Data Curation Layer for Wellness Platforms

Hafiz Syed Muhammad Bilal
Department of computer science and engineering,
Kyung Hee University,
Yongin, South Korea,
bilalrizvi@oslab.khu.ac.kr

Sungyoung Lee
Department of computer science and engineering,
Kyung Hee University,
Yongin, South Korea,
sylee@oslab.khu.ac.kr

***Abstract*—In this paper, we present our third iteration of Data Curation Layer (DCL) for wellness platforms. DCL performs device-independent accumulation of raw sensory data from multimodal data sources in real-time. Furthermore, it curates this data with associated user context over user life-log. DCL is equipped to monitor the lifelog instances for situations of interventions. For the permanent storage of large volume of raw sensory data, DCL uses big data platform. Persisted data is aligned to be reused for Analytics and visualization. DCL encapsulates all the computational complexity at the cloud keeping the client as thin as possible.**

Keywords—Data curation; Lifelog; Wellness platform; Data acquisition; Big data

I. Introduction

With the advent of smart devices, an opportunity has emerged for the healthcare providers and biomedical researchers to enable people for taking care of their health and wellness, by providing them timely, ubiquitous, and personalized support [1]. As a result, the penetration of fitness wearables with smartphone applications and systems supporting health and wellness have taken the market by storm [2]. Despite this enormous effort by the industry and research, most of the current solutions are single device focused on the limited scope and lack interoperability and performance [3]. Considering the limitations of existing efforts as an opportunity, we have proposed and implemented a comprehensive health and wellness platform called Mining Minds [4].

The core of the mining minds platform lies in its foundations built for accumulation, persistence, and monitoring of raw sensory data. These responsibilities are undertaken by the novel layer of Data Curation (DCL). This layer is designed for high-performance acquisition of a variety of user-based raw sensory data in real-time; furthermore, persistence as well as monitoring of context on this sensory data. Unlike other healthcare and wellness platforms where sensory data acquisition is device specific, DCL facilitates mining minds to communicate with a variety of multimodal data sources registered to a particular user for the sensory data acquisition making it compatible with an IoT-based health and wellness environment. Furthermore, DCL curates the sensory data in a user-based timeline model called user life-log. Dissimilar to modern healthcare and wellness platforms, this log provides continuous monitoring abilities to DCL, so that, it can identify any anomalies where a user might require automatic support or assistance from the platform. Moreover, DCL is equipped with big data support for permanent persistence of user's sensory data. This data is further used by other layers of mining minds platform for visualization and predictive analytics.

II. Methodology

As illustrated in fig 1, DCL consists of two primary components, i.e., Sensory Data Processing and Curation, and Non-volatile Sensory Data Persistence. Within the former, sensory data acquisition supports the acquisition and synchronization of raw sensory data obtained from multimodal data sources, both in real-time (active) and offline (passive) manner. Acquired sensory data is firstly synchronized based upon the user and the timestamp of data generation, and secondly sent to ICL [4] for context identification. In response, lifelog mapping and representation component receives the identified context from ICL and curate it by mapping the context instances to a time-based log registering the detected human activities and behaviors. This log is termed as user Life-Log or simply Lifelog and persisted in the Intermediate Database for the share-ability with participating layers and applications. The stream of lifelog instances is analyzed by a monitoring component called Lifelog Monitor (LLM). It is responsible for performing time-based monitoring of different user attributes and variables, hosted in the lifelog, cross-linked with their user profiles. Furthermore, LLM supports trigger-based mechanisms to notify SCL [4] for the occurrence of an abnormal or special event associated with a given user. This mechanism is the basis for push-based recommendation generation and notification to the user by the mining minds platform. Abnormal events monitored by LLM represent risky or unhealthy behaviors and are here defined as Situation-events or situations in general. These situations are described through various constraints (-e.g., age, gender, medical conditions) and monitor-able variables (-e.g., the intensity of a particular activity and its duration). Situation events can be generated both statically at design time and dynamically at run-time by a domain expert via KCL [4].

