



**Health, demographic change and wellbeing**  
**Personalising health and care: Advancing active and healthy ageing**  
**H2020-PHC-19-2014**  
**Research and Innovation Action**



***Deliverable D4.4***  
***Social Activity Repository (Final)***

|   |   |
|---|---|
| <b>Deliverable due date: May.2017</b>   | <b>Actual submission date: 29.05.2017</b> |
| <b>Start date of project: September 1, 2011</b>                                     | <b>Duration: 36 months</b>                |
| <b>Lead beneficiary for this deliverable: ATOS</b>                                  | <b>Revision: ---</b>                      |
| <b>Authors: Ivo Ramos (ATOS), Ingo Brauckhoff (ATOS), Maurizio Marchese (UNITN)</b> |   |
| <b>Internal reviewer: Andrew McNeill</b>  |   |

|  |  |          |
|--|--|----------|
| <b>The research leading to these results has received funding from the European Union's H2020 Research and Innovation Programme - Societal Challenge 1 (DG CONNECT/H) under grant agreement n°643644</b> |  |          |
| <b>Dissemination Level</b>   |  |          |
| <b>DEM</b>   | Demonstrator   | <b>X</b> |
| <b>CO</b>  | Confidential, only for members of the consortium (including the Commission Services) |          |

The contents of this deliverable reflect only the authors' views and the European Union is not liable for any use that may be made of the information contained therein.



# Contents

- EXECUTIVE SUMMARY .....4
- 1. ACTIVITY MODEL .....5
- 2. SOCIAL ACTIVITY REPOSITORY .....7
  - 2.1. ARCHITECTURE AND TECHNOLOGIES.....9
- RELATION WITH OTHER WORKPACKAGES .....13
- BIBLIOGRAPHY .....14
- ANNEX .....15

## Executive Summary

This document describes the final implementation of the social activity repository based on the ultimate version of the social activity models and also on the harvesting of open data services such as local and dedicated news/events feeds that produce a usable set of information of potentially interesting social activities (e.g. movies, concerts, cultural events, exhibitions, talks, meetings, etc.). This component makes part of the Cyber Physical Social Network (CPSN) architecture and data flow. The architecture changed a little since the preliminary version (see D4.3 – “*Social Activity Repository (preliminary)*”), with the exclusion of Kafka and Connect and the inclusion of Mahout and Hadoop HDFS. In this last version, the social activity repository includes dynamic information harvesting from local and dedicated news/events feeds providing potentially interesting social activities. The social activity is a concept related to group activities, like having a walk in the park with someone, going shopping, or going to an art exhibition. It is important to understand the possible characteristics of social activities that could eventually be relevant for better recommendations and exploitability of the activities.

First, in section 1, we present the final version of the social activity models based on the version presented in D2.4 – “*User, activity and environmental description: Final release models for user, activity and environment*” [2]. The social activity model is presented through a graph representation. In section 2, we present an analysis of the limitations executing social activity harvesting from known social media platforms (*Facebook*, *Google +*, etc.), the final version of the Social Activities Recommendation System Architecture, and we demonstrate how the Activity Harvester works.

# Chapter 1

## Activity Model

Figure 1 depicts the current status of the structure of a Social **Activity** model in the ACANTO KnowledgeBase together with some of its relationships as presented and detailed in D2.4 “User, activity and environmental description: Final release models for user, activity and environment” [1].

