【별첨 4】

Design Document

for

Development of Mining Minds Core Technology Exploiting Personal Big Data

(Project No. 10049079)

Professor Sungyoung Lee

Ubiquitous Computing Laboratory
Kyung Hee University, Korea

Kyung Hee University

April 24, 2015

Contents

Chapter 1
Mining Minds Ver. 1.5 Overview

Abstract

The world is witnessing a spectacular shift in the delivery of health and wellness care. The key ingredient of this transformation consists in the use of revolutionary digital technologies to empower people in their self-management as well as to enhance traditional care procedures. While substantial domain-specific contributions have been provided to that end in the recent years, there is a clear lack of platforms that may orchestrate, and intelligently leverage, all the data, information and knowledge generated through these technologies. Mining Minds presents an innovative framework that builds on the core ideas of the digital health and wellness paradigms to enable the provision of personalized healthcare and wellness support. Mining Minds embraces some of the currently most prominent digital technologies, ranging from Big Data and Cloud Computing to Wearables and Internet of Things, and state-of- the-art concepts and methods, such as Context-Awareness, Knowledge Bases or Analytics, among others. A limited set of these technologies are demonstrated in the initial prototype version of Mining Minds. The purpose of the prototype MM V1.5 is to build Mining Minds infrastructure and implement physical activities scenario for providing personalized services to the users. The scope of this version is limited to accelerometer, GPS, and smart watch wearable device with services only provided under the umbrella of physical activities. The proposed framework aims at thoroughly describing the efficient and rational combination and interoperation of these modern technologies and methods through Mining Minds, while meeting the essential requirements for personalized health and wellness support.

1. Introduction

The provision of healthcare and wellness services is expected to drastically change in the upcoming years. This change has particularly been accelerated by the global socioeconomic situation. Cuts in government spending, an increasing population of pensioners and a growing unemployment rate are critical challenges for most transitioning and developed countries that add urgency to the need of finding new healthcare solutions. Although the efficient delivery of healthcare and assistance is found of key importance, it is equally or more crucial to reduce as much as possible the need of care. In fact, it is well-known that most prevalent diseases are partly caused by lifestyle choices that people make during their daily living. For example, epidemic illnesses such as obesity are essentially due to unwholesome diets and lack of physical activity. Accordingly, a strong interest has lately been shown by healthcare stakeholders in people's lifestyle management and personal self-care. Bringing these lifestyle diseases under control may have a great impact on healthcare and assistance spending, and certainly on health itself. To support this new healthcare and wellness era a new concept has recently emerged. Commonly known as "Digital Health", it refers to a renovated and updated view of what was already cataloged in the early 2000s as Digital Health Care, which embraces an enormous variety of social, scientific and technological disciplines to empower people to better track, manage, and improve their health and well-being. Likewise, Digital Health joins medical and social knowledge with cutting-edge technologies to reduce inefficiencies in healthcare delivery, improve access, reduce costs, increase quality, and make medicine more personalized and precise. Key concepts supporting this digital health revolution include, but are not limited to, Mobile Health, Connected Health, Wearable Computing, Ubiquitous Computing, Big Data, Cloud Computing and Gamification, among others. One of the most important challenges posed by the Digital Health refers to the use of these technologies in a rational and comprehensive manner, as well as their integration with current and future personalized health services and business.

1.1. Definition

"Mining Minds is a collection of services, tools and techniques working collaboratively to investigate on human's daily routines to provide a personalized well-being and health-care support"

1.2. Motivation

The idea of Mining Minds covers multiple domains and set of personalized services that can benefit different stakeholders. To cover all aspects related to this innovative platform, and incremental strategy for services, components development, and infrastructure building is employed. Multi versions strategy is devised to achieve the overall objective of the Mining Minds. Our motivation for MM V1.5 is to develop the infrastructure to support physical activities scenario with smartphone sensors and smart watch as communicating devices. The physical activity scenario targets non patient as end users and the delivery mechanism is personalized wellness services.

1.3. Scope of Mining Mind Ver. 1.0

The objective of Mining Minds Ver.1 is to develop the infrastructure of Mining Minds in constrained environment. The services scope is related to physical activities that recognize only five activities using smartphone's accelerometer and GPS sensors. The recommendations to the end users are calories based personalized physical activities; personalization handled by the preference of activities information. The summary of the details of this version is shown in Table 1.1.

Table 1.1: Scope of Mining Minds Ver. 1.0

Input Data	Activities Recognized	Business Model	Services	Technical Contribution
Smartphone-based Accelerometer and GPS	1. Walk 2. Jog 3. Rest 4. Ride the bus 5. Take the subway	Business to consumer (B2C)	Calories-based physical activities recommendations	• Curation DCL : Data Representation and Mapping SCL : Rule-based Reasoning • Token-based Authentication • UI • Admin View • Communication Request-based SOAP Web services

1.3.1. Mining Minds Ver. 1.0 Architecture

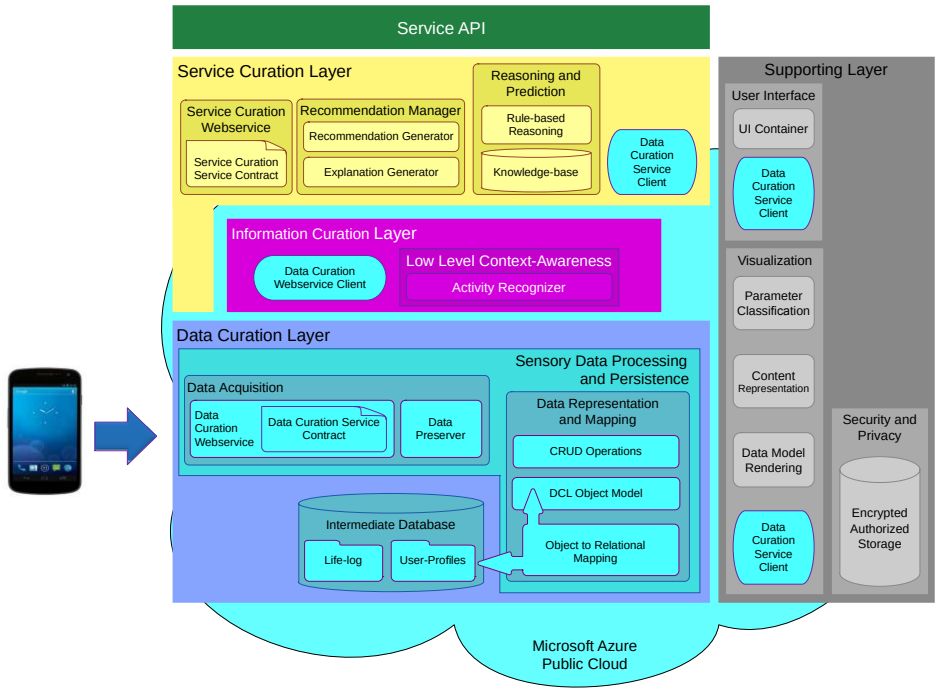

Figure 1.1: Mining Mind Ver. 1.0

1.3.2. Technical Limitation

- Poor Communication Performance due to SOAP-based Web services
- Stress on RDBMS due to excess of data storage and persistence
- Primitive amount of Activities were identified (5 only)
- Request-based Recommendation generation
- Message integrity can be compromised during communication

1.4. Mining Minds Ver. 1.5

The objective of Mining Minds Ver. 1.5 is to develop the infrastructure of Mining Minds in controlled environment. The services scope is related to physical activities that recognize only fifteen activities using smartphone's accelerometer and GPS sensors along with wearable smart watch. The recommendations to the end users are educational facts, healthy habits induction and calories based personalized physical activities; personalization handled by the preference of activities information. The summary of the details of this version is shown in Table 1.2.

Table 1.2: Scope of Mining Minds Ver. 1.5

Input Data	Activities Recognized	Business Model	Services	Technical Contribution
Smartphone-based Accelerometer & GPS, and Wearable device	1. Standing 2. Walking 3. Sitting 4. Jogging 5. Running 6. Hiking 7. Elevator up 8. Elevator Down 9. Escalator Up 10. Escalator Down 11. Eating 12. Dancing 13. sweeping 14. Laying 15. Exercising	Business to consumer (B2C)	• Calories-based physical activities recommendations • Educational facts • Healthy habits induction	• Curation DCL : Big data and Life-log Representation & Mapping KCL : Rule-authoring SCL : Generalized and Personalized Recommendations • Security and Privacy Access model for domain experts Data integrity check during communication • UI Expert View Analytics • Communication Real-time socket based communication Restful services with cloud push model

1.4.1. Mining Minds Ver. 1.5 Architecture

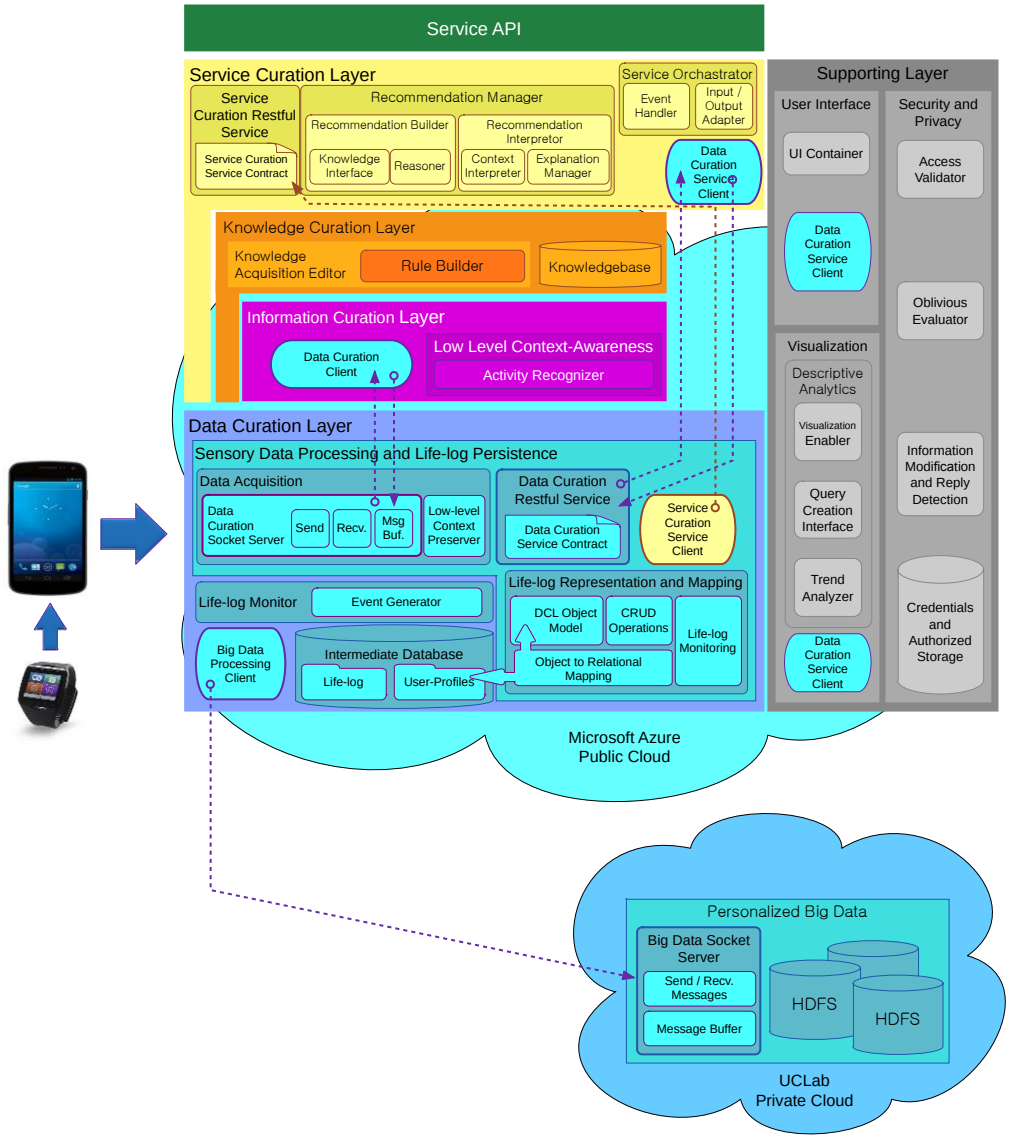

Figure 1.2: Mining Minds Ver. 1.5

1.4.2. Technical Limitation

- Object Oriented and Relational life-log models
- Big data is only used for persistence
- Lack of High-level context determination
- Only static rules as knowledge
- Anonymized Access

1.5. Functional View

The Mining Minds platform Ver. 1.5 is divided into four layers: Data Curation, Information Curation, Knowledge Curation, and Service Curation. The data curation layers is responsible for accessing the real time data and curated through our proposed method and storing the information into Hadoop-based distributed file system as Big data. The data stored in the Hadoop-based distributed file system is processed for conversion into structured format for the consumption of upper layers, i.e., Information Curation, Knowledge Curation, and Service Curation. This structured data is persisted as human life-log in the form of a relational database termed as intermediate data. The information curation layer utilizes the information from the data layer and process the low level context for activities recognition. The knowledge curation layer creates the rules required to generate recommendations with the help of a knowledge authoring tool and a knowledgebase. The service curation layer process the data for recommendation generation based upon the rules generated by knowledge curation layer. UI/UX Authoring Tool presents these recommendations to the user in a more elaborated form. The evolution of the knowledge in the system is maintained by taking user feedback and all the layers data; information as well as processes are made secure by providing security and privacy. The high-level architecture of Mining Minds Ver. 1.5 is illustrated in Figure 1.3.

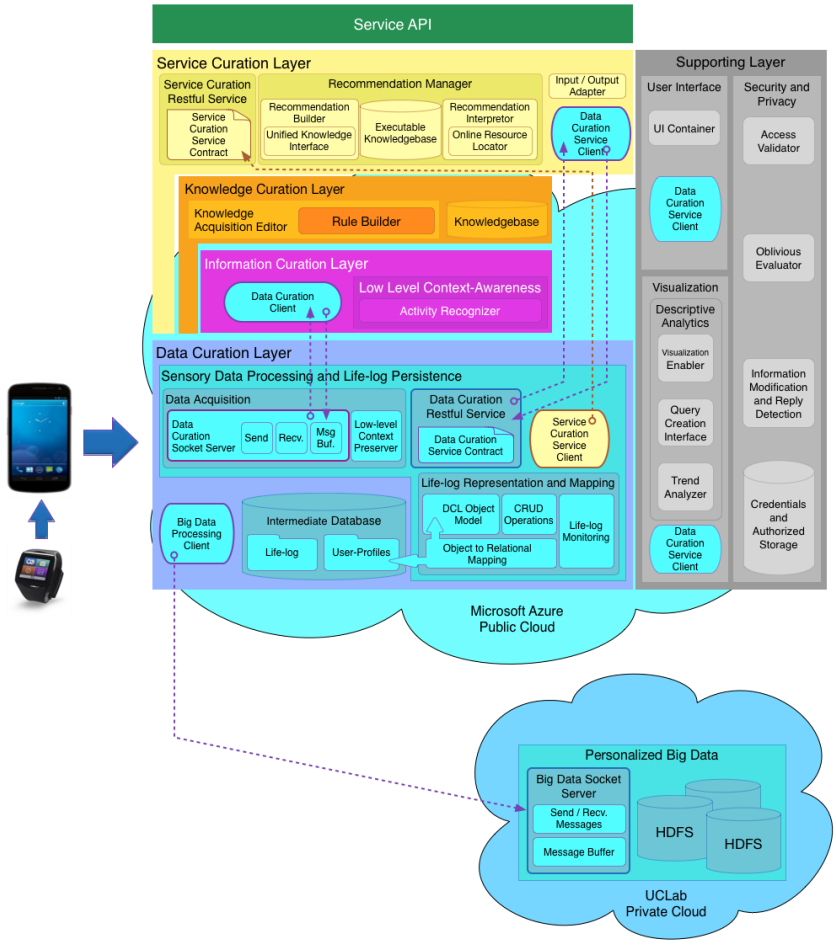

Figure 1.3: Mining Minds Platform Ver. 1.5

1.6. Use Case Diagram

The overall use case diagram of the proposed Mining Minds platform is shown in the following Figure 1.4.

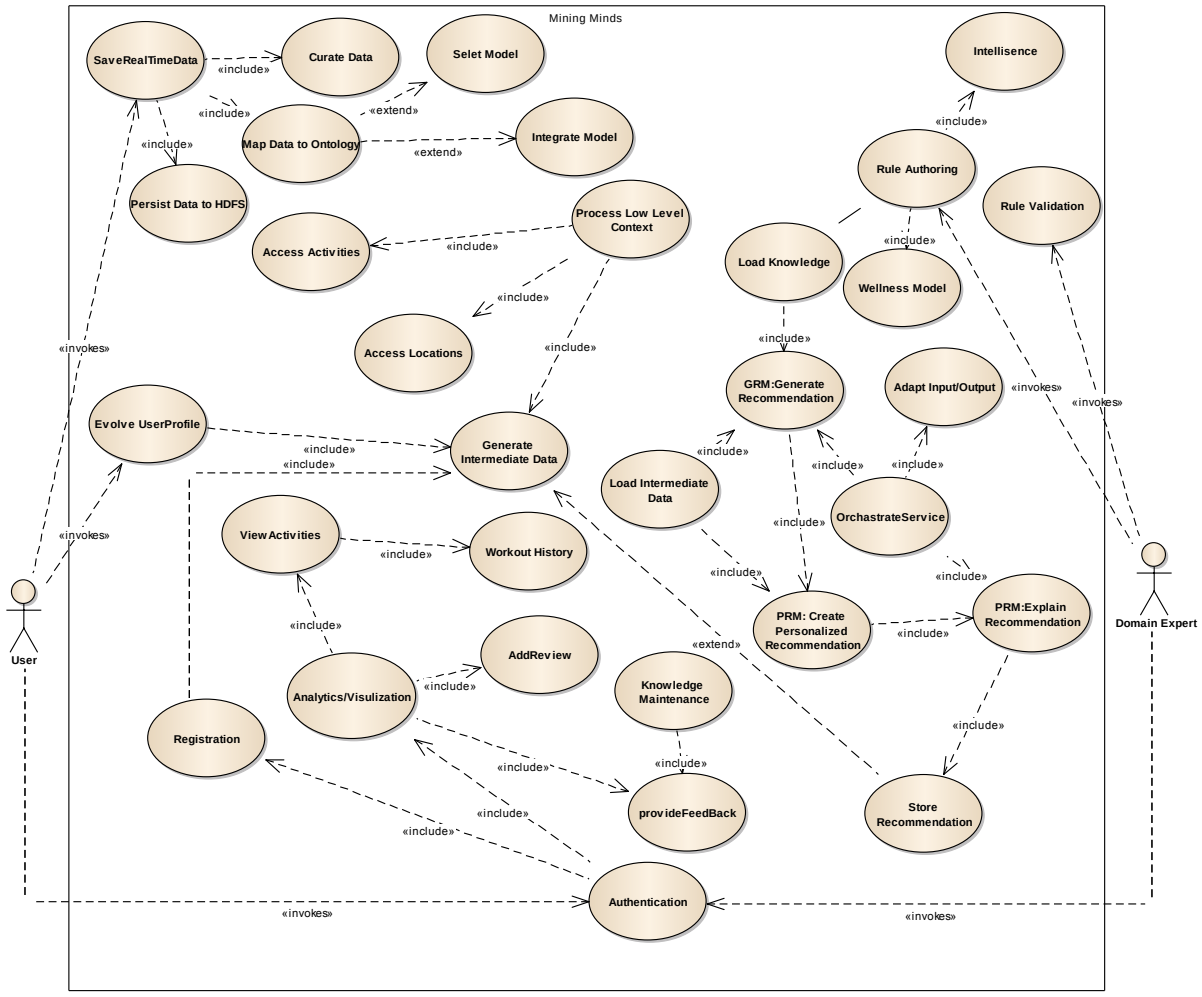

Figure 1.4: Overall Use Case Diagram

Sequence Diagram

The overall sequence diagram of the proposed Mining Minds platform is shown in the following Figure 1.5.

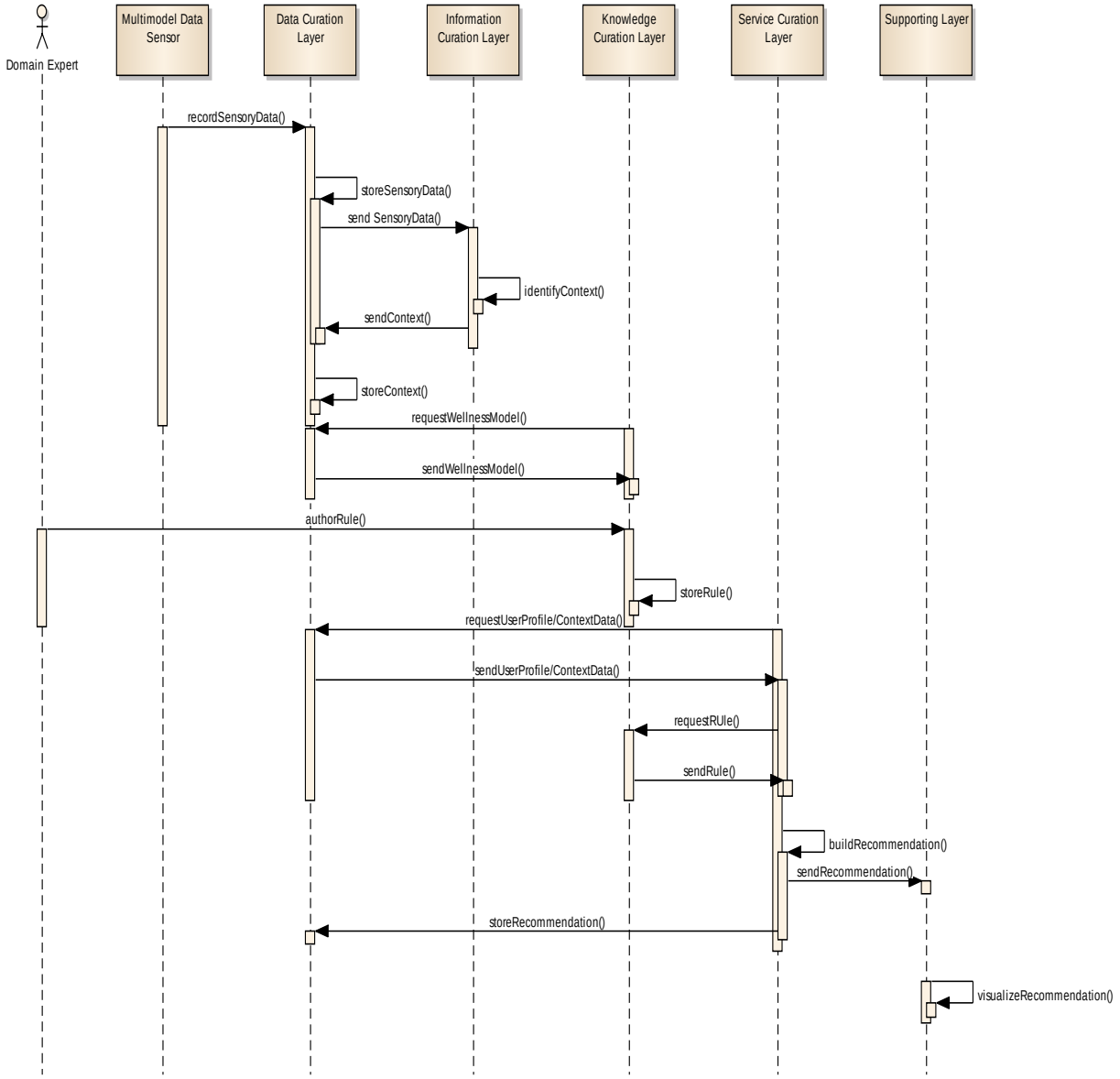

Figure1.5: Over all Sequence Diagram

Deployment Diagram

Our Proposed platform is deployed on a distributed hybrid cloud-computing environment (illustrated in above figure), where 5 families of computing virtual machines are collaborating as participating nodes. Each family of nodes hosts a respective core from mining minds architecture, Node 1 provides abstraction over the internal deployments of data curation. Data for big data storage is deployed over a private cloud infrastructure, illustrated as Node 2 to Node 5. Information curation is deployed over Node 6. Node 7 hosts knowledge curation layer. Service curation layer is deployed over Node 8.

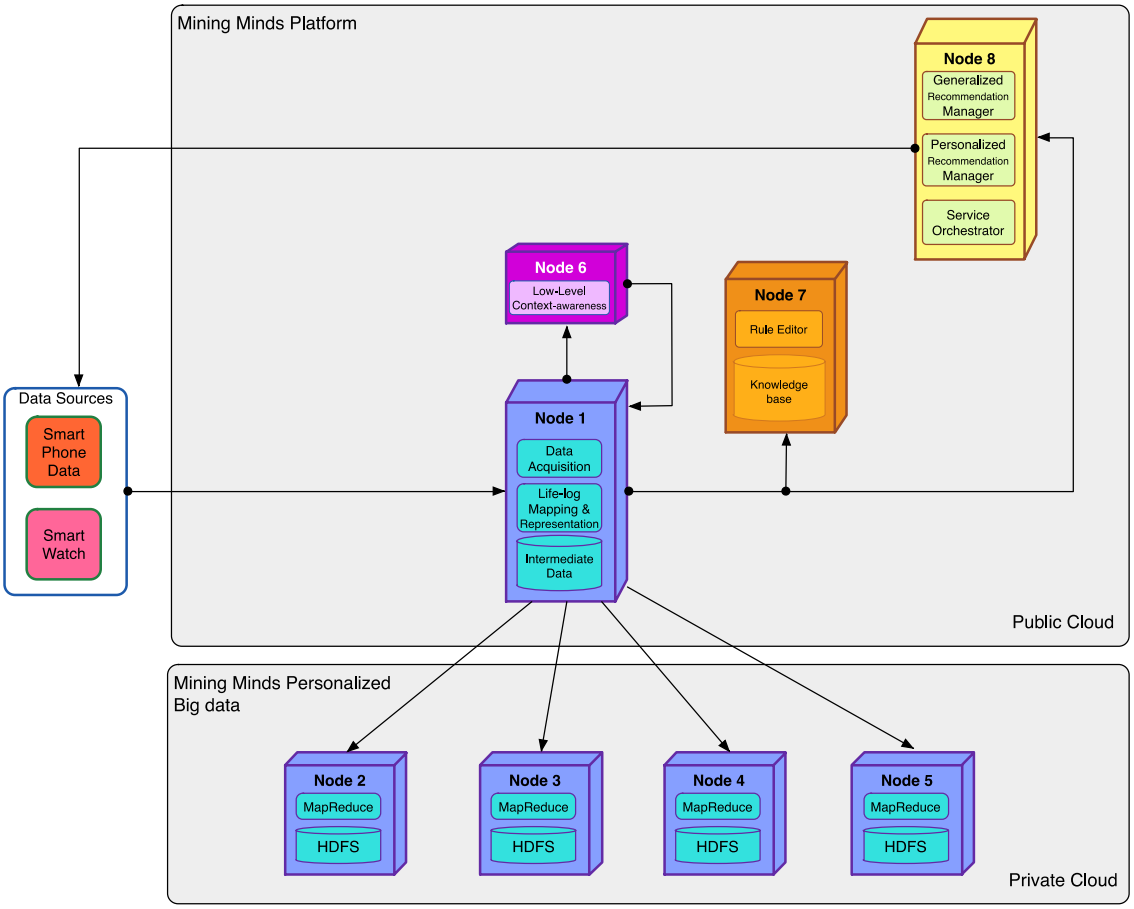

Figure 1.6: Deployment Diagram

1.7. Development Environment (Tools)

Mining Minds Ver 1.5. is built utilizing Java 1.8 programming platform over Microsoft Azure Public cloud infrastructure. The private cloud environment is built using Xen Hypervisor based Virtual machines. For real-time communication java socket libraries are used. For request-based communication and cloud-based alerts, restful web services are implemented. For development Eclipse IDE is utilized.

Chapter 2
Related Work

2. Related Work

The user awareness about personal health has been revolutionized in past few years which is mainly due to the proliferation of smart phones and tons of health applications available on mobile app stores. This privilege of using health related systems was limited with health care professionals in early 90s, who used PDAs for maintaining the healthcare records and kept this handy device in their pockets in addition to the cell-phones of that time [R.C.Wu 2010]. This specialized trend of one device for one task has now been transformed into generalized and convenient public usage with sophisticated devices e.g. smartphones. According to the global digital statistics [Simon2014], the smartphone penetration has been reported around 62% of the worlds' population. The popularity of health applications on these devices has been reported 4.83% from 75 billion downloads from AppleTM app store alone. It is also estimated that by year 2017, almost 50% of smartphone users will have downloaded the health application [J.Pennic2013].

The rationale behind targeting health and wellness systems for a user is to promote healthy life style, detecting early signs of ailments, reducing medical care cost, preventing un-healthy habits and overall monitoring of daily activities such as walking, jogging, eating and sleeping. While exhibiting these features certain applications like NoomCoach [Noom 2015], Argus [Azumio2015], Runtastic[Runtasticc 2015] , RunKeeper[RunKeeper 2015], Zombie!Run [R. Zombies 2015], Digifit i Cardio[Digifit 2015], Fitocracy [Fitocracy 2015], and Map my Fitness [Mapmyfitness 2015] have been recorded with millions of downloads. The variety of data collected by these or similar applications exploits the built-in sensors like proximity and light sensors, gyroscope, accelerometer, and magnetometer. This data is then used for setting and achieving certain user centric milestones like weight loss, exercise, and women's health during pregnancy [J.Pennic2013]. The importance of data related to user activities and health triggered a new wave of realization where it is believed that Future of Health Care is in the Data [C.M.News 2015]. Due to these factors, we surveyed and present few systems followed by certain additional features through a proposed system that delivers unique features.

