



**Institute for a
Broadband-Enabled Society**

The Informatics of Personal Data Management for Health and Fitness



May 2013

Authors

Manal Almalki, Fernando Martin-Sanchez, Kathleen Gray
Health and Biomedical Informatics Centre, University of Melbourne

Acknowledgements

This paper was written as part of a project exploring the concept of Self-Quantification in health and fitness applications and systems. Support for the project, *Self-omics: addressing the information and communication needs of the 'quantified individual' for enabling participatory and personalized medicine*, was provided by the Institute for a Broadband-Enabled Society (IBES).

Further Information

contact@broadband.unimelb.edu.au

Institute for a Broadband-Enabled Society

Level 4, Building 193
The University of Melbourne, Victoria 3010, Australia

ISBN 978 0 7340 4831 8

© The University of Melbourne 2013

This work is copyright. Apart from any use as permitted under the *Copyright Act 1968* (Cth), no part may be produced by any process without prior written permission from the University of Melbourne.

Executive Summary

With advances in Self-Quantification applications and systems, it is now possible to capture and record data about nearly all aspects of human health and fitness, including mental, emotional, physical, social and spiritual dimensions. By analysing these numbers, people have a better understanding of their health status and their relationship to the world around them. Furthermore, huge advances in sensor technology – in conjunction with widespread availability of wireless networks – have helped self-trackers to collect data whenever and wherever they want.

The amount of data that is being captured is growing at exponential rates. This large volume of data needs to be taken into consideration by health and biomedical informaticians, as such data are difficult to manage in the context of organising, accessing, using, sharing, and analysing in aggregate form. There are clear implications for the use of high capacity broadband to transmit health data.

This report aims to summarise the present state-of-the-art of Self-Quantification in health and fitness applications. The report begins by providing a classification of selected tools and data flows in Self-Quantification systems. It also identifies key directories with more extensive examples of tools currently available for public use. Next, it highlights Open mHealth and Health Level 7 (HL7) standards for dealing with the problem of data isolation. Finally, it profiles three types of big-data analytical tools. The report concludes with a summary of the main challenges facing Self-Quantification systems, and offers some possible solutions.

Contents

Executive Summary.....	2
Contents.....	2
1 Introduction.....	3
1.1 Classification of Self-Quantification Systems.....	4
2 Primary Self-Quantification Systems.....	5
2.1 Typology of the Primary Self-Quantification Systems.....	5
2.2 Examples of Primary Self-Quantification Systems.....	7
2.3 Informatics Aspects of Primary Self-Quantification Systems.....	23
3 Secondary Self-Quantification Systems.....	26
3.1 Typology of Secondary Self-Quantification Systems.....	26
3.2 Examples of Secondary Self-Quantification Systems.....	27
3.3 Informatics Aspects of Secondary Self-Quantification Systems.....	30
4 Directories of Apps.....	31
4.1 Quantified Self Guide.....	31
4.2 Happtique.....	31
4.3 European Directory of Health Apps 2012-2013.....	33
5 Interoperability Standards in Self-Quantification Systems.....	34
5.1 The Current Status of Mobile Health Applications.....	34
5.2 Stovepipe Architecture vs Open mHealth Architecture.....	34
5.3 Health Level 7 (HL7).....	41
5.4 Data Storage Locations in mHealth Systems.....	43
6 Big Data Analytics.....	46
6.1 General Data Analytics.....	47
6.2 Data Intensive Processing and Analysis.....	49
6.3 Online analytics Tools.....	51
7 Conclusion.....	53
8 References.....	54
8.1 Image Credits.....	56
9 Appendix.....	57

1 Introduction

In recent years the general public has become more health-conscious, due in part to network and sensor technologies that enable the non-expert to easily capture and share significant health-related information on a daily basis. As these technologies have become more widely available, they have given rise to a concept called the Quantified Self. The Quantified Self movement is concerned with capturing, recording and sharing personal health data. A smartphone or mobile device is typically used to track these health metrics; the collected readings are then communicated to friends, carers or healthcare professionals.

The practice of Self-Quantification has been associated with health and fitness maintenance. Most Self-Quantification applications are able to simultaneously track several factors that could be associated with a certain disease or health condition, such as weight and ambient temperature, and correlate these data with other metrics such as how many steps have been taken in a day. However, the capacity to track such a wide range of data does not necessarily translate into a coherent message for the consumer. Many of the data are scattered and do not make sense.

This report explores this issue. It also illustrates the data flow stages in various types of Self-Quantification tools, and sheds light on some user experience issues and technical problems (e.g., data integration) with these tools. The report offers a classification of self-tracking devices based on our study of many Self-Quantification tools, and provides a brief introduction to common interoperability standards for addressing the issue of isolated self-tracking systems. The final section of the report explores various options for analysing and deriving information from the raw data generated by self-quantifying systems.

The report is organised as follows:

- The first section provides a classification of self-tracking tools, illustrates the data flow stages in several Self-Quantification applications, and explains how typical self-tracking systems work.
- The second section offers a comparison between Open mHealth and HL7 standards for dealing with the issue of data isolation.
- The final section describes three types of big-data analytical tools, and provides examples of each.

1.1 Classification of Self-Quantification Systems

Quantified-self systems can be classified into two groups: primary quantified-self systems and secondary quantified-self systems.

The primary quantified-self systems can be described as a single tool or app for collecting one-to-several health-related metrics (see section 2). For example, Fitbit and Zeo are primary quantified-self systems.

The secondary quantified-self systems can be described as a single tool or app for aggregating or integrating the collected data by a primary QS tool (see section 3). For example, Wikilife and Digifit are secondary quantified-self systems.

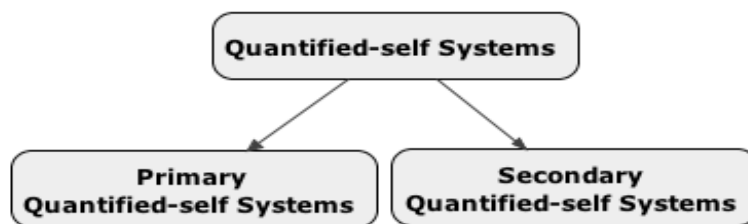


Figure 1: Primary Self-Quantification Systems classification

2 Primary Self-Quantification Systems

The primary Self-Quantification systems consist of a tool (e.g., Fitbit) or sensor (e.g., Zeo headband) for capturing and recording personal health data, and an app (e.g., Zeo Sleep Manager) for analysing, visualising and sharing the collected data. This section provides an analysis and some examples of primary Self-Quantification systems.

2.1 Typology of the Primary Self-Quantification Systems

The primary Self-Quantification systems can be classified into two groups: mobile and fixed. This classification is based on the location of the sensor(s) used by the system.

2.1.1 *Mobile and Fixed Self-Quantification Systems*

Taking into account the sensor's location, Self-Quantification systems can be classified into two categories: mobile and fixed systems. In mobile Self-Quantification, the sensor is collecting data while it is installed on a moving object such as a person (e.g., wearable sensors), vehicle, etc. In fixed Self-Quantification, the sensor is collecting data while it is installed in a fixed place such as an office, home, clinic, etc.

2.1.2 *Fixed Self-Quantification Systems*

In addition, taking into account the data types captured in Self-Quantification applications, fixed Self-Quantification systems can be divided into two groups: environmental sensor tools, and touchless sensor tools. Environmental sensor tools are concerned with measuring and recording environmental and climate factors such as temperature, precipitation rates, humidity, pollution percentage, etc. On the other hand, touchless sensors are concerned with taking unobtrusive measurements of the user's biomedical signals and activities. For example, a sensor can be attached to a user's bed for measuring ECG signals, weight, body movement, and snoring during sleep (Choi, Choi, Seo, Sohn, Ryu, Yi & Park, 2004). The captured data from environmental sensors can be correlated with users' other personal data to provide a comprehensive view of their health status.

2.1.3 *Mobile Self-Quantification Systems*

Mobile Self-Quantification systems can also be partitioned into two groups: invasive sensors (in-contact sensors) and non-invasive sensors (on-body sensors). This classification is based on whether the measurement tool pierces the skin. Invasive sensors involve tools for frequent pricking of the skin by patients – for example, taking blood samples in glucose testing (Vashist, 2012). Implantable sensors (e.g., insulin pump) and swallowable sensors (e.g., swallowable pill for sensing a biological condition within a body) also can be considered as invasive sensors. On the other hand, non-invasive sensors do not pierce the skin when collecting measurements. Wearable sensors are considered of this category.

2.1.4 Standalone, Smartphone, or Hybrid Self-Quantification Systems

Invasive and non-invasive sensors can be further classified into one of three groups of systems based on the information-processing unit type into: standalone, smartphone or hybrid systems (Gupta & Jilla, 2011). The information-processing unit could be a smartphone or computer.

2.1.4.1 Standalone Self-Quantification Systems

Standalone tools capture and display data in real time and store it in internal memory. The user can see the measurements on the tool's screen. In case the sensor is separated from the information-processing unit, the captured data by the sensor are sent automatically/wirelessly or manually/wired to the information-processing unit. Examples of on-body standalone tools are Garmin, Suunto, and Fitbit, whereas the example of in-body tool is Thermometer.

2.1.4.2 Smartphone Self-Quantification Systems

In the second group, data are captured by smartphone-based applications, using built-in device capabilities such as the phone's camera, Global Positioning System (GPS) and accelerators to sense the user's motion, and keyboard for entering data (e.g. in the Moodpanda app). Data are displayed on the phone's screen through the app's interface. The iOS and Android app stores have thousands of smartphone-based applications; iTreadMill, RunKeeper, and Endomondo are examples of on-body smartphone-based application that keep track of walking, running, cycling, and other physical activities. On the other hand, using an insulin pump with an app for managing diabetes is an example of in-body smartphone-based application.

2.1.4.3 Hybrid Self-Quantification Systems

In hybrid systems, external sensors capture data and sync with a smartphone and/or computer. With these types of sensors, users mostly have no way to view all of their aggregated data except via this secondary device (smartphone or computer). Nike+, and Adidas MiCoach are examples of on-body hybrid systems. These systems are taking advantage of smartphone capabilities and the accuracy of external sensors (e.g., pedometers). On the other hand, iBGstar tool and app for managing diabetes is an example of in-body smartphone-based application.

As noted earlier, standalone sensors are supplied with a tiny screen for displaying the taken measurements. However, the difference here is that in the hybrid system, the secondary device is essential for presenting all the collected data altogether. For example, Fitbit allows the user to see some measurements on its small screen; while as, Garmin watch can display all collected data on its screen, including heart rate and pulse measurements. Thus, there is no need for a smartphone or a computer for presenting the readings. Most of these tools have a web-based application for visualising the recorded measurements (e.g., pie chart, scatter chart, table, line chart, area chart, etc.).

It has been noted that smartphones are becoming more widely used than computers (HIMSS, 2012). This high rate of adoption is due to several factors, including their greater connectivity and mobility than personal computers. However, Turisco and Garzone (2011) believe that ease of use of smartphone-based app is the leading reason for smartphone adoption.

In summary, standalone systems, smartphones and hybrid systems fall into the mobile Self-Quantification category. Environmental sensor tools and touchless sensors fall into the fixed Self-Quantification category. The following diagram illustrates this classification.

