A Holistic Multipurpose Life-log Framework

Reza Rawassizadeh
Vienna University of Technology
Favoritenstrasse 9-11/188 A-1040 Vienna,
Austria
rrawassizadeh@acm.org

ABSTRACT
Life-log systems have a wide range of usages from memory augmentation to health monitoring. Recent advances in pervasive devices and sensor networks enable us to create tools that can continuously sense information from surrounding context of users and perform life logging. In this research we propose a life-log framework which is flexible to configure existing sensors and extend-able to add a new sensor or remove existing sensors. Additionally this framework provides facilities for long-term archiving, annotating and sharing life-log information. These features help users to benefit from this framework for different use cases.

Author Keywords
Life log, sensor network, personal information, personal archive.

ACM Classification Keywords
D.2.11 Software Architectures: Domain-specific architectures.

General Terms
Design, Human Factors

PROBLEM DESCRIPTION
Until the 18th century, medicine students used to study some part of *Canon of Medicine* book series in order to finish their studies and receive a doctorate degree in medicine. Nowadays, students should study huge books (such as Harrison's internal medicine) just as a "basis" for internal medicine, in addition to other books for each specific field in medicine to finish their medicinal doctorate. Likewise it might be possible that in the future medicine students might have to study more subjects such as nano-technology, software, etc. This simple example indicates that even the studying activity, which plays an important role in our life, will require more resources in the course of time and we are consistently suffering from lack of resources such as time. As a result, there is growing need to find new ways for optimizing our life, altering our lifestyle and augmenting our memory. Life-logs are supposed to be a big step toward achieving these goals and thus affecting our life style.

UbiComp '10, Sep 26-Sep 29, 2010, Copenhagen, Denmark.

BACKGROUND AND HISTORY
Wasting time, forgetting, physical and psychological health problems, and even uncluttered mind [3] are some of the challenges that we are facing in our life. Scientific achievements enable us to monitor ourselves and record our personal information, e.g. medical devices help us to stay in a healthy condition or prevent diseases. In addition to the memory augmentation, life-logs can help us to monitor our behavior by observing the past activities. In 1945, Vaneevar Bush [2] described an imaginary device (Memex), in his famous article "As We May Think", that was supposed to record all of the user's books, audio records, microfilm, etc. Memex could link and index information in order to facilitate information retrieval. Gordon Bell started to store his personal information digitally from 1998. In 2004 he and Jim Gemell launched a workshop at ACM Multimedia. They named it CARPE (Continues Archival and Retrieval of Personal Experiences). This workshop was held annually for three years. Gordon Bell is a pioneer in this area and predicted [1] that life-log systems would make dramatic changes in our life. They are going to be widely used in the near future and we will adopt them in our life like mobile phones and Internet. These tools can benefits us in both personal and social domains.

RESEARCH METHODOLOGY
This dissertation is built based on a novel platform that provides a set of open and extendable life-log tools, which are able to archive life-log information in long-term [19], enable users to share their information, and consider the security and privacy aspect of end users. The dissertation is divided into three parts. In the first part we define a conceptual data-model which is technology independent and extendable enough to enable users to add or remove sensors to/from the framework. Moreover by using this data model, users are able to apply security and privacy policies on any single information object in their life-log dataset. This is due to the fact that each life event (life-log record) has its own access scope. In the second part a sensor classification and an open generic framework for life-log systems will be proposed. For each sensor class, based on the generic framework, a prototype will be provided. In the third part we discuss different evaluation methods to evaluate prototypes and the proposed model. Evaluation methods are the implementation of some use-cases that use life-log dataset. Another evaluation ap-

proach is to assess the quality of prototypes by benchmarking the resource usage of a prototype [16]. Third evaluation method is to study controversial issues of these systems such as anonymization of the life-log dataset or security considerations in life-logs [18]. We have discussed ethical and privacy related issues of sharing the life-log information with society, in detail, in a submitted paper [17] (It is still under the review).

Since there are lots of sensors available and more sensors will enter into the market in the future, the project aims to provide an extendable framework which can accept any kind of sensors. The more sensors information in the dataset, the merrier the quality of the dataset can be achieved. It might be argued that the term sensor is not appropriate when we are monitoring desktop activities. However we refer to the term sensor for any component that sense information and can get plugged or unplugged to/from the life-log framework.

Data Model Definition: We are living in a spatio-temporal world. Meaning all of our life events, except dreams, happen in a specific location and at a specific date-time. Based on the current available technologies it is not always possible to sense the location, because location sensors such as GPS are not functional in every environment. For instance GPS can not work indoor. There are other approaches such as A-GPS (Assisted GPS) to solve this problem, but they can not always sense the location and they are imprecise. On the other hand most operating systems have date-time which is accessible since the target device has not been turned off. This means most devices with computing capabilities can provide timestamp. Therefore we conclude that date-time is a necessary field for any life-log record and all life-log information objects will be stored with the timestamp.