2.1. Commercial Projects and Frameworks

In present era, numerous solutions are being released for promoting the health and wellness services. The main rational behind these new technologies, systems, and frameworks is to bring humans more close to their health monitoring and to improve the sedentary lifestyle. Empowering such systems will cast bigger and positive impact on health, medical care cost, and well-being. MyOSCAR 3.0 [D. Chan 2007] and Know2Act [C. Health 2012] work as a health information hub and information exchange respectively. Know2Act facilitates the patient to receive recommendations from trusted circle of care givers and exploits PHR1 data. In the growing trend of internet of things and Bigdata, few systems exploit data generated from heterogeneous sources like sensors, wearable and social media e.g. NexJ[J. Butler 2013], Dancer Wellness Project [W.Shakespeare], 100K Wellness project[L. Hood 2014] , and Digital Health Pilot[I. R. I.] .

LiveWellNYU[L.W.N.Y2015] is a comprehensive framework to integrate evidence based public health and innovative strategies for engagement of students at New York university. Health Data

Mining (HDM) [T.U. Technology] is aimed to extract behavior of working aged population through retrospective analysis of variety of Bigdata using innovative mining techniques. The Melon Headband [TheMelon] developed by IDEO measures the cognitive performance of a human being brain and recognizes the focus. It helps to support the human to improve its focus by monitoring EEG and Blue- tooth 4.0LE and connects it to the mobile phone. Although aforementioned systems target the healthcare and relevant solutions; however, they lack providing a comprehensive framework for personalized healthcare data monitoring on large scale and providing recommendations, monitoring human activities and representing in lifelog such as provided in Mining Minds [Banos2015].

2.2. Commercial Platforms

From the available shelf of popular commercial plat- forms we have selected few for discussion including GoogleFit[Google2015], Samsung S Health[Samsung 2015], Microsoft Health [Microsoft2015] , Health Kit[Apple2015] , FitBit [Fitbit2015], SAMI [Sami2015], and Open Mhealth [O.Health2015].

2.2.1. Google Fit

Google Fit is an open ecosystem that lets developers to upload fitness data to a central repository where users can access their data from different devices and apps in one location. It helps to monitor body weight and heart rate and also manages data input from wide range of sources and uses restful API to assist Android applications.

2.2.2. Samsung S Health

It is a personal fitness tracker that helps through user fitness routines on daily basis. It allow to set goals and achieve them with the help of pedometer. Other features include continuous tracking on walking distance, calories consumption, speed, and activity duration. It lacks in local data storage (on user device) also it is available for Android platforms only.

2.2.3. Microsoft Health

Microsoft Health is a cloud-based service that provides actionable insights based on data gathered from the fitness devices and apps. Using this data it provides valuable, personal insights to achieve fitness goals. Microsoft Health offers only three tracker e.g. running, diet, and cardio tracker.

2.2.4. HealthKit

The HealthKit framework provides a structure that apps can use to share health and fitness data. HealthKit is designed to manage data from a wide range of sources and automatically merging this data according to user's preferences. Apps can also access the raw data for each source and let the app perform its own merging. HealthKit also works directly with health and fitness devices.

2.2.5. FitBit

Fitbit is about smart monitoring, deep analytics, and motivational tools to change people lifestyles. It keeps tracking of your all key stats. It provides personal fitness dashboard, 30-day trends, hour-by-hour stats, and logging food & other activities like sleep monitoring.

2.2.6. SAMI

SAMI is a data exchange platform that enables any device or sensor to push its data to the cloud. Applications, services and devices can then use that data through simple APIs. SAMI is open and agnostic.

2.2.7. Open M Health

It is aimed to create a universal language for mobile mental health software, enabling all applications in the field to work together. It develops open data schemas that is used to provide guidelines for structuring different types of digital health data optimally for clinical use. It enables personally generated data to be integrated and used alongside clinically generated data from electronic health records (EHRs) and other health information technology. It is a suite of specifications for data interoperability and software reusability.

This era of exponentially growing data from diverse input sources has given rise to the concept of big data. To handle such big data, Hadoop Distributed File System (HDFS) is used as storage framework and MapReduce as its processing framework. However, MapReduce is a batch processing framework and well-suited for offline processing. To overcome the problem of batch processing and support soft real-time and online processing, streaming and communication component provides real-time and online access to the data in HDFS and generates intermediate data for other components use.

2.3. Commercial Applications

To start with, a brief list of applications is given in Table I. These applications can log and record user activities as they happen. For extended functionalities, certain activities that cannot be detected automatically can be entered manually to enrich the input data sources. Usually these applications rely on the built-in sensors embedded in smart phones; however, with ever increased number of wearable sensors, these applications are also compatible with external devices.

Table 2.3: Commercial Applications

Sno	Application	Platform	Features
1	Argus	Microsoft & Apple	70 million downloads(integrates with other apps)
2	Every Move	Apple & Andriod	Most influential in a corporate office (office network)
3	Human	Apple	Detects Walk, run, skip, gallop, dance, jump rope (calory breakup)
4	Map my Fitness	Microsoft & Apple & Android	1 million downloads. The app allows users to explore and manage more than 600 fitness activities running, rock climbing, playing tennis, or swimming
5	Moves	Apple	Records any walking, cycling, and running (uses GPS)
6	Pumatrac	Apple	Running (history, influence of external factors on performance i.e., weather)
7	RunKeeper	Apple & Android	keeps track of the awe-inspiring sites along your running route by letting your share snapshots with friends. RunKeepers GPS can log walks, bike rides, hikes, and more

			(10 million downloads
8	Runtastic	Microsoft & Apple & An- droid &Black- berry	Activity, Fitness and Sleep tracker (GPS). Geo tagging while exercise
9	Fitocracy	Apple & Android	Can sync with RunKeeper (100,00 downloads)
10	Hot5Fitness	Apple & Android	Predefined set of workouts+ can be added at runtime (can sync with MyFitnesspal)
11	Johnson & Johnson	Apple & Android	Detects Walk, run, skip, gallop, dance, jump rope (calorie breakup)
12	Digifit iCardio	Apple & Android	Detects various activity using heart sensor (MIO Link)
13	Endomondo	Apple & Android	GPS (running, cycling, jogging skating)
14	FitBit	Apple & Andriod	Fitbit app runs without using the Fitbit wrist band
15	MyFitness Pal	Apple & Android	keeps track of the awe-inspiring sites along your running route by letting your share snapshots with friends. RunKeepers GPS can log walks, bike rides, hikes, and more (10 million downloads)

2.4. Proposed System

After reviewing the existing platforms, frameworks and commercial health applications we inherit certain features and address the highlighted issues.

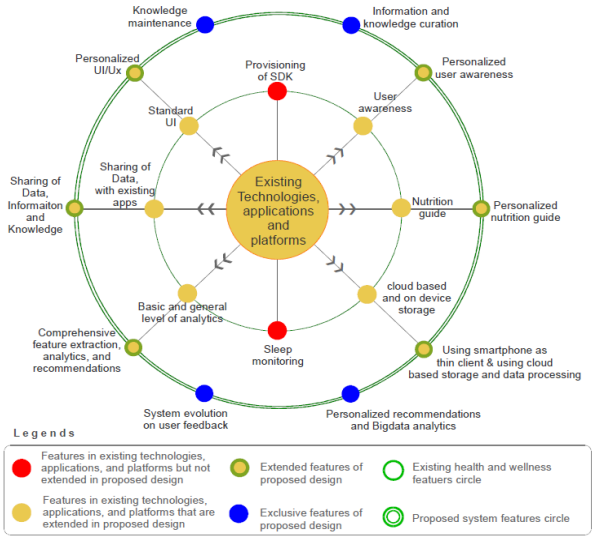

Figure 2.1: Example: Finite set world space

The proposed healthcare data curation system i.e., Mining Minds (MM) provides comprehensive personalized solution to the users using state of the art technologies. For this we propose a layered architecture comprising curation layers of data, information and services. First layer of data curation collects data from variety of sources such as multimodel sensors, social networking sites, wearables, and audio/video streams to acquire knowledge and generate personalized services for the end user. The acquired heterogeneous data is semantically linked to the user identity and mapped to an ontological representation. Data from multimodal sensors are curated to extract low level context information such as user activities (walking, jogging, running, sleeping etc.) and high level context information such as behavior modeling and context recognition. The data acquisition phase where data collection, synchronization, and storage are handled, is performed in the data curation layer. The context information and life-log creation based on personalized data are performed in the information curation layer. To extract knowledge and provide services to the end users; low level data from data curation, and contextual information from information curation layer are processed using state of the art techniques in knowledge curation and service curation layers. User interfaces provided in the literature are static and fixed. However, in MM, we provide adaptive user interface where the UI changes according the user preferences and feedback. Bigdata framework (Hadoop Distributed File System) is used for heterogeneous data storage. Compared to existing MM maintains the knowledge acquired through information curation, data curation, and context recognition to improve the knowledge base and evolve the learning process.

2.5. Summary

An energetic lifestyle is the procedure of enabling technology to improve a person's health and wellbeing in a proactive manner. There is an increasing current desire for societies to engage with technology-based solutions that can prevent conditions from arising and to avoid a deterioration in health-related conditions. Most of the existing models and frameworks focused either on physical, mental or social health. Nevertheless, a gap exists to model physical, mental and social factors all together for a personal smart care system. To overcome the limitations of existing platforms, there is the need for an alternative state-of-the-art personalized platform to fill in the gap by bridging physical, mental and social interaction in order to promote an active lifestyle in society. Therefore, the aim of this study is to promote active lifestyle and wellbeing by identifying the underline connections between physical, mental, and social health primitives. The physical health is related to exercise routines include sleep, mental health to a person’s feelings and emotional states, and social-wellbeing to the level of outdoor visited places.

Chapter 3
Data Curation Layer

3. Data Curation Layer

The data curation layer acquires the raw sensory data and securely stores into Hadoop with the help of data representation and mapping presenting the user's life-log. The details of the key methods and technologies are mentioned in the following Figure 3.1 and details are provided in the sub-sequent sections.

3.1. Data Curation Layer Functional Diagram

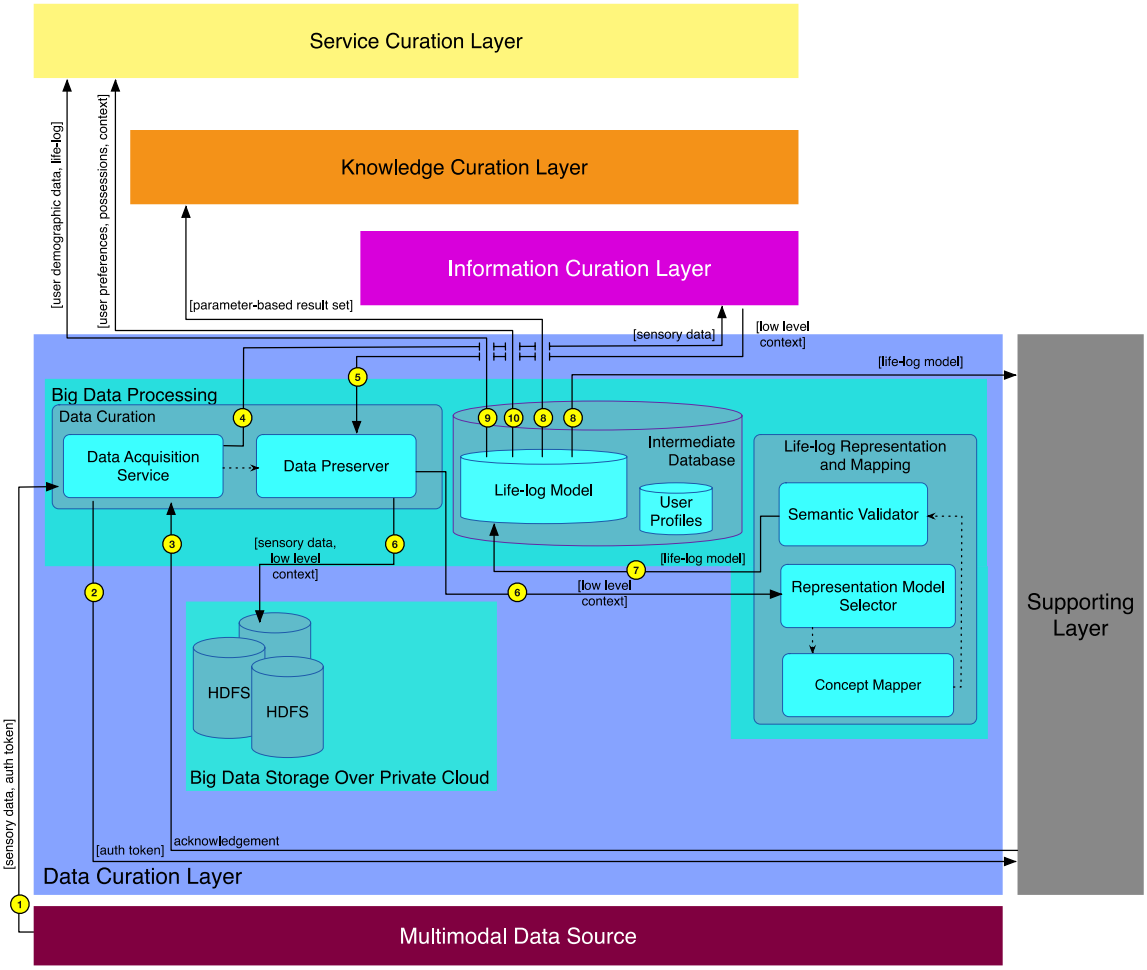

Figure 3.1: Data curation layer diagram

3.2. Data Curation Layer Use Case Diagram

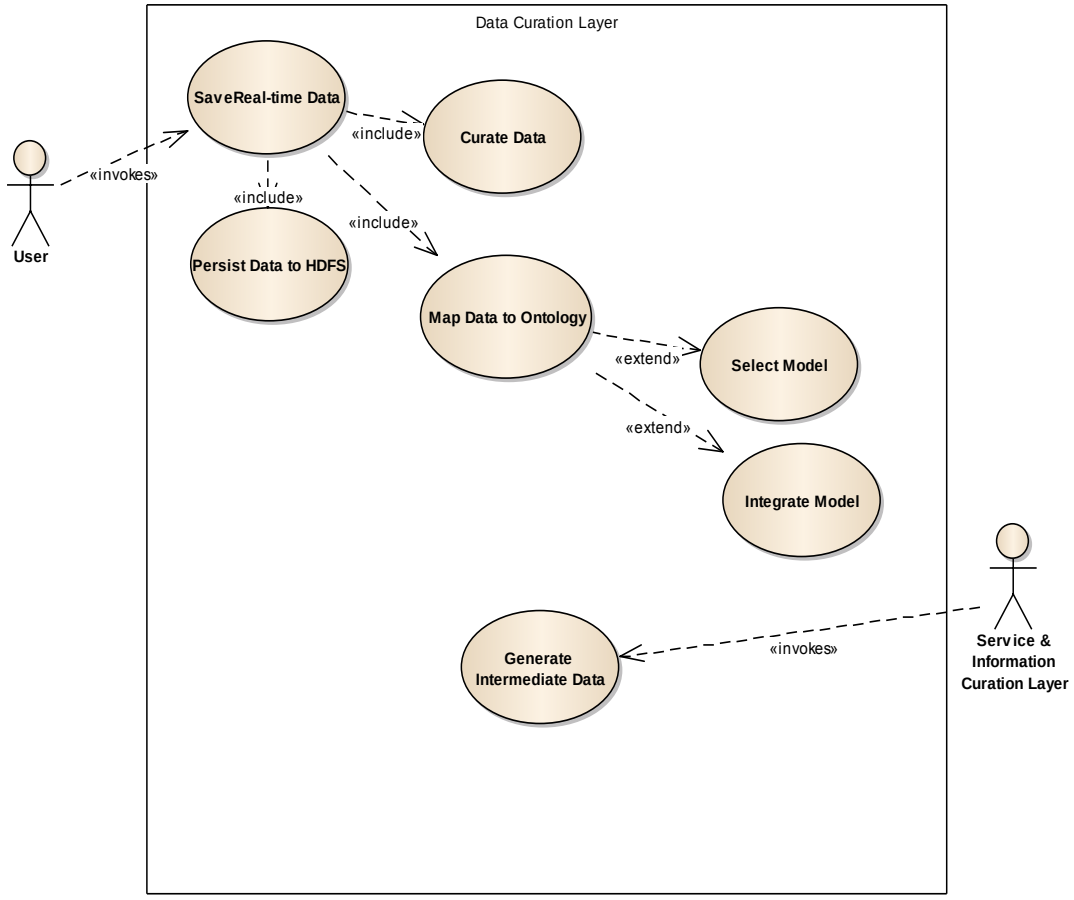

Figure 3.2: Use case diagram of data curation layer

3.3. Data Curation Layer Sequence Diagram

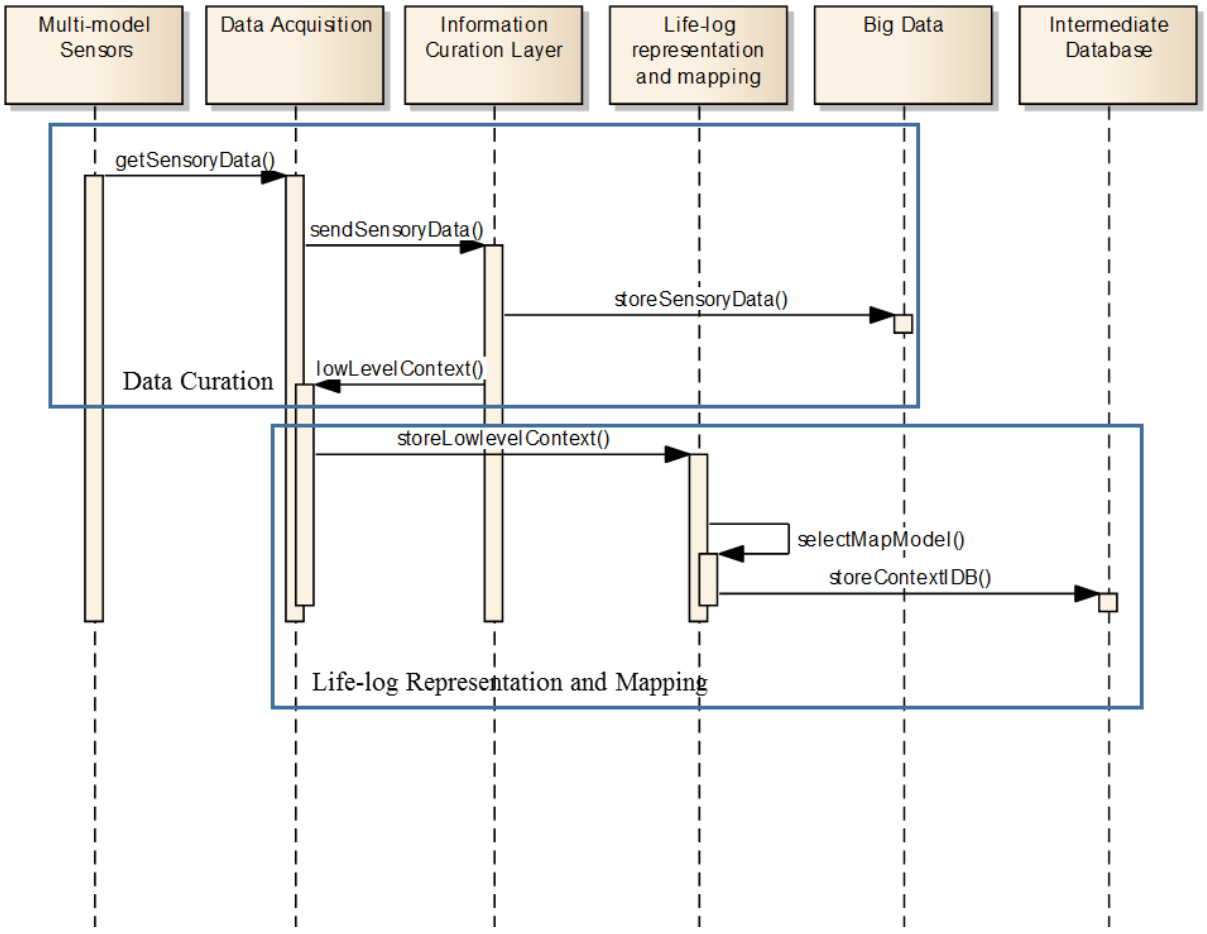

Figure 3.3: Sequence diagram of data curation layer

3.4. Data Curation Layer Component Diagram

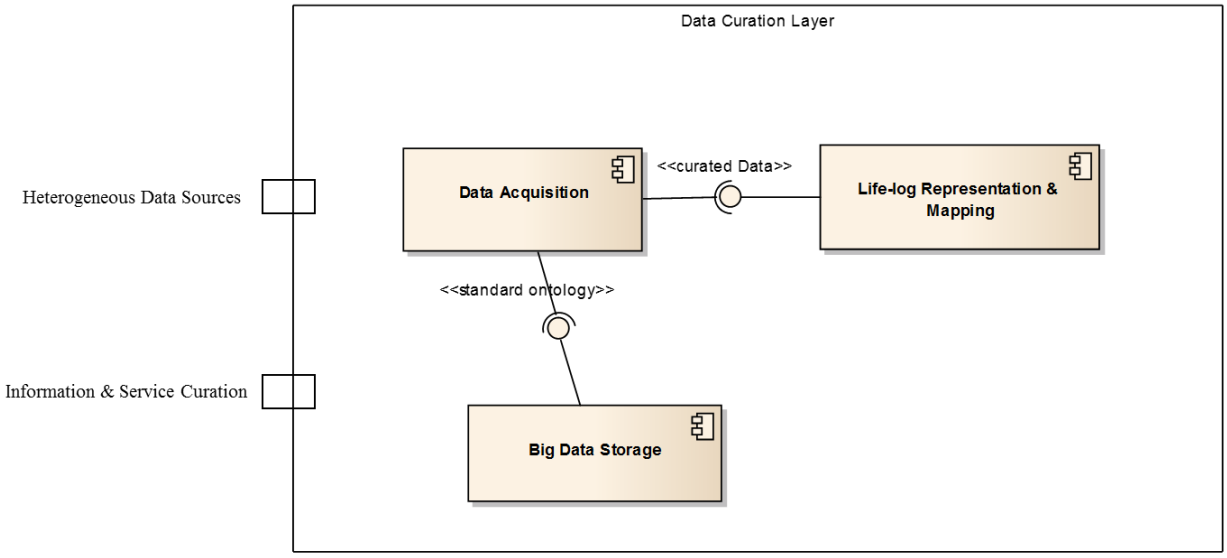

Figure 3.4: Data curation components Diagram

3.5. Life-Log Representation and Mapping

3.5.1. Introduction

Every system contains and handles some type of data or information in different manners. All the systems that maintain the log of user's different activities in their life, that needs a Life-log to handle it. Life-log mapping & representation is emerging and integral part of a system after the data acquisition phase. After data acquisition the data passes through different phases of the data curation process and information curation process to the life-log mapping and representation module to validate, verify and persist the recognized life activities. This module provides unified format of heterogeneous data to the information preservance. That preserves those information for further processing in upper layers and to maintain life-log history of user. The complete process from data curation , information curation and to life-log mapping and representation depicts that this process model is working in sequential manner. First the heterogeneous data is collecting from various sensors then processed data passes to the activity recognition and context aware modules of information curation layer. It recognized the user's activity and its behavioral context and then passes to life-log mapping and representation to persist and maintain it. Representation and Mapping component identify, classify and dynamically select an appropriate representation model and conform, validate the data accordingly.

3.5.2. Related Work

- Data representation and mapping is emerging and integral part of every system after the data acquisition phase. Scalable and interoperable data representation models are playing vital role in context-aware systems, smart homes applications, and mobile environment and in many more systems.
- Different representation models like Ontological, Graphical, Tuple based, Object Oriented, Spatial, Domain Focused, Hierarchical, Hybrid are all usable but in latest systems the ontological representation model have high preferences.

- L. Chen et al. [Chen2012] have presented a generic system architecture using ontology-based recognition process. Semantic information representation of the contextual information provides the formal specifications and is most appropriate method of context modelling [Khattak2011, Chen2012]. Ontologies provide the base for design of the contextual information representation. Ontology-based models of context information exploit the representation and reasoning power of logics [Ertel2011] for multiple purposes.
- M. Hussain [Hussain2013] has used an ontological based knowledge repository to persist and manage different kind of data. The knowledge repository can easily evolve with arriving new activities of a person and social media information. This ontological model is able to collect different sensors and social media data in raw form and able to stores the real time data after recognition that data by activity recognition module.
- [Kofod-Petersen2005] divides knowledge into two types, domain and non-domain knowledge. Domain knowledge taxonomic structure consists of Environment information, Personal information, Social information, Task information, and Spatio-temporal information. The model enables the system to infer relationships between concepts by constructing context-dependent paths between them.
- As ontological model, importing ontologies that defined by others, into a system is a normal and mostly happening task. Usually a system needs only some specific module of the importing ontology but there is no way in OWL to import that specific module instead of whole ontology. In each model a privacy model is needed to prevent different context information like personal information, location information and health information by little dissemination.

3.5.3. Components Description

- Input: Recognized Information (Structured/Unstructured/Semi structured)
- Output: Validated Object oriented Model with Instances and relational database Storage (Relational Schema and Tables of Recognized Information)
- Description: Life-log Mapping and Representation contains three major components (Representation, Information Persistence and Information Retrieval.
- The Representation module represents four different Representation Modules according to the different layers of Mining Minds.
    - Data Representation provides interfaces to the Data Curation Layer of the system to persist and retrieve data or information needed.
    - Information Representation provides interfaces to the Information Curation Layer to store and retrieve the recognized activities and locations.
    - Service Representation provides interfaces to communicate with Service Curation Layer. Through these interface the SCL stores and retrieves the data and information that it needs.
    - Supporting Layer also have an interface for its communication to stores and retrieve that needs for supporting layer like for visualization, analytics and for security activities.
- Information Persistence contains two different sub components. Such as Mapper and Storage Verifier to persist the data and information to Life-log Database.
- Information Retrieval component contains two sub components Query manager and Data Retrieval to provide the desired data or information to the corresponding layer.

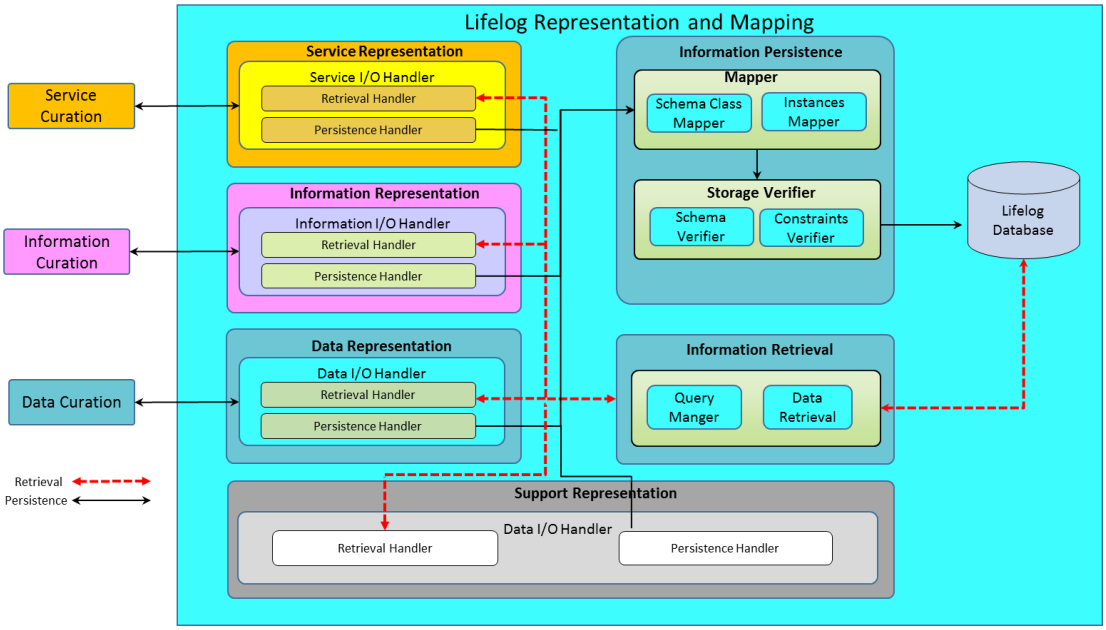

Figure 3.5 Architecture of Life-log Representation and mapping

Data, Information, Service and Support Representation:

- Input: Recognized Data (XML/JSON)
- Output: Selected Representation Model (Selected Schema)
- Description: Firstly, this component decides the retrieval and persistence process of corresponding layer. This component identifies, classifies the recognized data and selects appropriate representation model based on source and nature of data from Models.
- Retrieval Handler component activates the Information Retrieval module for the requested layer with corresponding data.
- Persistence Handler persists the data or information for the requested layers using the Information Persistence Module.

Ontology Persistence:

Mapper:

- Input: Selected Representation Model, Recognized Data/Information
- Output: Mapped Data (Classes, attributes with instances)
- Description: Extracts objects, attributes and relationships from input XML/JSON data as resources.
- Identifies and maps the classes for the extracted resources.
- Maps and finds the attributes, relationships of the classes and puts the instances accordingly.

Storage Verifier:

- Input: Mapped Data (Classes, attributes, relationships with instances)
- Output: Verified Data Model (Complete table (Schema and Instances))
- Description: This component validates the table data semantically with available schema.

- The different types of constraints are using on different tables, keys, values, primary keys, foreign keys and many others. To check these constraints the system is using Constraints Verifier

Information Retrieval:

- Input: Queried parameters (Attributes with desired Instances)
- Output: tables with instances (Tabular format)
- Description: This component gets the desired attributes and conditions to fetch the required data/information for the particular layer.
- The Query Manager component formulates the TSQL Queries according to the input attributes and conditions that runs on the Life-log database.
- The Data Retrieval component manipulates the retrieved data for the different layer according to their required format.