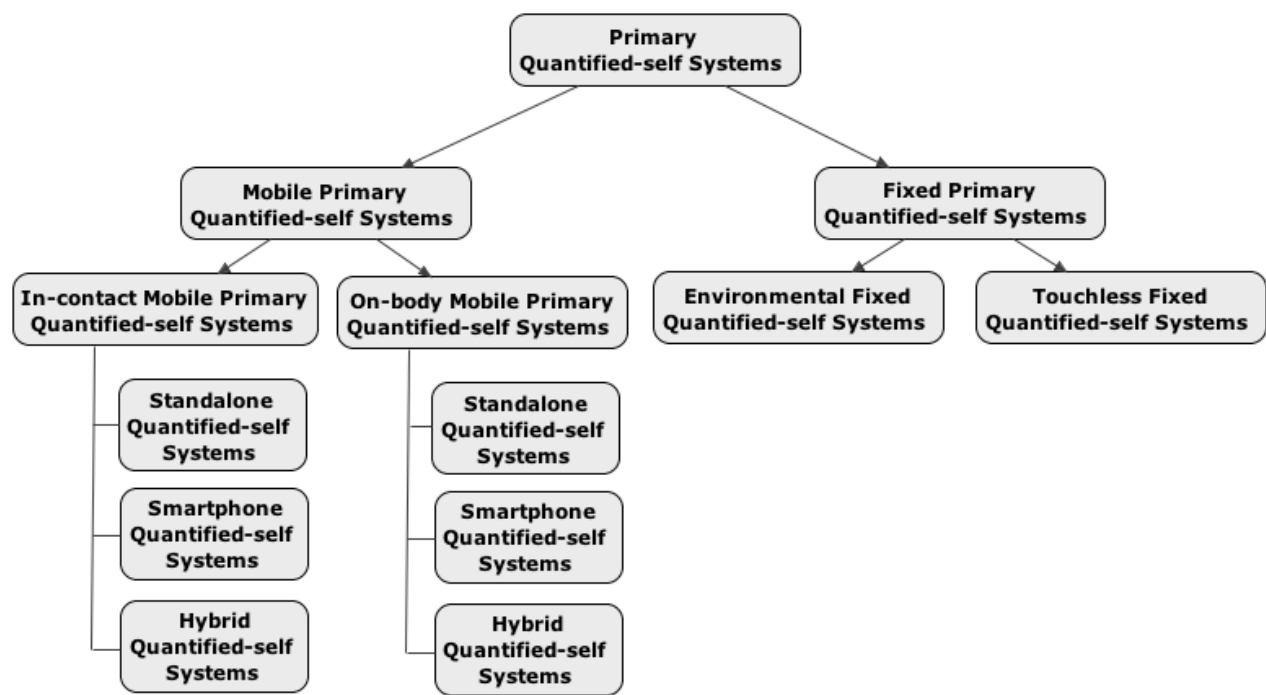


Figure 2: Primary Self-Quantification Systems classification

2.2 Examples of Primary Self-Quantification Systems

This section provides several examples of primary Self-Quantification systems, and includes detailed analysis of each system's data flow. Also, challenges in some Self-Quantification systems are illustrated.

2.2.1 Zeo Sleep Manager

Zeo Sleep Manager tracks information on the amount of hours slept and different sleep states. It measures the brain’s electrical signals and provides a quantitative sleep quality value called the “Z score” (Chang, 2012). The measured signals are used to indicate the four different stages of sleep (REM, deep sleep, light sleep, and waking) (Zeo, 2012).

2.2.1.1 Classification of Zeo Systems

Zeo Inc. offers both standalone and hybrid self-tracking systems. Within its product range, the Zeo Sleep Manager is an example of a standalone system. It consists of a headband paired with a bedside-clock device that displays the collected data, indicating which stage of sleep the user is currently experiencing. In the hybrid category, the Zeo Sleep Manager Pro consists of a wireless headband that transmits data to a smartphone (e.g., iPhone or Android). The smartphone is used for displaying and sharing the collected data.

2.2.1.2 Challenge in Zeo Systems

Based on user reviews, some issues have been experienced with Zeo systems. The main issue is that interpreting the generated data is confusing. “Customers are requesting more coaching support. Some Zeo users are taking their charts to their doctors, but their doctors are unable to offer much interpretation or recommendation based on such data” (Mehta, 2011).

Zeo Sleep Manager	Classification	Picture
Zeo wireless headband and bedside display	Standalone tool	 <p data-bbox="991 972 1401 1023">Image 1: Zeo bedside display and wireless headband</p>
Zeo wireless headband and Zeo Sleep Manager application	Hybrid tool	 <p data-bbox="991 1693 1378 1722">Image 2: Zeo Sleep Manager application</p>

Table 1: Zeo Sleep Manager classification

2.2.1.3 Data Flow Stages in Zeo Sleep Manager

The following table demonstrates the data flow stages in Zeo systems.

Data Flow Stage	Methods	
Data type	Electrical signals from the brain	
Collect/aggregate data	Collecting method	Wearable sensor (headband)
	How	Sleep data are captured via a lightweight headband, worn while sleeping
Transmitting data type	SD memory card with USB adapter Bluetooth, up to 25 feet	
	<p>In the case of standalone Zeo systems, data are sent from the headband to bedside-clock device where it is saved temporarily in SD card memory. Data are then transmitted through a USB adapter to the user's computer.</p> <p>In the case of hybrid Zeo systems, data are sent to the paired smartphone or computer via Bluetooth and temporarily saved. Once the user logs on to their account, data are synced with the smartphone or PC. Internet access and user login are required to send data to Zeo's server.</p>	
Saving data (temporary storage)	SD memory card on the bedside-clock device, or the smartphone's memory storage	
Analysing data	<p>Analysing tools are available in Zeo's mobile app and at mysleep.myezo.com. For example, ZQ Sleep Score summarises the sleep quality in a single objective number.</p> <p>Then, an expert sleep-coaching program is provided for users based on their quality of sleep.</p>	
Visualising data	<p>Bedside-clock device</p> <p>Zeo's mobile app to illustrate sleep patterns in charts</p> <p>Zeo's web interface</p>	
Storing data (permanent storage)	Internal flash storage (temporary storage), then the user's computer or Zeo's server	
Sharing data	Users can share the collected data by using Zeo's mobile app or Zeo's web interface at mysleep.myezo.com	

Table 2: Data flow stages in Zeo Sleep Manager

2.2.2 Fitbit

Fitbit tracks movement, showing the exact steps taken, stairs climbed, distance travelled, and calories burned (Fitbit, 2012). It also can track hours of sleep. It consists of a wearable sensor, a base station attached to a PC or Mac, and a web-based application used to record, visualise, and analyse collected data.

2.2.2.1 Classification of Fitbit Tool

According to our classification, Fitbit can be either a standalone or hybrid tool. As a standalone tool, the Fitbit tracker can be used exclusively for collecting and displaying data. The data are automatically uploaded to the user's computer via the tool's base station, which is connected via USB to the user's computer. The uploading happens wirelessly whenever the tracking tool comes within range (approx. 15 feet) of the base station.

Fitbit can also be used as a hybrid tool, if the user pairs the tracker tool with a smartphone that displays the collected data. Fitbit Inc. currently offers apps for both iOS and Android that allow users to log activities like walking, yoga, or weight lifting. However, activities measured by the tracking tool, such as steps taken, must still be uploaded by the Fitbit base station attached to the user's computer (Fitbit, 2012).

2.2.2.2 Challenges in Fitbit

User reviews report some issues with the Fitbit system. These issues are:

Data should be presented together: People who are interested in tracking their health status tend to track many factors that could influence their health, such as heart rate, calories burned, sleep quality, eating habits, etc. Therefore, people are using different tracking tools to capture all of this information. The main problem with this approach is that people are experiencing confusion in exploring multiple types of data and difficulty in understanding the influences of various factors on their health. Even when data are stored in the same tool, users have to look at different graphs separately. For example, a participant in a field study (Li, Dey & Forlizzi, 2011) used Fitbit for both sleep and physical activity tracking, so she was able to explore both types of data together. However, she also collected other types of data using Daytum and your.flowingdata, which she could not easily review along with her Fitbit data.

There is no data sharing: Users complain that they need to go to different applications and/or websites to answer their questions because the current Self-Quantification systems do not share data with other systems. Users express the wish that they could explore their data in a single interface (Li, Dey & Forlizzi, 2011).

Finding relationships and correlations in various collected data is challenging: As data are collected from multiple sources and not presented in one place, finding the relationships between different kinds of personal data is difficult. For example, a participant in a field study (Li, Dey & Forlizzi, 2011) used Fitbit and Zeo to automatically record his physical activity and sleep quality. He used the data from these tools to figure out the relationship between his physical activity and sleep, and his blood sugar level. As there are no means for determining the correlation between these three factors, some

people use written notes (on paper or online) to remind themselves of important events that happened at a particular time.

2.2.2.3 Fitbit Partners

To overcome the above-mentioned issues, Fitbit and other popular fitness programs such as Loselt!, RunKeeper, Microsoft HealthVault, and Zeo recently partnered. This can be seen as a positive movement because people now can use different tools to collect and share data (Fitbit, 2012). However, it does not go far enough as many tools are still not interoperable (Li, Dey & Forlizzi, 2011).


Fitbit	Classification	Picture
Fitbit.com and tracker (clip)	Hybrid tool	 Image 3: FitBit.com and tracker clip

Table 3: Fitbit classification

2.2.2.4 Data Flow Stages in Fitbit

The following table demonstrates the data flow in Fitbit tools.

Data Flow Stage	Methods	
Data type	Steps taken, calories burned, distance travelled, and hours of sleep	
Collect/aggregate data	Collecting method	Wearable sensor (tracker) contains accelerometer (tracking activity), altimeter (tracking stairs climbed)
	How	Tracker can be attached to the user's pocket or wrist. It captures data while the user is moving.
Transmitting data type	WiFi (2.4Ghz radio frequency), data are sent from the tracker when it comes within 15 feet of the base station plugged into Mac or PC.	
	Data sync automatically with the smartphone or PC once the tracker is within range of the base station. Internet access and user login are required to send data to Fitbit's server. Alternatively, the user can plug the Fitbit tracker directly into the computer to upload data.	
Saving data (temporary storage)	Internal memory storage on the Fitbit tracker.	
Analysing data	There are no analysing tools. Activity status updates are only illustrated in charts.	
Visualising data	Fitbit tracker has a small OLED screen to display measurements. Fitbit's mobile app illustrates daily or historical collected data in different charts, and provides a percentage-based view of how much the users have achieved of their goals. Web-based application at Fitbit.com .	
Storing data (permanent storage)	Data are stored at Fitbit.com . Data can also be stored on the user's PC as an XML file.	
Sharing data	Users can share the collected data by using: Fitbit's mobile website at m.fitbit.com Fitbit's web-based application at fitbit.com	

Table 4: Data flow stages in Fitbit

2.2.3 Fitlinxx Actipressure

Actipressure tracks blood pressure readings. Users can see the history of their collected data on activePressure, which is a component of the ActiHealth website. ActivePressure illustrates the blood pressure readings taken with the ActiPressure. The device connects to the user's computer wirelessly via an ActiLink Personal Access Point that plugs into a computer's USB port. It is usually used with ActiScale to monitor weight, and Actiped to track steps taken, calories burned, and distance travelled (Fitlinxx, 2012a).