Life-log data types vary based on the sensor output. Data can be a text or a binary object such as a movie or a picture. We assume all records of a life-log dataset have a date-time without considering the data-type of the record.

Each life event is a data entity or record in the dataset. Each life-log data entity contains Information object, timestamp, annotation and social scope. Timestamp can be continuous or discrete; if continuous, it will be the start timestamp and the end timestamp. Information object can be binary data e.g. image, audio or textual data e.g. GPS location, micro blog content, etc. We defined three access scopes, *Private*, *Public* and *Friend*. Gross et al. [12] identified these three scopes and named them, private, semi-public and open-ended. Private means that the data entity is not shared and nobody other than the owner can access this information object. Public means that everybody in this social domain can access this information object and no access limitation has been defined for it. Friend defines users who can access this information object. The user can define which user(s) or group of users from the social domain can access this object. A finite set of users or group of users can access users' information objects. In respect to security and privacy a shared information object should contain an expiration timestamp and access scope. Access scope is the list of friends who can access this information object (if it is not public or private) and expiration date used to disable any access to that information object after the specified timestamp. A social relation description models such as FOAF [1] can be used to extract the list of users or the users who can see this data entity or it can be specified manually by the user. How to extract the user list is not in the scope of this research. Figure 1, shows a data object that has been created at time t0 by User A and then shared at time t1 with User B and a group of Users (User Group C). Access for User B to data object A will be expired at time t3 and the users in Group C can not access this object after time t5.

Sesnor Classification: Life-log sensors have been classified

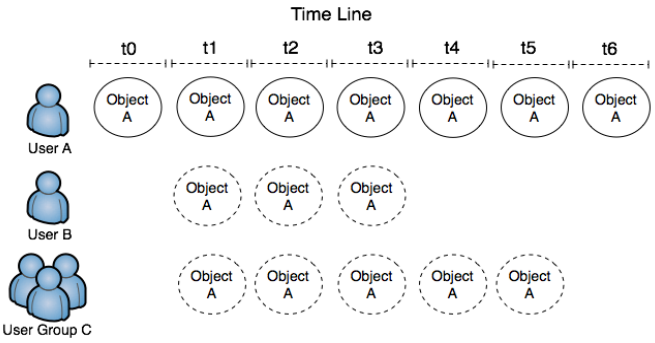

Figure 1. Access expiration definition for a shareable information object.

based on their physical proximity in three different classes: *Mobile Sensors*, *Proximity Sensors* and *Desktop Sensors*. The border between sensor classes is not clear. For instance a weather sensor may appear in both desktop and mobile classes. Furthermore a device can host sensors from different classes. Mobile sensors such as mobile phone and bio-sensors are always carried by users. Proximity sensors such as a wall mounted camera, an activity monitor on an exercise device, etc. reside in the surrounding environment of users. Desktop sensors are applications or sensors that record users' activities on personal computer such as monitoring visited websites, chat, etc. In order to evaluate the classification a prototype implementation (Data Reader) will be created for each sensor and then the prototype is evaluated based on the life-log use-cases, application quality and controversial issues such as privacy.

We have built a mobile phone based life-log tool which provides some ubiquitous sensors. It is important to note that mobile sensors are very appropriate for life logging, are always carried by users.

Framework Architecture: The technical architecture of the framework is composed of two main components, "Data Reader" and "Reliable Storage Media" or "Server". Data Reader contains sensors or it resides near external sensors. "Reliable Storage Media" (RSM) hosts the information which has been gathered from Data Readers. Information flows to the life-log dataset from multiple sensors in a continuous manner. Sensors information have different data formats and there is no guarantee for the sequence of information and the sequence varies over the time. These properties indicate that a life-log dataset is a data stream. In order to retrieve information from this kind of dataset an annotation mechanism

[1] http://www.foaf-project.org

is required. Annotation can be done both by Data Readers or RSM. If the Data Reader is a pervasive device then it can not host large processes. Since annotation could be a process-intensive task, we need to hand over annotations to the RSM.
We have created a Data Reader for mobile phones that its architecture has been inspired by context-aware systems. This life-log tool has been being used for about seven month and it sends data to a personal computer which plays the role of RSM. As the next step, we are working on an additional data reader which is a semi-intelligent wall mounted camera. At the very end some datasets will be created that contain rich information about life-log sensors.