3.5.4. UML Diagrams

- Sequence Diagram

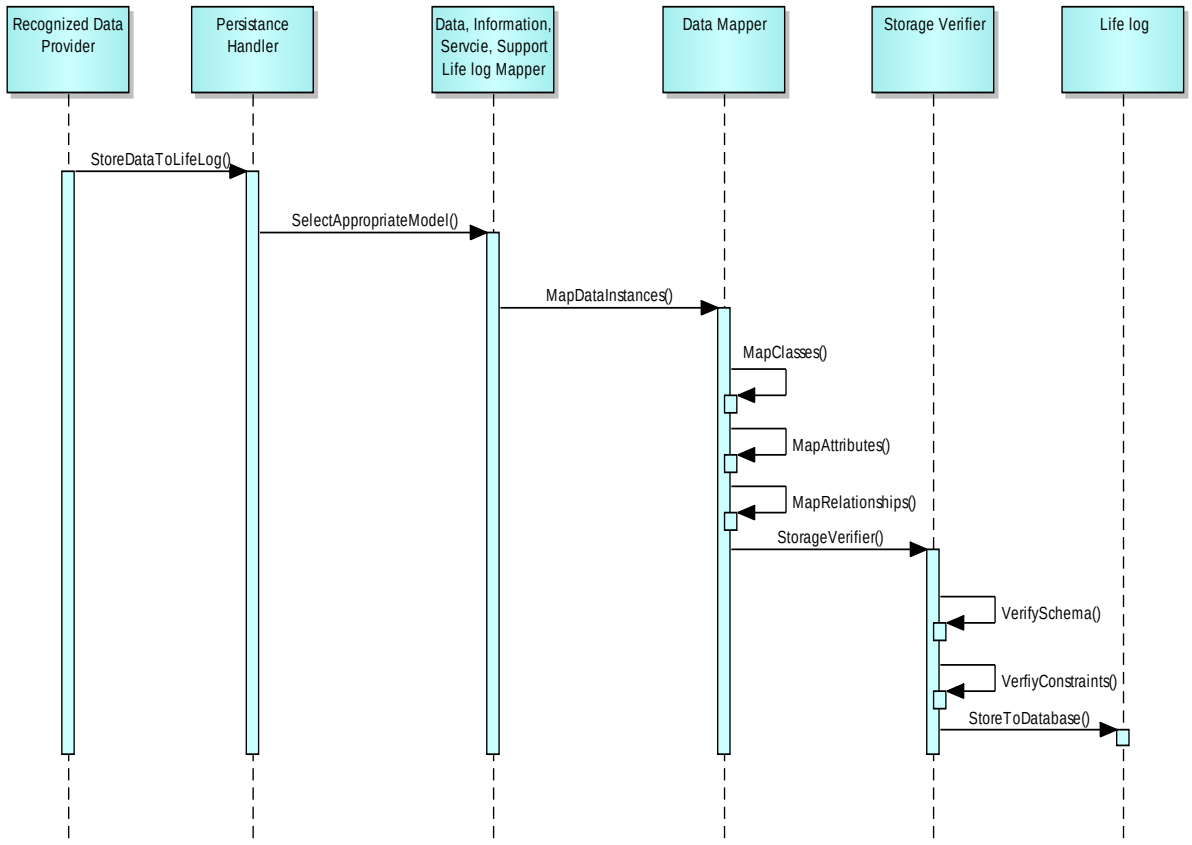

Figure 3.6 Sequence diagram of Life-log Representation and Mapping

- Component Diagram

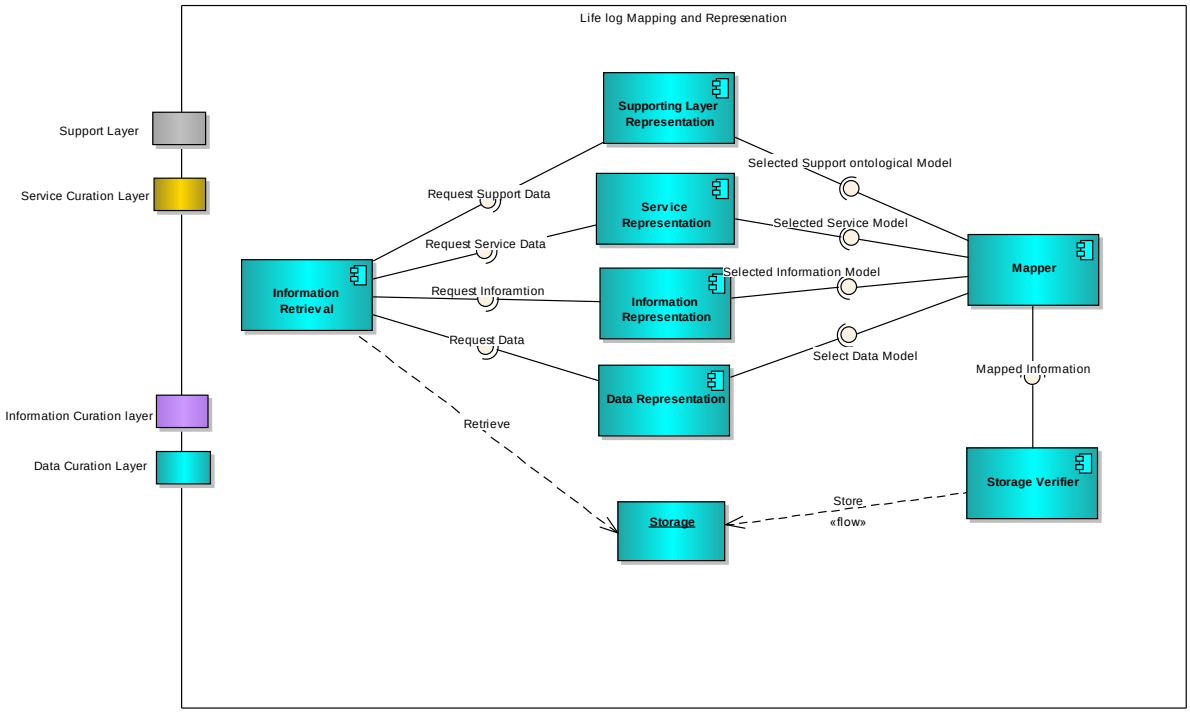

Figure 3.7: Component Diagram of Life-log Representation and Mapping

3.5.5. Uniqueness

- Comprehensive Lifelog database Engineering.
- Object oriented Representation models for heterogeneous sources of data.
- Mapping of data streams with continuous semantics management.

3.5.6. Summary

Human's daily activities and context are managed using Lifelog. Lifelog is represented in object oriented model and stored in relational database storage. We will transform this object oriented model into ontological model in next version. The Ontological representation is more expressive and highly flexible representation model. Model ontological representation is more scalable model to represent multiple source data.

3.6. Data Acquisition

3.6.1. Introduction

Among the primary goals of data curation layer is the acquisition of multi-modal sensory data from various sources per user in real-time, keeping the volume in consideration. Data acquisition is one such component that caters the need of high performance data acquisition from sensors and gateways route them for context identification and persist them for permanent storage in a big data store. Data acquisition also routes the identified context to life-log representation and mapping for daily life activity curation. For Ver. 1.5, role of data acquisition is encapsulated in to primary components.

- Data Acquisition Service,

- Data Preserver.

3.6.2. Related Work

- Curation of data so far only deals with persisted and stale data for intermediate data processing or service enabling. The opportunity of data acquisition for curation in real-time has largely been missed.
- The nature of data that has been under the investigation is so far homogenous i.e., data either belongs to same type or same scope. However, due to the nature of today's personal data, it is obtained from multiple sources and is of various types. Thus traditional data curtion is unfit for today's heterogeneous data.
- Although most of the research under the umbrella of data curation is focused on data quality and provenance, data acquisition has been missed.
- Data curation has been actively done of biomedical and bioinformatics persisted data, due to its scope homogeneity and its volume.
- This includes works over Protein Data (Research Collaboratory for Structural Bioinformatics, Rutgers University USA), Genetics Data (OpenHelix, University of California, Santa Cruz Genome Bioinformatics Group), and DNA Variations Data (DNAVaxDB, University of Michigan Medical School, Ann Arbor, MI).
- PDB Data Curation [Wang06]
    - Only support Protein Database files in XML-like formats.
    - No support for heterogeneous data sources and real-time data analysis.
- iRODS [Mark07]
    - Intended for Data Preservation
    - No Data Analysis and Labeling Support
    - No support for heterogeneous data sources and real-time data analysis.
- Data Curation for DNA Variation [Sophia11]
    - Only support Gene centered data and databases
    - No support for heterogeneous data sources and real-time data analysis.
- The Green Computing Observatory [Andrew09]
    - Build for collecting usage data of datacenters and clusters.
    - Built for datacenter data interfaces, no support for heterogeneity.
- Effective curation of cancer biomarker research data [Cecile11]
    - Build for U.S. National Cancer Institute's Early Detection Research Network (EDRN).
    - No support for heterogeneity

3.6.3. Components Description

- *Input*: Sensory Data from Smartphone and Smart watch as a Data Sources
- *Output*: Sensory data is persisted in Big data, context information is provided to Life-log Mapping and Representation.
- *Description*: In proposed resolution, data curation is evaluated over raw sensory data from the smartphone and smart watch. Data Acquisition is responsible for acquiring the stream of raw sensory data and routes it for context identification. Upon receiving the context information from information curation layer, it is forwarded to life-log representation and mapping component of data curation layer.

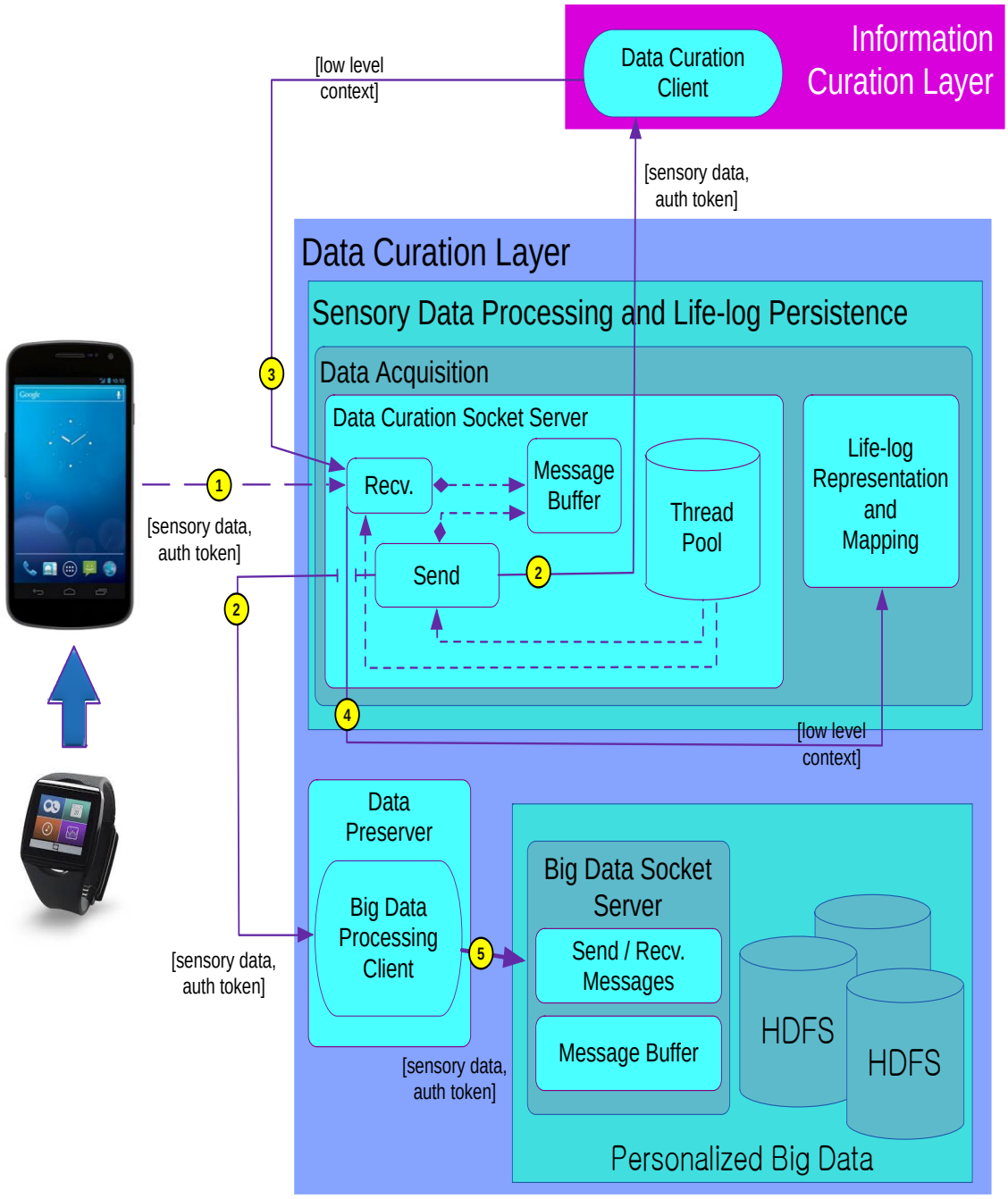

Figure 3.8: Data Acquisition Architecture

Data Preserver

- *Input*: Raw sensory data received by data acquisition component.
- *Output*: Raw data is persisted into HDFS.
- *Description*: Raw sensory data from users data source must be persisted for evaluation in big data. Data preserver takes this responsibility and persists all received raw sensory data for every Mining Mind user in big data storage.

3.6.4. UML Diagrams

- Sequence Diagram

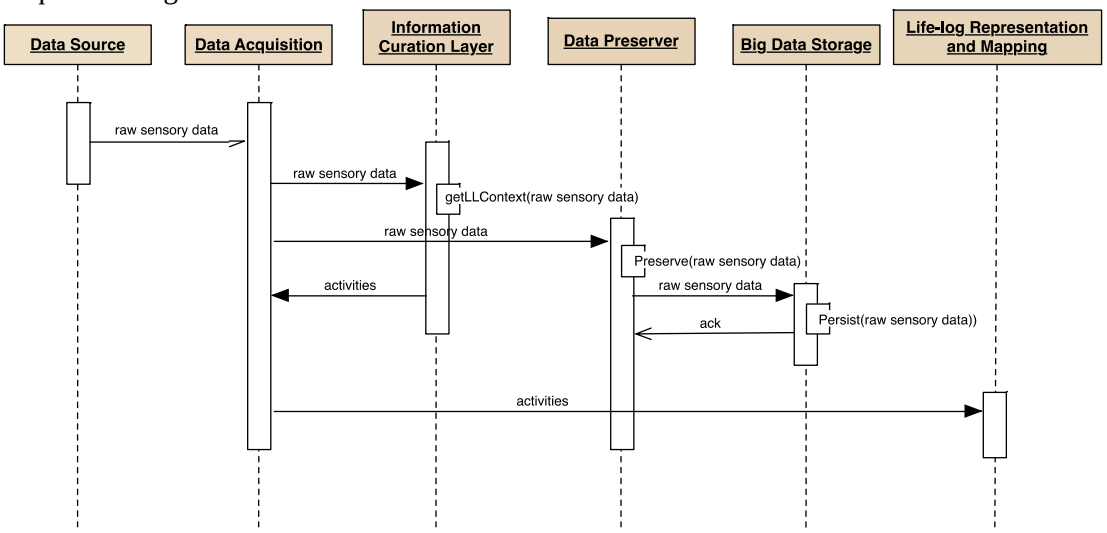

Figure 3.9: Sequence Diagram

- Component Diagram

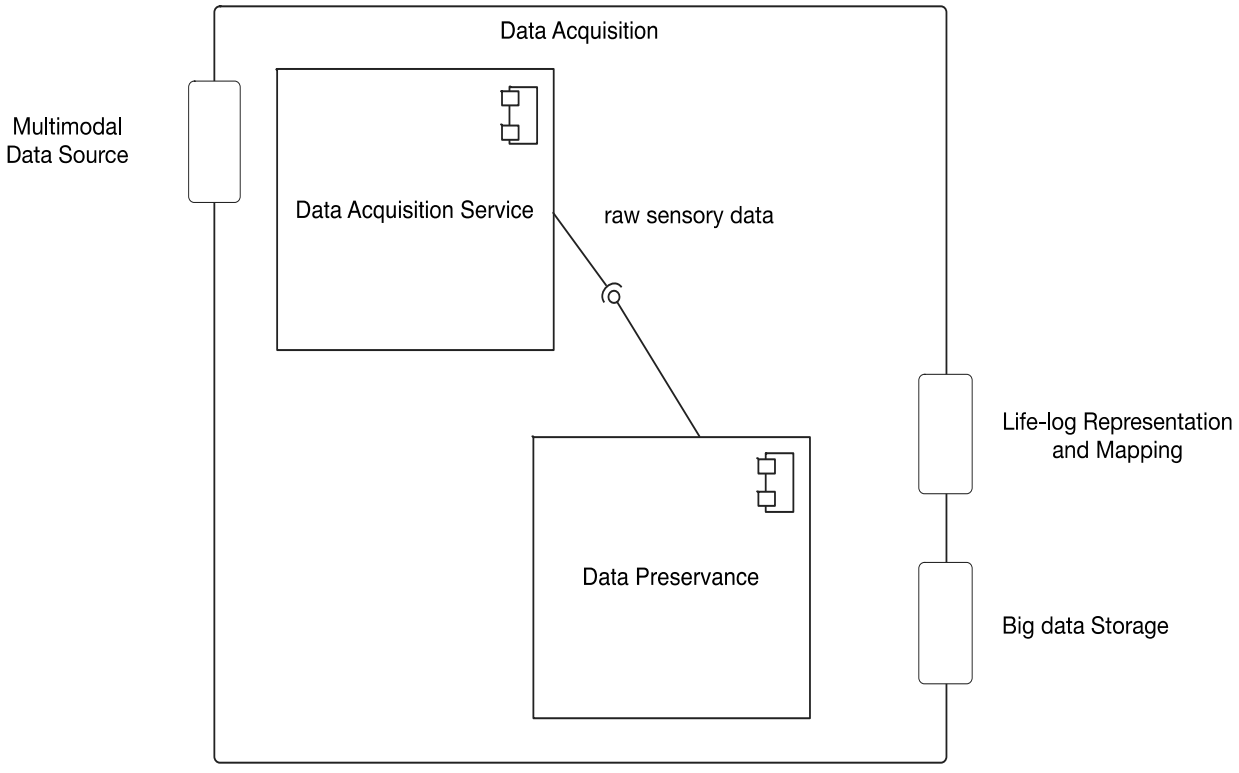

Figure 3.10: Component Diagram

3.6.5. Uniqueness

- Comprehensive Curation
    - **Customized Data Analysis** filters for **offline and online labeled** data processing
    - **Data confidence by source priority tables** providing redundancy and heterogeneity resolution
    - Checker module for **data flagging** in case of inconsistency
    - Machine learning based checker module for data conflict resolution
- Performance
    - Improved analysis performance **decommissioning stale data**
- Conformance
    - Conformance to Standard **Ontologies for Data Preservation**

3.6.6. Summary

- Data Curation component is built to handle the real-time inflow of heterogeneous data, in larger volume with reliability.
- It enables mining mind system to predict on real-time data and curate it according to the data sources. This data is further streamed and stored in HDFS.

Chapter 4
Information Curation Layer

4. Introduction to Information Curation Layer

The information curation layer curates the information to extract the low-level and high-level information. Furthermore, it also analyzes the long-term and short-term user behavior to provide quality of services. The details of the key methods and our unique component (i.e. highlighted red) are mention in the following figure and details are provided in the sub-sequent sections.

4.1. Information Curation Layer Functional View

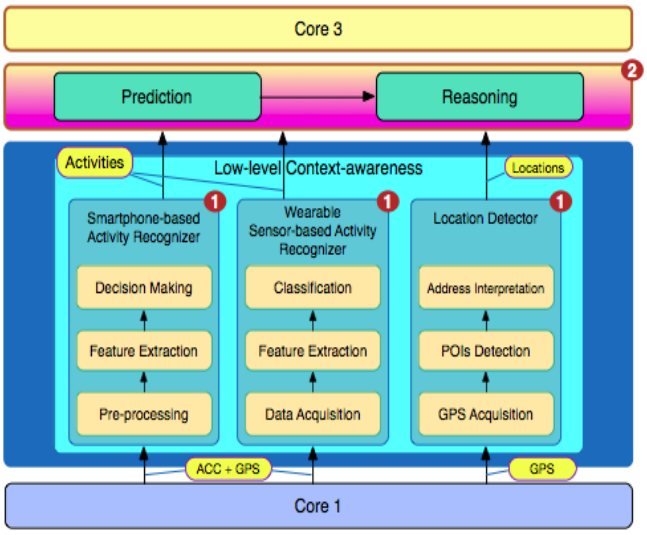

Figure 4.1: Information Curation Layer Functional View

4.2. Information Curation Layer Use Case Diagram

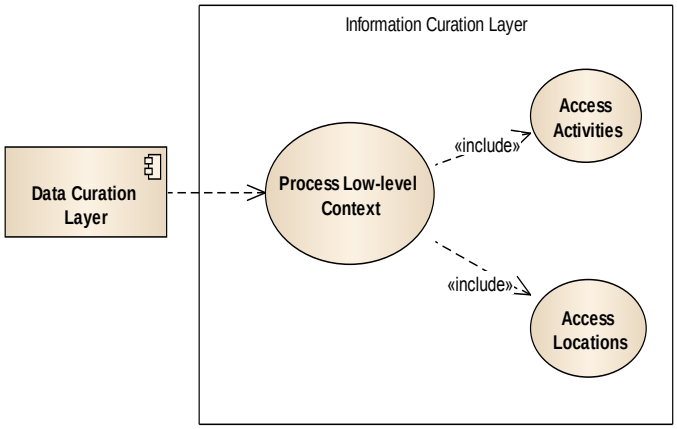

Figure 4.2: Information Curation Layer Use Case Diagram

4.3. Information Curation Layer Sequence Diagram

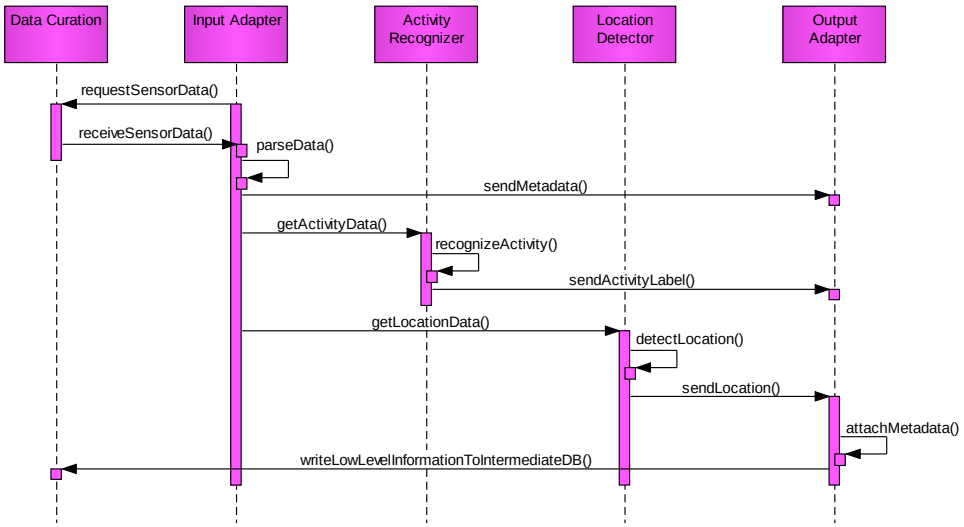

Figure 4.3: Information Curation Layer Sequence Diagram

4.4. Information Curation Layer Component Diagram

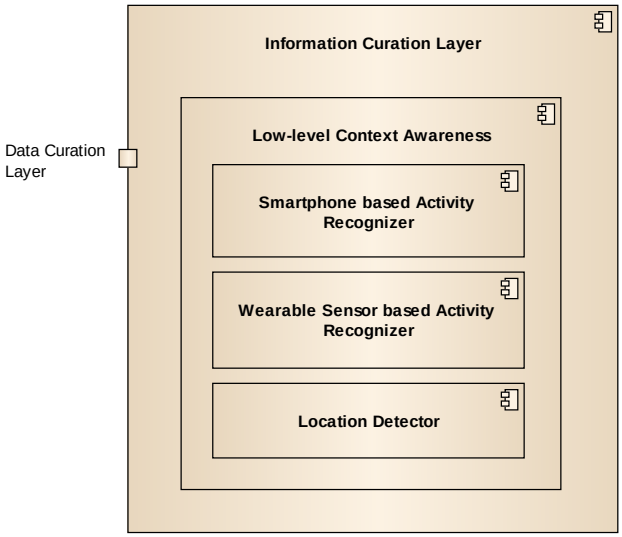

Figure 4.4: Information Curation Layer Component Diagram

4.5. Low Level Context Awareness

4.5.1. Introduction

Physical activity recognizer method is introduced to solve the challenging problem of intrusiveness of sensor devices and server side processing. We present a novel method for real-time physical activity recognition, position independent, and light weight classifier inside the smartphone environment. It is based on embedded accelerometer, ambient light and proximity sensors of the smartphone. Furthermore, our model introduces a device independent data acquisition method because data frequency is dependent on the sensor manufacturer and provided sensor manager API.

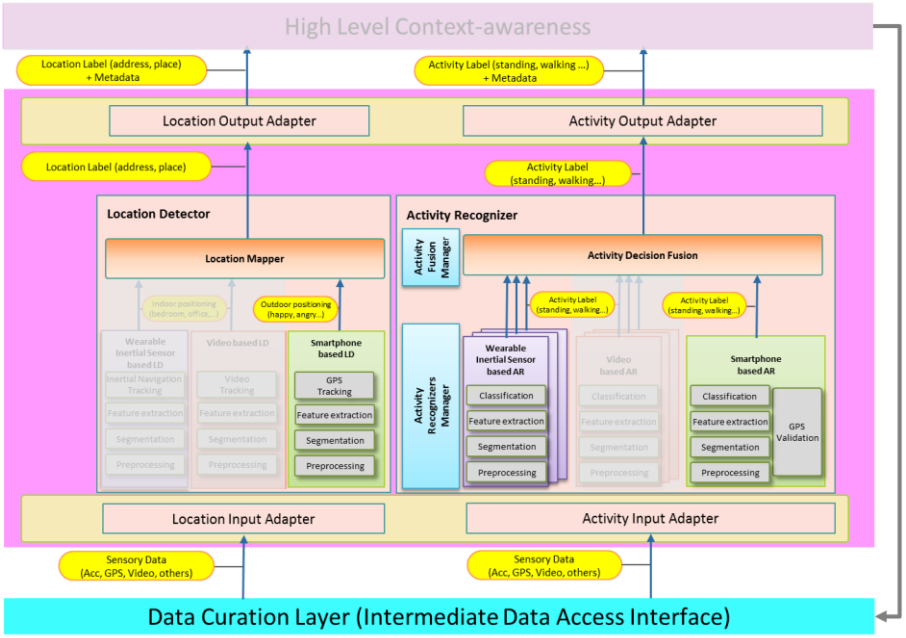

Figure 4.5: Architecture view of Low Level Context Awareness for Mining Mind

4.5.2. Related Work

- Activity Recognition

Wearable based Activity Recognition

Tapia et al. [Tapia2007] coupled heart rate sensor with five accelerometers attached to the human body for recognizing the daily life activities. They analyzed the performance of Naïve Bayes Decision Tree (DT) and C4.5 tree family algorithm over the physical activities.

Ayua et al. [Ayu2012] compared and evaluated different classifier algorithms for mobile's phone accelerometer-based activity recognition. They utilized the smartphone as an obtuse agent and their pre constrain is fixed location.

In classification algorithms, nearest neighbor is one the most used algorithm in many applications domain and rated in the list of top ten best algorithms in data mining [Garcia 2012]. However, the nearest neighbor suffers from several drawbacks such as high storage requirement, low efficiency in classification response, and low noise tolerance [Xindong2007].

Smartphone based Activity Recognition

Bao et al. [Bao2004] worked with 2-axis accelerometer. They attached sensors to forearm, wrist, pelvis, knee, and calf fixed, and recognized walking, jogging, stay, etc. They calculated average, energy, frequency domain entropy, and correlation on each sensor, and classified with different classifier.

There were some experiments using audio to recognize vehicles. Lee et al. [Lee2008] recognized environmental sound using significant feature vector automatically extracted when 3GPP2 Seletable Mode Vocoder is coded.

Activity recognition research was still actively undergoing even after the emergence of smartphone. In [Wang2010], they used 3-axis accelerometer of Nokia's N82 mobile phone and recognized 6 activities, such as walking, subway, bus, car, cycling, and stay.

[Han2012] tried to overcome the limitation of accelerometer based activity recognition, and used audio, GPS, wifi, etc. He used accelerometer which has big differences such as walking, running, and stay, and used audio, GPS and wifi to recognize bus and subway which has small differences of signal.

- Location Detector

With the dramatic evolution of smartphone, there are many sensor technologies are integrated in smartphone such as accelerometer, gyroscope, GPS etc. Smartphones become a convenient device supporting the personal localization.

Lan et al. implemented pedestrian dead reckoning (PDR) for indoor localization. With a waist-mounted PDR based system on a smart-phone, they estimated the user's step length that utilizes the height change of the waist based on the Pythagorean Theorem. This method does not require training to develop the step length model. Exploiting the geometric similarity between the user trajectory and the floor map, the map matching algorithm includes three different filters to calibrate the direction errors from the gyro using building floor plans [Lan2014].