2.2.3.1 Classification of Fitlinxx Actipressure Tool

According to our classification, Actipressure can be used as a standalone tool to collect and display data. When the user wants to see a history of his blood pressure readings, he logs in and the collected data is automatically uploaded. Uploading occurs whenever the tracking device comes within range of the Fitlinxx base station.

2.2.3.2 Challenge in Fitlinxx

Based on user reviews, there are some issues with the Fitlinxx system. Fitlinxx wearable wireless activity monitors do not use standardized protocols such as Bluetooth for transmitting data. Fitlinxx Actipressure uses the BodyLAN Wireless Protocol, a patented ultra-low power wireless network. The BodyLAN Wireless Protocol can automatically and securely transmit data from devices to the web (Fitlinxx, 2012b). However, standardized protocols such as Bluetooth Smart and ZigBee are gaining wider adoption than BodyLAN Wireless Protocol. Newer Fitlinxx products such as the Pebble activity tracker continue to use Actilink for transmitting data instead of the standard Bluetooth or WiFi.

Fitlinxx Product	Classification	Picture
Actipressure	Standalone tool	Fitlinxx image permission withheld

Table 5: Fitlinxx Actipressure classification

2.2.3.3 Data Flow Stages in Fitlinxx Actipressure

The following table demonstrates the data flow in the Fitlinxx Actipressure system.

Data Flow Stage	Methods	
Data Type	Blood pressure and pulse rate	
Collect/aggregate data	Collecting method	Sensor
	How	One-button start operation and pressure rating indicator
Transmitting data type	Wireless, data are sent from the Actipressure device when within range of the ActiLink base station plugged into Mac or PC. Up to 200 feet.	
	Data are sent from the Actipressure device that is within range of the ActiLink base station plugged into Mac or PC. Once the Actipressure device is within range, data sync automatically with the Mac or PC. Internet access and user login are required to send data to the Fitlinxx server.	
Saving data (temporary storage)	Internal memory storage on the Actipressure device saves up to 51 measurements.	
Analysing data	There are no analysing tools. History of collected data is visualised at ActiHealth.com	
Visualising data	Actipressure device has a display screen to present measurements.	
	Web-based application at ActiHealth.com illustrates the history of the user's blood pressure.	
Storing data (permanent storage)	Data are stored at ActiHealth.com once the user uploads them.	
Sharing data	Users can share the collected data by using the ActiHealth web-based application.	

Table 6: Data flow stages in Fitlinxx Actipressure

2.2.4 23andMe

23andMe is a genetic test for DNA analysis. Users send a sample of their saliva or a cheek swab to 23andMe's lab in the United States, and within 4 to 6 weeks they receive a report based on their sample (Ng, Murray, Levy & Venter, 2009). The report identifies many of the genes and genetic variants that could be associated with risk of diseases, and provides some information about the person's ancestors. There are two ultimate aims of DNA analysis: first, predicting diseases that may affect a person in the future (Carlson, 2008); and second, prescribing a more personalized treatment based on the analysis (Ng, Murray, Levy & Venter, 2009).

2.2.4.1 Classification of 23andMe Tool

According to our classification, 23andMe is a hybrid tool – consisting of a kit that is used by the customer and sent back to the lab for analysis, and a website that allows the user to log in and find out about their genes.

2.2.4.2 Challenge in 23andMe

In a study titled “Concordance Study of 3 Direct-to-Consumer Genetic-Testing Services”, researchers state that there is a large variation in relative disease risks reported by 23andMe, deCODEme and Navigenics. This leads to the possibility of a misleading risk assessment and reduces the validity of the result (Imai, Kricka & Fortina, 2011).


23andme	Classification	Picture
23andMe kit (tube) + computer (website)	Hybrid tool	 <p>Image 4: 23andMe (©23andMe, Inc. 2007-2013. All rights reserved; distributed pursuant to a Limited License from 23andMe)</p>

Table 7: 23andMe classification

2.2.4.3 Data Flow Stages in 23andMe

The following table shows the data flow in 23andMe.

Data Flow Stage	Methods	
Data Type	DNA	
Collect/aggregate data	Collecting method	Tube
	How	A sample of saliva is obtained and sent to the 23andMe lab.
Transmitting data type	By the user (postal mail)	
	The user needs to send the tube back to the lab after registering it.	
Saving data (temporary storage)	No temporary saving	
Analysing data	The 23andMe lab, using a microchip by Illumina for analysing the DNA (Illumina OmniExpress Plus Genotyping BeadChip)	
Visualising data	Users log in to 23andMe.com to see their genome test result.	
Storing data (permanent storage)	Data are stored at 23andme.com Data can be stored on the user's computer by saving the generated report (storing data must be done manually by the user).	
Sharing data	Users can print the generated report for sharing with their doctors.	

Table 8: Data flow stages in 23andMe

2.2.5 uBiome

uBiome provides an analysis of the microbes that exist in the skin, ears, mouth, sinuses, genitals and gut. uBiome also provides personal analysis tools and data viewers so that users can anonymously compare their own data with crowd data as well as with the latest scientific research.

2.2.5.1 Classification of uBiome

According to our classification, uBiome falls in the hybrid category. It consists of a kit that is used by the customer and then sent back to uBiome for analysis, and a website that allows the user to log in and view their results.

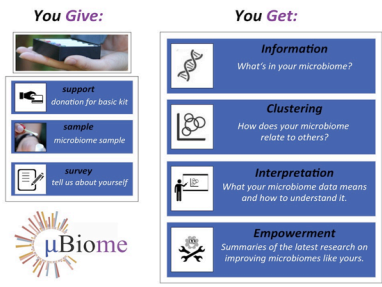
uBiome	Classification	Picture
uBiome kit + computer (website)	Hybrid tool	 <p>You Give:</p> <ul style="list-style-type: none"> support donation for basic kit sample microbiome sample survey tell us about yourself <p>You Get:</p> <ul style="list-style-type: none"> Information What's in your microbiome? Clustering How does your microbiome relate to others? Interpretation What your microbiome data means and how to understand it. Empowerment Summaries of the latest research on improving microbiomes like yours. <p>Image 5: uBiome (©uBiome)</p>

Table 9: uBiome classification

2.2.5.2 Data Flow Stages in uBiome

The following table shows the data flow in uBiome.

Data Flow Stage	Methods	
Data Type	Microbes	
Collect/aggregate data	Collecting method	uBiome kit
	How	Samples are obtained from skin, ears, mouth, sinuses, genitals and gut, and sent to uBiome.
Transmitting data type	By the user (postal mail)	
	The user receives a participant ID upon ordering the kit. The uBiome kit associated with this ID is then sent back to the lab for analysis.	
Saving data (temporary storage)	No temporary saving	
Analysing data	The sample is analysed by scientific researchers.	
Visualising data	Users can log in to uBiome.com to see their results.	
Storing data (permanent storage)	Data are stored at uBiome.com.	
	Data can be stored on the user's computer by saving the generated report (storing data must be done manually by the user).	
Sharing data	uBiome supports anonymous data sharing.	

Table 10: Data flow stages in uBiome

2.2.6 Moodpanda

Moodpanda is a mobile application for iOS and Andriod that allows the user to rate and track their happiness on a scale of 0-10. Users can also add a brief comment about what is influencing their mood and share it with friends.

2.2.6.1 Classification of Moodpanda

According to our provided classification, Moodpanda falls in the smartphone category. Users enter their data manually via the keyboard on their mobile device.


Moodpanda	Classification	Picture
Moodpanda app	Smartphone	<div></div> <div>Image 6: MoodPanda</div>

Table 11: Moodpanda classification

2.2.6.2 Data Flow Stages in Moodpanda

The following table shows the data flow in Moodpanda.

Data Flow Stage	Methods	
Data Type	Perceived happiness	
Collect/aggregate data	Collecting method	Keyboard of mobile device
	How	Users rate their happiness on a 0-10 scale
Transmitting data type	No transmitting data method; data are entered directly into the app by the user via the mobile device keyboard.	
	Data are saved directly to the smartphone's memory storage. Internet access and user login are required to send data to the Moodpanda server.	
Saving data (temporary storage)	The smartphone's memory storage	
Analysing data	There are no analysing tools. Collected data are illustrated in two charts, one illustrating the user mood and the other illustrating the world mood.	
Visualising data	Mood charts can be seen on: Moodpanda's mobile app Web-based application at Moodpanda.com	
Storing data (permanent storage)	Data are stored at Moodpanda.com	
Sharing data	Users can share the collected data by using: Moodpanda's mobile app has a button that allows the user to post data on Facebook or Twitter Web-based application at Moodpanda.com	

Table 12: Data flow stages in Moodpanda

2.2.7 iBGStar

The iBGStar is a blood glucose meter for displaying, managing, and communicating diabetes information. It consists of a blood glucose meter that can be used on its own or connected to an iPhone or iPod touch, and the iBGStar Diabetes Manager App to track diabetes and influential factors. Once the app is launched, the results are automatically logged in the app. If the meter is used alone, the data are saved on the meter's memory and loaded onto the mobile app at next connection. The app also allows the user to email these collected readings (to healthcare professionals, e.g.), or transfer them to a computer (iBGStar, 2012).

2.2.7.1 Classification of iBGStar Diabetes Manager App

iBGStar can be classified under invasive tools. Data are collected by using a lancing tool to obtain a blood sample, which is placed on a test strip for measuring blood glucose levels.

2.2.7.2 Challenge in iBGStar

Based on user reviews, some issues have been experienced with the iBGStar system. The main issue is that data interpretation tools are missing.

One iBGStar user complains that the app is not as smart as it claims to be. It does not flag trends or provide recommendations about managing glucose levels. Various data are tracked by the app; however, the app does nothing but present them back. To figure out what is causing fluctuations in the readings, the user must analyse the data or seek assistance in analysing the data (AmyT, 2011).


iBGStar	Classification	Picture
The iBGStar Diabetes Manager App	Invasive tool Hybrid	 <p>Image 7: iBGStar Diabetes Manager app</p>

Table 13: iBGStar classification

2.2.7.3 Data Flow Stages in iBGStar

The following table shows the data flow in iBGStar Diabetes Manager App.

Data Flow Stage	Methods	
Data type	Blood glucose level	
Collect/aggregate data	Collecting method	External smartphone-attached sensors
	How	A blood sample is obtained by a lancing tool and placed on the test strip for measuring the blood glucose level.
Transmitting data type	Direct connection by attaching the iBGStar blood glucose meter to an iPhone or iPod touch	
	Once iBGStar Diabetes Manager App is launched, data sync automatically with the smartphone. Internet access and user login are required to send data to the iBGStar server.	
Saving data (temporary storage)	The smartphone's memory storage	
Analysing data	There are no analysis tools. Test results saved on the meter's memory are only visualised in charts.	
Visualising data	iBGStar's glucose meter (a small display screen)	
	iBGStar mobile app presents the history of data collected in three different data viewing options (Trend Chart, Logbook, Statistics)	
Storing data (permanent storage)	Data can be stored on the user's PC by saving the contents of the logbook	
Sharing data	The application has a share button that allows the user to email the contents of the logbook to a healthcare professional	

Table 14: Data flow stages in iBGStar

2.2.8 Sensaris Senspod

Sensaris Senspod is an environmental data sensor that captures data in real-time and sends them via Bluetooth to a smartphone. It captures noise, humidity, temperature, and carbon monoxide and nitrogen oxide levels. All Senspods are provided with an Android application and access to the web interface at www.Sensdots.com (Sensaris Senspod, 2012). Users can log in through their smartphones that are paired with a Senspod, then read collected data and send them to the web.