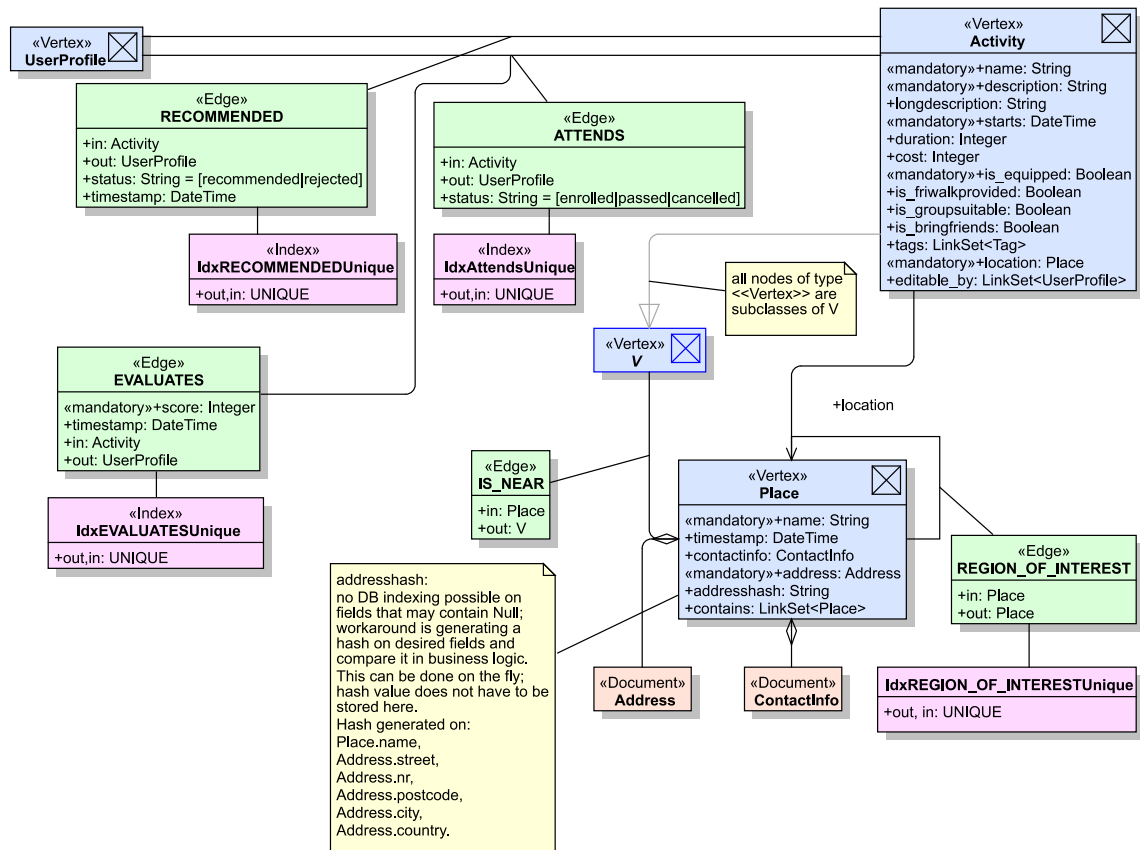


Figure 1: Social Activity schema

We recall here, briefly the main elements of the Social Activity model: every vertex of type **Activity** has a series of properties to describe the nature of a social event. Moreover it contains information on the suitability of the activity for group visits and – eventually – its compatibility with the *FriWalk* and whether *FriWalks* are available on site. The location of an **Activity** points to a **Place** node, which holds information about address, name of the place and its geolocation (latitude and longitude) as optional data.

The relationships between **Activities** and **UserProfiles** are described using edges of type **ATTENDS** and **RECOMMENDED**. The latter of which is being created by the recommender system. Once a user interacts with a recommended **Activity** an **ATTENDS** edge is created, in case he enrolls to attend the activity, or the **RECOMMENDED** edge's property *status* is set to 'rejected' if the user chose to reject the proposed activity.

**EVALUATES** edges are created once the related activity has been performed and the user has confirmed its level of satisfaction about the event giving a **score** from one to five stars, being one star the least and five stars the most satisfactory experience. These edges will be then used by the recommender system to update and improve the recommendations.

The **IS\_NEAR** edges are available to mark the proximity of another object; the type of the pointed to (out) vertex is the generic class *V* and allows therefore to connect all subtypes of *V*. The only node types of interest are *Place* (say a Toilet or Bus stop), *UserProfile* (another person) and maybe *Circle* to indicate the proximity of a group of people. The Social Activity schema is the main model description of the data representation and the relationships for activities in ACANTO.