Montoliu in [Montoliu2010] presented a new framework to discover places-of-interest from multimodal mobile phone data. Mobile phones have been used as sensors to obtain location information from users' real lives. Two levels of clustering are used to obtain places of interest. First, user location points are grouped using a time-based clustering technique which discovers stay points while dealing with missing location data. The second level performs clustering on the stay points to obtain stay regions.

4.5.3. Components Description

Activity Recognition

Wearable based Activity Recognition

- *Input:* Sensory data such as accelerometer, proximity, physiological information
- *Output:* Activity label (walking, running, sitting, standing, …)
- Current approaches require server side processing and predefined location of the devices that limits their applicability in real-world applications.

- In addition, we introduce smartphone independent accelerometer sensing method that can be work well with android OS regardless of its manufacture sensor frequency rate.

Smartphone based Activity Recognition

- *Input:* Accelerometer, Gyroscope, GPS
- *Output:* Activity labels: walking, running, standing, bus, subway
- We recognize walking, jogging, stay, bus and subway independent to smartphone's position and orientation.
- Accelerometer data is revised with gyroscope data, and acquire fixed signal vector from front and rear, left and right, up and down independent to smartphone's orientation.
- Features are extracted from this signal to get similar signals that have identical activity.
- In addition, data is collected from different positions that we may carry on real life such as in top and bottom cloth, bag, hand, etc.

Location Detector

- *Input:* GPS coordinates (longitude and latitude), speed, time and other support sensor data (accelerometer, wifi)
- *Output:* Public location and personal location
- The public locations are fetched by using Google Map or Naver Map API to map GPS coordinates with predefined locations.
- The personal locations are constructed firstly by collecting the users' place of interest, and then prompting and labeling by users to build a personal map for each user, finally mapping to get the location.
- Other support sensor data such as accelerometer is useful in order to recognize location-based context such as transportation, gym, etc.

4.5.4. UML Diagrams

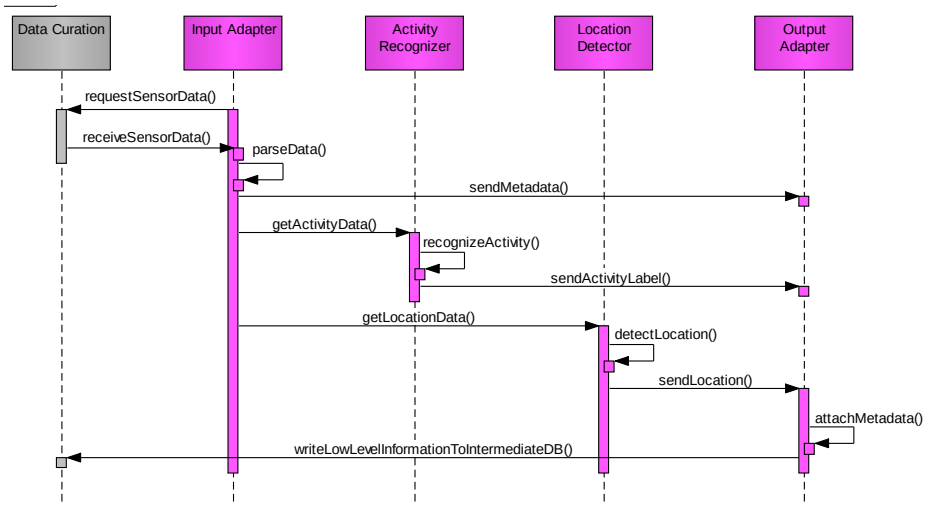

Figure 4.6: Sequence Diagram of Low Level Context Awareness for Mining Mind

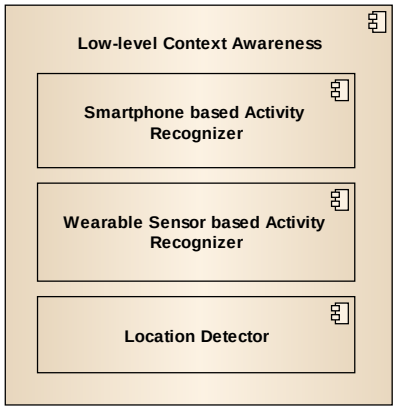

Figure 4.7: Component Diagram of Low Level Context Awareness for Mining Mind

4.5.5. Uniqueness

Activity Recognition

Wearable based Activity Recognition

- No existing method of physical activity recognition that gives the maximum flexibility to carry the phone according to one's own choice either in pocket or hand.
- Processing sensory data inside the mobile environment to save the communication cost. Therefore it is an objective of the present recognizer to introduce a novel physical activity method to utilize the embedded sensors of smartphone.

Smartphone based Activity Recognition

- We recognized bus, subway, and stay extracting unique vibration from fixed axis signal by accelerometer independent from smartphone's orientation. With this, we could recognize walking, jogging, stay, bus and subway by only using accelerometer with high accuracy.
- A revision algorithm is proposed to show high accuracy in real-field, and confirmed it by experiment.
- The usage of battery is minimized by only using accelerometer and GPS among all other sensors.

4.5.6. Summary

- We provide low level context awareness module that is able to provide useful information from heterogeneous data sources including of social networking, wearable sensors, physiological sensors and smartphone based sensors.
- Extracted information covers many aspects of daily life such as the behavior, activities and emotions of users.
- Output from this module is exploited to predict short-term and long-term behaviors.

Chapter 5
Knowledge Curation Layer

5. Introduction to Knowledge Curation Layer

- For effective knowledge base it is critical that knowledge base should not go stale and old. As we know that in every field of life, knowledge plays a vital role and all intelligent decisions are made based on knowledge. So knowledge should be accurate and updated. For producing and maintaining accurate knowledge, various machine learning and evolutionary techniques are used to update and maintain the knowledge base.
- The knowledge curation layer creates and maintains the knowledge using data-driven and knowledge-driven approaches to facilitate the service curation layer for better quality of service. In data-driven approach, knowledge is extracted from curated information that is generated by data curation and information curation layers. While in knowledge-driven approach, domain expert creates and maintains the knowledge through authoring environment with minimum intervention of knowledge engineer.
- Dynamic construction of the knowledge and evolutionary knowledge maintenance through our Knowledge Maintenance Tool using both data-driven and knowledge-driven makes this tool unique and distinguishable. The abstract architecture is shown below, while the details of Knowledge Acquisition Editor are provided in the sub-sequent sections.

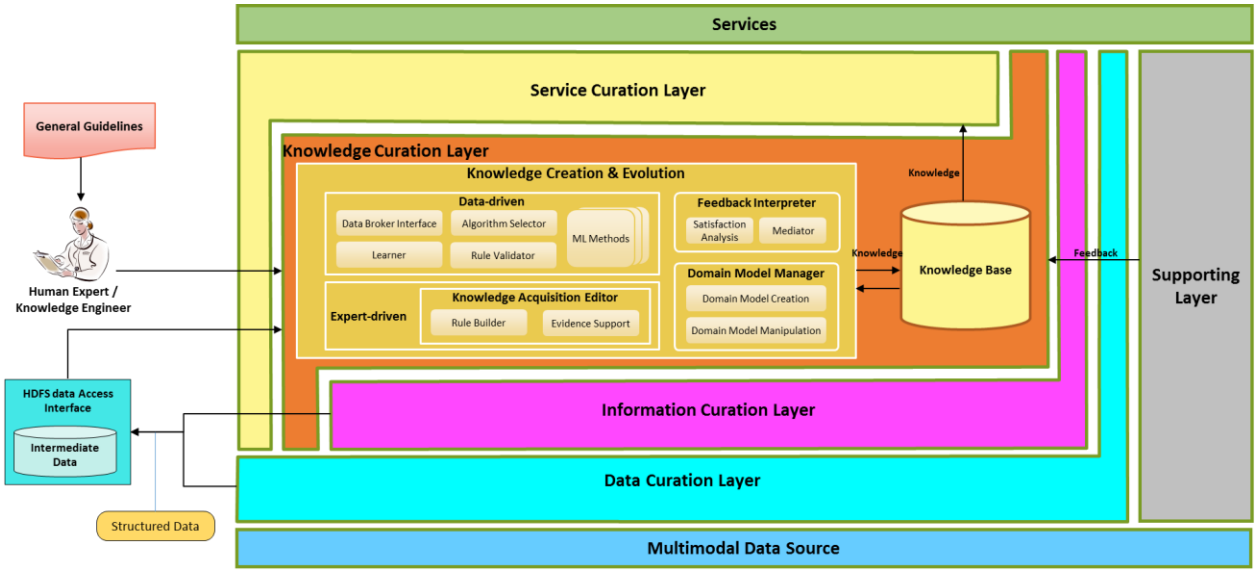

Figure 5.1 Abstract Architecture of Knowledge Curation Layer

5.1. Knowledge Acquisition Editor Functional Diagram

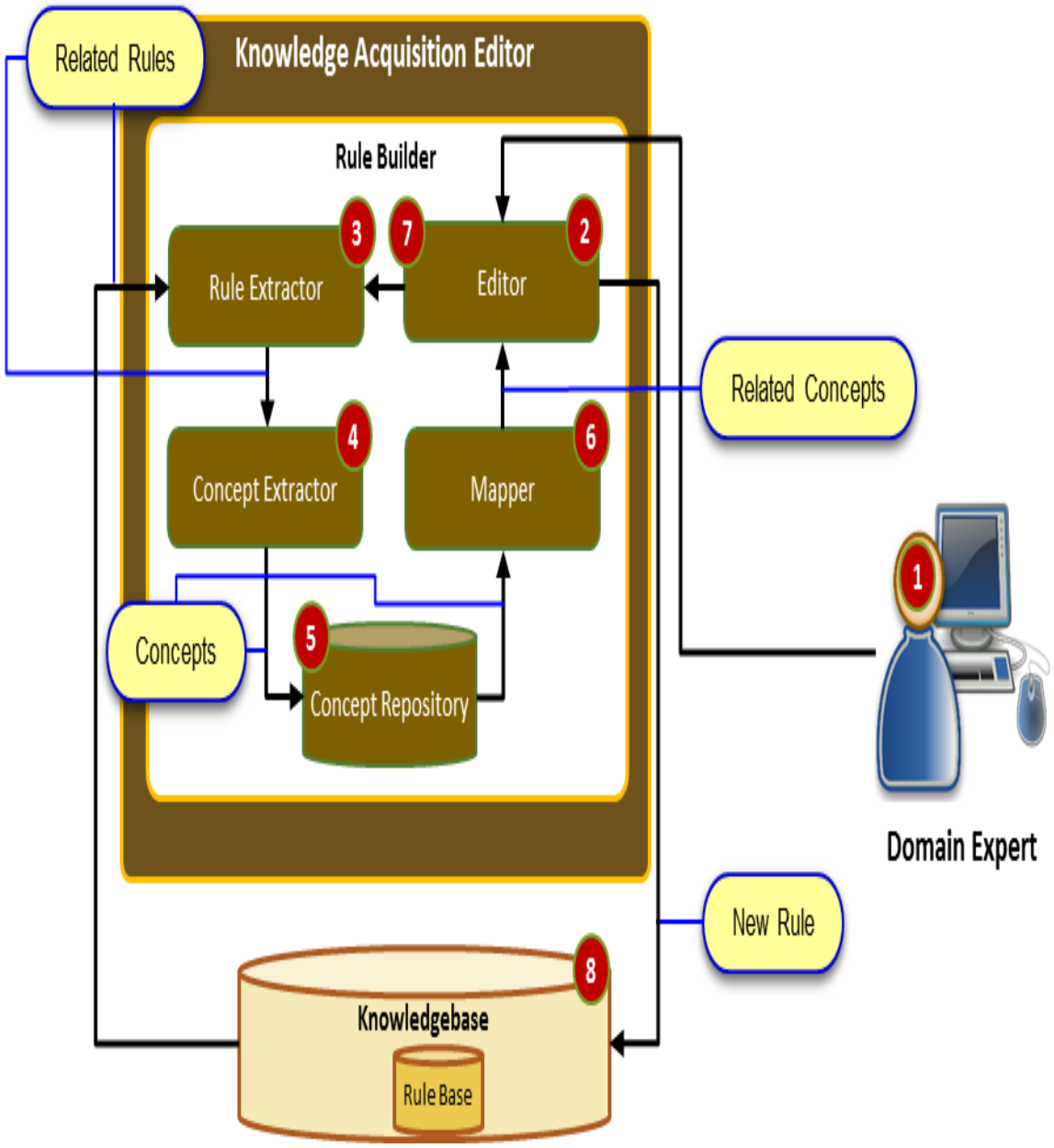

Figure 5.2: Knowledge Acquisition Editor Functional Diagram

5.2. Knowledge Acquisition Editor Use Case Diagram

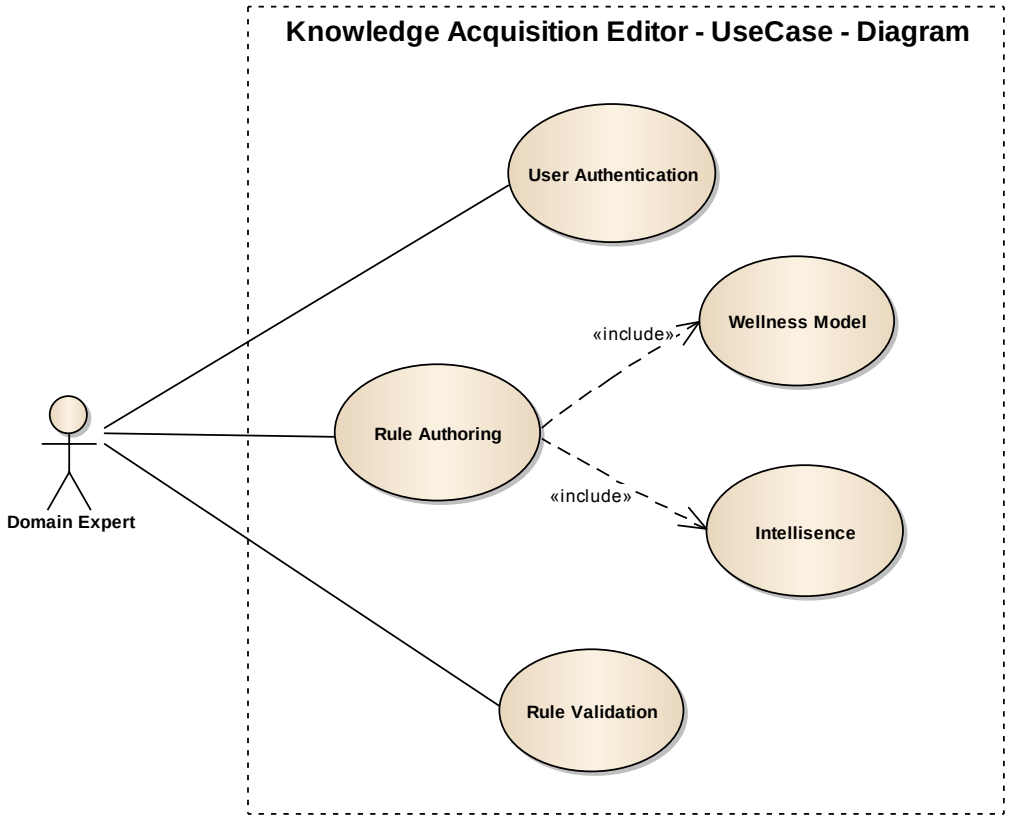

Figure 5.3: Use case Diagram

5.3. Knowledge Acquisition Editor Sequence Diagram

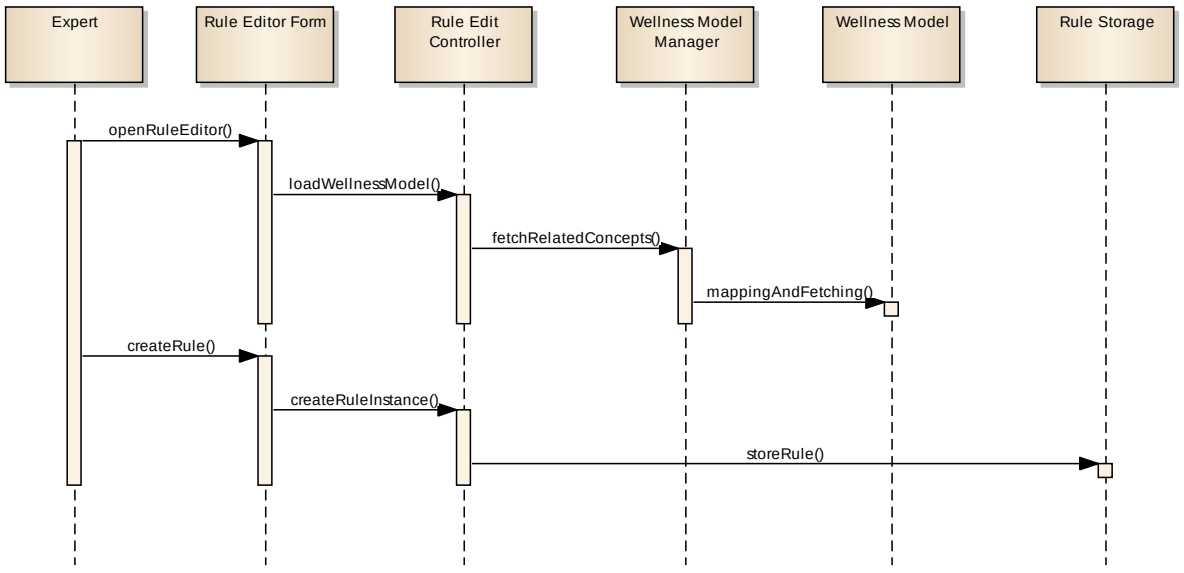

Figure 5.3: Sequence Diagram

5.4. Knowledge Acquisition Editor Component Diagram

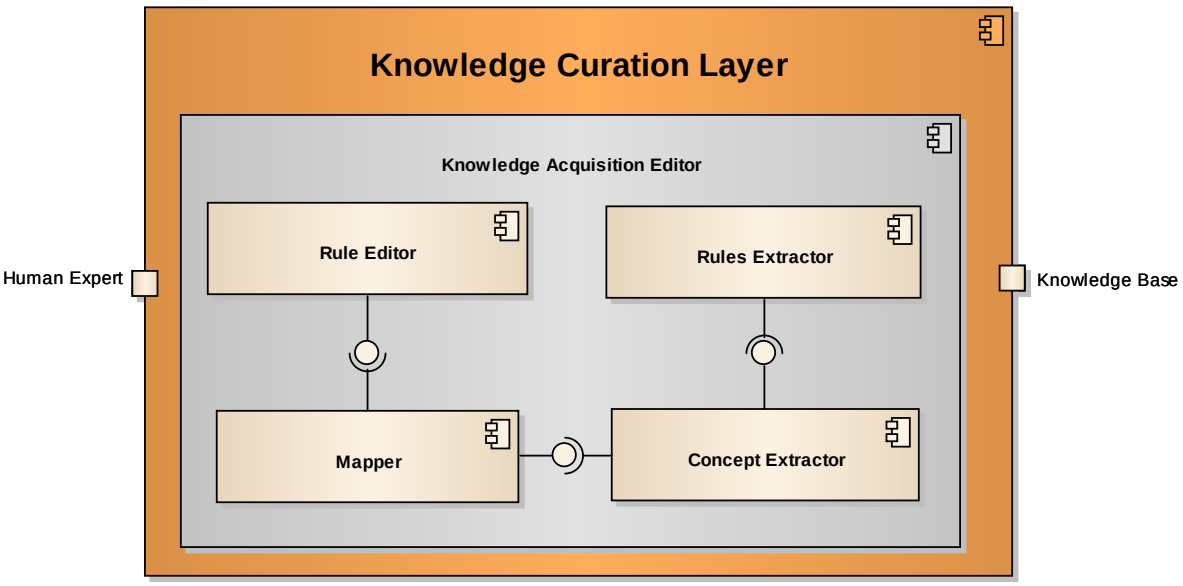

Figure 5.4: Component Diagram

5.5. Knowledge Acquisition Editor

5.5.1. Introduction

- The *Knowledge Acquisition Editor* creates the knowledge using knowledge-driven approaches to facilitate the service curation layer for better quality of service. In knowledge-driven approach, domain expert creates the knowledge through authoring environment with minimum intervention of knowledge engineer. It provides an efficient rule authoring environment.

5.5.2. Related Work

- More expressive and semantically as well as syntactically complex knowledge rules need to be edit with rule editing systems that help in maintaining the knowledge base [Kaljurand2008]
- According to [Regier2009], the main challenge is to keep the knowledge bases up to date and decrease the turnaround time for logical changes in the existing rules of knowledge base in every expert system.
- The success of any contemporary CDSS is based on KB upon which it is built [Hulse2005]. The crucial aspect for KB is the richness with respect to domain knowledge which is mainly dependent on domain experts.

5.5.3. Components Description

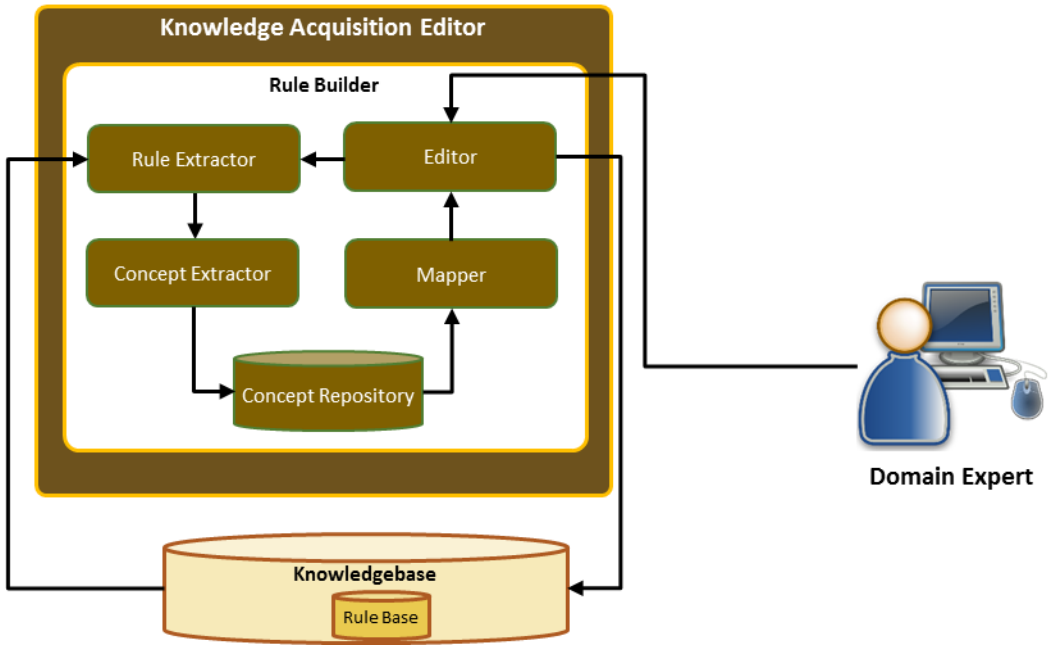

Figure 5.5: Architecture of Knowledge Acquisition Editor

Editor

- Input: Expert Knowledge
- Output: Rules
- Description: This component creates, validates, and stores the rules into Knowledgebase. It provides a user friendly interface to knowledge experts to create or edit the existing rules in the knowledge base.

Rule Extractor

- Input: Expert Knowledge
- Output: Related Rules
  Description: This component provides the facility to extract all related existing rules from knowledgebase to help domain expert to easily create the new rule.

Concept Extractor

- Input: Rules
- Output: Concepts
  Description: This component extracts concepts from all extracted rules and then store into concept repository.

Mapper

- Input: Concepts
- Output: Intelli-sense
  Description: This component provides the Intelli-sense functionality that helps in creating and editing the rules in a rapid and easy manner. This functionality increases the recall of concepts for the experts and decreases the ratio of errors during the editing of rules.

5.5.4. Knowledge Curation Layer Uniqueness

- High quality knowledge creation using *dynamic selected algorithms*
- Integration of *data-driven* and *knowledge-driven* knowledge management
- Tool support and evidence support for expert for *easy maintenance*
- Knowledge development with *minimum intervention* of knowledge engineer
- Support of *MCRDR* for *easy* and *rapid* knowledge maintenance

5.5.5. Summary

Knowledge plays a vital role in every field of life and all intelligent decisions are made based on knowledge. So knowledge should be accurate and updated. For producing and maintaining accurate knowledge, various techniques are used to update and maintain the knowledge base. Not always but most of the times automatic generation rules creates anomalies in the rules which need to be corrected from the domain expert. The *Knowledge Acquisition Editor* is an effective rule authoring interface through which domain expert can easily create and modify rules with minimum interception of knowledge engineer. Dynamic knowledge acquisition using data-driven approach is our future research direction, that how dynamically best algorithm is selected that can produce high quality rules/models/case-base to provide better services.

Chapter 6
Service Curation Layer

6. Introduction to Service Curation Layer

The objectives of service curation layer is to provide timely and accurate personalized cross-domain recommendation using domain knowledge and users preferences/context and schedule. The layer curates services based on the user's curated information and different unhealthy/risky behavior. It consists of service orchestrator, recommendation builder and recommendation interpreter. The main purpose of this layer in MM is to handle events, curate services and generate personalized recommendations according to the user's interests and preferences. The contents of recommendations for each service are generated by the recommendation builder while the personalization are done by the recommendation interpreter. The interpreter deals with how, when and where to generate recommendations. The details of key methods and our unique component are mentioned in Figure 6.1.