2.2.8.1 Classification of Sensaris Senspod

Sensaris Senspod can be classified as a hybrid tool. It consists of a sensor and a smartphone-based application (MobiSense) or a web interface at sensaris.com.


Sensaris Senspod	Classification	Picture
Sensaris Senspod sensor, and web interface or smartphone-based application	Hybrid tool	 Image 8: Sensaris sensor and app

Table 15: Sensaris Senspod classification

2.2.8.2 Data Flow Stages in Sensaris Senspod

The following table shows the data flow in Sensaris Senspod system.

Data Flow Stage	Methods
Data type	Carbon monoxide (CO), nitrogen oxide (NOx), noise, temperature, and humidity
Collect/aggregate data	Collecting method: Touchless sensors
	How: Sensaris Senspod can be installed in a fixed place such as an office. It captures data in real-time and sends them to a smartphone.
Transmitting data type	Bluetooth, up to 30m
	Data are sent from the Senspod to the paired smartphone or computer when it is within range. Internet access and user login are required to send data to the Sensaris server.
Saving data (temporary storage)	SD card internal memory storage on the Senspod (2GB or 4GB)
Analysing data	There are no analysis tools. Mobile app and web-based app are only for displaying the measurements
Visualising data	Mobile app (MobiSense) presents the measurements in real time
	Web-based app at sensaris.com , which is integrated with Google maps Collected data can be presented in two different data viewing options: charts and graphs. Also, users can select a period of time to study events in detail
Storing data (permanent storage)	Data can be stored locally (SD card) or sent to a server
	Data can be exported using CSV or RSS format
Sharing data	Users can export data as CSV or RSS format and share them. Sensaris has a web interface that can be accessed globally and where users can upload the collected raw data.

Table 16: Data flow stages in Sensaris Senspod

2.3 Informatics Aspects of Primary Self-Quantification Systems

The following chart shows the mapping between the provided examples of primary Self-Quantification systems and their classification.

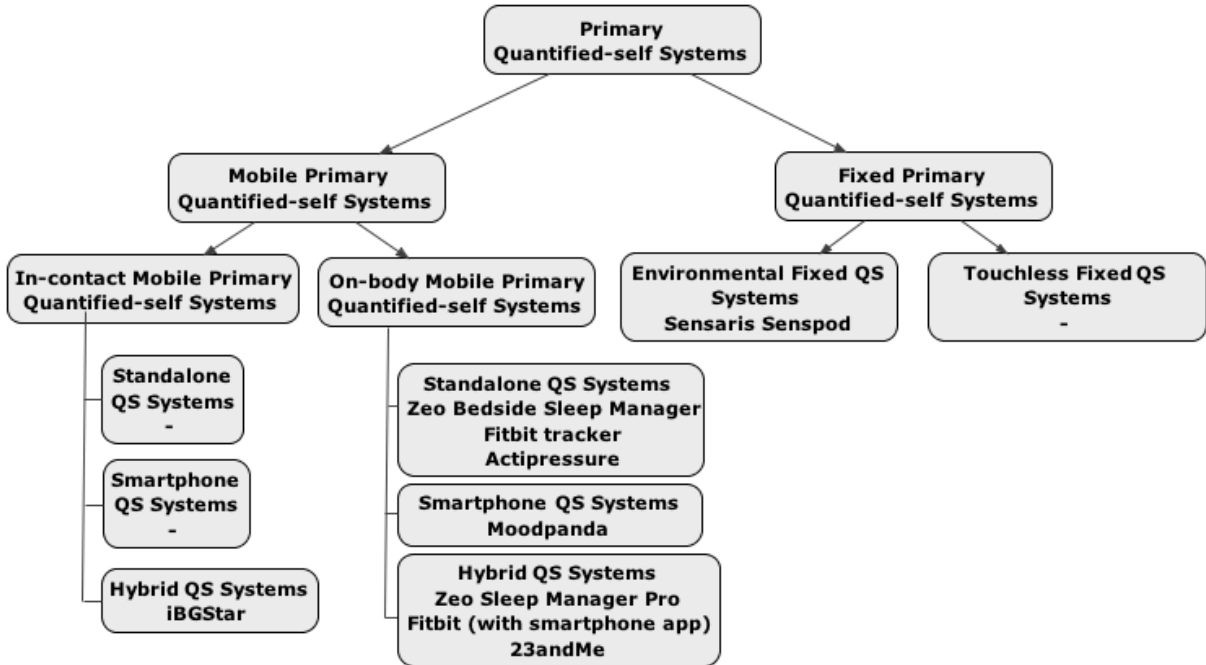


Figure 3: Primary Self-Quantification systems classification with examples

2.3.1 Data Types in Self-Quantification Systems

In Self-Quantification systems, data are grouped into three categories: exposome, phenome, and genome. The term 'exposome' has been coined to refer to the lifelong exposure of an individual to environmental risk factors. The term 'phenome' refers to "the overall expression of a person's characteristics and traits as determined by the interaction of genetics and environment" (Young, 2012). The term 'genome' refers to the hereditary instructions of a life form. In a human being, these instructions are encoded in the DNA. Human DNA consists of about 3 billion bases, and more than 99% of those bases are the same in all people. However, the order or sequence of these bases determines the variances in hereditary instructions between people (Bandyopadhyay & Kumar, 2011). Most human diseases are the result of a complex interplay between exposome, phenome, and genome factors. These three types of data are not constant in nature. They are subject to modifications as a result of exposure to several things, such as environmental changes, diet, and stress levels (Payne, 2012). The following table illustrates the data type categories of several Self-Quantification systems profiled in this report.

Group	Measure	Self-Quantification Systems
Exposome	Sleep	Zeo
	Physical activity	Fitbit
	CO, NOx, noise, temperature, and humidity	Sensaris Senspod
	Microbes	uBiome
Phenome	Mood	Moodpanda
	Blood pressure	Actipressure
	Blood glucose	iBGStar
Genome	SNPs (Single nucleotide polymorphisms, or genetic variations)	23andMe

Table 17: Data types in Self-Quantification systems

2.3.2 Summary of How Typical Primary Self-Quantification Systems Work

Most Self-Quantification systems typically work as follows:

- A tool that has a sensor or data input mechanism (e.g., keyboard) is used to collect the required data.
- Collected data are saved locally in the tracking tool, or on a smartphone or computer connected directly or wirelessly to the sensor.
- Collected data can be synced automatically or manually with the user's computer or smartphone.
- Internet access is necessary to send the collected data to the service-provider's server. The computer or the smartphone that runs the app is used to enable the user to login and synchronize the data with the service-provider server.
- Collected data are stored at the service-provider side to be analysed later.
- Data analysis can be done to interpret the user's data, extract patterns, and find correlation among collected data.
- The generated results are illustrated in different data viewing options (e.g., charts, logbook, etc.) describing the user's health status.
- An action will be taken upon the generated results such as sharing data with a healthcare professional, social networking, or doing more exercise to improve blood pressure.

2.3.3 Summary of Data Flow in Self-Quantification Systems

The following table shows the summary of data flow in Self-Quantification systems:

Data Flow Stage	Description of the Stage	Methods
Collect Data	It can be described as the process of capturing data by a method.	Wearable sensor (Zeo headband) Accelerometer and Altimeter (Fitbit) Tube (23andme) Smartphone keyboard (Moodpanda) External smartphone-attached sensors (iBGStar) Touchless sensors (Sensaris Senspod)
Transmitting Data	It can be described as the process of sending data from a sensor or tracking tool to reside temporarily or permanently in a storage place.	Bluetooth (Zeo headband) WiFi (Fitbit) Postal mail by the user (23andme) Direct connection between an iPhone or iPod touch and the sensor (iBGStar) SD memory card and USB adapter (Zeo bedside)
Saving Data	It can be described as the process of storing data in a temporary storage place, and then offloading data automatically with the paired device when the tracking tool within the range of the access point.	Internal flash storage or SD card on the tracking device Smartphone memory storage
Storing Data	It can be described as the process of keeping data permanently, either on the user side or service provider side, or both.	Data can be imported and stored on the user's computer in different file formats e.g., XML, CSV, RSS. Data can also be stored on the service provider's server.
Analysing Data	It can be described as the process of drawing conclusion out of the collected data or the presented measurements.	A system such as the Zeo Sleep Manager use a proprietary algorithm for analysing data, and then provides a customized programme based on the user's Z score. Third party apps for data analysis (see section 6).
Visualising Data	It can be described as the process of presenting back the collected data in a sort of graphical illustration such as trend chart, logbook, and statistics.	The tracker tool may have a small display screen to present the measurements. A smartphone-based app to illustrate the user's data history in different viewing options. Here data and the app are all on the smartphone. Data sync with the app when there is Internet access. A web-based application to present the user's data in different viewing options.
Sharing Data (by user)	It can be described as information and data exchange within a network of care providers, family members, and other care and support providers for preventative, promotive and curative objectives through a range of devices and communication networking tools.	Most systems allow the user to export data or share it via social networks such as Facebook or Twitter. Section 5 discusses sharing or exchanging data among service providers.

Table 18: Summary of data flow stages in Self-Quantification systems

3 Secondary Self-Quantification Systems

The secondary quantified-self (QS) systems can be described as a single tool or app for aggregating or integrating the collected data by a primary QS tool. For example, BodyTrack is a secondary quantified-self system. It is able to integrate personal measurement readings that are delivered by different devices. We can also refer to secondary quantified-self systems as Personal Informatics.

Secondary Self-Quantification systems are mostly focused on helping the user to better reflect on their data. Illustrating this point, BodyTrack founder Anne Wright said the main goal of the BodyTrack system is to “empower individuals to explore potential environment/health interactions (food sensitivities, asthma or migraine triggers, sleep problems, etc.) and better assess strategies they think might help” (Wright, Kemmler & Gibson, 2012).

3.1 Typology of Secondary Self-Quantification Systems

Secondary QS systems can be classified into two groups: software-based secondary QS systems, and hardware-based secondary QS systems. A software-based secondary QS system has mainly a web-based or smartphone-based application for integrating, visualising, and sharing tracked data. BodyTrack and Wikilife are examples of software-based secondary quantified-self systems. However, a hardware-based secondary QS system consists of a connector for integrating data that are captured by primary tracking tools, and a web-based or smartphone-based application for visualising, and sharing tracked data. Digifit is an example of hardware-based secondary quantified-self systems.