## Chapter 2

### Social Activity repository

The design and implementation of the Social Activity repository was based on the social activity model, the possibility of harvesting existing social networks and local and dedicated news/events feeds with potentially interesting social activities (e.g. movies, concerts, cultural events, exhibitions, talks, meetings, etc.). The purpose is to collect social activities information at specific social environments based on the local communities, social networks and local providers. The social activity observation is obtained through a layer of applications that get indoor and outdoor information into the mobility record. The goal is to connect formal caregivers, informal caregivers, and Medical Doctors, provide helpful information to the users, and to exploit the information coming from the user's daily activities and provide this context to the professionals. These topics will be analysed in the next sections.

#### KNOWLEDGE BASE

The KnowledgeBase aggregates data from user profiles, circles, tags, evaluations, environments and activities, as shown in the Figure below and it is used, amongst other things, to provide the collected information to the recommender system. The KnowledgeBase diagram did not undergo changes since the version presented in D2.4 [2]. Therefore, in order to avoid the duplication of writing, we recommend reading that document for further details.

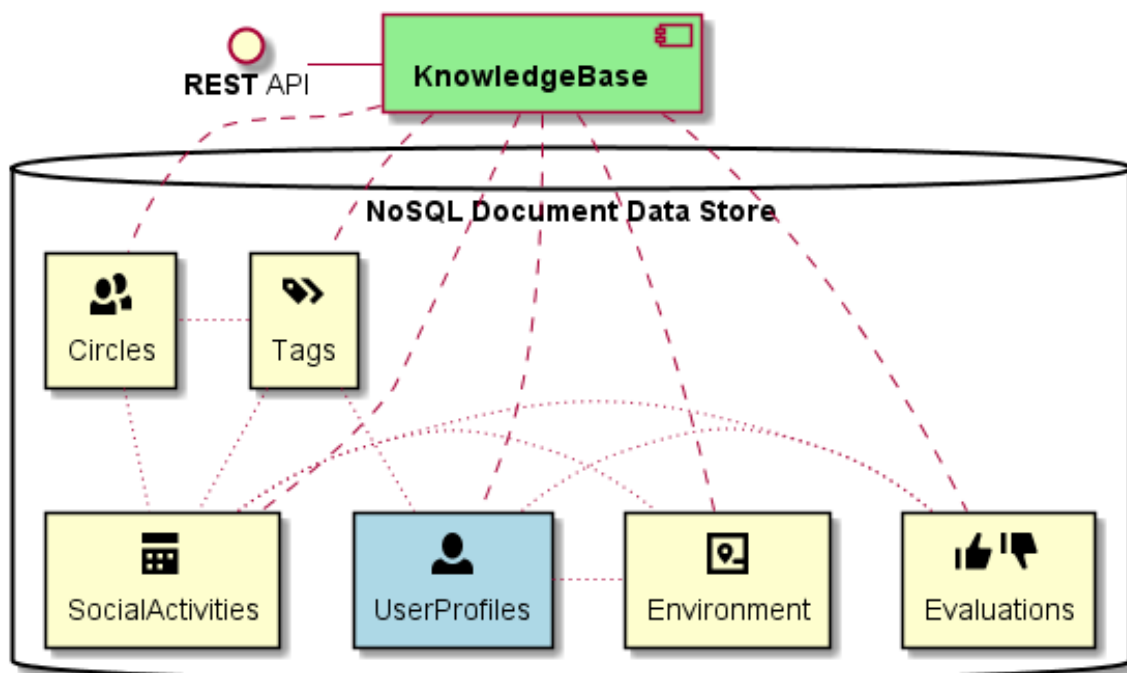


Figure 2: Knowledgebase diagram

Figure 3 represents an example of the Orient DB vertex classes created for ACANTO (Activity, Circle, Evaluation, Place, and Tag).

Vertex Classes

+ NEW VERTEX

| Name       | Color | SuperClasses | Alias | Abstract                 | Clusters      | Default Cluster | Cluster Selection | Records | Actions  |
|------------|-------|--------------|-------|--------------------------|---------------|-----------------|-------------------|---------|--|
| Activity   |       | V            |       | <input type="checkbox"/> | [65,66,67,68] | 65              | round-robin       | 6,000   | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| Circle     |       | V            |       | <input type="checkbox"/> | [61,62,63,64] | 61              | round-robin       | 3,182   | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| Evaluation |       | V            |       | <input type="checkbox"/> | [69,70,71,72] | 69              | round-robin       | 0       | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| Place      |       | V            |       | <input type="checkbox"/> | [57,58,59,60] | 57              | round-robin       | 12      | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| Tag        |       | V            |       | <input type="checkbox"/> | [21,22,23,24] | 21              | round-robin       | 142     | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |

1

2

Figure 3: Example of the Orient DB vertex classes

The Figure 4 represents an example of the Orient DB Edge classes created for ACANTO (ATTENDS, CAREGIVER\_OF, CONNECTED\_TO, E, EVALUATES).