6.1. Service Curation Layer Functional Diagram

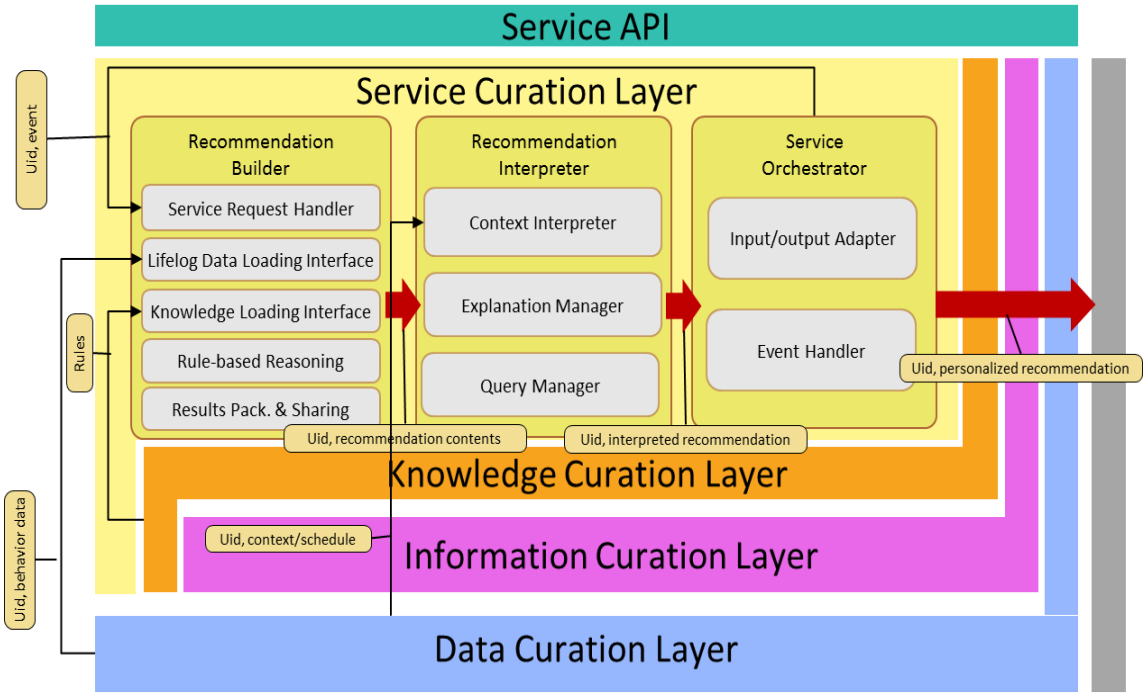

Figure 6.1: Service Curation Layer Functional View

6.2. Use Case Diagram

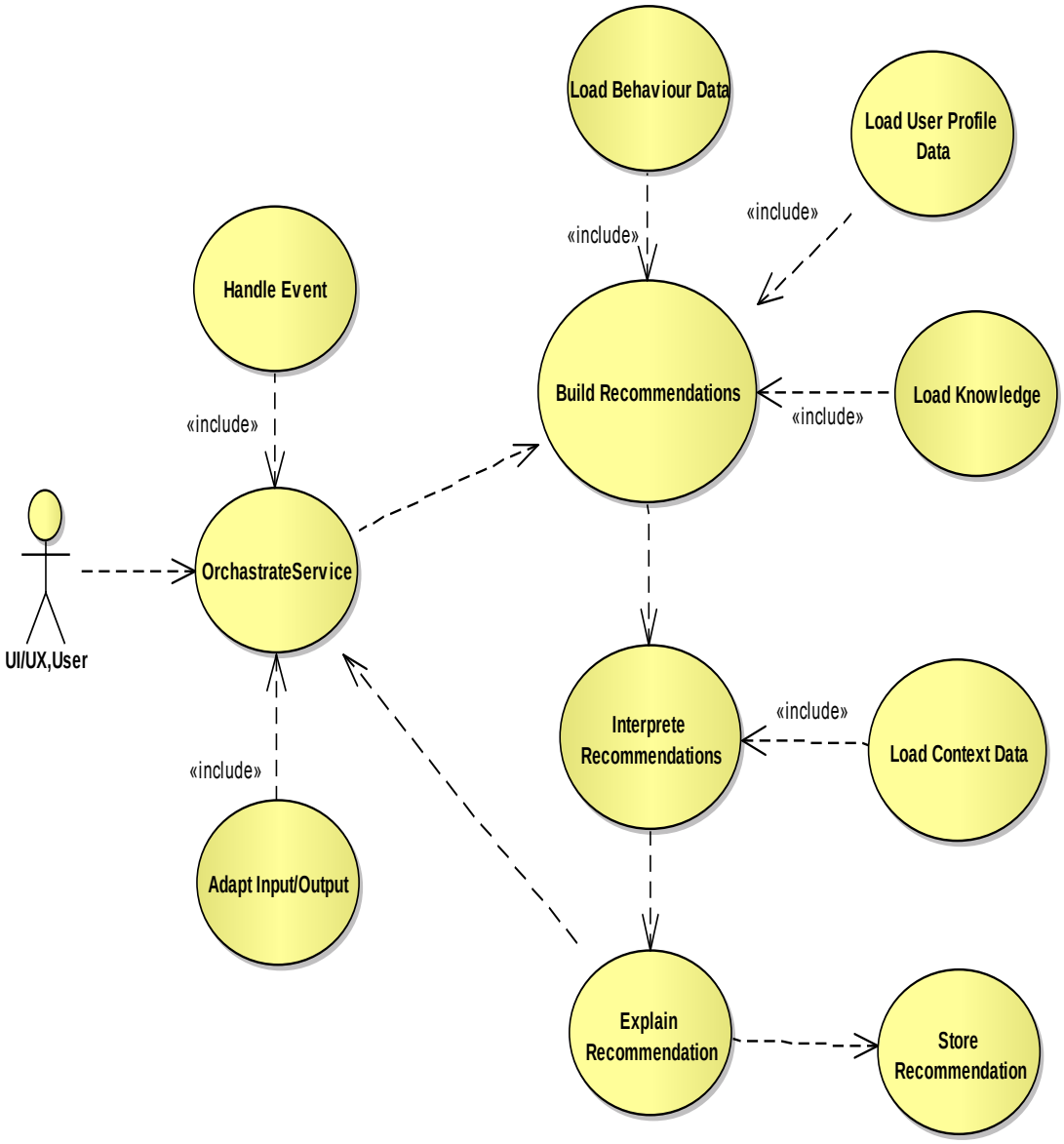

Figure 6.2: Service Curation Layer Use Case Diagram

6.3. Service Curation Layer Sequence Diagram

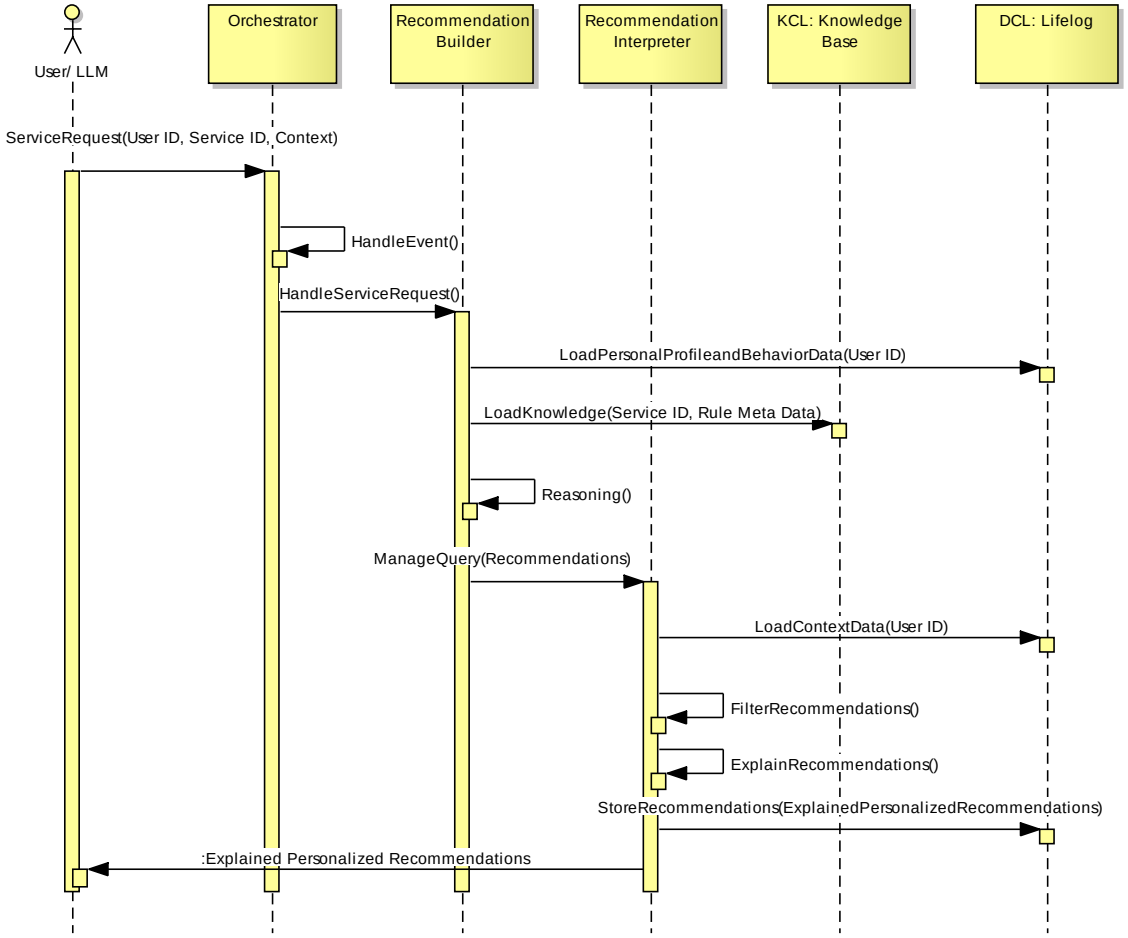

Figure 6.3: Service Curation Layer Sequence Diagram

6.4. Service Curation Layer Component Diagram

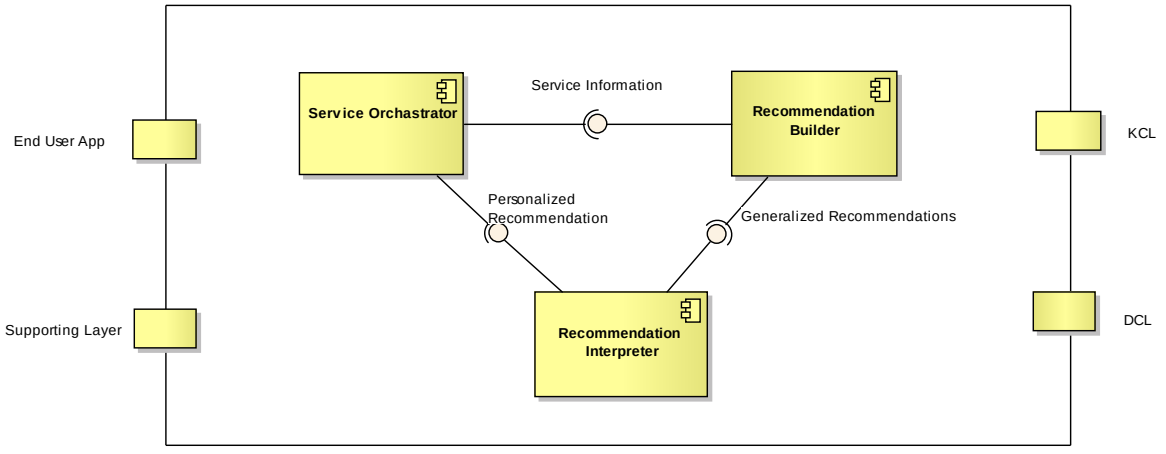

Figure 6.4: Service Curation Layer Component Diagram

6.5. Service Orchestrator

6.5.1. Introduction

As Mining Minds is a multi-service platform designed in a way to generate services for different applications, therefore a façade type of service is needed to activate appropriate type of service for user's requests. Service orchestration is the component that performs the same task and coordinates or integrates several services and exposing them as a single service. In MM version 1.5, the role of service orchestrator is restricted to handling events generated by the data curation layer scheduler. The event handler takes the event and handle it by directing other components of the layer to generate the appropriate service.

6.5.2. Components Description

- Input: Event and Interpreted Recommendations
- Output: Invocation of recommendation builder and UI/UX
- Description: The Service Orchestrator Service orchestration performs the task of handling the event provided by the data curation layer and pass the personalized interpreted recommendations to the UI/UX.

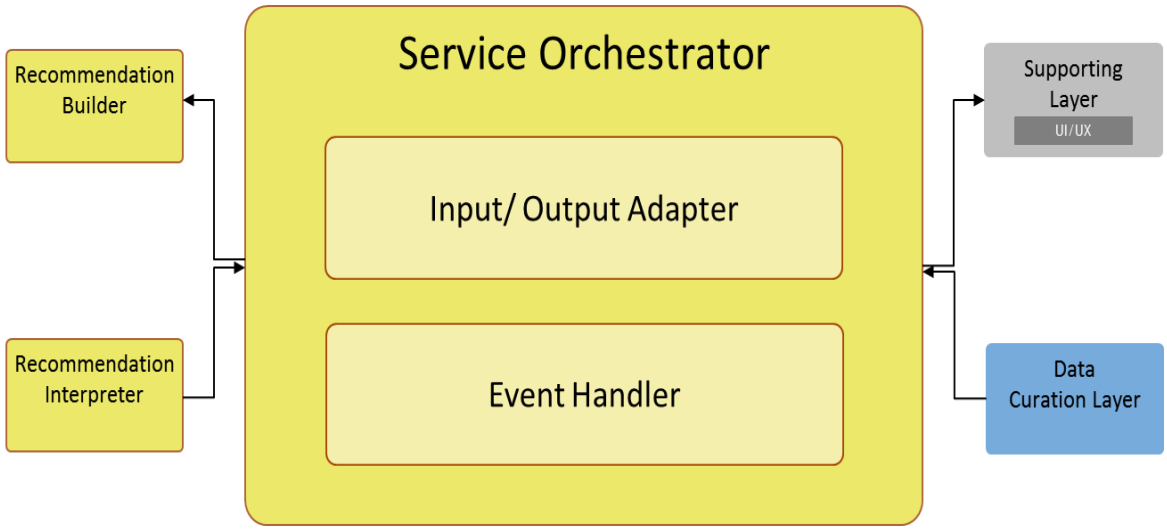

Figure 6.5: Service Orchestrator for Mining Minds

Input/output Adaptor

- Input: Event from DCL and interpreted recommendations from recommendations interpreter
- Output: Activation of event handler and UI/UX
- Description: It catch the event thrown by the data curation layer and the output/personalized recommendations generated by the recommendations interpreter.

Event Handler

- Input: Event request from I/o adapter
- Output: Invocation of recommendations builder
- Description: the purpose of this component is to handle the event collected from the I/o adapter and activates the recommendations builder to generate contents of recommendations for this event.

6.5.3. UML Diagrams

- Sequence diagram

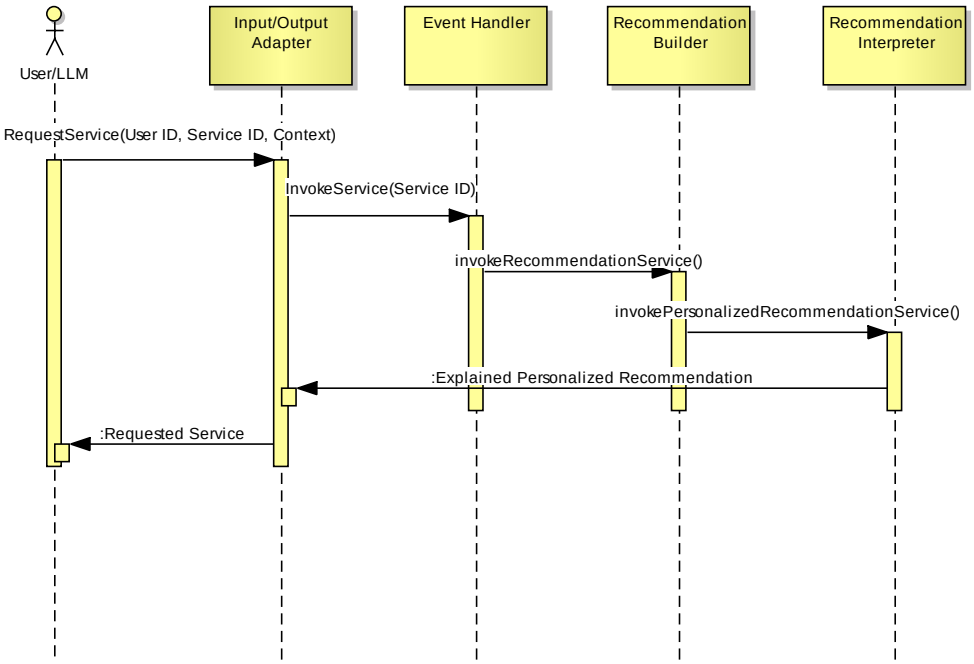

Figure 6.6: Service orchestrator sequence diagram for Mining Minds

- Component diagram

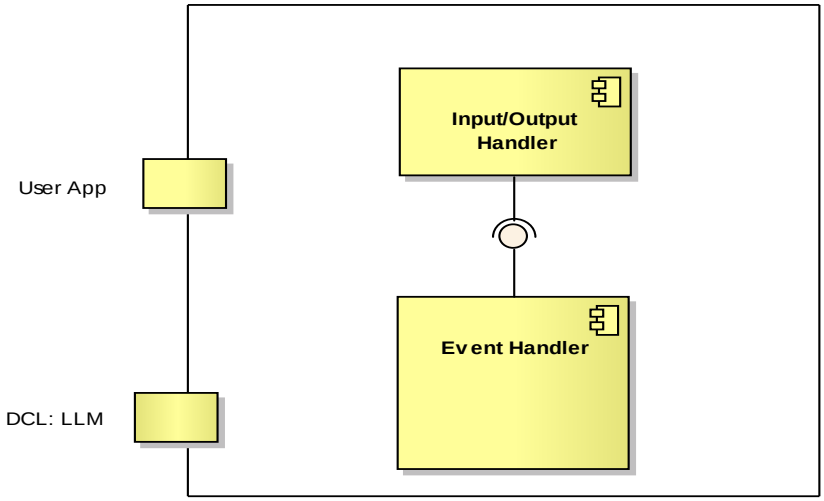

Figure 6.7: Service orchestrator component diagram for Mining Minds

6.5.4. Uniqueness

- Dynamic Invocation of appropriate service based on dynamically generated events by DCL for the user behaviors.
- Implementation of the PUSH model for service provisioning.

6.5.5. Summary

This component is used to handle the events generated by the LLM in DCL. Based on different criteria, the event is passed to corresponding component i.e., recommendations builder to handle the vent. It also presents the personalized service to UI/UX for displaying.

6.6. Recommendation Builder

6.6.1. Introduction

Every recommender system is based on reasoning methodology to generate recommendations. This reasoning is a process whose computational function is to generate conclusions (recommendations) for user query from available knowledge using different intelligent techniques. The main constituents of reasoning includes: knowledge/rules, data, execution engine/interpreter and input/output interfaces. In Mining Minds, the purpose of this module is to automatically generate and build recommendations contents for the user's requests by utilizing the user's behavior and personal profile data. The services are generated by executing knowledge created by the knowledge curation layer from general healthcare guidelines and user historical data. The recommendations are provided to the recommendation interpreter for further filtration and interpretation to personalize. In mining mind version 1.5, we are focusing on the rule-based reasoning that is activated by the service orchestrator based on the unhealthy event generated in data curation layer. The RBR methodology exploits heterogeneous nature of data from the user personal profile and behavior along and apply rules intelligently accessed from the KCL. The module uses two type of execution models: push model where DCL event triggers the reasoned to generate the service and pull model where the user request for the service. The proposed reasoner generates only the contents of the recommendations which are then interpreted by the recommendations interpreter for further interpretation.

6.6.2. Related Work

- Reasoning and prediction systems have widely been studied and used in domain, such as medical (e.g. PIP, INTERNIST, CASNET, and MYCIN), business markets for predicting stock prices, weather forecasting, fault proneness in software's systems, lightning detection and prediction (e.g. SkyScan, THOR GUARD) and control system engineering.
- [Yuan 2014] have developed a healthcare system that is used for at home monitoring of the elderly people. The services they have provided are personalized healthcare for and the reasoning framework used is CARA that uses CBR-RBR approach.
- In medical, the Pharmacogenomics project (2006-2008), the viral genomics data is integrated with clinical data to predict responses to anti-HIV treatment to the clinicians [Zazzi2010]. Several prediction engines are designed to predict the efficiency of possible drug combinations and to recommend an optimal treatment by combining the results from all together.
- Hybrid reasoning systems, such as "a personalized wellness recommendation system" [Husain2010], for predicting therapy and Context-Aware Real-time Assistant (CARA) [Yuan2014] for personalized healthcare services of the elderly people have been studied using case-based and rule-based approaches. Similarly, probabilistic approaches (e.g., Bayesian and Apriori) have been used for predictions of hepatitis infection [Drăgulescu2007], and diabetes [Pandey2012].
- In existing reasoning and prediction systems, the predictions are based on knowledge base, learned in its first phase, with minimum support for data abstraction. These systems either have reasoner or predictor as the main engine and the both. Likewise, they use a single learning and conflict resolution method for learning rules and resolving conflict during the inference process. We propose an integrated system with a high-level data abstraction, reasoning and prediction components using a hybrid approach for providing accurate and precise recommendations.

6.6.3. Components Description

- Input: event, profile data, behavior data
- Output: Recommendations contents
- Description: recommendations builder provides recommendations contents for the user's requests and unhealthy events triggered from the data curation layer through service orchestrator. These events and requests are passed to the recommendations builder where RBR methodology is used for generation of recommendation contents. These contents are passed to the recommendation interpreter for further personalization and interpretation so that the user can easily understand.
- Service request handler takes request from the orchestrator and pass to the data loader interface and reasoning module.
- The data loader interface loads the profile and behavior data and pass to the reasoner module.
- Once data is prepared, unified rules interfaces uses meta-knowledge of the rules and loads rules form the knowledge curation layer for the reasoning.
- In the reasoned, RBR methodology is used for generating recommendations contents.
- The recommendations contents are passed to the results packaging and sharing interface to pass to the recommendation interpreter.

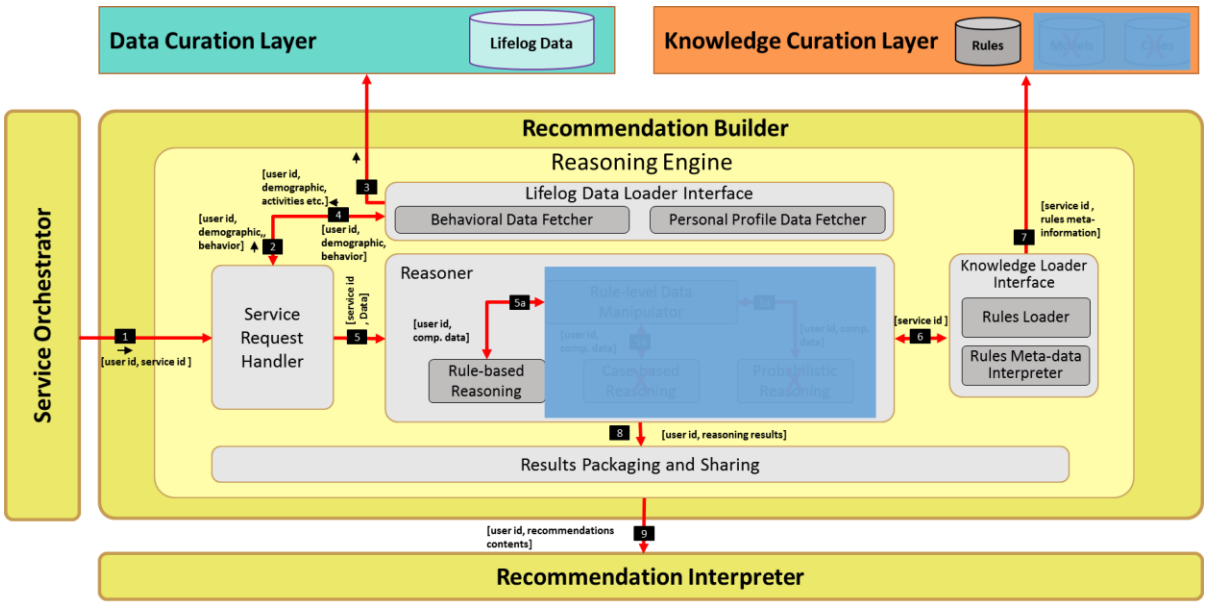

Figure 6.8: Architecture of Recommendation Builder for Mining Mind

Service Request Handler

- Input: User's request, Event
- Output: Activation of lifelog data loader interface and reasoning methodology
- Description: The aim of this module is to collect the service request from the orchestrator, analyze it, and activate the appropriate reasoning methodology and the necessary behavior data and profile information.

Lifelog data loader interface

- Input: Service request
- Output: Profile and behavior data

- Description: The aim of this module is to load behavior and personal profile data from the intermediate database for the users id which is passed as a parameter.

Reasoner

- Input: Service request, loaded data
- Output: Recommendation Contents
- Description: The reasoning process for generating recommendations contents is activated and started. The rule-based reasoning methodologies is used in the MM version 1.5. The rationales are that we have only rules for controlling user unhealthy habits.

Knowledge Loader Interfaces

- Input: Service id
- Output: Rules loaded from KCL
- Description: The purpose of this component is to load rules from KCL to the reasoner so that to start the reasoning process. During the loading process, rules meta-information is used to retrieve specific category of rules from the KCL knowledge base.

Rules Packaging and Sharing

- Input: recommendations contents
- Output: Packaged recommendations contents
- Description: The purpose of this component is to collect all the individuals recommendations generated by the reasoner, package them together and share with recommendations interpreter to be personalize and interpret.

6.6.4. UML Diagrams

- Sequence diagram

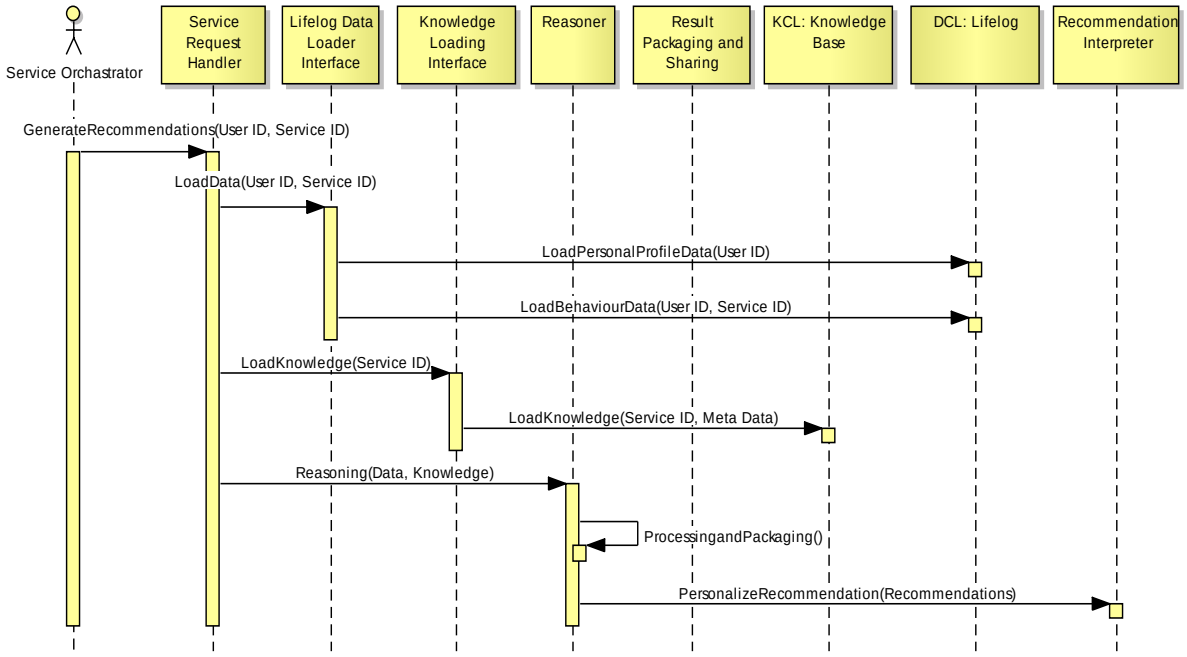

Figure 6.9: Sequence Diagram of Recommendation Builder for Mining Minds

- Component diagram

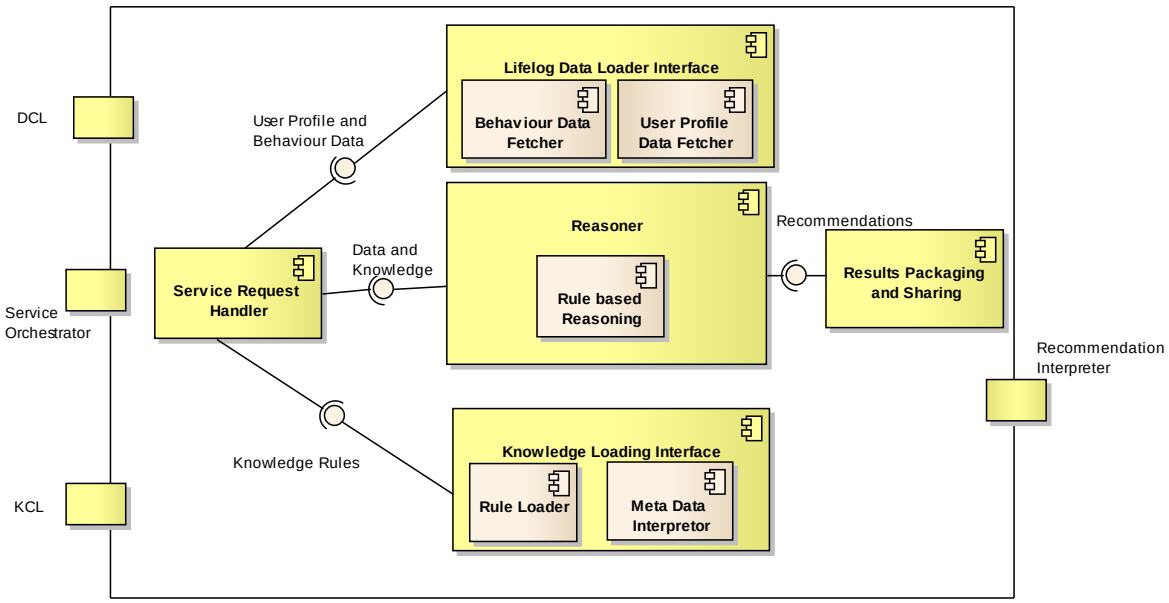

Figure 6.10: Sequence Diagram of Recommendation Builder for Mining Minds

6.6.5. Uniqueness

- Multiple-level reasoning: use of reasoning functions in multiple level to handle the service at more granular level and enhance the recommendations
- Flexible and extensible reasoning framework: design and development of flexible reasoning framework for supporting diverse services, such as physical activities, education and habituation of unhealthy habits etc.
- Handling the diverse nature of data from the user personal profile to the behavior data and intelligently selecting the rules from the knowledge base for reasoning and generating services.

6.6.6. Summary

The recommendations builder module of the Service Curation layer uses multi-level rule-based reasoning methodology in its reasoning components that exploits heterogeneous nature of data from the user personal profile and behavior along with rules from KCL. The knowledge is intelligently curated from the KCL knowledge base for different services generation. There are two model of execution of this module: one is based on the push model of the data curation layer where event triggers the reasoned and the pull model where the user by himself request for the service. The reasoner generates the contents of the recommendations and presents to the recommendations interpreter.

6.7. Recommendation Interpreter

6.7.1. Introduction

Recommendations built with recommendation builder component are needed to deliver according to user context and situation. User context i.e. the daily schedule should be checked and decision to be made whether the recommendation to be forwarded or not.

Context Interpreter is composed of three major components; Query Manager, Context Interpreter and Explanation Manger. Query Manger collects contents of recommendations form

Recommendation Builder and loads context/schedule information from the DCL. The context interpreter has sub-component of time context handler takes care of checking user schedule and act accordingly. If user is available the recommendations are forwarded to next component otherwise the recommendation are held and the lifelog monitor component of DCL is informed with this context information accordingly. Recommendation Explanation component generates explanations upon recommendations by exploiting the contents of recommendations. It also handles the contents translation as by adding alternate formats (text, audio, or video) to the original recommendation according to the user condition.

6.7.2. Related Work

- Researchers have indicated that integrated medical information systems are becoming an essential part of the modern healthcare systems [Duan2011].
- There is increasing awareness in recommender systems research of the need to make the recommendation process more transparent to users. Explanation provides answers to "How" and "Why" questions. How you come up with this recommendation? And why should I accept it? Explanations in recommender systems can be generally understood as a form of communication between the recommender system and the user [Jannach2010].
- Authors of [Tintarev2007] and [Masthoff2007] discussed the set of characteristics that can be associated with explanations such as transparency, comprehensibility, validity, trustworthiness, persuasiveness, effectiveness and education. The explanation helps the users not only to comprehend about the recommended task, item or service but also get educated.
- In reasoning model proposed by Friedrich, Gerhard, and Markus Zanker, the authors explained two models; white box and black box. A White-box explanation describes the underlying conceptual model of the recommendation engine, while black-box explanations do not disclose the functioning of the system to the user [Friedrich2011].
- The multi-dimensional knowledgeable explanation is required to generate alongside the recommendation in order to increase the level of satisfaction and belief of users on the system.