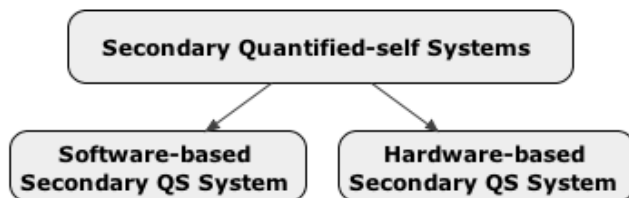


Figure 4: Secondary Self-Quantification systems classification

3.2 Examples of Secondary Self-Quantification Systems

3.2.1 *BodyTrack*

BodyTrack integrates data on activities, environmental and food inputs, and health status that are delivered by different devices (such as Fitbit and Zeo) over time in order to get an overall picture of health and derive intelligence from the collected data (BodyTrack, 2013).

3.2.1.1 *Data flow Stages in BodyTrack*

The following steps show how data flows in BodyTrack.

- **Collecting data:** Collect data by using primary Self-Quantification system. The generated data from such systems will be stored on the user's computer.
- **Storing data:** Data can be stored on either the user's computer, or the service provider's server, or both.
- **Integrating data:** Upload the collected data into the BodyTrack website. The BodyTrack website provides a variety of visualisation tools for data presentation. BodyTrack also allows the user to explore relationships between different datasets and scale the timeline from milliseconds to decades reading down the graph (Johnfass, 2012).
- **Sharing data:** BodyTrack supports a data-sharing feature.

3.2.2 Wikilife

Wikilife integrates lifestyle information such as exercise, health, psychological state, nutrition, milestones (important events during an individual's lifetime), work, education, beauty, travel, spirituality, and physiological data. It enables data integration by allowing the user to export data, results and statistics generated by health tracking devices to the Wikilife website.

3.2.2.1 Data flow Stages in Wikilife

The following steps show how data flows in Wikilife.

- Collecting data: Collect data by using primary Self-Quantification system. The generated data from such systems will be stored on the user's computer.
- Storing data: Data can be stored on either the user's computer, or the service provider's server, or both.
- Integrating data: Upload the collected data to the Wikilife website. The Wikilife website provides a variety of visualisation tools for data presentation. Wikilife also allows the user to explore relationships between different datasets.
- Sharing data: Wikilife supports the anonymised sharing of data.

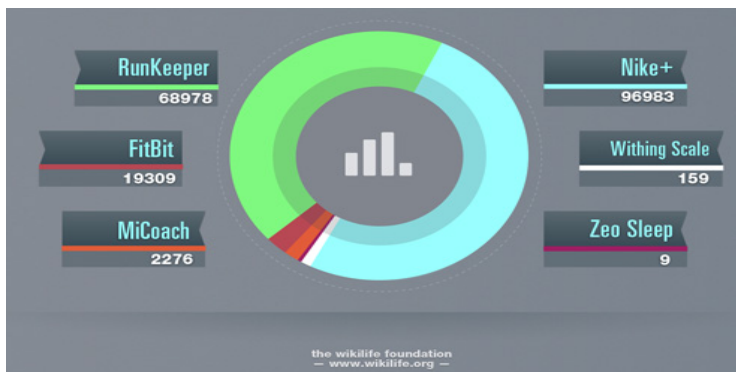


Image 9: Wikilife

3.2.3 Digifit

Digifit is a cardio fitness tool that is compatible with about 80 ANT+ sensors, Zeo, Fitbit, Garmin, Adidas, Withings (for tracking weight) and more (ANT+ is an interoperability protocol that uses primarily for designing collection and transfer of sensor data (Wikipedia, 2013)). Digifit can integrate heart rate and all runs, rides, spinning and cardio on a single device. It consists of Digifit app and connector attached to the smartphone (Digifit, 2012).

3.2.3.1 Data Flow Data flow Stages in Digifit

The following steps show how data flows in Digifit.

- Collecting data: Collect data by using primary Self-Quantification systems that are compatible with ANT+ sensors.
- Transmitting data: Digifit connects wirelessly to the smartphone with ANT+ health and fitness sensors.
- Integrating and storing data: The Digifit connector integrates multiple datasets into one dataset and stores it in the smartphone's memory.
- Visualisation: Upload/sync the single aggregated dataset to the Digifit website. The Digifit website provides a variety of visualisation tools for data presentation. Digifit also allows the user to explore relationships among different datasets. Although it may combine collected data to be presented in one place, it also provides separate paragraphs for each measurement.
- Sharing data: The user can share workout data on my.Digifit.com, Facebook, Twitter, by email and via other sharing options.



Image 10: Digifit

3.3 Informatics Aspects of Secondary Self-Quantification Systems

The following picture shows the mapping between the examples provided of secondary Self-Quantification systems and their classification.

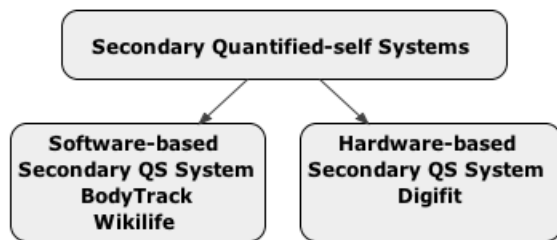


Figure 6: Secondary Self-Quantification systems classification with examples

3.3.1 Summary of How Typical Secondary Self-Quantification Systems Work

Most secondary Self-Quantification systems work as follows:

- Primary Self-Quantification systems are used for collecting data.
- Data may be stored on the user's computer, on the service provider's server, or both. However, the data generated from such systems will be stored locally for the next step.
- The user can then upload the collected data to the service provider's website.
- Data will be analysed and/or visualised on the service provider's website. The website usually provides a variety of analytics and visualisation tools for data presentation and exploring relationships between different datasets.

The user takes a subsequent action, such as sharing data with a healthcare professional or social network, or doing more exercise to improve blood pressure.

3.3.2 Summary of Data Flow Stages in the Secondary Self-Quantification Systems

Following is a summary of the data flow in secondary Self-Quantification systems.

Data Flow	Description
Data Collection/Aggregation	It can be described as the process of capturing data through the use of primary QS systems, and aggregating or integrating them by using secondary QS systems.
Data Storing	It can be described as the process of keeping data permanently, either on the user side or service provider side, or both.
Data Integration	It can be described as the process of uploading the generated data into a smartphone-based or web-based application where data integration happens.
Data visualisation	It can be described as the process of using a variety of visualisation tools for exploring or interacting with data presentation or information visualisations.
Data Sharing	When a QS system supports data sharing, users can share workout data on the service provider website (e.g., my.Digifit.com), Facebook, Twitter, email and more. Sharing data could also happen in an anonymous way.

Table 19: The data flow stages in secondary Self-Quantification systems

4 Directories of Apps

App stores exist across all mobile platforms, including the Apple Store, Google Play (formerly the Android Market) and BlackBerry World. Each of these stores offers apps in a variety of categories, including games, music, lifestyle, fitness, medication, finance, etc. More than other apps, health-related apps pose specific concerns in regard to quality, accuracy, privacy, security and so on, because they collect sensitive personal data. In addition, searching for the right health-related app can be time-consuming for both patients and healthcare providers, due to a lack of categorization of these apps.

4.1 Quantified Self Guide

The Quantified Self movement is a group of people interested in tracking data about themselves and using this data to change and improve their lives. Their focus is a kind of self-experimentation to see what and how a variable of interest can be improved. For example, a user could measure the impact of variations in diet on productivity or happiness over the course of a year. It can almost be described as an “individualized evidence-based” approach (Rossouw, 2012).

The website quantifiedself.com provides a complete guide for self-trackers. The guide includes a collection of tools, apps, and projects for self-tracking/logging. The guide categorizes apps based on meta tag classification and the app’s price.

4.2 Happtique

There are more than 23,000 apps available in mobile app stores. If a diabetic person were to search for an app that monitors blood sugar, the hundreds of apps from various developers and software companies would likely overwhelm them, each claiming to offer the best solution for managing diabetes. It would be difficult for the patient to know which apps, if any, are useful and safe to use. A company called Happtique believes it has the solution to this problem.

Happtique is the first mobile app store developed by healthcare professionals for healthcare professionals and patients. Happtique was founded in 2010 by the venture arm of the Greater New York Hospital Association. Happtique offers a platform for the curation, certification, and prescribing of mobile health apps. This is accomplished by providing three products:

- The hApp Catalog: Where apps are categorized using the vocabulary of healthcare professionals
- Enterprise Application Sub-Stores: So that healthcare organisations can create and customize their own secure mobile application store.
- Happtique’s mHealth Community: A space for healthcare professionals to share ideas about mobile health.

For an easy way to find relevant apps, Happtique offers a comprehensive catalogue that “uses a healthcare app indexing method designed to be intuitive to industry professionals and patients” (Happtique, 2011). There are more than 300 categories in Happtique’s catalogue. The app classification is not based on subjective standards such as popularity; rather, physicians classify apps from the perspective of appropriateness of use with their patients. The hApp Catalog categorizes apps using techniques similar to those used to organise medical libraries.

For certification of mobile health apps, Happtique has launched its App Certification Program to help physicians, patients, and other mHealth consumers identify apps that have reliable content and meet high operability, privacy, and security standards. Any app developer can apply for certification. Once a developer submits an app, the app will first undergo testing to determine its compliance with technical standards. Apps that pass the technical standards assessment will proceed to content review by a professional from the relevant field or discipline. By relying on these evaluation standards, users will be able to identify and locate reliable apps for their needs. Happtique has provided further details about their app certification standards on their website www.happtique.com.

For prescribing of mobile health apps, Happtique has announced the commencement of its pilot program of mRx. Happtique’s patent-pending solution enables physicians and other health practitioners to electronically prescribe medical, health, and fitness apps to their patients. The pilot will focus particularly on cardiology, rheumatology, endocrinology, orthopaedics, physical therapy, and fitness training. Such practice is identified as an app therapy, according to Ben Chodor, CEO of Happtique. It has been claimed that mRx is the first program to enable doctors to prescribe mHealth apps to patients. It also enables physicians to know whether a prescribed app was downloaded, if the app is Android or HTML5-based.



Image 11: hApp is a mobile medical/health app available through the Happtique store.

4.3 European Directory of Health Apps 2012–2013

The European Directory of Health Apps 2012–2013 is the first directory of apps that are recommended by patient groups and empowered consumers. It is published by PatientView (www.patientview.com) and contains about 200 mobile health-related apps categorised by: service provided to the patient/consumer, language(s) in which the app is offered, price, and platform (Android, Apple, BlackBerry, Nokia, Windows Phone). Each app has a one-page entry in the directory, which includes the patient group/consumer recommendations, the cost of the app, the identity of the developers, and a link to the webpage where the app was originally published (Madelin, 2012).

Additionally, PatientView provides online surveys for consumers and developers. Consumers can use a survey to recommend an app for inclusion in the next directory. Developers can use a different survey to leave details of the apps they have created; patients or consumers then review these apps and the developer gets feedback.

5 Interoperability Standards in Self-Quantification Systems

One of the main issues in Self-Quantification systems is sharing data or exchanging data among various systems. In this section, we explore the ways in which two different organisations handle this data-sharing challenge. These organisations are Open mHealth and Health Level 7 (HL7).

5.1 The Current Status of Mobile Health Applications

According to Estrin and Sim (2010), current mobile health applications follow a stovepipe architecture, where “each app is built as a closed application with its own proprietary data format, management, and analysis”. So what does stovepipe architecture mean, and why are stovepipe systems bad? How are Open mHealth and HL7 (FHIR) overcoming the defects of the stovepipe?