Edge Classes

+ NEW EDGE

| Name         | Color | SuperClasses | Alias | Abstract                 | Clusters          | Default Cluster | Cluster Selection | Records | Actions  |
|--------------|-------|--------------|-------|--------------------------|-------------------|-----------------|-------------------|---------|--|
| ATTENDS      |       | E            |       | <input type="checkbox"/> | [77,78,79,80]     | 77              | round-robin       | 15      | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| CAREGIVER_OF |       | E            |       | <input type="checkbox"/> | [93,94,95,96]     | 93              | round-robin       | 3       | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| CONNECTED_TO |       | E            |       | <input type="checkbox"/> | [101,102,103,104] | 101             | round-robin       | 0       | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| E            |       |              |       | <input type="checkbox"/> | [13,14,15,16]     | 13              | round-robin       | 128,654 | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |
| EVALUATES    |       | E            |       | <input type="checkbox"/> | [109,110,111,112] | 109             | round-robin       | 30,001  | <a href="#">RENAME</a> <a href="#">QUERY ALL</a> <a href="#">+ NEW RECORD</a> <a href="#">DROP</a> |

1

2

3

Figure 4: Example of Orient DB Edge classes

Figure 5, exemplifies the relation between social activities, user profiles and circles. It also represents how social activities are recommended to a circle, to the users that belong to that specific circle, how the activities are evaluated and the CAREGIVER\_OF relation between users.

