V.Cyst	Number	
L- Bone marrow lesions (BMLs) and cyst: Other subregions	Number	
L- Articular cartilage	Number	
L- Articular cartilage: Femur: central medial: Cartilage loss % (full + partial)	Number	
L- Articular cartilage: Femur: central medial: Cartilage loss % (full)	Number	
L- Articular cartilage: Trochlea medial: Cartilage loss % (full + partial)	Number	
L- Articular cartilage: Trochlea medial: Cartilage loss % (full)	Number	
L- Articular cartilage: Patella lateral: Cartilage loss % (full + partial)	Number	
L- Articular cartilage: Patella lateral: Cartilage loss % (full)	Number	
L- Articular cartilage: Patella medial: Cartilage loss % (full + partial)	Number	
L- Articular cartilage: Patella medial: Cartilage loss % (full)	Number	
L- Articular cartilage: other areas	Number	
L- Osteophytes: score	Number	
L- Osteophytes: superior patela	Number	
L- Osteophytes: inferior patela	Number	
L- Osteophytes: other subregions	Number	
L- Meniscal extrusion: score	Number	

L- Meniscal extrusion: medial meniscus: medial extrusion	Number	
L- Meniscal extrusion: medial meniscus: anterior extrusion	Number	
L- Meniscal extrusion: lateral meniscus	Number	
L- Meniscal morphology	Number	
L- Meniscal morphology: Lateral	Number	
L- Meniscal morphology: Lateral: Anterior: Signal	Number	
L- Meniscal morphology: Lateral: Anterior: Meniscal cyst	Number	
L- Meniscal morphology: Lateral: Anterior: Other menisci morfology	Number	
L- Meniscal morphology: Lateral: Body: Signal	Number	
L- Meniscal morphology: Lateral: Body: Meniscal cyst	Number	
L- Meniscal morphology: Lateral: Body: Meniscal hipertrophy	Number	
L- Meniscal morphology: Lateral: Body: Other menisci morfology	Number	
L- Meniscal morphology: Lateral: Posterior	Number	
L- Meniscal morphology: Medial	Number	
L- Meniscal morphology: Medial: Body: Signal	Number	
L- Meniscal morphology: Medial: Body: Tear	Number	
L- Meniscal morphology: Medial: Body: Parcial maceration	Number	
L- Meniscal morphology: Medial: Body: Meniscal cyst	Number	
L- Meniscal morphology: Medial: Body: Meniscal hypertophy	Number	
L- Meniscal morphology: Medial: Posterior horn: Signal	Number	
L- Meniscal morphology: Medial: Posterior horn: Vertical tear	Number	
L- Meniscal morphology: Medial: Posterior: Horizontal tear	Number	
L- Meniscal morphology: Medial: Posterior: Radial tear	Number	
L- Meniscal morphology: Medial: Posterior horn: Root tear	Number	

L- Meniscal morphology: Medial: Posterior horn: Meniscal cyst	Number	
L- Meniscal morphology: Medial: Posterior horn: Meniscal hypertrophy	Number	
L- Meniscal morphology: Medial: Other meniscal morphology	Number	
L- Meniscal morphology: Medial: Anterior horn	Number	
L- Ligaments and tendons: score	Number	
L- Ligaments and tendons: ACL and PCL: score	Number	
L- Ligaments and tendons: BML/cyst	Number	
L- Ligaments and tendons: repair	Number	
L- Ligaments and tendons: Patellar tendon	Number	
L- Periarticular features:	Number	
L- Periarticular features: Infrapatelar bursa signal:	Number	
L- Periarticular features: Popliteal cyst	Number	
L- Periarticular features: Other periarticular features	Number	
L- Hoffa's fat synovitis	Number	
L- Synovitis / effsuion	Number	

7. Conclusions and future work

In this document we presented the current data management infrastructure of OActive project. This is based on the current flow of data samples and we have concluded that Airtable is the best solution that we could employ to handle their current volume.

In the future, as already mentioned in the description of work (DoW), another online infrastructure and in particular VPH-Share[1] could be used given that the volume of available data samples would be increased. VPH-Share is an online infrastructure that offers *data services, tools, metadata management and security* for sharing and accessing biomedical data.

In terms of data services, VPH-Share supports handling of *structured data* and *Large Binary Data.* Structured data are record collections that share a common schema such as those found in comma-separated values (CSV) files and relational databases. VPH-Share could access and query those interfaces via SPARQL Protocol and RDF (Resource Description Framework) Query Language and SQL (Structured Query Language). These two interfaces sit on top of MySQL databases for each individual data set. *Large Binary Data,* such as images, computational models and physiological signals are stored in a cloud based file-system and could be accessed through the Web Distributed Authoring and Versioning (WebDAV) protocol.

In terms of tool provisioning, VPH-Share cloud-based approach supports the sharing of tools and applications using virtualisation technologies, in particular those used in the OpenStack11 cloud model.

Metadata management for VPH-Share is relatively simple and integration with it, is a relatively simple task. There are many search services offered over this data in terms of semantic search, free text search, and faceted search for specific resource types.

Finally, VPH-Share's security has been developed as its own proprietary model based on the requirements of its user base. The model is distributed in its enforcement, simple to use and precluded the use of strong authentication such as certificate based access. This is due to the fact that VPH-Share should not contain any sensitive information, especially regarding clinical data. The data publication process should remove this before it reaches the network. VPH-Share supports the concept of a security proxy that sits in front of every service and evaluates an access policy against the security ticket contained in the web service request. This ticket contains all of the roles the user has in the VPH-Share network and access policies are distributed across the network to these proxies from a central repository. This approach of using a proxy means that the developers of applications and tools are completely unaware that their applications are being protected and have no integration overheads from a security perspective, which is highly desirable. It also means that access policies are made local to the resource, thereby removing the dependence of a central authority making the decisions, which, on a large scale, makes the system perform much better.

---

[1] http://www.vph-share.eu/