6.7.3. Components Description

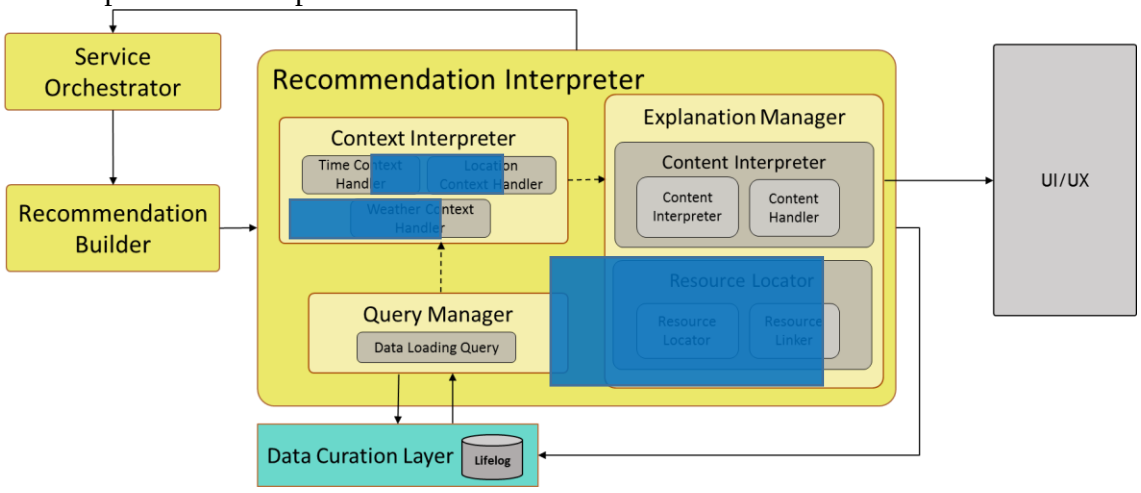

Figure 6.11: Recommendation Interpreter Component of the Mining Minds

Query Manager

- Input: Recommendation Contents, context/schedule

- Output: Query
- Description: Query manager component of the Recommendation Interpreter collects recommendations from the Recommendation Builder and loads the contextual information form the data Curation layer.

Context Interpreter

- Input: Recommendations contents, Context
- Output: Context Interpreted Recommendations
- Description: The purpose of this component is to interpret the user context and deal with the recommendations accordingly. Each of the applicable contexts such as time, location, and weather are checked before forwarding the recommendation to the user.

Explanation Manager

- Input: Filtered Recommendations
- Output: Explained Recommendation
- Description: While Recommendation Explanation component generates explanations by understanding the recommendation contents. It add the appropriate form of recommendation (text, audio, or video) according to the condition of the user.

6.7.4. UML Diagrams

- Sequence Diagram

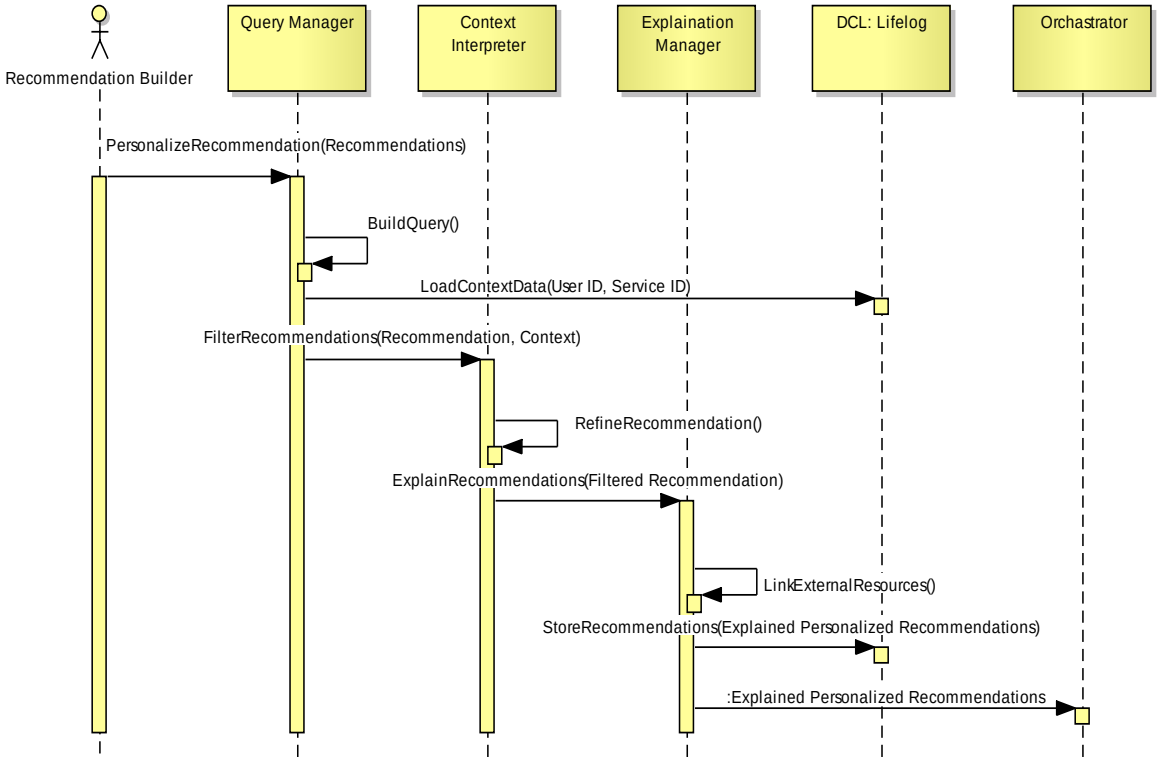

Figure 6.12: Recommendation Interpreter Sequence Diagram

- Component Diagram

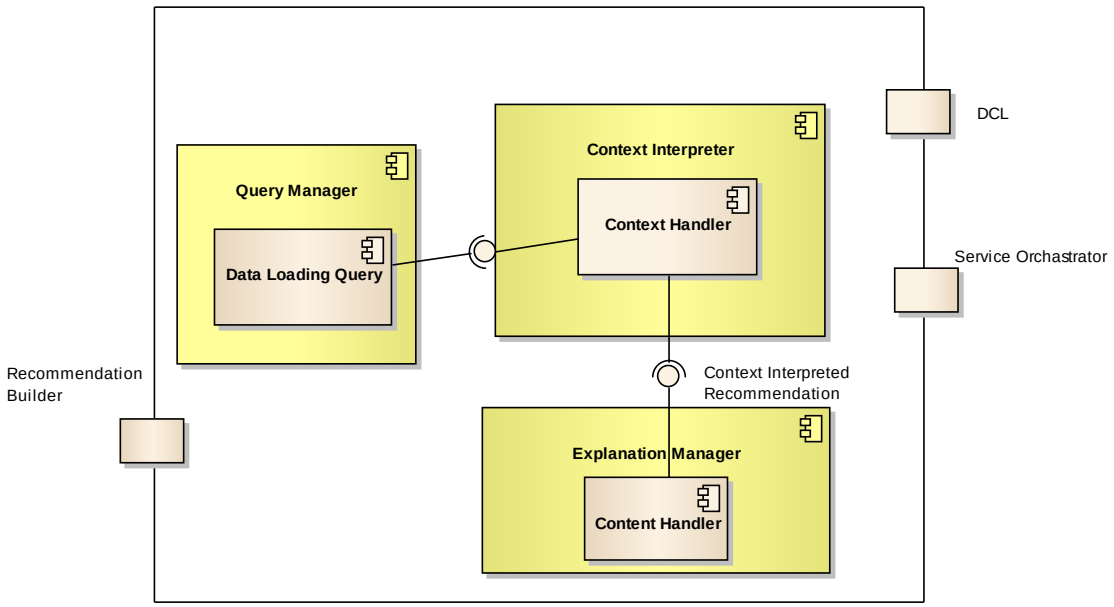

Figure 6.13: Recommendation Interpreter Component Diagram

6.7.5. Uniqueness

- Context-aware recommendation: deliver recommendation according to the context and situation. Checking more than one contexts before delivering the recommendation to the user.
- User-aware content addition: Based on the interpretation of recommendation, appropriate contents are added to the original recommendation to satisfy the user requirements.

6.7.6. Summary

The purpose of this module is to interpret the recommendation contents and provide them to the end user's in an intuitive format. This is done by presenting the information in the how, when to whom model. Sometimes the explanation become difficult for the users to understand due to logic complexities. Unless users provided with proper explanation along with the original recommendations, it is hard to believe the decision made by the system. This module proposes to interpret the user context and recommendation contents and deliver to the user accordingly.

6.8. Service Curation Layer Overall Uniqueness

- Dynamic Invocation of appropriate service based on dynamically generated events by DCL for the user behaviors.
- Implementation of the PUSH model for service provisioning.
- Multiple-level reasoning: use of reasoning functions in multiple level to handle the service at more granular level and enhance the recommendations

- Flexible and extensible reasoning framework: design and development of flexible reasoning framework for supporting diverse services, such as physical activities, education and habituation of unhealthy habits etc.
- Handling the diverse nature of data from the user personal profile to the behavior data and intelligently selecting the rules from the knowledge base for reasoning and generating services.
- Context-aware recommendation: deliver recommendation according to the context and situation. Checking more than one contexts before delivering the recommendation to the user.
- User-aware content addition: Based on the interpretation of recommendation, appropriate contents are added to the original recommendation to satisfy the user requirements.

Chapter 7
Supporting Layer

7. Introduction to Supporting Layer

The supporting layer is in charge of user interface, visualization, analytics and security. The supporting layer create a unique interface which has the adaptive and personalized approach towards building and managing the user interfaces. The analytics gives user and experts different insights into the habits, activities and different classification of the application.

7.1. Functional Diagram for Supporting Layer

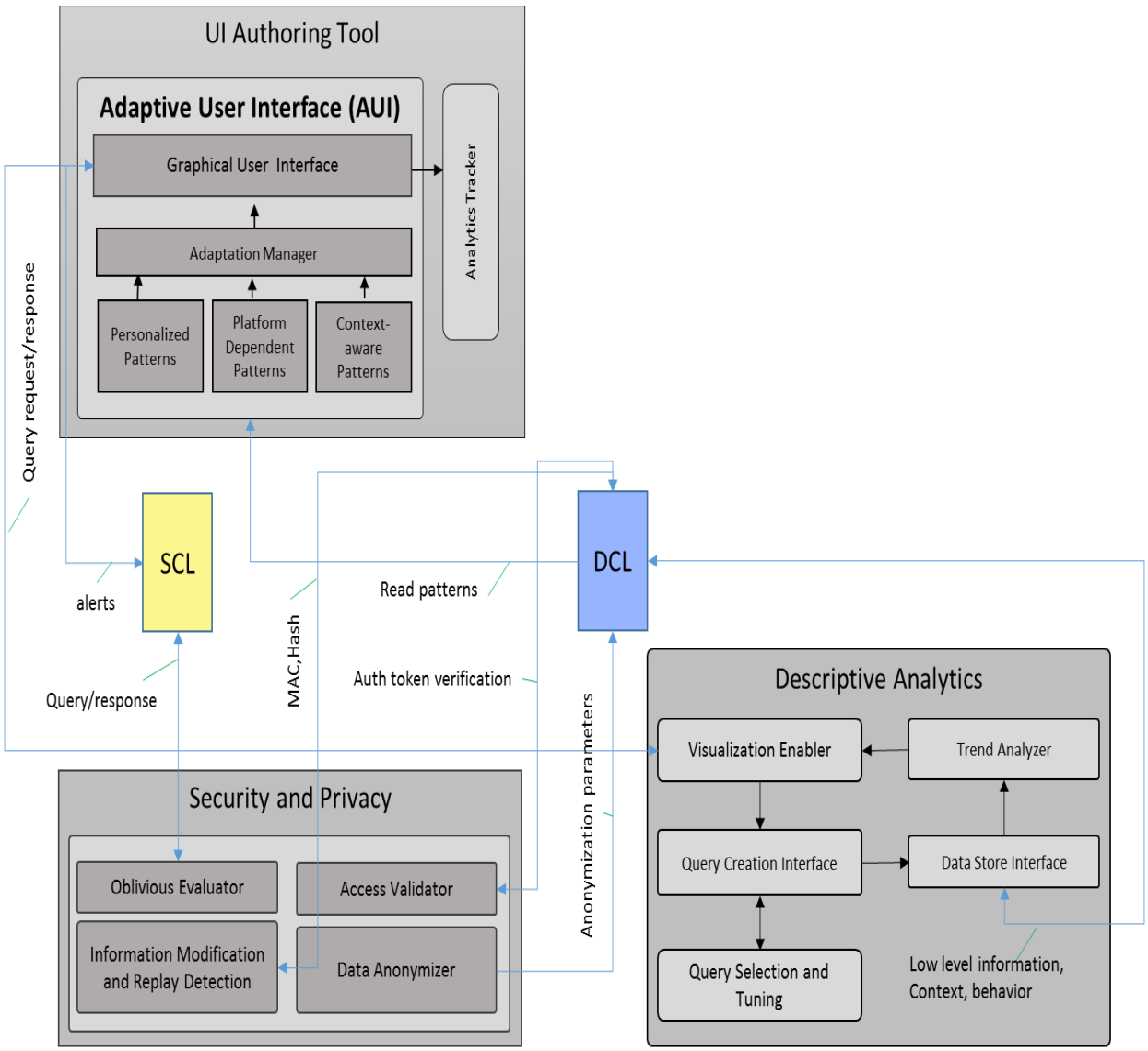

Figure 7.1: Functional Diagram for Supporting Layer

7.2. User Case Diagram for Supporting Layer

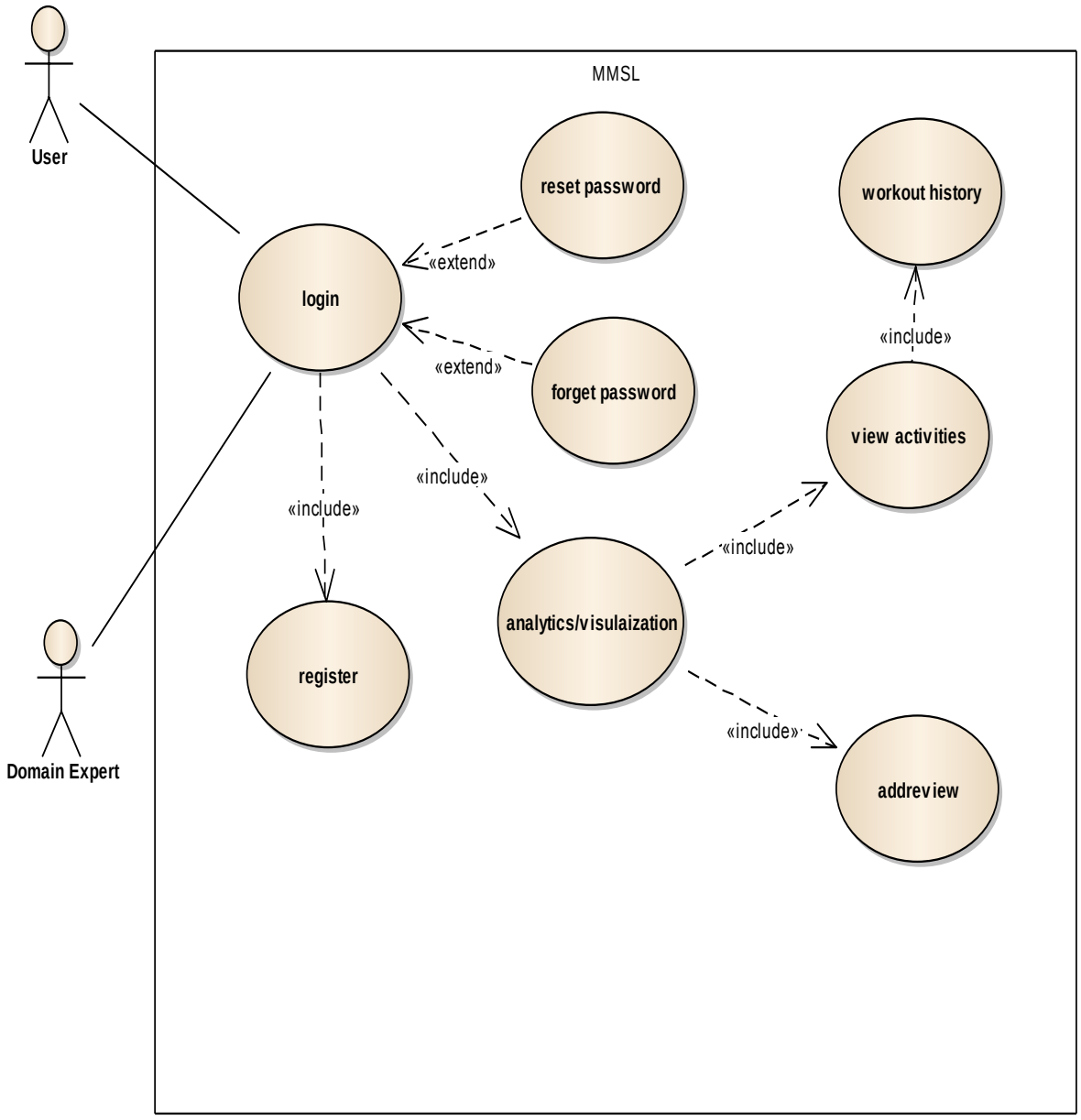

Figure 7.2: Use case Diagram of supporting Layer

7.3. Component Diagram for Supporting Layer

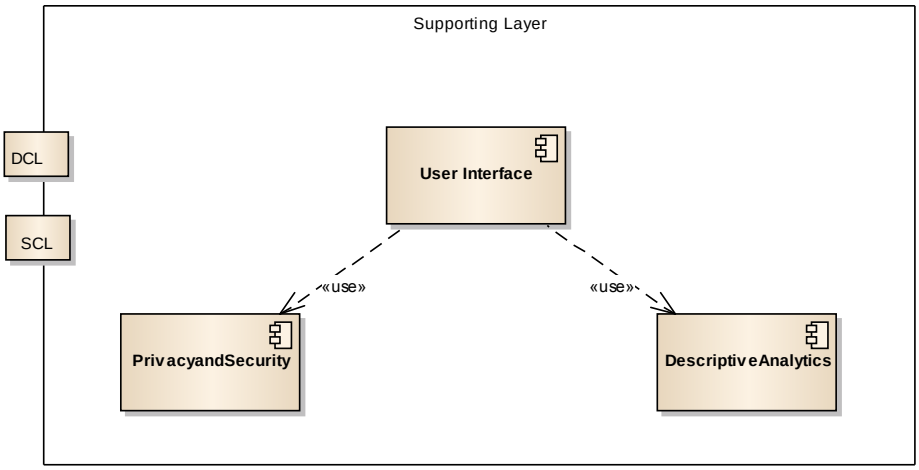

Figure 7.3: Component diagram of supporting layer

7.4. Sequence Diagram for supporting layer

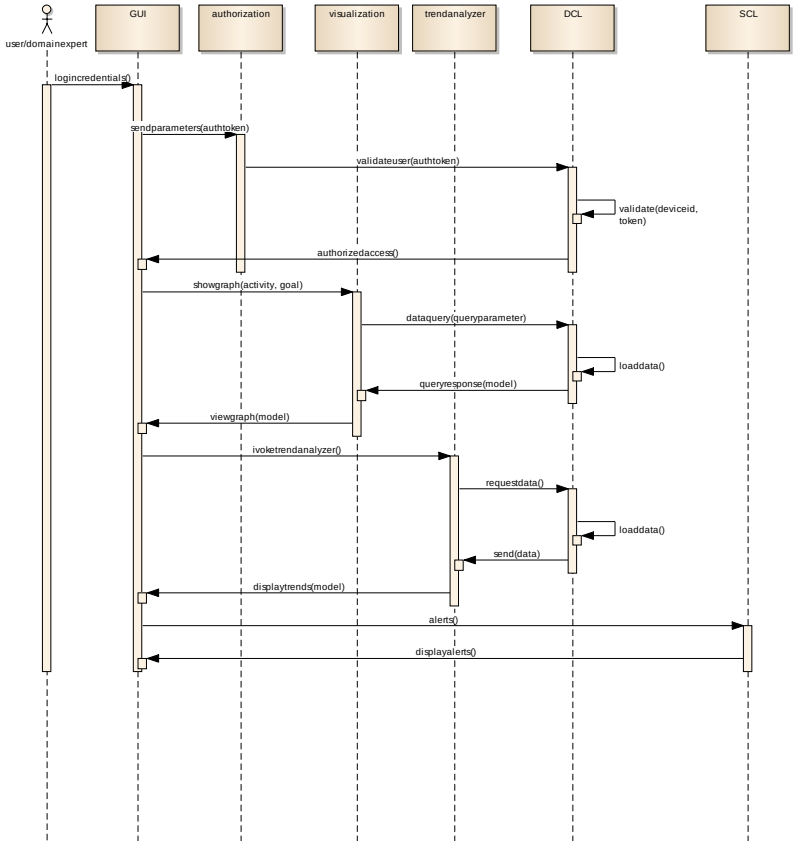

Figure 7.4 Sequence Diagram of Supporting Layer

7.5. UI Authoring Tool

7.5.1. Introduction

UI provides the interface to the user where all the services are provided in a personalized manner. The different constructs of the MM are displayed in the UI at the most appropriate location for getting the user experience. It contains the use profile information and the lifelog view of the user.

7.5.2. Components Description

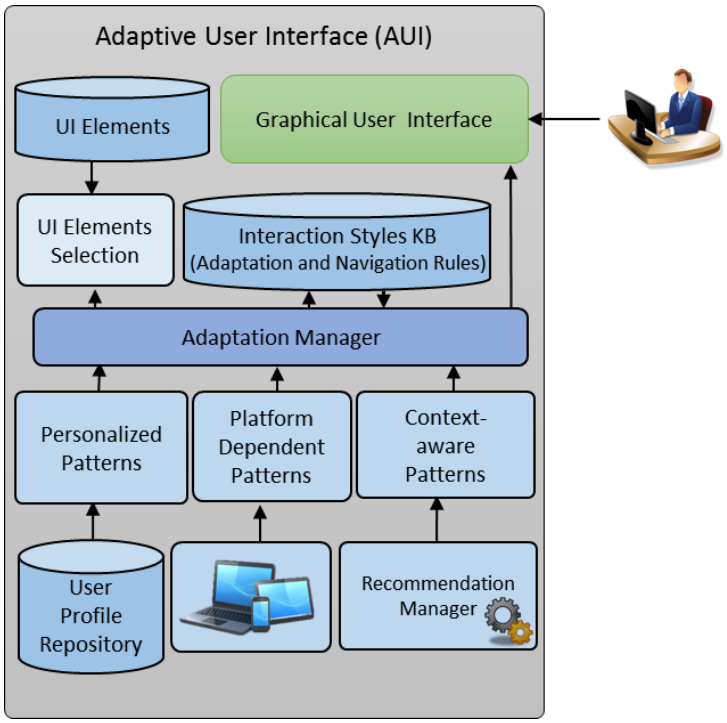

Figure 7.5: AUI Architecture

- **Adaptive User Interface (AUI)**

Personalized Repositories

- Input: User personalized data
- Output: User preferences, platforms used, and contextual information
- Description: This layer consists of three repositories: User Profile Repository, Platform Information Repository, and Context Analysis Repository. User Profile Repository is used for storing the user preferences in his daily life activities. Platform Information Repository stores the platform related information of the user. Context Analysis Repository stores the current context about the user.

Personalized Patterns

- Input: User preferences, platforms used, and contextual information
- Output: Personalized preferences, platform, and contextual patterns

- Description: This layer consists of three modules: Personalized Patterns, Platform Dependent Patterns, and Context Aware Patterns. Personalized Patterns are based on the preferences information stored in the repository. Platform Dependent Patterns are the patterns derived from the specific platforms and its specifications used by the users. Context Aware Patterns are the patterns based on the current contextual information of the user.

Interactive Styles KB

- Input: Personalized preferences, platform, and contextual patterns
- Output: Individualized Pattern
- Description: This Knowledge Base stores the adaption and navigation rules that are personalized to specific users. These are derived from the personalized patterns. These are used for adaption of the user interface and also navigation between different graphical user interfaces.

UI Elements Selection

- Input: Individualized Pattern
- Output: UI Elements
- Description: This component retrieves the UI elements from the UI Elements repository. This is used by the Adaption Manager for building the graphical user interface.

Adaptation Manager

- Description: This component behaves as a coordinator between different internal components. It takes initial input from the Personalized Patterns layer and forwards it to the Interactive KB. It then takes input from the Interactive KB and provides it to UI Elements Selection. It finally takes output of the UI Elements Selection and builds Graphical User Interface accordingly.

7.5.3. UML Diagrams

- Sequence Diagram

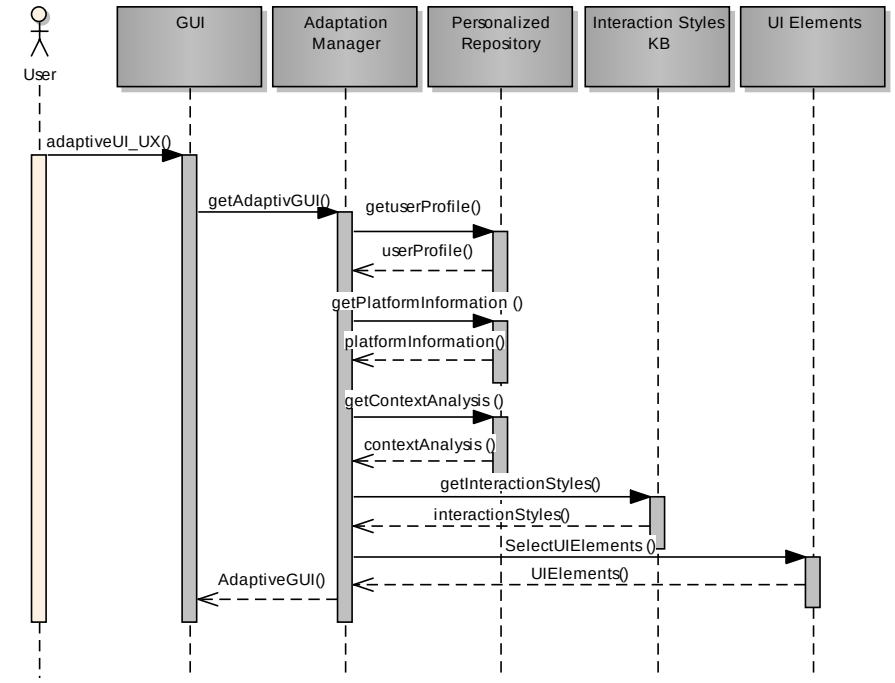

Figure 7.6: AUI Sequence Diagram

- **Component Diagram**

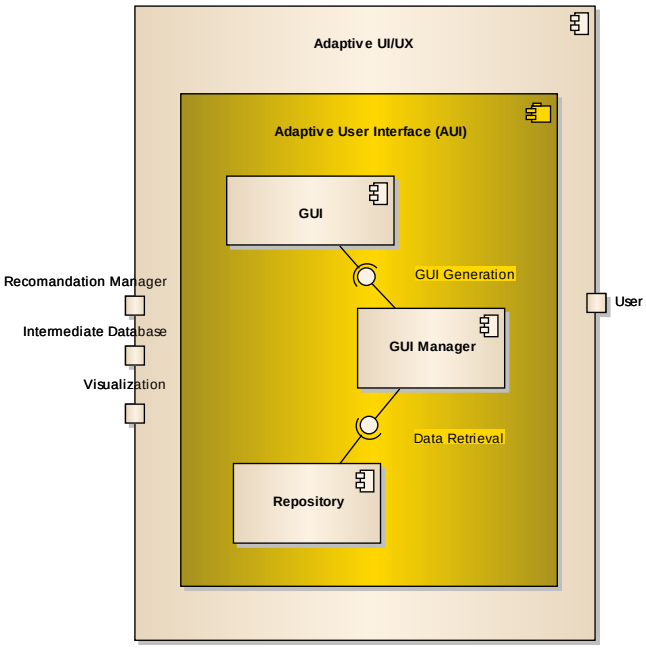

Figure 7.7: AUI Component Diagram

7.5.4. Summary

AUI module takes the personal preferences and contextual information of the user into account for adapting the user interface. User experience controls the evolutionary process of the adaptation of the user interfaces and also maintaining the personalization aspect.

7.6. Descriptive Analytics

7.6.1. Introduction

- Big data is creating many opportunities for different and diverse fields to achieve deeper and faster insights that can enhance the decision making.
- Descriptive analytics mines data to provide trending information on past or current events that can give the context needed for future actions. It is basically a quantitative summary that describes the main characteristics of a data collection.
- Descriptive analytics, such as reporting, dashboards, and data visualization is being widely used now with emergence of big data.
- These field are increasingly turning to visualization based tools has 30 percent growth rate in 2015. We can provide interactive and easily understandable visual formats to improve the overall user experience.

7.6.2. Related Work

- A novel network traffic visualization scheme is proposed based on exploiting frequent item-set mining with the aim to visualize traffic patterns extracted from communication logs using hypergraphs [Glatz2014].
- A Intra-user and inter-user social media from a big data framework is proposed. It analyzes social media and visualization from the server and display to a user. It gets the signal fusion from heterogeneous sensors [Chang2013].
- Kandogan ET all, created a feature ranking and annotation method. They do annotation interaction to help support understanding of the structure of data [Kandogan2012].
- The new emerging directions are highlighted to create clear, meaningful and integrated visualizations that give biological insight, without being overwhelmed by the intrinsic complexity of the data [Gehlenborg2010].
- A novel Massive Model Visualization via a rendering approach called MMDr using spatial hierarchies, with the goal of a better understanding, to achieve interactive frame rates on extremely large data sets, and help to accomplish more advanced research into Massive Model Visualization [Bennett2009].
- A tuning parameter taxonomy is proposed which consists of behavioral, Numerical statistically and temporal classifications and are further sub divided. This taxonomy help analyze different attributes and dimensions of the data being visualized [Dancy2008].
- An overview of current real-time massive model visualization technology, with the goal of providing readers with a high level understanding of the domain [Dietrich2007]. They have included simulations and different rasterization algorithms for rendering.
- A novel approach to automatic image annotation based on two key components: (a) an adaptive visual feature representation of image contents based on matching pursuit algorithms; and (b) an adaptive two-level segmentation method [Shi2004].