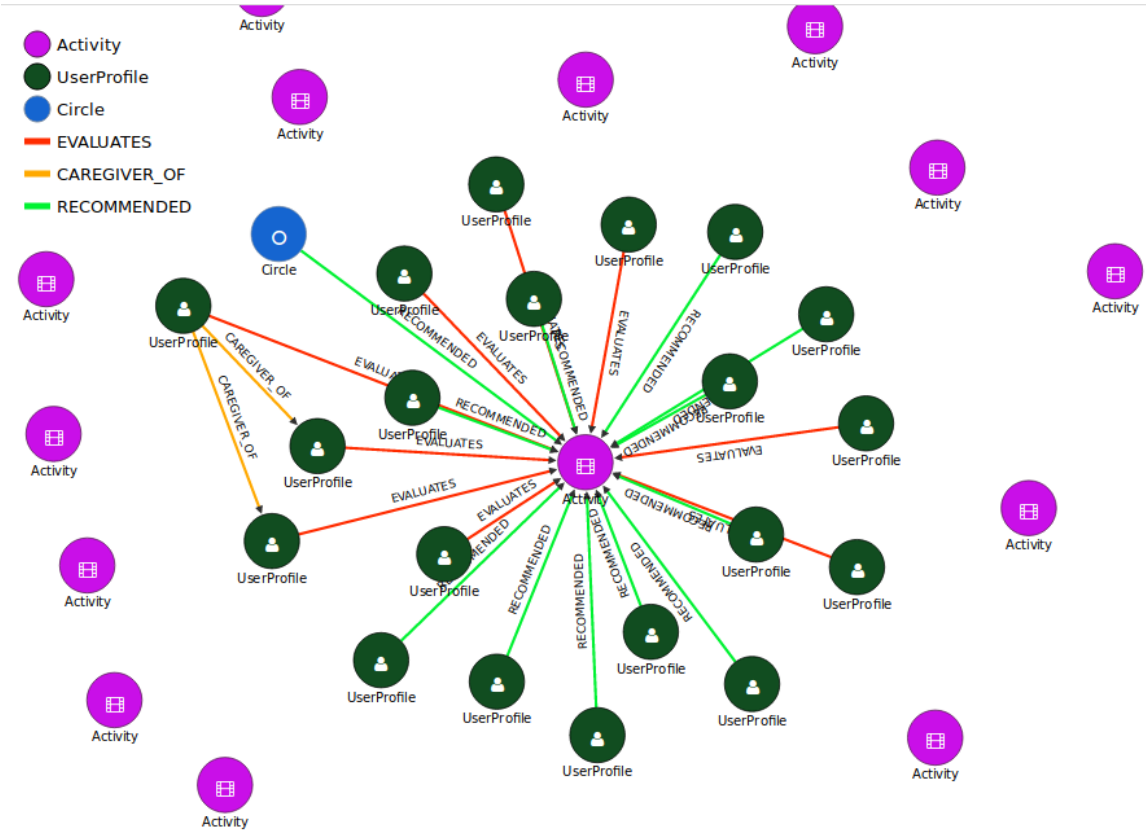


Figure 5: Graph representation example of the relation between social activities, user profiles and circles



## Activity Harvester

During the design phase an analysis on the most common social networks was made in order to study the feasibility to “harvest” social activities or events from known social networks and import them into the Cyber Physical Social Network (CPSN).

### Social Activities harvesting from known social media platforms:

In our first approach, we considered the possibility to use popular social networks, such as *Facebook*, *Google+*, etc., to “harvest” information about social activities or events for reuse in the CPSN, but privacy restrictions, lack of control over the activities data and their relations, and the fact that data is not available freely anymore, and additionally is not searchable in a regional context (geolocation), made none of them suitable for social activities harvesting. For instance, *Facebook* increased privacy restrictions some time ago and is basically a commercial platform, where there is almost no control of the content and associated users. Also, we found that the data access of the Application Programming Interface (API) is in constant change and *Facebook* gives no warranty for stability for more than two years. In the case of *Google+*, as a competing platform to *Facebook* it is a dying project. We found the same restrictions as with *Facebook* and the *Google+* ‘events’ already have been discontinued.

Consequently, we concluded that the use of the Really Simple Syndication (RSS) [4] of regional official websites is the option that better fits the requirements of the social activities harvester. These websites generate a usable amount of dedicated news/events feeds about interesting social activities (e.g. movies, concerts, cultural events, exhibitions, talks, meetings et al.) and they are considered safe. But, the addition of these external social activities should be monitored by a different user role: the “Activity Creator”.

RSS allows website owners to organize their website content. It allows an easy and light way to share and view information and content. Files can be automatically updated and it is based on the eXtensible Markup Language (XML). RSS was designed to process selected data automatically. Without RSS, users would have to check for updated website contents every day. In the ACANTO CPSN, we have created a specific role (Activity Creator) responsible for selection and control of thematic social activities from specific websites. This means that without RSS, the Activity Creator would need to daily check the websites for new updates and to create new Activities manually. Since RSS data is light and fast-loading, it can easily be used with services for cell phones or tablets.

RSS is used by web sites that are updated frequently, such as:

- News sites - Lists news with title, date and descriptions
- Companies - Lists news and new products
- Calendars - Lists upcoming events and important days
- Site changes - Lists changed pages or new pages

The Figure below shows an example of the list of events obtained from a feed about cultural social activities on the regional website journal of Trento [5].

## Harvest Activity

## List of events in channel [CULTURA]

**La terza e ultima giornata del Festival delle Lingue**

La terza e ultima giornata del Festival delle Lingue di Rovereto si è aperta con due spettacoli al Cinema Teatro Rosmini di Rovereto, che hanno visto protagonisti i ragazzi trentini. Da un lato il musical "L'elisir della bontà", messo in scena dagli scolari dell'IC Pergine 1, dall'altro la prima tappa dell'Odissea, nell'ambito del progetto "Lingue in scena", che scaturisce da una collaborazione fra Iprase e la Città di Torino. Al mattino, ad animare il Festival, alcuni relatori interessanti come Nicola Gardini, che insegna letteratura italiana ad Oxford e ha parlato dell'etimologia delle parole, e Marianella Sclavi, già docente al Politecnico di Milano, che ha divertito il pubblico nel suo seminario legato alla "conversazione e dinamiche dell'umorismo". Fra gli approfondimenti anche la tavola rotonda sull'alternanza-scuola lavoro all'estero. Per tutto il giorno il pubblico ha affollato il Festival e visitato gli stand allestiti nel piazzale antistante la sede di Iprase, confermando il successo di questa seconda edizione. Come infatti ha evidenziato il presidente del Comitato scientifico di Iprase, Mario Giacomo Dutto, oggi si chiude il Festival ma è solo un arrivederci alla terza edizione: "Siamo soddisfatti per la collaborazione che abbiamo trovato, per il programma ricco e articolato che guarda al futuro, inoltre la risonanza che questo evento ha avuto anche a livello nazionale, ci conferma che sul tema delle lingue vale la pena lavorare e investire".