5.2 Stovepipe Architecture vs Open mHealth Architecture

In this section we provide a brief comparison between stovepipe and Open mHealth architectures.

5.2.1 Stovepipe Architecture

Stovepipe refers to an architecture that does not have the ability to share data or functionality with other systems (Wikipedia, 2012). Stovepipe systems produce their own silos of information. The reason for building such systems is that the initial intention of developers did not include further developments on top of the system. Thus, each stovepipe had to act as if nothing else in the world existed, and was built to function in isolation (Inmon, 2003). Image 14 illustrates how stovepipe systems work.

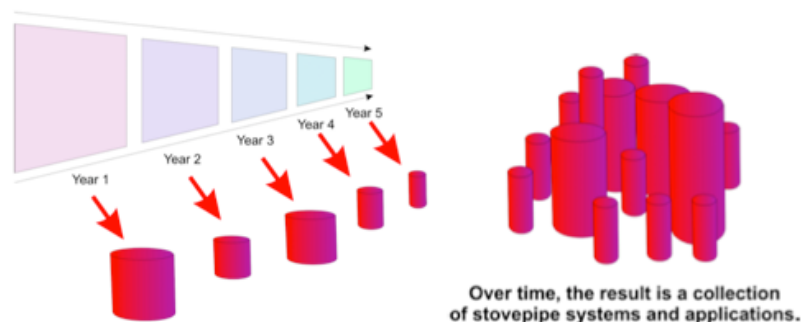


Image 12: Stovepipe system (Inmon, 2003)

5.2.2 Issues in Stovepipe Systems

According to Inmon, there are many deficiencies in stovepipe architecture. First, there is redundancy in functionality as depicted in Image 15. The same or very similar functions can be found in many places. Because previous efforts of developers are ignored, similar functions are built, rebuilt, and rebuilt once again but in an inconsistent manner. Consequently, the rebuilding of the same information infrastructure in different forms results in wasted storage and wasted development, execution and maintenance time.



Image 13: Same or similar functions are found in many places (Inmon, 2003)

Second, the collected data cannot be used outside of the silo. “As a simple example of this tremendous overlap, how many times does a government agency require personal information, such as gender, age, place of birth, education, current job grade, and so forth? Practically every stovepipe system gathers the same information that every other stovepipe system has already gathered.” (Inmon, 2003)

Third, sharing data is difficult due to integrity and redundancy issues, as illustrated in Image 16. As an example of this issue, Mary is listed in one database as a registered healthcare practitioner, in another place she is shown as a student in a medical school, and in another place as a PhD. None of these entries are associated with a date or other clues. In this case, how can we get an accurate view of Mary’s education and qualifications?

Fourth, stovepipes represent short-term information architecture, because they are built with no consideration for further developments on top of the system. When a new function is needed, adding it is difficult (see Image 16).

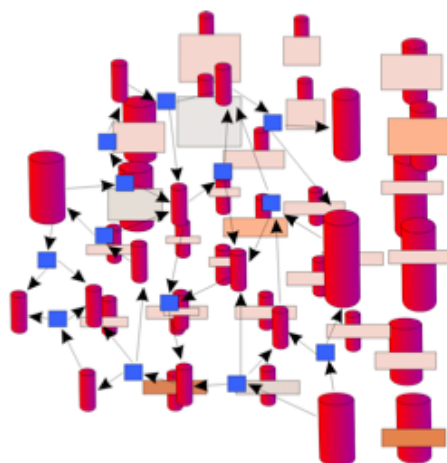


Image 14: Exchanging data in a stovepipe system

5.2.2.1 Summary of Issues in Stovepipe Systems

The following table summarises the issues identified in stovepipe systems.

Stovepipe System Defects
Redundancy in functionality
Previous efforts of developers are ignored
The collected data cannot be used outside of the silo
Sharing data is difficult due to integrity and redundancy
Short-term information architecture

Table 20: Summary of issues in stovepipe systems

5.2.3 Open mHealth Organisation

Open mHealth is a non-profit organisation that aims to catalyze the transition of current medical health application development to an open ecosystem (Chen, Haddad, Selsky, Hoffman, Kravitz, Estrin & Sim, 2012). Open mHealth was established in 2011 by Deborah Estrin and Ida Sim, researchers from the University of California, San Francisco and the University of California, Los Angeles.

5.2.4 Open mHealth Architecture

In stovepipe architecture, the data flow from the collection stage to the database and then back to the user to be visualised on a smartphone or computer screen is happening in a fractured manner. In contrast, Open mHealth architecture supports a fully integrated data flow at all points in the data ecosystem, as illustrated in Image 17.

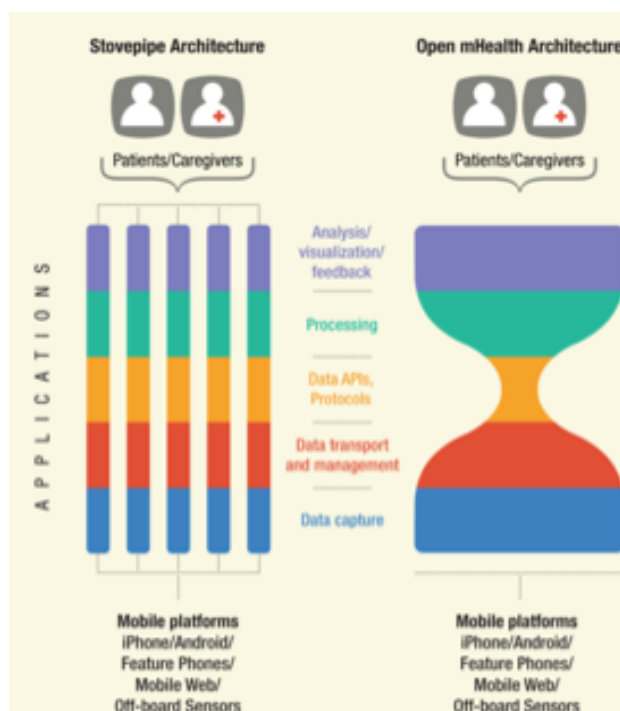


Image 15: The differences between stovepipe and Open mHealth architecture

5.2.5 Open mHealth Architecture Features

According to Estrin (2011), open mHealth architecture consists of easy-to-use software modules and Application Programming Interfaces (APIs) that programmers can incorporate directly into product development. This requires building two kinds of infrastructure: technical infrastructure and social infrastructure. The technical infrastructure is needed for data-driven iterative learning and sharing effective methods. The social infrastructure is comprised of stakeholder communities who build and reuse modules. The technical and social infrastructure are built with the following points taken into account (Chen, Haddad, Selsky, Hoffman, Kravitz, Estrin & Sim, 2012):

- **Iteration:** Delivers efficient reuse through collaborative cycles of development. In other words, reusable components enable rapid authoring, integration, and evaluation of personal data capture for clinical care and research. Therefore, mHealth apps can be iteratively modified to configure customized apps “e.g., what symptoms to monitor, when, where, and how, or what data sources to incorporate”.
- **Flexible architecture:** Recognizes both the limits and the utility of existing closed systems and is designed to maximize participation from all players. In other words, existing components should be re-used as much as possible and be able to be implemented in ways other than originally intended. This allows interested parties to expand the functionality of the system without modifying existing components.
- **Scalable solutions:** Offers mass customization of applications and evidence, from personal to population. In other words, tools and methods should be applicable across a broad range of health conditions (e.g., mental health, diabetes), technical platforms (e.g., iOS, Android; various electronic and personal health record platforms – EHRs and PHRs), and user contexts (e.g., self-care, specialist care). Also, all components must follow interoperability specifications for data interchange. Some of these interoperability standards are defined and refined in the Open mHealth GitHub Repository at: github.com/openmhealth
- **Shared learning:** Uses the strongest appropriate methods, matched to the evidence needs and the rapid pace of technological advances in mHealth. The Open mHealth goal is that mHealth becomes a learning community that effectively innovates, shares, and deploys best technology and best practices for improving individual and population health.
- **Community:** Must be multidisciplinary, safe and collaborative. The community consists of patients, clinicians, family and others who can be involved in mHealth application design. The community is a tool for providing daily real-world health data. The community also supports collaborative development and testing of apps, which increases the quality and adoption of open mHealth.

In summary, the quality of an open architecture depends on the modularity and reusability of common functions, and the simplicity and legibility of the APIs (Estrin, 2011).

5.2.6 InfoVis

Open mHealth researchers state that common standards along with open access APIs can meet the requirements of open mHealth architecture. Therefore, in an effort to catalyze this open access approach, these researchers have developed a framework called InfoVis for use by this new community. InfoVis is a module for analysing and visualising data. It consists of two components: data processing units (DPUs) and data visualisation units (DVUs). DPUs extract patterns from the data while DVUs present these extracted patterns. Figure 14 depicts the InfoVis architecture and how it can be used within a third-party application [Appendix A shows the full details of the Open mHealth system].

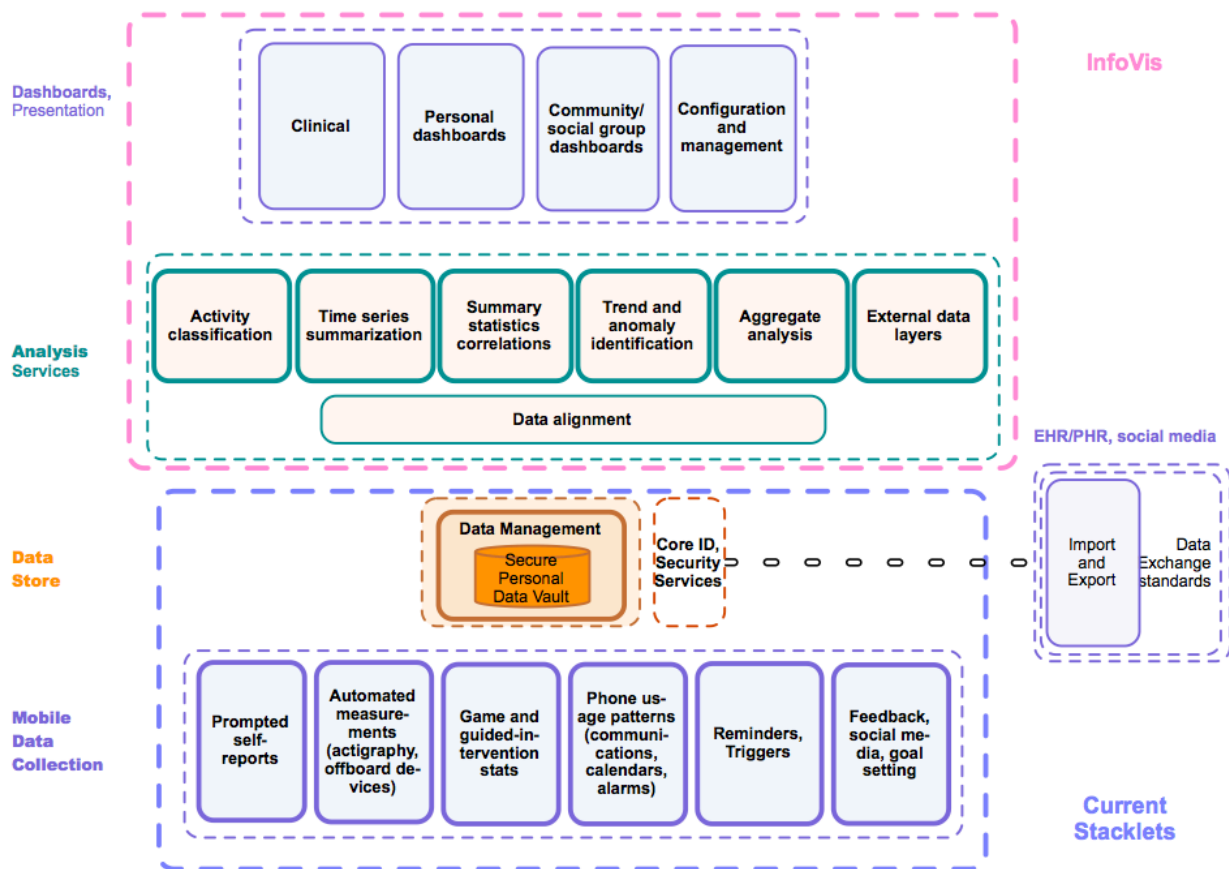


Image 16: InfoVis architecture and use within a third-party application (Estrin & Sim, 2011)

5.2.7 How InfoVis Meets the Open mHealth Requirements

The following table shows how InfoVis meets the Open mHealth requirements.