- For parameter tuning McAdams et al uses a subtle, but significant, change in the design: the addition of a tuning parameter in place of an increase in component precision. Statistical models are used to develop a framework for the tuning parameter design method in [McAdams2000].

7.6.3. Components Description

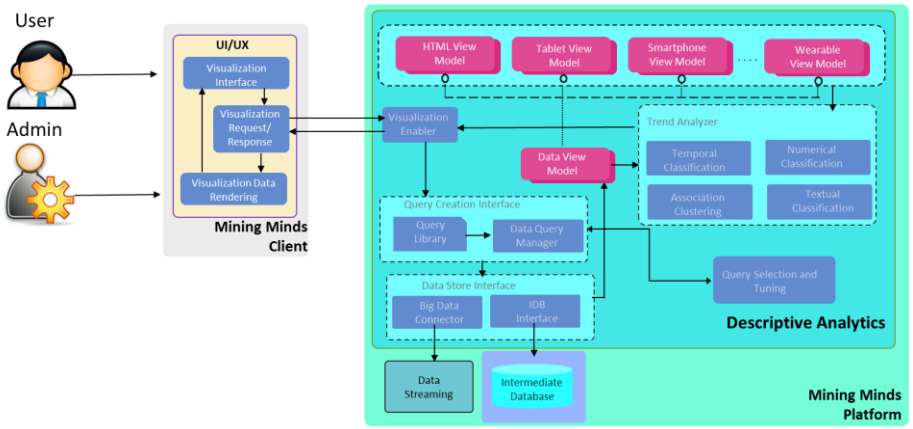

Figure 7. 8: Architecture of Descriptive Analytics

Trend Analyzer

- Input: Data from Data Store Interface.
- Output: Trends and summaries to Visualization Enabler
- Description: In Trend Analyzer, Numerical Classification and Temporal Classification are performed. In Numerical Classification, tuning will be done by considering capacity, counter, size and threshold parameters. In association clustering statistical features like Minimum, Maximum, and Average features that are used to tune the parameters. This approach finds the ranks of features as high ranked features produces more accurate models.

Query Creation Interface

- Input: Query from Visualization Enabler.
- Output: Mapped Query with parameters
- Description: It maps the query from visualization enabler. The data query manager takes the query from query library. If the query is not in the library or needs to be tuned/changed manually, the query selection and tuning module is triggered to change the query.

Visualization Enabler

- Input: a) Request from UI. b) Trends and graph data from trend analyzer
- Output: a) Query for Query interface. b) Summaries and graphs for UI
- Description: a) The visualization enabler has a request manager which forwards the query from UI module. This query is then passed to query creation interface. b) The graph evaluation modules takes graph data from the trend analyzer and maps it on a graph template. The visualization context takes into account the data and profile of the user

7.6.4. UML Diagrams

- Sequence Diagrams

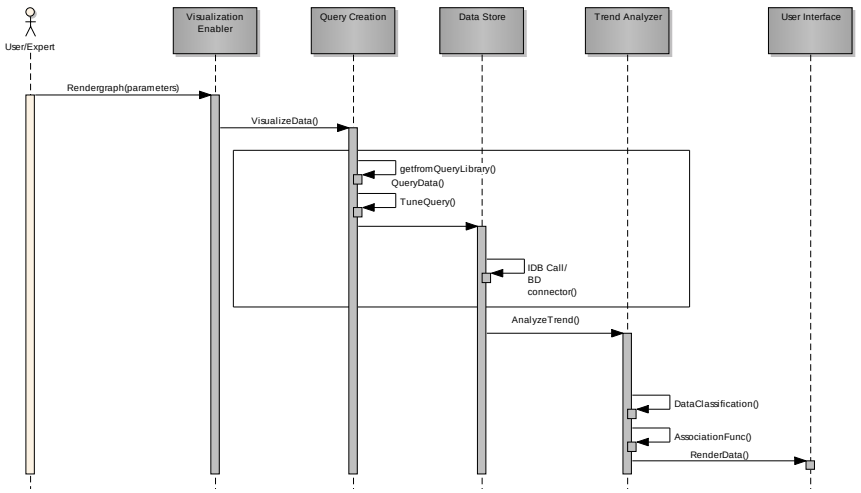

Figure 7.9: Sequence Diagram of Descriptive Analytics

- Component Diagram

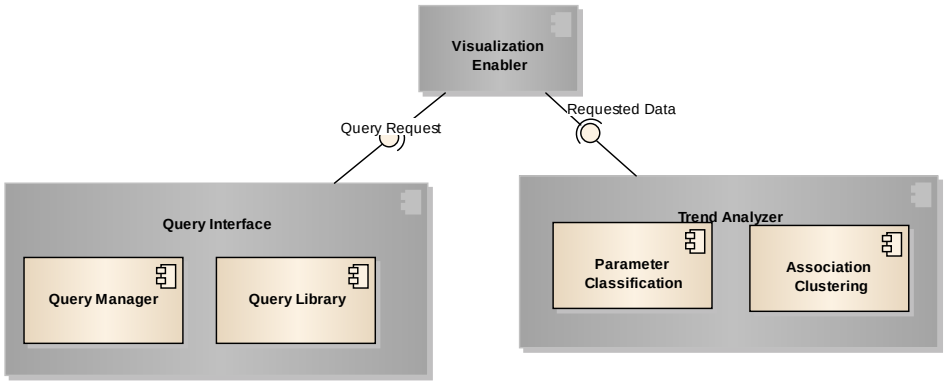

Figure 7.40: Component Diagram of Descriptive Analytics

7.6.5. Summary

Big data analytics and visualization go hand in hand as we need an effective way to display the data intuitively for the users and the developers. Interactive data analysis, infographics and data mashups are the latest trends in displaying big data graphics in social networks. The parameters vary from social attributes, temporal attributes and location based. We will take the data from the big data and show visualization in different layers to different users which vary from service users to domain experts.

7.7. Security and Privacy

7.7.1. Introduction

Privacy and security are the inevitable components of any system dealing with sensitive contents. The sensitive contents mainly include personal information, data related with health or finance and the media files. The core idea behind Mining Mind is to collect data from multiple streams like sensors, wearable devices, cameras and smart devices like mobile phone and television. After

collecting data from various streams the curation process is applied on that to discover hidden knowledge out of it. This knowledge which otherwise not possible without using the curation process, is then induced and shared back to its owner to assist in better and healthy living style. The orientation of this data is purely related with an individual; therefore concerns for protecting this data and information are very high.

7.7.2. Related Work

- For secure storage, public key encryption will be used and to protect further concerns of user query tracing or pattern discovery, paradigm of oblivious computing will be used. For oblivious computation, we have already developed oblivious term matching and oblivious access policies while dealing with the public cloud infrastructure [Zeeshan2012] [Zeeshan2013].
- While working with the encrypted storage, the data utility gets compromised and for this reason various searchable encryption schemes [Michel2005], [Mihir2007] have been proposed to deal in this situation. Besides using the oblivious term matching, searchable encryption are also helpful to deal with the encrypted storage yet utilizing the data utility at optimal level.
- Private matching is another way that is useful to protect the condition of any query until it is not satisfied. Recently, we have also proposed a light weight protocol for the private matching. The idea behind this protocol is to randomize the output result even repeated with same values again and again and that too without using the encryption. The motivation behind this approach is to mislead the intrusion or inference process with least instrumentation. If this approach is further wrapped with encryption, the overall system will become more deterrent against the malicious activity.
- With completion of curation and inferred knowledge, it is time to deliver it with the intended user. This step can be done by using the public key cryptography where only authorized user would be able to recover the information. To further enhance knowledge delivery, the output can be tied with some dependency of biometric or physical devices like smart phone or geographical area.
- The idea behind Mining Mind is to refine and alleviate the life style of humans especially the old aged society. If the information is useful for an individual, the same information can be shared as collective knowledge with the research community. For this purpose anonymization [Roberto2005]. [Ashwin2007] can be used that help to preserve the individual identity. At further level of refined access, the secure function [Louis2006] evaluation or private matching [Mahmood2013] are yet other available options.

7.7.3. Components Description

- Access Validator: To ensure authorization, the incoming fresh request is validated against user credentials. To validate a user, a unique authentication token is used which is sent along the user request. Upon receiving the token and login credentials, the access validator contacts with the credentials and authorized storage information. After the request initiator is validated, the access is granter or turned down otherwise.
- Data Anonymizer: To share data with subscribed users, data anonymizer invokes standard techniques of k-anonymization and l-diversity while transforming the Microdata into anonymized data. The purpose of utilizing services of this component is to hide individual details from the shared information. Also its services are utilized to minimize the linking attack possibility.
- Oblivious Evaluator. In this component, mainly the oblivious term matching (OTM) technique is used. Through OTM, it is possible to evaluate encrypted user queries on encrypted data. While utilizing the services of a public cloud, enforcing its features will preserve privacy of user queries as well as response which is prepared obliviously by the cloud. Besides OTM, Searchable

Encryption: Searchable encryption (SE) is a standard that is used to work directly with the encrypted data without decrypting it.

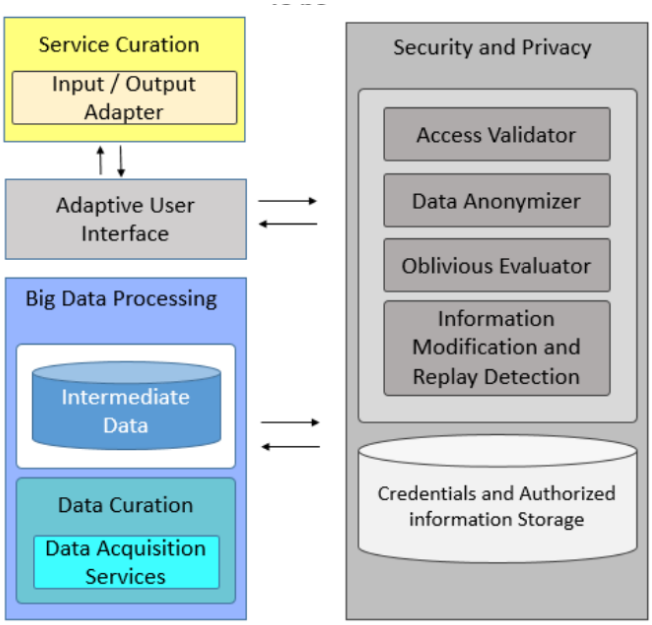

Figure 7.51: Architecture for Security and Privacy in Mining Minds

- Information Medication and Replay detection. To safeguard against inference attack and man in the middle attack, a new module has been added in the security and privacy component of supporting layer. The purpose of this component is to validate the integrity check received from the user side. After the integrity check (by employing the MAC), the corresponding data is make persistent for further processing. In case of mismatched MAC the integrity is declared suspicious and data packets are discarded. In this case, user credentials are refreshed and reflected in relevant components.

7.7.4. UML Diagrams

- Sequence diagram

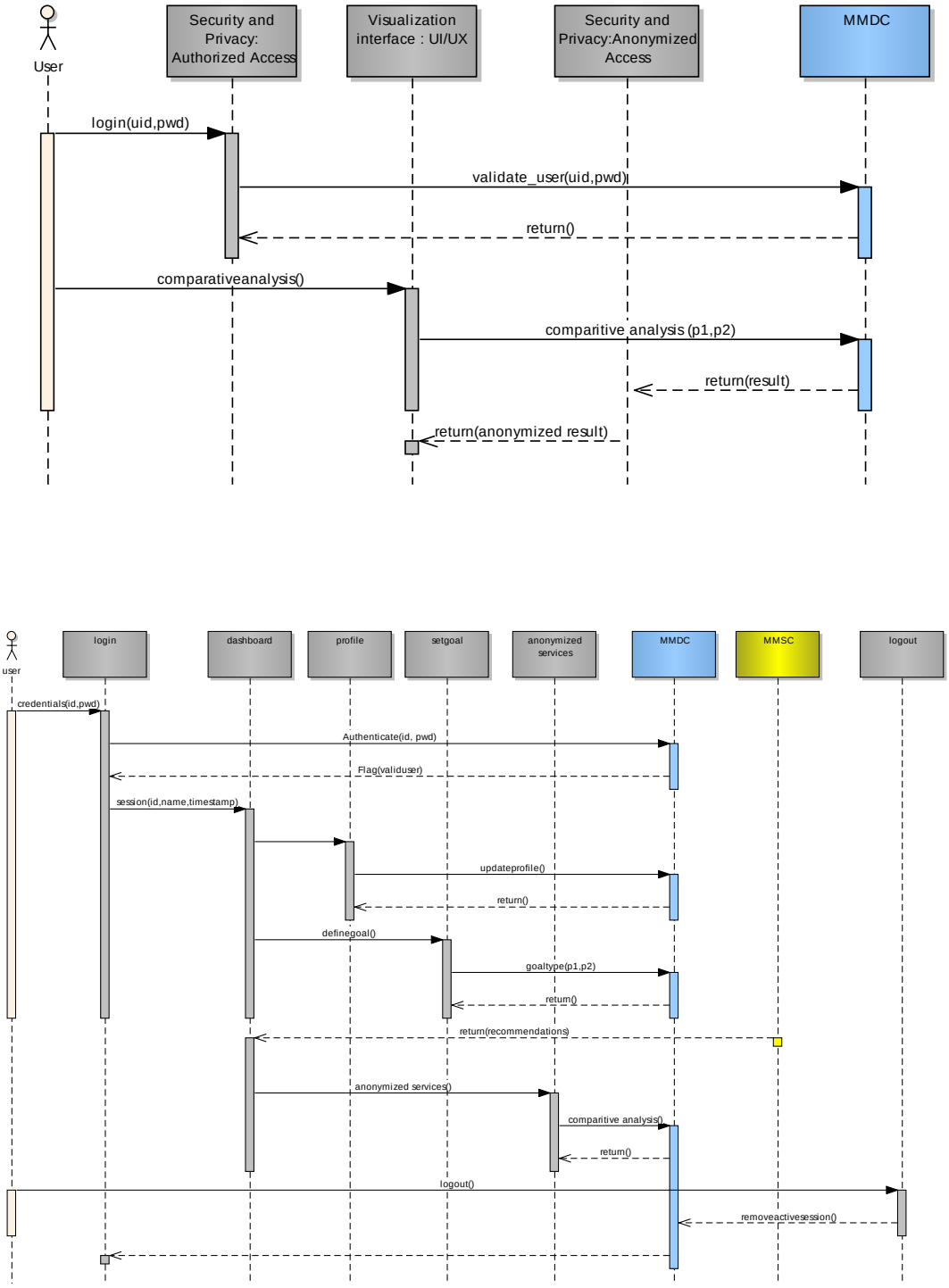

Figure 7.62: Interaction of MMDC and SC with Supporting Layer

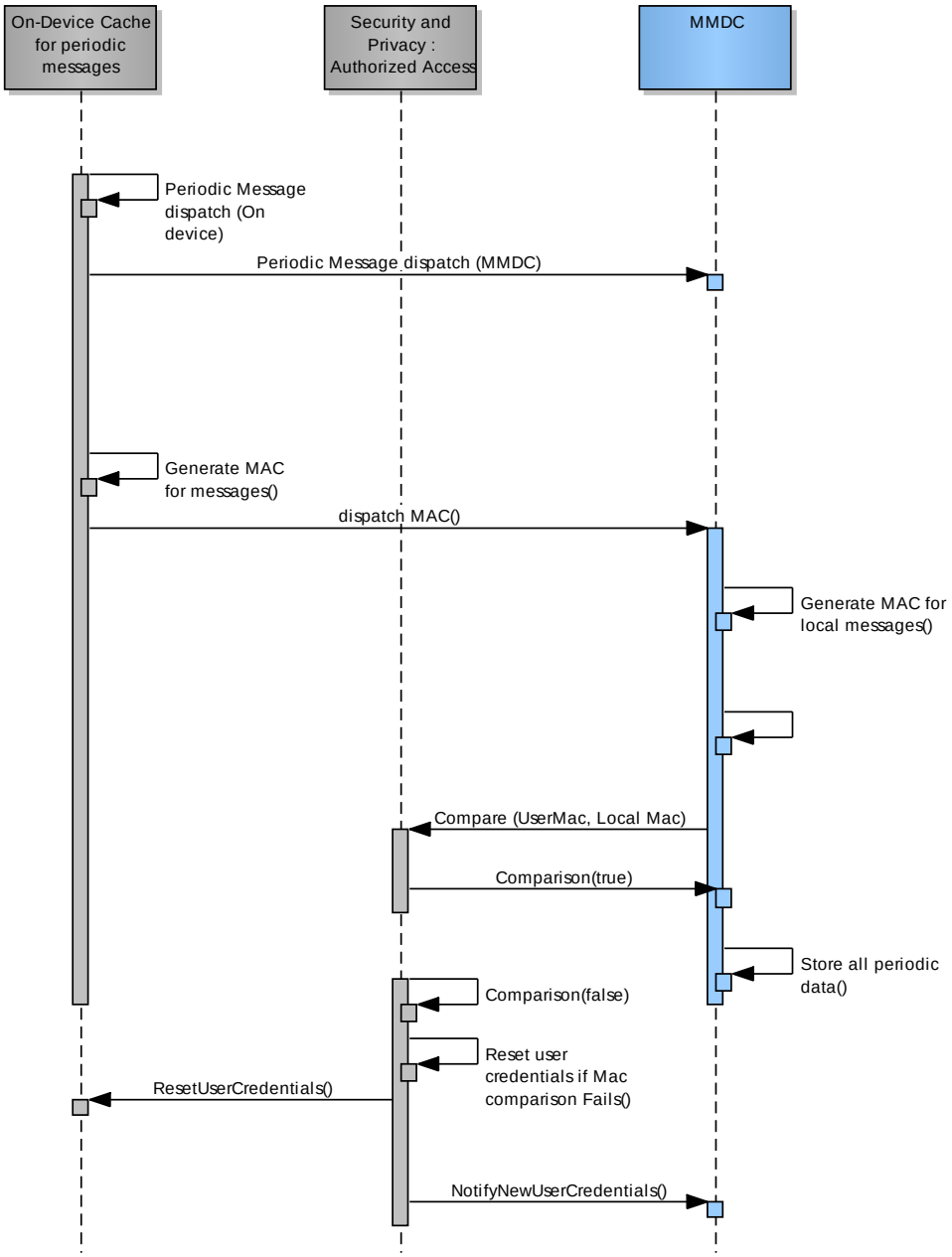

Figure 7.73: Detecting Information Modification and Replay attack

- Component Diagram

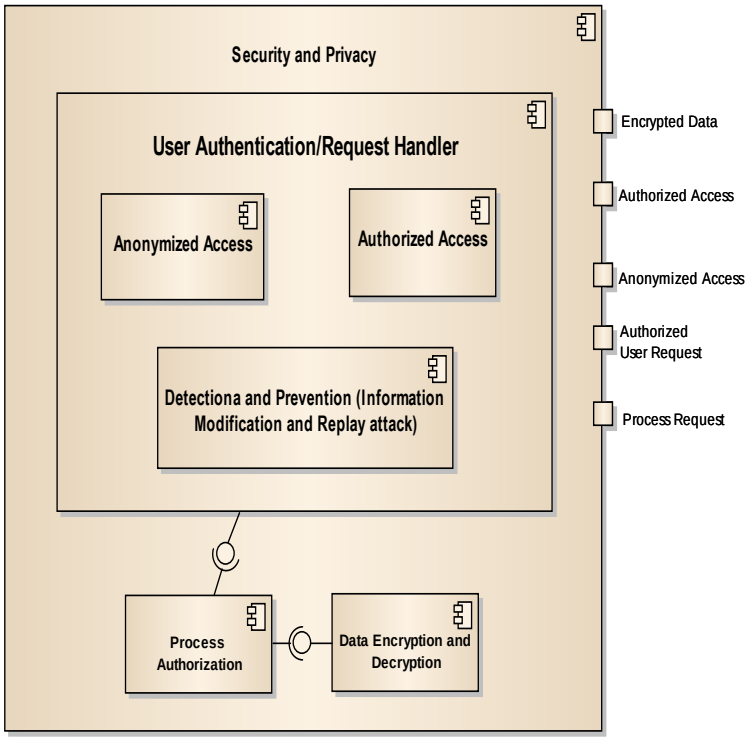

Figure 7.84: Component Diagram (Security and Privacy) for Mining Minds

7.7.5. Summary

Privacy has two fold aspect to safeguard it. First is directly related with data and second is related with its usage. This usage is again subdivided into two categories. The usage within the system and second is by the user. While availing the services of cloud infrastructure it is desired that a cloud service provider must not lean anything else beyond expected. The expected is the user request and its reply. Within cloud infrastructure the output result can be made independent from the fact that either the quires have been met with the given condition or not. Considering the second perspective with respect to user, her identity as well as access pattern must remain hidden as an additional leakage of information.

For secure storage, public key encryption will be used and to protect further concerns of user query tracing or pattern discovery, paradigm of oblivious computing will be used. For oblivious computation, we have already developed oblivious term matching and oblivious access policies while dealing with the public cloud infrastructure

7.8. Supporting Layer Overall Uniqueness

- Continuous evolution of AUI with contextual information change
- User satisfaction as a feedback metric
- Personalized, Platform and contextual information are combined for AUI
- Parameter Classification and structuring attributes w.r.t. the query
- Grouping and association techniques
- The prompt labeling mechanism to get the user's feedback.
- Two scheduler methods for reflecting the user feedback.
- High entropy solution with efficient execution time named "Reflection".
- Our Design which works obliviously while working with the encrypted storage by utilizing the cloud resources optimally.

Chapter 8
Services Scenarios

8. Service Scenarios

Our proposed Mining Mind platform can provide diverse amount of services to the user. We are explaining the weight management service scenario to get an idea of improved services by our platform.

8.1. Service Scenarios

- Weight Management and Monitoring Services

Mining Mind project is a proactive approach to adopt healthy lifestyle in our daily routines. For instance, daily exercise, diet, sleep and social relationships are the wellbeing indicators. A progressive health effects can be observed if they are well managed. Weight is one of the alarming factor among most of the healthy communities. Management and monitoring of weight is one of the scenario of our Mining Minds project. This is performed by monitoring the daily routine as well as personalized daily life activities of the person and providing him recommendations in management of the weight.

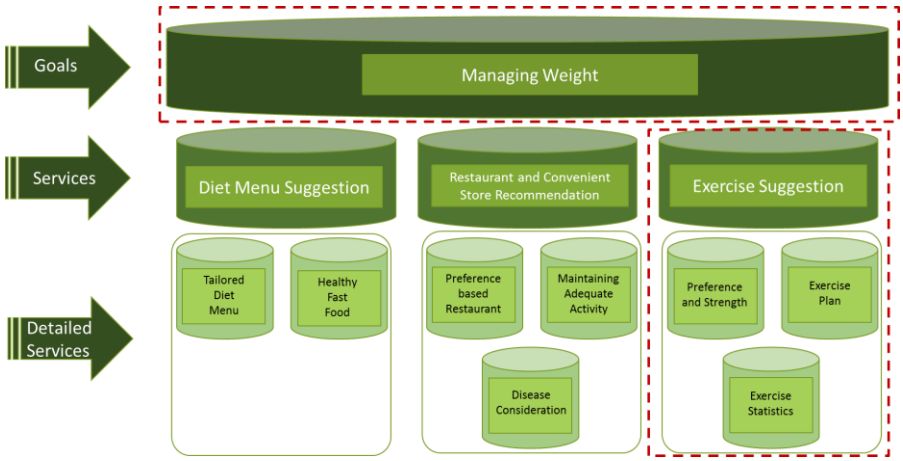

Figure 8.1: Weight Management and Monitoring Services

Followings are the features that our system provides

- Management and Monitoring of personalized weight
- Personalized analytics to find abnormalities in the weight
- Know the situation before and after the activities performed to burn calories
- Smart Interfaces
- Using feedback, improve the quality of services
- Smart Analytics to visualize the personalized weight pattern

- Use case Scenario: Weight Management

Mining Minds project recommends about the activities that are most feasible at a particular circumstances for the user to perform such activities that can burn calories and help to manage weight. The user scenario is presented in Figure 8.2.

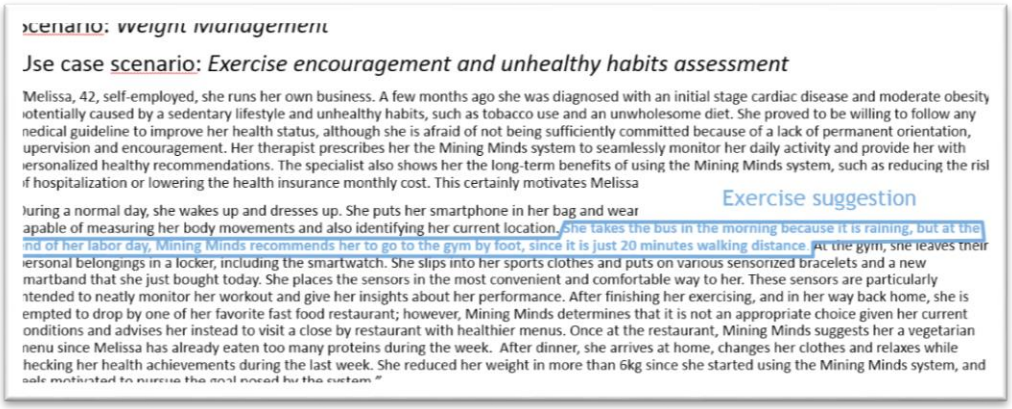

Figure 8.2: weight management scenario

The details of the use case are elaborated in the table given below:

Use Case Name:	*Exercise encouragement and unhealthy habits assessment*
Actors:	User /Customer
Description:	***User takes the bus in the morning because it is raining, but at the end of his/her labor day, Mining Minds recommends him/her to go to the gym by foot, since it is just 20 minutes walking distance." (EXERCISE RECOMMENDATION)***
Trigger:	Activity recognized
Preconditions:	1) The **user signs up** in the Mining Minds system and **fills in** the requested profile information (age, height, weight, diet habits) for the first time **2)** The **user logs into** the Mining Minds system
Postconditions:	1) At the end of the day, the **user is asked** to weight herself and update her weight in the system 2) Afterwards, the **user checks** the evolution of her weight, progresses made, and goal
Normal Flow:	1) The **user initiates** the behavior monitoring process 2) The registered data (ACC+GPS) is stored in the **Intermediate DB** 3) The user starts performing the first devised activity, here, "take the bus" 4) The **activity recognition** functionality detects that the user is in the bus 5) The detected activity is stored in the **Intermediate DB** 6) At lunchtime, and based on the user's cumulative previous activity, the **recommender** suggests her to go to the restaurant by foot

Exceptions:	Not recognized the user Communication connection lost

- Mockups of Scenario: Weight Management

Interactive and attractive interface is required to maintain the interest and catch the attention of the user towards tedious tasks. Minding minds will give highly attractive user interface to users to evaluate progress in activities towards healthy life. Mockup mentioned in the Figure 8.3 reflects recommended activities according to the weight.

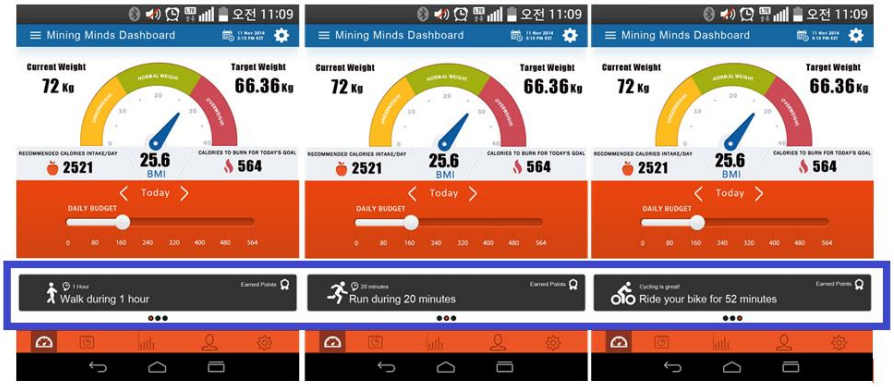

Figure 8.3: User Interface Mockups

These performed activities will have impact on user's weight, the progress can be visualized in the mockup as mentioned at Figure 8.4.

Figure 8.4: Analytical View Mockup

- Mockups of Scenario: Educational facts

Mining Minds Ver. 1.5 focuses on education for healthy lifestyle with interactive tutorials. Interface provides the facts and recommendation on timeline bases. User can view the archive of facts, recommendations and workout on the basis of weeks. Consider the recommendation is about the stretch legs and arms the tutorial video will be available for stretching technique and user can view it by clicking the popup as shown in Figure. 8.5.

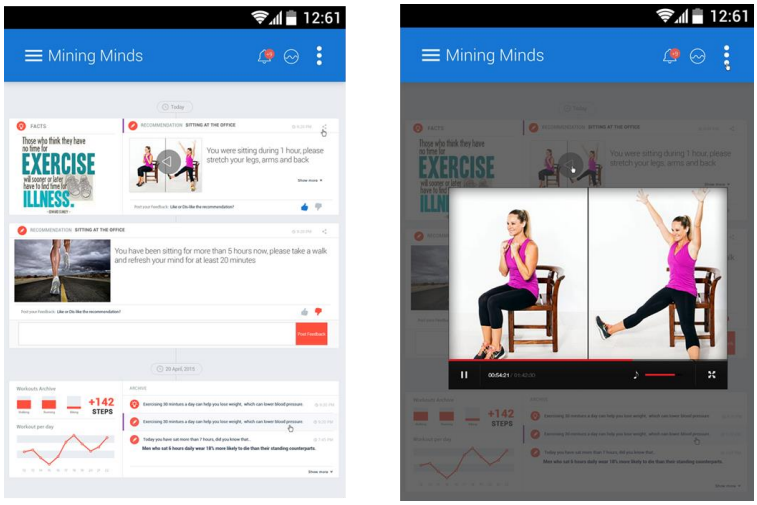

Figure 8.5: Mockups for educational tutorial

The designing of the mockups is according to the google design guide line and two menus options are providing to the user for his convenient and attraction. The sidebar menu is for application setting, general feedback and edit the user profile. While dotted menu option is to refresh the dashboard and sign out facility. In feedback option user can record his/her like and dislike in both free text feedback style as well as template base style as shown in Figure.8.6.