[no tags available](#)

[open](#) [copy to activity list](#)

---

**Al Festival delle Lingue Sabatini e le "Lezioni di italiano"**

Due sono stati i "big" che hanno animato il pomeriggio del secondo giorno del Festival delle Lingue, a Rovereto. Da un lato la scrittrice Andrea Marcolongo, autrice di "La lingua geniale. 9 ragioni per amare il greco" (editori Laterza), vero e proprio caso letterario con 80.000 copie vendute, dall'altro Francesco Sabatini, presidente emerito dell'Accademia della Crusca, che al Festival ha portato le sue "Lezioni di italiano. Grammatica, storia, buon uso" (Mondadori) e ha spiegato come la lingua sia impressa nel cervello di ognuno di noi fin dai primi anni di vita. Tante persone, nel pomeriggio, hanno poi seguito non solo i seminari e workshop del Festival, i percorsi didattici e i laboratori, ma hanno affollato anche gli stand allestiti nel piazzale antistante la sede di Iprase, curati da musei, istituti ed enti, scuole, associazioni, nonché da case editrici

[change URL](#)

Figure 6: List of social activities related with culture, taken from the website [www.ufficiostampa.provincia.tn.it](http://www.ufficiostampa.provincia.tn.it)

The Activity Creator will be able to select and add the social activity into the CPSN by pressing the link "[copy to activity list](#)". It is also possible to edit and add more ACANTO related information to the social activity, for instance, if the location is equipped with FriWalks, or if it is suitable for group visits, etc. (see figure below).

## Harvest Activity

## Modify ACANTO Social Activity .

Name (Required)

La terza e ultima giornata del Fes

Description (Required)

Please enter a short description for the Activity.

## Detailed description

Rich text editor toolbar with icons for bold, italic, underline, strikethrough, bulleted list, numbered list, link, unlink, image, video, table, undo, redo, and other formatting options. The text area contains the same content as Figure 6, including the description of the Festival delle Lingue.

[more...](#)

body

Is equipped?

☒ Yes ☐ No

Suitable for group visits?

☒ Yes ☐ No

FriWalkers available?

☒ Yes ☐ No

Can bring friends?

☒ Yes ☐ No

Figure 7: Example of a social activity selection and editing

## 2.1 Architecture and technologies

The Social Activity repository, as well as the CPSN Architecture and Data Flow, did not undergo any changes since the last version presented in D4.8 – “*Social Activities Recommendations*” [1]. In particular, the social activity repository is related to number 6 in the figure, i.e. Data ingestion from diverse sources, and number 5, i.e. Apache Mahout + Apache Spark MLlib [6].