Open mHealth Features	InfoVis
Iteration	InfoVis can be re-implemented in another app regardless of the initial purpose for which it was developed. For example, InfoVis is incorporated into ohmage (see next section). Ohmage is used in seven independent studies, each addressing a different population such as breast cancer survivors, new mothers, and at-risk HIV+ men. "This rich feedback has driven the key features included in the ohmage platform" (Chen, Haddad, Selsky, Hoffman, Kravitz, Estrin & Sim, 2012).
Flexible architecture	<p>DPU and DVU can be embedded in a plug-and-play fashion within the application. Thus, InfoVis can be directly incorporated with old modules as well as with newly added modules, facilitating the growth of the system structure.</p> <p>DPU and DVU can be composed to produce higher-level functions.</p> <p>DPU and DVU can be incorporated into applications as libraries or can be invoked using JavaScript object notation over hypertext transfer protocol if they are developed with a Web service wrapper.</p> <p>Applications with embedded DPU and DVU can run on any system, ranging from the Android operating system, for example, to full-featured platforms such as those of large telecoms services providers.</p>
Scalable solutions	InfoVis DPU and DVU can be used across the range of diseases and health conditions.
Shared learning	As InfoVis architecture allows developers to compose its components to produce higher-level functions, developers can share the updated modules with the Open mHealth community, which may result in propagating these updates across Open mHealth systems.
Community	DPU and DVU will process data from data storage units that access a wide range of third-party data applications and stores. This will build a strong community to complement innovations to maximize the overall impact of mHealth.

Table 21: How InfoVis meets the Open mHealth requirements

5.2.8 Ohmage: An Example of Incorporating InfoVis Into an App

Ohmage (AndWellness) is an app developed on an Open mHealth architecture for personal data collection systems and self-management of health (Ramanathan, Alquaddoomi, Falaki, George, Hsieh, Jenkins, Ketcham, Longstaff, Ooms, Selsky, Tangmunarunkit & Estrin, 2012). Ohmage contains three subsystems: an application to collect data on an Android mobile device, a server to configure studies and store collected data, and a dashboard to display users' statistics and data (Hicks, Ramanathan, Kim, Monibi, Selsky, Hansen & Estrin, 2010). The ohmage system is illustrated in Image 19.

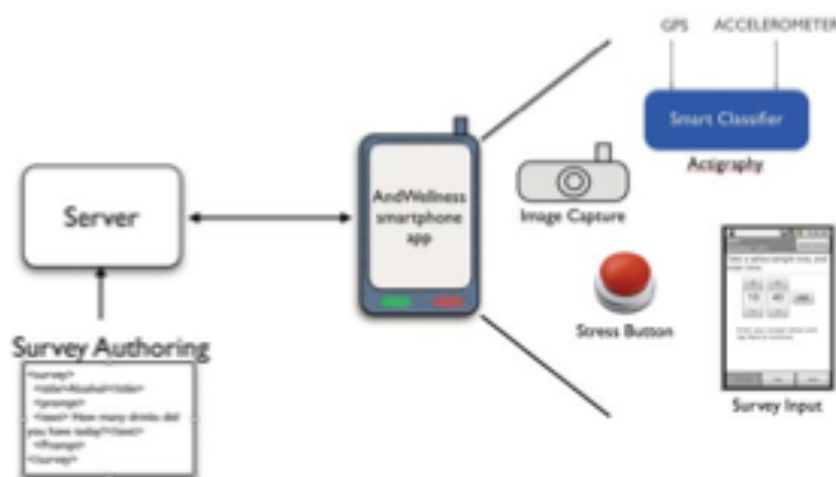


Image 17: Ohmage contains three subsystems: app, server, and dashboard for displaying data

5.2.8.1 Data flow Stages in Ohmage

The following table shows the data flow in ohmage.

Data Flow Stage	Ohmage
Data type	Tracking users' behaviours to help them design customized interventions
Collect/aggregate data	Survey responses, images and sensor readings (GPS and accelerometer)
Transmitting data	Internet: HTTP, JSON (JavaScript object notation)
Storing data	Personal Data Vault
Analysing data	InfoVis, Data Processing Units (DPUs)
Visualising data	InfoVis, Data Visualisation Units (DVUs) Data are visualised: Personal dashboard for the user Another dashboard for clinicians/researchers
Sharing data	Open mHealth Community
Reusing stored data	Personal Data Vault, see appendix B.

Table 22: Data flow stages in ohmage

5.3 Health Level 7 (HL7)

HL7 is a non-profit organisation, founded in 1987 and headquartered in Ann Arbor, Michigan, with more than 55 affiliate organisations worldwide. It was originally founded to develop a standard for hospital information systems, and has developed several standards such as V2, V3, CDA, HDF and FHIR. “The standards address message and data exchange, decision support, rules syntax, visual integration of applications, insurance claims, clinical documents such as discharge summaries, product labels for prescription medication, electronic health records and personal health records” (Healthcare IT news, 2012). The main aim of developing these standards is to help various healthcare systems to communicate with each other, share information and process data in a consistent manner.

In addition, “HL7 encompasses the complete life cycle of a standards specification including the development, adoption, market recognition, utilization, and adherence” (HL7, 2012). HL7 promotes the use of its standards through collaboration with other standards organisations (e.g., ANSI and ISO) and also collaborates with healthcare information technology users to ensure that HL7 standards meet real-world requirements.

5.3.1 Overview of HL7 Standards

Following is a summary of some common HL7 standards:

- V2 is a HL7 standard for messaging but does not support XML format.
- V3 is a HL7 standard for messaging and supports XML format.
- Clinical Document Architecture (CDA) is a HL7 standard for clinical document processing/parsing to makes these documents human readable, machine processable, and exchangeable by using the XML format. CDA is also used in electronic health records projects to provide a standard format for entry, retrieval and storage of health information (HL7 Australia, 2012).
- Reference Information Model (RIM) and HL7 Version 3 Development Framework (HDF) are the main standards in HL7 V3. “They include specification of information models, data types, and vocabularies; messaging, clinical documents, and context management standards; and implementation technology, profile, and conformance specifications” (HL7, 2012). RIM and HDF are based on the Unified Modeling Language (UML).
- FHIR is a RESTful framework for messaging.

5.3.2 Fast Healthcare Interoperability Resources (FHIR)

HL7 has exploited the open Internet standards for developing FHIR. FHIR uses XML standards and an HTTP-based RESTful protocol to document and exchange data. Exchanged data are primarily represented as resources. Resources may refer to persons, patients, prescriptions, etc. (see Image 20). Each resource has a unique ID, and a URL that is derived from the ID, the type, and the local base URL. For implementers, the UML class diagram and the matched XML file are available for each resource. Each XML or UML file is associated with a reference that can be used by other standards; therefore, resources can be exchanged.

5.3.2.1 Resources in FHIR's RESTful framework

Resources may refer to persons, patients, prescriptions or any object in the system (HL7, 2012). Image 20 depicts many examples of resources.

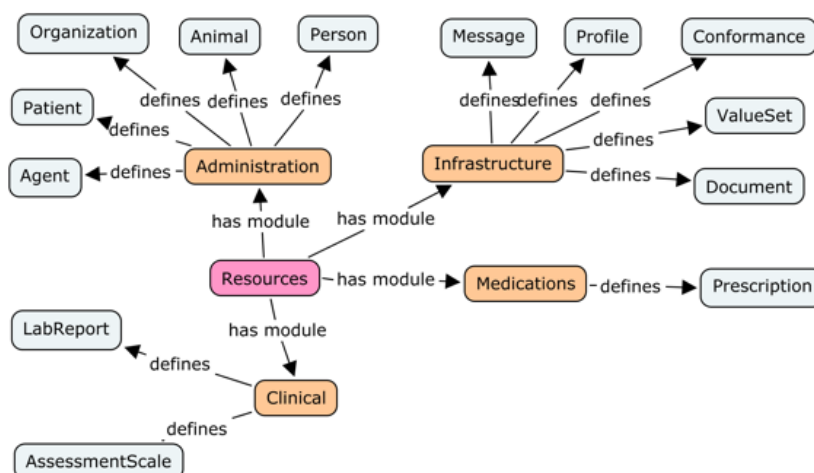


Image 18: Resources in FHIR's RESTful framework

Image 21 illustrates an example of a resource format used in FHIR. The user enters the required URL, and an XML file is returned that contains the result details.

<http://example.com/customers/1234>

```

<order self='http://example.com/customers/1234' >
  <amount>23</amount>
  <product ref='http://example.com/products/4554' />
  <customer ref='http://example.com/customers/1234' />
</order>
  
```

Image 19: Instance of resource format in FHIR v0.05

5.3.2.2 Resource Features in FHIR

The exchangeable resources should be:

- Granular: They are the smallest unit of operation and have a transaction scope of their own.
- Independent: The content of a resource can be understood without reference to other resources.
- Simple: Each resource is easy to understand and implement without needing tooling or infrastructure (though that can be used if desired).
- RESTful: Resources are able to be used in a RESTful exchange context.
- Flexible: Resources can also be used in other contexts, such as messaging or service oriented architectures (SOA), and moved in and out of RESTful paradigms as convenient.
- Extensible: Resources can be extended to cater for local requirements without impacting on interoperability.

- Web-enabled: Where possible or appropriate, open Internet standards are used for data representation.
- Free for use: The FHIR specification itself is open – anyone can implement FHIR or derive related specifications without any IP restrictions.

5.3.2.3 *How FHIR is Used in Resource Exchanging*

For exchanging a particular resource, the XML and HTTP-based RESTful protocol are used as following:

- Each resource is invoked by its URL and transmitted by using a RESTful framework that is based on a HTTP request/response. The exchanged message could be a collection of all the resources that are aggregated and sent in an atom feed/ bundle. For example, an Order resource might be composed of order items, an address and many other attributes but will not expose these as individually identifiable resources (in the appearance of the URL).
- The URL is sent to the server using a simpler GET request, and the HTTP reply. The reply is an XML file of the data result.