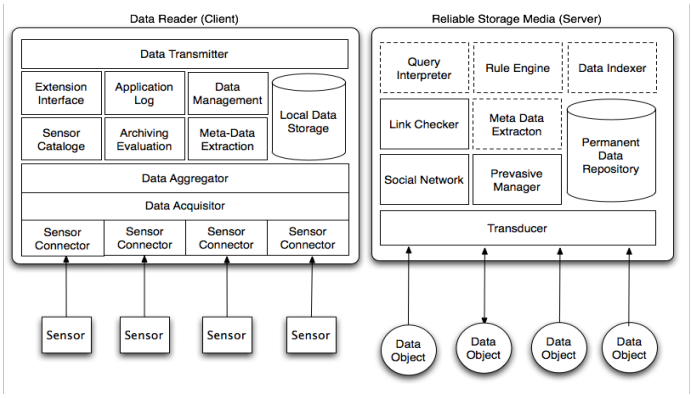

Figure 2. Architecture of the Framework. The right figure depicts the RSM architecture and the left figure shows the Data Reader architecture which resides near sensors.

CONTRIBUTIONS

Followings are the main contributions of this research:
Design of a Life-log Framework:
This part focuses on design of a flexible and extensible life log framework which will sense and record individuals life events. Our life-log framework contains a data model and some prototype implementations for different kind of sensor classes. This framework will be open to add and configure sensors. Openness is one of the major novelties of this framework, because it might assists users in better memory augmentation, multipurpose use-cases and better user modeling. It is important to highlight the differences between life-log tools and Context-aware tools. Context-aware tools are not designed for life logging purposes. Information is being sent to context-aware systems manually. On the other hands, life-logs are those types of applications that should run 24/7 in the background of the hosting device. Furthermore, life-log dataset needs to be archived for long term, shared with society while considering privacy and annotation. Despite all similarities, these facts lead us to conclude that context-aware tools are not enough for life logging purposes.
Annotation and Digital Preservation:
Life-log data are worth to maintain at least during the life of its owner. Users can allow their heirs to access their life-log dataset, which can be used for scientific purposes, historical studies or even a memorial for the family. In order to make life-log information accessible in long term, life-log tools should store information in appropriate data formats and in case sensors do not provide a long-term archive-able format, Data reader or RSM should be able to change the format to a long-term archive-able one. Hardware preservation is not in the scope of this dissertation.
In order to enable users reflecting on their life-log information, they should be able to search and browse their information. Data which will be acquired by sensors is in raw format and this indicates the need for annotation. Annotations are very important to facilitate further information retrieval, especially for life-log datasets which are a kind of datastream.
Sharing life-log information with Society:
Sharing life-log information with society will benefit both users and society in historical studies, social patterns analysis, group behavior learning [8], matchmaking, recommendation systems, health and medical studies and sousveillance [17]. Although social medias enable users to define share limitation for their information, because of the life-log data structure, current sharing models are not able to handle life-log information while maintaining users privacy. We perform risks and benefit assessment for sharing life-log information with society and based on the identified risks we propose a data model for sharing life-log information.
Security, Privacy and Ethical Issues:
Lior J. Strahilevitz [20] claimed most private information consists of sensitive personal matters such as sexual encounters and bodily functions, sensitive medial information, knowledge of owners fundamental weaknesses, etc. Life-log tools can sense and record this type of information, and therefore in terms of individual privacy, life-log dataset is considered as a very sensitive object. We will discuss and propose methods that reduce risks of sharing life-log information with society and reduce security related risks. These methods might suggest ethical considerations to service providers and service consumers, anonymization and pseudonimization approaches that could be performed on the life-log dataset, and securing the sensing and recording processes of life-log systems [18].

RELATED WORK

As related works we can refer to more than only life-log systems. There are tools such as PIM (Personal Information Management) systems or context-aware tools which are not called life-log systems but similar to life-logs, they read and record information. iMemex [6], UbiFit [5], ContextPhone [15], MyExperience [9] are some examples of those systems which are not life-log but they read and record information similar to life-log systems. MyLifeBits [10] is the largest scientific effort toward providing a life-log system that can record desktop activities of users and uses SenceCam [13] as a body mounted camera to capture pictures from users' environment. Nokia LifeBlog [14] is a life-log application that runs on the S60 series of Nokia phones. Reality Mining [7] uses a mobile phone based life-log system to study social behavior of a group of users. To our knowledge this is the only project that considers social aspects of life-log systems. Brian Patrick Clarkson proposed "I Sensed" [4] which uses a wearable microphone and a video recording apparatus to record his contextual data in order to learn human's

life pattern. iRemember [21] uses an external microphone on a PDA to record audio communications of users. Audio data stayed temporary on the PDA, and then they were sent to a large capacity server. Every Sense of life [11] monitors users' health status by recording continuously their health factors with bio sensors such as tools to measure heart beats, body temperature galvanic skin responses, etc.