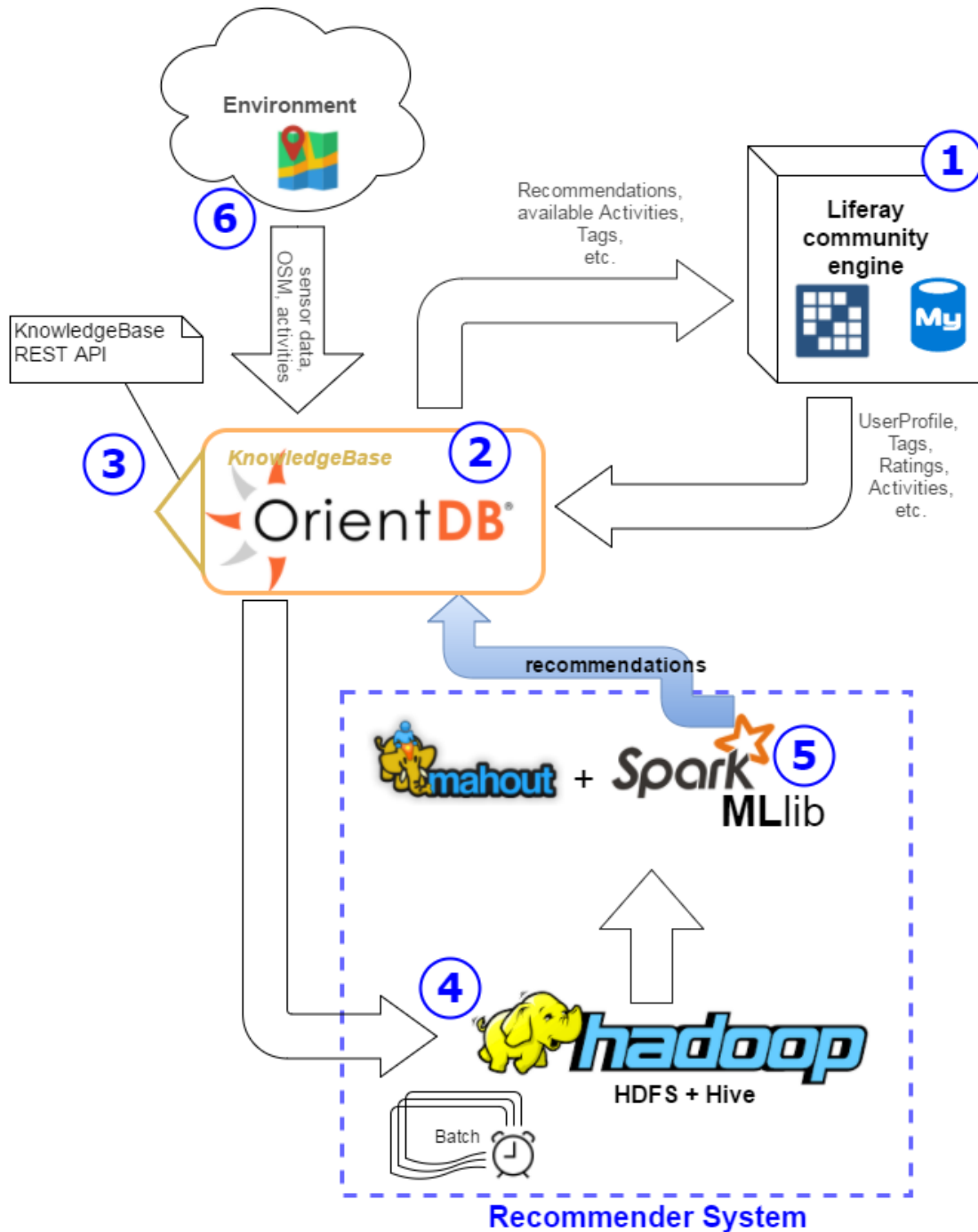


Figure 8: CPSN Architecture and Data Flow

We consider that the repetition of the description of the CPSN Architecture and Data Flow would not be needed. For further details, we recommend reading the description in D4.8 – “*Social Activities Recommendations*” [1].

Our artificial dataset has 6000 social activities located at different states of the USA. In order to get more “realistic” data we have decided to create an “Activity Creation” form and circulate it among our project partners, especially universities (UNITN, UNAN). The idea is to create a set of new activities with locations around Madrid, Trento and Northumbria, in order to support the recommendation system. A screenshot of this form is provided in the annex of this document.

## Relation with other work packages

The work presented in this document is the final version of the continuous model development carried out in WP2, which drew on user-input from WP1 to identify the kinds of activities users like and what information they want to know before attending an activity. Specifically, we have updated and implemented the latest models of Social Activities. However, it is important to refer that the models can undergo potential updates until the end of the project. It is also related with the WP5 – “Execution Support of Social Activities”, explicitly with task 5.2 (Activity Planning), where the output of the Recommender system is processed and translated to an executable plan. The current implementation of the Social Activities will also be the base for its evolution towards the inclusion of user information coming from dynamic sources (i.e. sensing data stored in the knowledge base by the WP3, at task 3.2 – “Perception of the environment” and displayed in the CPSN, as well as from the evolving social data coming directly from the users’ social network). Additionally, the components developed in this work have been designed and implemented in compliance with the overall cloud ACANTO infrastructure developed in WP7.