DCL's Non-volatile Sensory Data Persistence is responsible for providing permanent and distributed big data persistence to the raw sensory data. Non-volatile sensory data persistence provides mechanisms to access the persisted data as a response to an active or a passive request generated by participating layers and applications. For online requests by SL [4] for visualization and analytics on incoming raw sensory data, the subcomponent of Active Data Reader is utilized. For the training of models used for generation of rules by KCL, Passive Data Reader subcomponent is used. LLM further utilizes these rules for its trigger-based mechanism. To create periodic backups of lifelog, the subcomponents of Lifelog Sync is used. This method provides permanent storage to the user's lifelog which can be utilized in future for human behavior analysis.

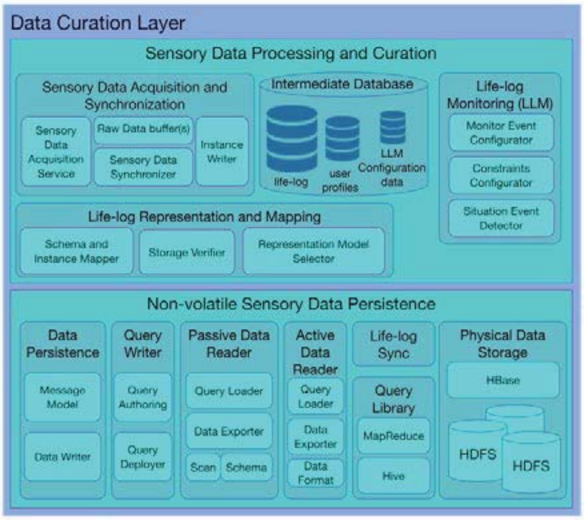

Fig. 1. Data Curation Layer (DCL) architecture

DCL is a dual-mode execution layer. Depending on the needs of the mining minds platform and its participating layers, DCL concurrently performs online and offline execution.

III. Conclusion

In this paper, we presented the core layer of mining minds platform called data curation (DCL). This layer is responsible for the acquisition of raw-sensory data in real-time. Its data source independent nature makes it more scalable and IoT compatible. DCL enables the multimodal sensory data sources to send data directly via REST service. Its logical clock synchronization enables the platform layers to identify the right context of the user at an instance of time. DCL maps the context derived by ICL from the raw-sensory data over a time-based user life-log. Furthermore, it monitors this lifelog of registered users for the detection of situations in which a push from the platform needs to be invoked to alert the user. This monitoring can integrate static, dynamic, and complex situations created by the data-driven approach of KCL. DCL incorporates multi-level abstraction on data depending upon its usage and persistence. Frequently required user lifelog and profiles data is persisted in an intermediate database hosted over an RDBMS; whereas, the historic and raw sensory data is persisted in a non-volatile storage provided by big data technologies. Active and passive data read of DCL facilitates the SL for analytics and visualization, and KCL for its model training to generate data-driven knowledge respectively.

Acknowledgment

This work was supported by the Industrial Core Technology Development Program (10049079, Develop of mining core technology exploiting personal big data) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea).

References

[1] Swan, M. Health 2050: the realization of personalized medicine through crowdsourcing, the quantified self, and the participatory biocitizen. Journal of Personalized Medicine 2012, *2(3)*, 93–118.

[2] Powell, A.C.; Landman, A.B.; Bates, D.W. In search of a few good apps. The Journal of the American Medical Association 2014, *311(18)*, 1851–1852.

[3] Amin, M.B; Banos, O.; Khan, W.A.; Bilal, H.S.M; Gong, J.; Bui, D.; Cho, S.H; Hussain, S.; Ali, T.; Akhtar, U.; Chung, T.C.; Lee, S.Y. On Curating Multimodal Sensory Data for Health and Wellness Platforms, Sensors 16(7) · July 2016.

[4] Banos, O., Bilal Amin, M., Ali Khan, W., Afzel, M., Hussain, M., Kang, B. H., Lee, S.The Mining Minds Digital Health and Wellness Framework. BioMedical Engineering OnLine, vol. 15, no. 1, pp. 165-186 (2016)