ROAD MAP

As it has been explained before, for each sensor class, an implementation will be proposed and each implementation will be evaluated. On the other hand we need to take important requirements and capabilities of life-log tools into account. These include sharing life-log with society, digital preservation, privacy of the life-log dataset, annotating life-log information and securing the life logging process. Mobile sensors, which are the most important sensors, have been implemented, and now we are in the process of submitting a paper for this implementation. A datamodel for securing the life-log information and other technical considerations has been proposed [17]. We have briefly discussed digital preservation of pervasive device information [19], and now we are about to submit another paper which describe our digital preservation tool in more details. Privacy aspects of sharing life-log information and security related issues about sharing this information have been described in another paper [18]. Creating a life-log tool for proximity sensor is an open topic which is in progress. Annotation and anonymization of life-log datasets are other open topics which require more investigation.

REFERENCES

1. G. Bell and J. Gemmell. *Total Recall: How the E-Memory Revolution Will Change Everything*. Dutton Adult, 2009.
2. V. Bush. As we may think. *The Atlantic*, 1945.
3. N. Carr. Is Google making us stupid? *The Atlantic*, 302(1):56–63, 2008.
4. B. P. Clarkson. *Life Patterns: Structures of wearable sensors*. PhD thesis, Massachusetts Institute of Technology. Deptartment of Electrical Engineering and Computer Science, 2002.
5. S. Consolvo, D. W. McDonald, T. Toscos, M. Y. Chen, J. Froehlich, B. L. Harrison, P. V. Klasnja, A. LaMarca, L. LeGrand, R. Libby, I. E. Smith, and J. A. Landay. Activity Sensing in the Wild: a Field Trial of Ubifit Garden. In *CHI '08: Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems*, pages 1797–1806, 2008.
6. J. P. Dittrich, M. A. V. Salles, D. Kossmann, and L. Blunschi. iMeMex: Escapes from the Personal Information Jungle. In *VLDB '05: Proceedings of the 31st International Conference on Very Large Data Bases*, pages 1306–1309, 2005.
7. N. Eagle and A. Pentland. Social Serendipity: Mobilizing Social Software. *IEEE Pervasive Computing*, 4(2):28–34, 2005.
8. N. Eagle and A. S. Pentland. Reality Mining: Sensing Complex Social Systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.
9. J. Froehlich, M. Y. Chen, S. Consolvo, B. Harrison, and J. A. Landay. MyExperience: A System for In situ Tracing and Capturing of User Feedback on Mobile Phones. In *MobiSys*, pages 57–70. ACM, 2007.
10. J. Gemmell, R. Lueder, and G. Bell. The MyLifeBits Lifetime Store. In *ETP '03: Proceedings of the 2003 ACM SIGMM workshop on Experiential telepresence*, pages 80–83, 2003.
11. V. Gerasimov. *EVERY SIGN OF LIFE*. PhD thesis, Massachusetts Institute of Technology, 2003.
12. R. Gross, A. Acquisti, and H. Heinz III. Information Revelation and Privacy in Online Social Networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society*, pages 80–91, 2005.
13. S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood. SenseCam: A Retrospective Memory Aid. *UbiComp 2006: Ubiquitous Computing*, pages 177–193, 2006.
14. A. Myka. Nokia lifeblog - Towards a Truly Personal Multimedia Information System. In *Proc. of Workshop des GI-Arbeitkreises "Mobile Datenbanken and Informationsysteme*, 2005.
15. M. Raento, A. Oulasvirta, R. Petit, and H. Toivonen. ContextPhone: A Prototyping Platform for Context-Aware Mobile Applications. *IEEE Pervasive Computing*, 4(2):51–59, 2005.
16. R. Rawassizadeh. Mobile Application Benchmarking Based on the Resource Usage Monitoring. *International Journal of Mobile Computing and Multimedia Communications*, 1(4):64–75, 2009.
17. R. Rawassizadeh. Toward Sharing Life-log Information with Society. In Submission Process, 2010.
18. R. Rawassizadeh and A. M. Tjoa. Securing shareable life-logs. In *PASSAT 2010: First International Workshop on Privacy Aspects of Social Web and Cloud Computing (PASWeb-2010)*, 2010.
19. R. Rawassizadeh and M. Tomitsch. Towards Digital Preservation of Pervasive Device Information. In *CHI 2010 Extended Abstracts, Workshop on Know Thyself: Monitoring and Reflecting on Facets of One's Life*, 2010.
20. L. Strahilevitz. A Social Networks Theory of Privacy. *SSRN eLibrary*, 2004.
21. S. Vemuri, C. Schmandt, and W. Bender. iRemember: a Personal, Long-term Memory Prosthesis. In *CARPE '06: Proceedings of the 3rd ACM workshop on Continuous archival and retrival of personal experences*, pages 65–74, 2006.