Finally, the work of this WP will be integrated and used in the experimental validation of WP8 where real users will receive recommendations of real activities in their areas. Already the groundwork is being laid for this validation through the Activity Creation Form (Annex) by means of which the project partners are contributing data about relevant activities in their areas.

## Bibliography

- [1] I. B. M. M. Ivo Ramos, "D4.8 - Social Activities Recommendations," 2017.
- [2] I. B. Ivo Ramos, User, activity and environmental description: Final release models for user, activity and environment, 2017.
- [3] OrientDB, "OrientDB programming language bindings," 2017. [Online]. Available: <http://orientdb.com/docs/last/programming-language-bindings.html>.
- [4] Wikipedia, "Wikipedia," 2017. [Online]. Available: <https://en.wikipedia.org/wiki/RSS>. [Accessed 2017].
- [5] Provincia Autonoma di Trento, "Trentino - Quotidiano online della Provincia Autonoma di Trento," 2017. [Online]. Available: <https://www.ufficiostampa.provincia.tn.it/feed/rss/tag/117>. [Accessed 2017].
- [6] Apache, "Apache Spark MLlib," 2016. [Online]. Available: <http://spark.apache.org/mllib/>. [Accessed 8 May 2016].
- [7] L. P. Ivo Ramos, D2.4 - User, activity and environmental description: Final release models for user, activity and environment, 2017.

# ANNEX

## Acanto Activity creation form

\*Required

### 1. Name \*

The name of the activity, e.g. 'Dance under the Sea', 'Black Friday shopping', 'Star Wars Movie Night'

---

### 2. Description (short) \*

a short description of what the activity is about, e.g. 'watch 3 classic movies in a row'

---

### 3. Description (detailed) \*

a more elaborate and exhaustive description

---

---

---

---

---

## Categorization

### 4. Tags \*

Select the tags to categorise the type of activity. Mark all that apply.

*Tick all that apply.*

☐ Museum

☐ Cinema

☐ Walking

☐ Concert

☐ Picnic

☐ Theatre



- ☐ Table\_Games
- ☐ Fishing
- ☐ Shopping
- ☐ Cats
- ☐ Dogs
- ☐ Meeting\_People
- ☐ Music
- ☐ Cars
- ☐ Motorcycles

#### 5. Mobility Constraints \*

Select all possible constraints that may apply to this activity.

*Tick all that apply.*

- ☐ balance :: activity involves balance activities
- ☐ visually\_impaired :: activity suitable for visually impaired
- ☐ toilets\_nearby :: toilets available nearby
- ☐ hard\_of\_hearing :: activity suitable for those with bad hearing
- ☐ need\_Assistance :: provides assistant to users with physical limitations
- ☐ blind :: suitable for blind people
- ☐ wheelchair :: site of activity is wheelchair accessible
- ☐ FriWalker :: site of activity is FriWalker accessible
- ☐ deaf :: suitable for deaf people
- ☐ no\_walk\_longer\_distances :: does not involve long distance walking (less than 1km of walking)
- ☐ walk\_longer\_distances :: involves lots of walking (more than 1km)
- ☐ build\_muscle\_strength :: helps to build muscle strength

## When?

When does this activity take place?

#### 6. Startdate and time \*

*Example: 15 December 2012 11.03 a.m.*

#### 7. Event Duration \*

in minutes, for example 60 or 90 for events that take an hour or an hour and

a half

---

## Where?

Where does this activity take place?

### 8. Location Name \*

enter the name of the place the event takes place, for example: 'Fancy Shopping Mall', 'Fine Art Museum', 'Hennessey Park'

---

### 9. Street \*

---

### 10. Housenumber \*

as a whole number, in case a house number is not available enter 0

---

### 11. Postcode \*

---

### 12. City \*

---

### 13. Latitude [optional]

Decimal number as provided by Google Maps.

---

### 14. Longitude [optional]

Decimal number as provided by Google Maps.

---