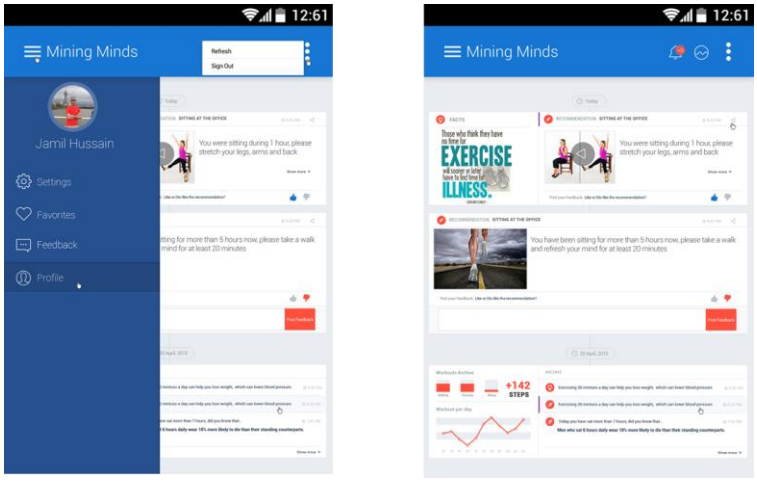

Figure 8.6: Mockups for feedback and Edit profile

Chapter 9
Conclusion

9. Conclusion

Mining Minds platform provides personalized services to users and it is based on high quality of content and high quality of services. It analyses the user's behaviors to provide customized services to the users. The quality of contents and presentation are two important aspects to use the system for a long duration. If the quality of content is rich but poorly presented to the users', then there is very low probability of system adoption in daily routines and vice versa. This objective can be achieved by providing virtual personal assistant to support person in better life-care management.

The proposed platform maintains the quality of content and provide quality of services as a virtual personal assistant or coach. Therefore, we introduce the layered architecture to support real time data streaming and processing. We separate the data, information, and service curation layers. Data curation layer deals with the real time data streaming and processing and store into the Hadoop based file system. Furthermore, intermediate data is generated to provide fast access and structured data when it is required by information and service curation layer. Information curation layer extracts the low-level contexts for human's life modeling. Service curation layer consists of reasoning module to provide the generalized recommendation. These recommendations are further refined in recommendation manager to compliance with the user's profile.

The Mining Minds, a novel digital platform for personalized healthcare and wellness support. The platform has been neatly designed taking into account crucial requirements of the digital health and wellness paradigm. This work has also described a unique architecture defined to provide the necessary functionality to enable curation and mining of data, information, knowledge and services for personalized health and wellness support. An initial realization of the key architectural components, as well as an exemplary application that showcases some of the benefits provided by Mining Minds, have also been presented. The work is ongoing to complete the implementation of the devised architecture with new additional components as well as to evaluate its services on a large scale testbed.

Privacy has two fold aspect to safeguard directly data and secondly its usage. The usage is subdivided into within the system usage and by the external user. While availing the services of cloud infrastructure it is desired that a cloud service provider must not lean anything else beyond expected. Within cloud infrastructure the output result can be made independent from the fact that either the quires have been met with the given condition or not. Considering the second perspective with respect to user, her/his identity as well as access pattern must remain hidden as an additional leakage of information.

The updated personal and contextual knowledge keep on updating in the knowledge bases for processing and providing new personalized services to the users. High quality of contents and personalized services are generated through our proposed platform. The scope of the services includes personalized life-quality improving services, silver business services, life-care services, and avail the proactive approach to control the chronic disease as well as active lifestyle.

References

Related Work

[R.C.Wu 2010] R. C. Wu, D. Morra, S. Quan, S. Lai, S. Zanjani, H. Abrams, and P. G. Rossos, "The use of smartphones for clinical communication on internal medicine wards," Journal of Hospital Medicine, vol. 5, no. 9, pp. 553–559, 2010.

[Simon2014] M. D. Simon Kemp, "Global digital statis- tics 2014," http://wearesocial.net/blog/2014/01/ social-digital-mobile-worldwide-2014/, 2014.

[J.Pennic2013] J.Pennic, "The rising popularity of mobile health and mhealth apps," http://hitconsultant.net/2013/08/21/infographic-the-rising-popularity-of-mobile-health-mhealth-apps/2013.

[Noom 2015] Noom, "Easy to use mobile solutions. powerful behavioral change," http://us.noom.com/health/, 2015.

[Azumio2015] Azumio, "Argus, improve people's wellness through mobile," http://www.azumio.com/, 2015.

[Runtasticc2015] Runtasticc, "Reliable and on-the-go tracking for better health," https://help.runtastic.com/hc/en-us, 2015.

[RunKeeper2015] RunKeeper, "Track, measure and share fitness activites," http://runkeeper.com/apps, 2015.

[R.Zombies2015] R. Zombies, "Getfit, escape zombies and become a hero," https://www.zombiesrungame.com/, 2015.

[Digifit2015] Digifit, "Tailored to your personal goals and attitudes," http: //www.digifit.com/personal-solutions.html, 2015.

[Fitocracy2015] Fitocracy, "Track workouts, develop healthy habbits and reach goals with the help of experts," https://www.fitocracy. com/, 2015.

[Mapmyfitness2015] Mapmyfitness, "24/7 activity tracking with ios 8," http://about.mapmyfitness.com/2014/09/ updates-to-mapmyfitness-with-ios-8/, 2015.

[C.M.News2015] C.M.News, "The future of health care is in the data," http://www.cmu.edu/news/stories/archives/2015/ march/the-future-of-health-care-is-in-the-data.html, 2015.

[D.Chan2007] D. Chan, C. Kirkham, J. Coward, and M. Howard, "Myoscar, a patient controlled personal health record," in Poster presented at: North American Primary Care Research Group Annual Meeting, 2007.

[C.Health2012] C. Health and W. Project, "Knowledge2action," 2012.

[J. Butler2013] J. Butler, "Nexj healthy recipe development," 2013.

[W.Shakespeare] W. Shakespeare, "Screening for improved dance function, http://www.dancerwellnessproject.com/information/summary.aspx.".

[L.Hood2014] L. Hood and N. Price, "Promoting wellness & demystifying disease: the 100k project," Clin. Omics Innov, vol. 3, pp.20–24, 2014.

[I.R.I.] I. R. I. for Wellness and Prevention, "Digital health pilot."

[L.W.N.Y2015] L. W.N.Y. University, “Livewell nyu” http://www.nyu.edu/content/dam/nyu/takeCareNYU/documents/live well policy document 09 17 2012.pdf, 2015.

[T.U.Technology] T. U. of technology, “Health data mining.”

[TheMelon] “The melon band.”

[Banos2015] O. e. a. Banos, “Mining minds: an innovative framework for personalized health and wellness support,” in Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth 2015), 2015.

[Google2015] Google, “Google fit,” https://fit.google.com/fit/u/0/, 2015.

[Samsung 2015] Samsung, “Samsung s health,” https://us-shealth.samsung. com/websvc/intro/intro.do, 2015.

[Microsoft2015] Microsoft, “Microsoft health,” https://www.microsoft.com/microsoft-health/en-us, 2015.

[Apple2015] Apple, “Healthkit,” https://developer.apple.com/healthkit/,2015.

[Fitbit2015] Fitbit, “Fitbit,” http://www.fitbit.com/uk/setup?platform=mac, 2015.

[Sami2015] Samsung, “Sami,” https://developer.samsungsami.io/, 2015.

[O.Health2015] O. health, “Open health,” http://www.openhealth.co.uk/,2015.

Life-log Representation and Mapping

[Chen2012] L. Chen, C. Nugent, H. Wang, "A Knowledge-Driven Approach to Activity Recognition in Smart Homes," Knowledge and Data Engineering, IEEE Transactions on , vol.PP, no.99, pp.1, 0, 2012.

[Khattak2011] A. M. Khattak, P. T. H. Truc, L. X. Hung, L. T. Vinh, V. H. Dang, D. Guan, Z. Pervez, M. H. Han, S. Y. Lee and Y. K. Lee, "Towards Smart Homes Using Low Level Sensory Data", Journal of Sensors (SCIE, IF 1.77), ISSN: 1424-8220, Vol. 11, No. 12, pp.11581-11604, 2011

[Hussain2013] M. Hussain, A. M. Khattak, W. A. Khan, I. Fatima, M. B. Amin, Z. Pervez, R. Batool, M. A. Saleem, M. Afzal, M. Faheem, M. H. Saddiqi, S. Y. Lee and K. Latif, "Cloud-based Smart CDSS for chronic diseases", Health and Technology (non-SCI), 2013

[Petersen2005] A. K. Petersen and M. Mikalsen. "Context: Representation and reasoning." Revue d'Intelligence Artificielle on Applying Context-Management (2005).

[Ertel2011] W. Ertel, "First-order Predicate Logic." Introduction to Artificial Intelligence. 31-55, Springer London, 2011

Data Acquisition

[Wang06] Wang. Y, Sunderraman, R. “PDB Data Curation,” Engineering in Medicine and Biology Society, 2006. EMBS '06. 28th Annual International Conference of the IEEE

[Mark07] Mark Hedges, Adil Hasan, Tobais Blanke, “Curation and Preservation of Research Data in an iRODS Data Grid,” Third IEEE International Conference on e-Science and Grid Computing 2007

[Sophia11] Sophia Zaimidou, Sjozefvan Baal, Timothy D. Smith, Konstantinos Mitropoulos, Mila Ljujic, Dragica Radojkovic, Richard G. Cotton, and George P. Patrinos “Development of a universal,

flexible and freely available database management system for gene-centered data collection, curation and display of DNA variation", 2011 IEEE International Geoscience and Remote Sensing Symposium (IGARSS).

[Andrew09] Andrew F. Hart, Chris A. Mattmann, John J. Tran, Daniel J. Crichton, J. Steven Hughes, Heather Kincaid, Sean Kelly, "Enabling Effective Curation of Cancer Biomarker Research Data", 22nd IEEE International Symposium of Computer-Based Medical Systems, 2009. CBMS 2009.

[Cecile11] Cecile Germain-Renaud, Frederic Furst, Michel Jouvin, Gilles Kassel, Julien Nauroy, Guillaume Philippon, "The Green Computing Observatory: a data curation approach for green IT", 2011 Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing.

Low-level Context Awareness

[EuropeanCommission2010] "Social networks Overview: Current Trends and Research Challenges," European Commission, Future Media Networks Cluster, October 2010

[Prakash2009] Prakash, B.A.; Seshadri, M.; Sridharan, A.; Machiraju, S.; Faloutsos, C.; , "EigenSpokes: Surprising Patterns and Scalable Community Chipping in Large Graphs," Data Mining Workshops, 2009. ICDMW '09. IEEE International Conference on, vol., no., pp.290-295, 6-6 Dec. 2009

[White2010] White, T.; Chu, W.; Salehi-Abari, A., "Media Monitoring Using Social Networks," Social Computing (SocialCom), 2010 IEEE Second International Conference on , vol., no., pp.661,668, 20-22 Aug. 2010

[Abel2011-3] Abel F, Gao Q, Houben G, Houben K (2011) Analyzing temporal dynamics in twitter profiles for personalized recommendations in the social web. In Proceedings of the ACM WebSci '11, 3rd International Conference on Web Science, Koblenz, Germany, pp. 1-8

[Juyoung2010] Juyoung K, Hwan-Seung Y (2010) Mining Spatio-Temporal Patterns in Trajectory Data. Journal of Information Processing Systems , Vol.6, No.4, pp. 521-536

[Christopher2003] Christopher SC, Maglio PP, Cozzi A, Dom B (2003) Expertise identification using email communications. In Proceedings of the 12th international conference on Information and knowledge management (CIKM '03). ACM, pp.528-531.

[Chen2010] Chen J, Nairn R, Nelson L, Bernstein M, Chi E (2010) Short and tweet: experiments on recommending content from information streams. In Proceedings of the 28th international conference on Human factors in computing systems. ACM, pp. 1185–1194.

[Celik2011] Celik I, Abel F, Houben G (2011) Learning semantic relationships between entities in twitter. In Proceedings of the Web Engineering, pp. 167–181

[Yang2009] Yang Z, Qingquan L, Qingzhou M (2009) Mining time-dependent attractive areas and movement patterns from taxi trajectory data. In Proceedings of the 17th International Conference on Geoinformatics, pp.1-6

[Zhu2011] Zhu F (2011) Mining ship spatial trajectory patterns from AIS database for maritime surveillance. In Proceedings of the 2nd IEEE International Conference on Emergency Management and Management Sciences (ICEMMS), Beijing,China, pp. 772 - 775

[Braga2011] Braga RB. Martin H (2011) CAPTAIN: A Context-Aware system based on Personal TrAckINg. In

Proceedings of the 17th International Conference on Distributed Multimedia Systems, Florence, Italy, pp. 130-133

[Lahiri2010] Lahiri M, Tanya YB (2010) Periodic subgraph mining in dynamic networks. Journal Knowledge and Information Systems, Volume 24, Issue 3, pp. 467-497

[Amit2006] Amit AN, Siva G, Gautam D, Dipanjan C, Koustuv D, Sougata M, Anupam J (2006) On the structural properties of massive telecom call graphs: findings and implications. In Proceedings of the 15th ACM international conference on Information and knowledge management, USA, pp. 435–444

[Hayley2011] Hayley H, Yan H, Friedland G, Gatica-Perez D (2011) Estimating Dominance in Multi-Party Meetings Using Speaker Diarization. IEEE Transactions on Audio, Speech & Language Processing 19(4), pp. 847-860

[Yingjie2010] Yingjie Z, Kenneth RF, William AW (2010) Automatic Text Analysis of Values in the Enron Email Dataset: Clustering a Social Network Using the Value Patterns of Actors. In Proceedings of the 2010 43rd Hawaii International Conference on System Sciences (HICSS '10). IEEE Computer Society, Washington, DC, USA, pp. 1-10.

[Pawel2012] Pawel L, Mikolaj M (2012) Measuring the Importance of Users in a Social Network Based on Email Communication Patterns. In Proceedings of the IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining ASONAM, pp. 86-90

[Tapia2007] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, R. Friedman, Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. 11th IEEE International Symposium on Wearable Computers, 2007, pp. 37-40.

[Ayua2012] M. A. Ayua, S. A. Ismail, A. F. Abdul Matina, T. Mantorob, A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition. In proceeding of Elsevier Journal of Procedia Engineering, vol. 41, 2012, pp. 224-229.

[Garcia2012] S. Garcia, J. Derrac, J. R. Cano, F. Herrera, Prototype Selection for Nearest Neighbor Classification: Taxonomy and Empirical Study. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(3), 2012.

[Xindong2007] W. Xindong, Kumar, Vipin, R. Quinlan, J. Ghosh, Joydeep, Yang, Qiang, Motoda, Hiroshi, McLachlan, J. Geoffrey, Angus, Liu, Bing, Yu, S. Philip, Zhou, Zhi-Hua, Steinbach, Michael, Hand, J. David, S. Dan, Top 10 algorithms in data mining. Journal of Knowledge and Information Systems, vol. 14(1), 2007, pp. 1-37.

[Bao2004] Bao, Ling, and Stephen S. Intille. "Activity recognition from user-annotated acceleration data.", Pervasive Computing. Springer Berlin Heidelberg, 2004. 1-17.

[Lee2008] Kye-Hwan Lee, Joon-Hyuk Chang and Hyoung-Gon Kim, "Acoustic Environment Classification Algorithm for Context-aware Mobile Phone", Telecommunications Review, vol. 18, no. 1, 139-149p, 2008.02.

[Wang2010] Shuangquan Wang, Canfeng Chen, Jian Ma, "Accelerometer based transportation mode recognition on mobile phones", APWCS '10 Proceedings of the 2010 Asia-Pacific Conference on Wearable Computing Systems, 44-46p

[Han2012] Manhyung Han, La The Vinh, Young-Koo Lee and Sungyoung Lee, "Comprehensive Context Recognizer Based on Multimodal Sensors in a Smartphone", Journal of Sensors, vol. 12, no. 9, pp. 12588-12605, 2012

[Pourjam2013] Pourjam, E., Ide, I., Deguchi, D., & Murase, H. Segmentation of Human Instances Using Grab-cut and Active Shape Model Feedback. In proceddings of MVA2013 IAPR International Conference on Machine Vision Applications, pp. 77–80, May 20–23, 2013.

[Atefian2013] Atefian, M., & Mahdavi-Nasab, H. (2013). A Robust Mean-Shift Tracking Using GMM Background Subtraction, J. Basic. Appl. Sci. Res., vol. 3, no. 4, 596–607, 2013.

[Takur2013] Takur, R., Mehan, N., Namitakakkar. Recognition of human actions using motion history information extracted from the compressed. Int. J. Comput. Vis. Image Process. (IJCVIP), vol. 3, no. 7, pp. 973–977, 2013.

[Everts2014] Everts, I.; Gemert, J.C.V.; Gevers, T. Evaluation of color spatio-temporal interest points for human action recognition. IEEE Trans. Image Process, vol. 23, no. 4, pp. 1861–1869, 2014.

[Wang2011] Wang, H.; Klaser, A.; Schmid, C.; Liu, C.L. Action recognition by dense trajectories. In Proceedings of the 2011 IEEE Conference on. Computer Vision and Pattern Recognition (CVPR), pp. 3169–3176, 20–25 June 2011.

[Ayadi2011] Ayadi, M.E., Kamel, M.S., Karray, F.: Survey on speech emotion recognition: Features, classification schemes, and databases. Pattern Recognition 44 (3), 572 - 587 (2011).

[Xiao2007] Z. Xiao, E. Dellandrea, W. Dou, and L. Chen. Automatic hierarchical classi_cation of emotional speech. In Multimedia Workshops, 2007. ISMW '07. Ninth IEEE International Symposium on, pages 291-296, 2007.

[Lee2011] C.-C. Lee, E. Mower, C. Busso, S. Lee, and S. Narayanan. Emotion recognition using a hierarchical binary decision tree approach. Speech Commun., 53(9-10):1162-1171, Nov. 2011.

[Mao2010] Q.-R. Mao and Y.-Z. Zhan. A novel hierarchical speech emotion recognition method based on improved ddagsvm. Comput. Sci. Inf. Syst., 7(1):211-222, 2010.

[Azami2013] Azami, H., Malekzadeh, M., and Sanei, S. A New Neural Network Approach for Face Recognition based on Conjugate Gradient Algorithms and Principal Component Analysis. Journal of Mathematics and Computer Science, vol. 6, pp. 166–175, 2013.

[Ghimire2013] Ghimire, D., & Lee, J. A Robust Face Detection Method Based on Skin Color and Edges. Journal of Information Processing Systems, vol. 9, no. 1, pp. 141–156, 2013.

[Beham2013] Beham, M. P., & Roomi, S. M. M. Face recognition using appearance based approach: A literature survey. International Journal of Computer Applications, vol. 12, pp. 16–21, 2012.

[Mistry2013] Mistry, V. J., & Goyani, M. M. A literature survey on facial expression recognition using global features. Int. J. Eng. Adv. Technol, vol. 2, pp. 653–657, 2013.

[Kalita2013] Kalita, J., & Das, K. (2013). Recognition of Facial Expression Using Eigenvector Based Distributed Features and Euclidean Distance Based Decision Making Technique. International Journal of Advanced Computer Science & Applications, vol. 4, no. 2, pp. 196–202, 2013.

[Ahsan2013] Ahsan, T., Jabid, T., & Chong, U. P. Facial Expression Recognition Using Local Transitional Pattern on Gabor Filtered Facial Images. IETE Technical Review, vol. 30, no. 1, pp. 47–52, 2013.

[Kim2004] K. Kim, S. Bang, and S. Kim, “Emotion recognition system using short-term monitoring of physiological signals,” Medical and Biological Engineering and Computing, vol.42, no. 3, pp. 419-427, 2004.

[Nasoz2003] F. Nasoz, C. L. Lisetti, K. Alvarez, and N. Finkelstein, “Emotion recognition from physiological

signals for user modeling of affect," in M'2003, 9th International Conference on User Model, PA, June 2003.

[Haag2004] A. Haag, S. Goronzy, J. Williams, and P. Schaich, "Emotion recognition using bio-sensors: first steps towards an automatic system," in Tutorial and Research Workshop on Affective Dialogues, vol. 1, June 2004, pp. 36 - 48.

[Setz2009] C. Setz, J. Schumm, C. Lorenz, B. Arnrich, and G. Tröster, "Using ensemble classifier systems for handling missing data in emotion recognition from physiology: one step towards a practical system" in International Conference on Affective Computing & Intelligent Interaction. Amsterdam, Netherlands, September 2009.

[Leon2007] E. Leon, G. Clarke, V. Callaghan, and F. Sepulveda.,"A user-independent real-time emotion recognition system for software agents in domestic environments," Engineering Applications of Artificial Intelligence, vol. 20, no. 3, pp. 337-345, April 2007.

Knowledge Curation

[Regier2009] Regier, R., Gurjar, R., & Rocha, R. A. (2009). A clinical rule editor in an electronic medical record setting: development, design, and implementation. InAMIA Annual Symposium Proceedings (Vol. 2009, p. 537). American Medical Informatics Association.

[Kaljurand2008] Kaljurand, K. (2008, October). ACE View---an Ontology and Rule Editor based on Attempto Controlled English. In OWLED.

[Hulse2005] N. Hulse, R. Rocha, G. Del Fiol, R. Bradshaw, T. Hanna, and L. Roemer. Kat: a flexible xml-based knowledge authoring environment. Journal of the American Medical Informatics Association, 12(4):418-430, 2005.

Service Curation Layer

[Zazzi2010] M. Zazzi, R. Kaiser, A. Sonnerborg, et al., Prediction of response to antiretroviral therapy by huma n experts and by the EuResist data-driven expert system (the EVE study). HIV Medicine, 2010.

[Husain2010] Husain, W. and P. Lim Thean, The development of personalized wellness therapy recommender sy stem using hybrid case-based reasoning. Computer Technology and Development (ICCTD), 2010 2nd International Conference on.

[Drăgulescu2007] Drăgulescu, Doina, and Adriana Albu. "Medical Predictions System." Acta Polytechnica Hungari ca 4, no. 3 (2007): 89-101.

[Pandey2012] Dhiraj Pandey, Santosh kumar, "prediction system to support medical information system using da ta mining approach", International Journal of Engineering Research and Applications (IJERA), Vo l. 2, Issue 3, May-Jun 2012, pp.1988-1996.

[Yuan2014] Yuan, Bingchuan, and John Herbert. "Context-aware hybrid reasoning framework for pervasive he althcare. Personal and Ubiquitous Computing: April 2014, Volume 18, Issue 4, pp 865-881.

[Jannach2010] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, Recommender Systems - An Introduction. Ca mbridge University Press (CUP), 2010.

[Tintarev2007] N. Tintarev, "Explanations of recommendations," in Proceedings of the 2007 ACM Conference on R ecommender Systems (RecSys'07). Minneapolis, MN, USA: ACM, 2007, pp. 203–206.

Tintarev2007] N. Tintarev and J. Masthoff, "Effective explanations of recommendations: user-centered design," in P

roceedings of the 2007 ACM Conference on Recommender Systems (RecSys'07). Minneapolis, M N, USA: ACM, 2007, pp. 153–156.

[Fogg1999] B. J. Fogg, "Persuasive technologies," Communications of the ACM, vol. 42, no. 5, pp. 26–29, 1999.

[Grabner2003] S. Grabner-Kr¨auter and E. A. Kaluscha, "Empirical research in on-line trust: a review and critical as sessment," International Journal of Human- Computer Studies, vol. 58, no. 6, pp. 783–812, 2003.

[Fernandez2011] Fernández-Tobías, Ignacio, et al. "A generic semantic-based framework for cross-domain recomme ndation." Proceedings of the 2nd International Workshop on Information Heterogeneity and Fusio n in Recommender Systems. ACM, 2011.

Descriptive Analytics

[Glatz2014] Glatz, Eduard, et al. "Visualizing big network traffic data using frequent pattern mining and hypergraphs." Computing 96.1 (2014): 27-38.

[Gehlenborg2010] Gehlenborg, Nils, et al. "Visualization of omics data for systems biology." Nature methods 7 (2010): S56-S68.

[Bennett2009] Bennett, Jeremy S., "Massive model visualization: An investigation into spatial partitioning" (2009). Graduate Theses and Dissertations. Paper 10546.

[Dancy2008] Elizabeth Dancy, Uliyana Markova "Autonomic computing in Canadian academia"

[Dietrich2007] Dietrich, Andreas, Enrico Gobbetti, and Sung-Eui Yoon. "Massive-model rendering techniques: a tutorial." IEEE Computer Graphics and Applications27.6 (2007): 20-34.

[Shi2004] Shi, Rui, et al. "An adaptive image content representation and segmentation approach to automatic image annotation." Image and Video Retrieval. Springer Berlin Heidelberg, 2004. 545-554.

[McAdams2000] McAdams, Daniel A., and Kristin L. Wood. "Tuning parameter tolerance design: foundations, methods, and measures." Research in Engineering Design 12.3 (2000): 152-162.

[Kandogan2012] Just-in-time annotation of clusters, outliers, and trends in point-based data visualizations," Visual Analytics Science and Technology (VAST), 2012 IEEE Conference on, vol., no., pp.73,82, 14-19 Oct. 2012

[Chang2013] Chang, Ya-Ting, and Shih-Wei Sun. "A real time interactive visualization system for knowledge transfer from social media in a big data." Information, Communications and Signal Processing (ICICS) 2013 9th International Conference on. IEEE, 2013.

UI Authoring Tool

[Liberman1995] H. Lieberman, "Letizia: An Agent That Assists Media Lab, 1995, http://web.media.mit.edu/~lieber/Lieberary/Letizia/Letizia-AAAI/Letizia.html.

[Vivacqua1997] Vivacqua, H. Lieberman, and N.V. Dyke, "Let's Browse: A Collaborative Web Browsing Agent," MIT Media Lab, 1997 http://web.media.mit.edu/~lieber/Lieberary/Lets-Browse/Lets-Browse-Intro.html

[Jiang2007] He, Jiang, and I-Ling Yen. "Adaptive user interface generation for web services." e-Business Engineering, 2007. ICEBE 2007. IEEE International Conference on. IEEE, 2007.

[Patent2010] "ADAPTIVE USER INTERFACE." U.S. Patent 20,100,097,331, issued April 22, 2010.

[Matthias2013] Peissner, Matthias, and Rob Edlin-White. "User Control in Adaptive User Interfaces for Accessibility." Human-Computer Interaction–INTERACT 2013. Springer Berlin Heidelberg, 2013. 623-640.

[Ibrahim2011] Ibrahim, Razale, and Rozilawati Razali. "A performance-oriented interface design model of web applications." Electrical Engineering and Informatics (ICEEI), 2011 International Conference on. IEEE, 2011.

[Wu2012] Wu, Wen-Yi. "Adaptive user interface." U.S. Patent Application 12/334,720.

Security and Privacy

[Zeeshan2013] Zeeshan Pervez, Ammar Ahmad Awan, Asad Masood Khattak, Sungyoung Lee, and Eui-Nam Huh. Privacy-aware searching with oblivious term matching for cloud storage. The Journal of Supercomputing, 63(2):538–560, 2013.

[Zeeshan2012] Zeeshan Pervez, Asad Masood Khattak, Sungyoung Lee, Young-Koo Lee, and Eui- Nam Huh. Oblivious access control policies for cloud based data sharing systems.Computing, 94(12):915–938, 2012.

[Mahmood2013] Mahmood Ahmad, Zeeshan Pervez, YongIK Yoon, Byeong Ho Kang, and Sungy-oung Lee. Reflection: A lightweight protocol for private matching. In Signal-ImageTechnology & Internet-Based Systems (SITIS), 2013 International Conference on,pages 673–678. IEEE, 2013.

[Michel2005] Michel Abdalla, Mihir Bellare, Dario Catalano, Eike Kiltz, Tadayoshi Kohno, Tanja Lange, John Malone-Lee, Gregory Neven, Pascal Paillier, and Haixia Shi. Searchable encryption revisited: Consistency properties, relation to anonymous ibe, and extensions. In Advances in Cryptology–CRYPTO 2005, pages 205–222. Springer,2005.

[Mihir2007] Mihir Bellare, Alexandra Boldyreva, and Adam ONeill. Deterministic and efficiently searchable encryption. In Advances in Cryptology-CRYPTO 2007, pages 535–552. Springer, 2007.

[Louis2006] Louis Kruger, Somesh Jha, Eu-Jin Goh, and Dan Boneh. Secure function evaluation with ordered binary decision diagrams. In Proceedings of the 13th ACM conference on Computer and communications security, pages 410–420. ACM, 2006.

[Roberto2005] Roberto J Bayardo and Rakesh Agrawal. Data privacy through optimal k-anonymization. In Data Engineering, 2005. ICDE 2005. Proceedings. 21st International Conference on, pages 217–228. IEEE, 2005.

[Ashwin2007] Ashwin Machanavajjhala, Daniel Kifer, Johannes Gehrke, and Muthuramakrishnan Venkitasubramaniam. l-diversity: Privacy beyond k-anonymity. ACM Transactions on Knowledge Discovery from Data (TKDD), 1(1):3, 2007.