5.4 Data Storage Locations in mHealth Systems

There are four types of databases in the mHealth ecosystem, each of which is a self-contained silo that cannot easily share data with other silos or systems:

- Self-Quantification service provider data repository. For example, Fitbit, Zeo, and iBGStar each generates its own data, which is stored in the service provider's own database. This currently is happening in a way corresponding to stovepipe systems. The data generated by these systems cannot be shared unless there is collaboration between service providers, such as that between Fitbit and Zeo.
- Health service provider electronic health records (EHR). These are currently mainly clinical repositories of patient data, with the collection of data done by healthcare professionals within a specific clinic or hospital. (Note: in Open mHealth systems, healthcare professionals are not the only data collectors. Australia's nationally shareable Personally Controlled Electronic Health Record, PCEHR, is open in this respect.)
- Research data repository which is only accessed by researchers. If a research organisation has its own repository, only researchers employed by this organisation can access it. If research data linkage arrangements are in place, they are still unlikely to include Self-Quantification data.
- Government agencies' population aggregated data repositories which may be accessible to policy-makers, researchers and the public for various purposes.

5.4.1 Solution provided by FHIR for solving data exchange between various repositories

FHIR uses the RESTful principles and framework as the architecture for its systems. Doing so ensures that the system will be coherent and its components consistent with each other. For example, instead of building the system as illustrated in Image 22, the developers should implement RESTful principles as depicted in Image 23.

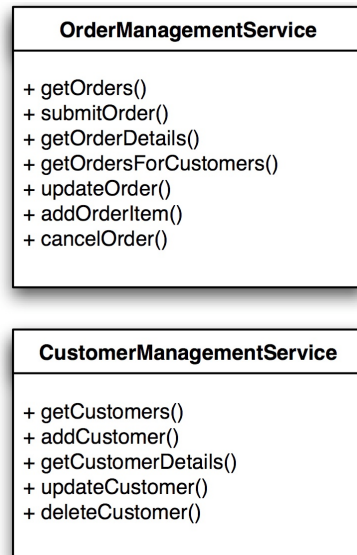


Image 20: Building a service class without RESTful standards (InfoQ, 2007)

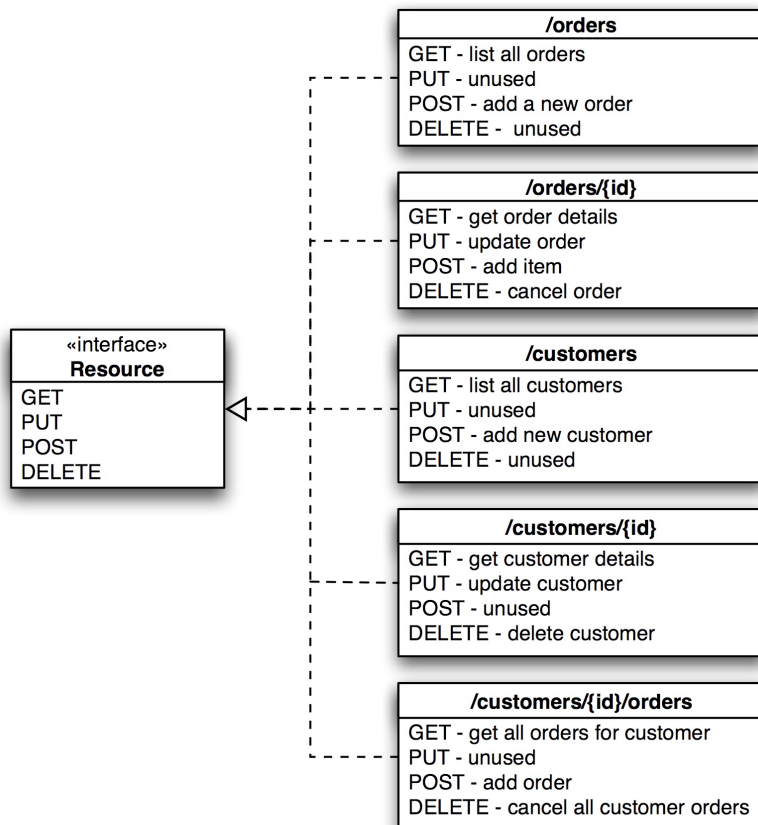


Image 21: Using RESTful standards (InfoQ, 2007)

5.4.2 Solution provided by Open mHealth for solving data exchange between various repositories

Open mHealth encourages building modules that can be incorporated within the application. These modules can access the data repository and do a particular processing job. The processed data can be used for messaging, and analysing with features such as extraction algorithms and visualisations. This developed module needs to meet Open mHealth specified features as mentioned earlier (e.g., it should be flexible, scalable, etc.). An example of a suggested solution by Open mHealth is Personal Data Vault (PDV), detailed in Appendix C.

6 Big Data Analytics


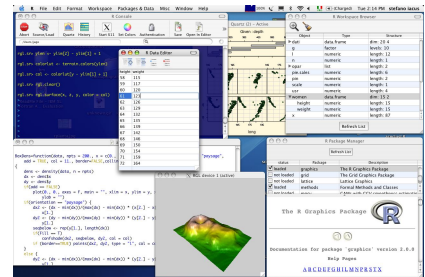

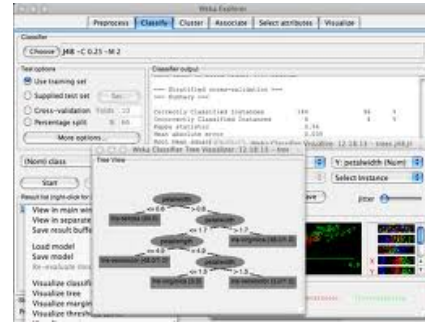



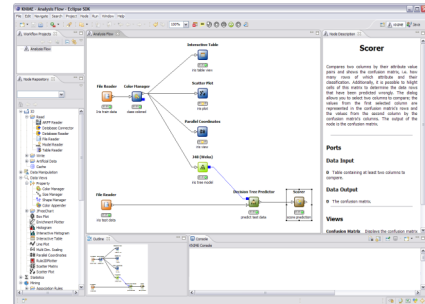
A broad range of mHealth applications is being developed at a rapid pace. Some of these apps are able to capture and record several factors that could be associated with a certain disease or health condition, such as weight and ambient temperature, and correlate these data with other metrics such as how many steps have been taken in a day. With these sophisticated tracking capabilities, smartphones and medical monitoring devices are capturing a huge volume of data, placing them in the realm of big data generators.

Big data is known for three Vs: volume, variety and velocity. Kim, Moon, Lee & Bae, (2012) define personal big data as “data created by the user’s activity that has the attribute of big data”. Personal big data records have volume, in the sense that these data are recorded over a lifetime. The devices capture a variety of personal data such as heart rate, weight, blood pressure, walking distance, calories, time, and sleep patterns. These data are used to provide personalized services in real time, and are created in streams, hence the velocity attribute.

In this section we introduce three kinds of data analytics tools that could help in understanding the generated big data: general data analytics tools, data intensive processing and analysis tools, and online analytics tools.

6.1 General Data Analytics

The following table shows the most popular big data analytics that use a visual programming approach (Sinkovits, Cicotti, Strande, Tatineni, Rodriguez, Wolter & Balac, 2011). In this type of analytics tool, a dataset is extracted from the database and exposed to the analytics application. The datasets and databases reside in separate places.

Tool	Description	Screenshot
	<p>R is an open source tool for programming, statistical computing, and data mining. It provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering) and graphical techniques, and is highly extensible (r-project, 2012).</p> <p>http://www.r-project.org</p>	
	<p>Waikato Environment for Knowledge Analysis (WEKA) is open source software for data mining and machine learning. Weka, along with R, is amongst the most popular open source software for this task. Weka is a Java-based language and includes a GUI for interacting with data files and producing visual results and graphs. WEKA is also extendable, so developers can provide additional functionality to the basic software.</p> <p>www.cs.waikato.ac.nz/ml/weka/</p>	
	<p>RapidMiner, formerly known as YALE (Yet Another Learning Environment), is an environment for machine learning, data mining, predictive analytics, and graphical representation of results. It is also extendable.</p> <p>http://rapid-i.com/content/view/181/190/</p>	
	<p>Konstanz Information Miner (KNIME) is open source software for data mining and machine learning. It allows data modelling, data analysis, and visualisation. KNIME is written in Java and has the extension feature, so developers can add plugins to provide additional functionality.</p> <p>www.knime.org</p>	

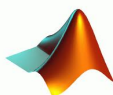
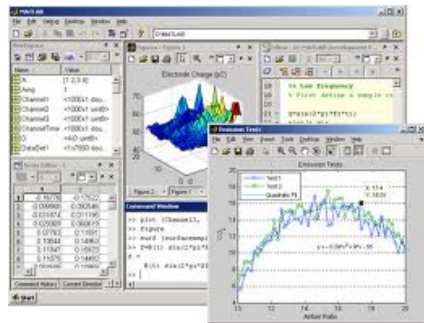

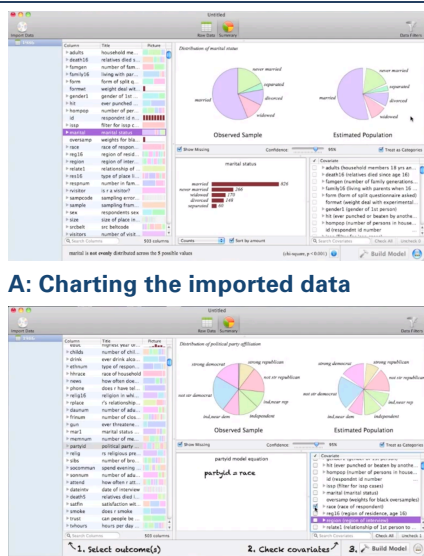
Tool	Description	Screenshot
	<p>MATLAB is statistical computing software developed by MathWorks. MATLAB allows matrix manipulations, plotting of functions and data, implementation of algorithms, visualisation, and programming. It is also extendable, so developers can provide additional functionality such as statistics, image processing and bioinformatics. www.mathworks.com.au/</p>	
 <p>Wizard for Mac A powerful but easy-to-use tool for statistics and data analysis.</p>	<p>Wizard positions itself as the first statistics program designed to make data analysis easy and fun. Researchers import their prepared datasets (e.g., csv, txt, xls) and Wizard generates various kinds of visualisations (scatterplots, pie charts, histograms) and performs simple regressions (ordinary least squares, probit and logit, and several models for count data). It can also build a learning model from the datasets. Wizard allows the researcher or analyst to perform fast analysis on prepared data and make sense out of the results.</p> <p>wizard.evanmiller.org/</p>	 <p>A: Charting the imported data</p> <p>B: Building a model</p>

Table 23: General big data analytics tools

6.2 Data Intensive Processing and Analysis

In this category, analytic logic has been implemented inside the database itself rather than in individual business intelligence or analytic applications, as in the previous type. This integration allows faster results on larger datasets; also, data do not need to be cleansed and transferred to a separate destination (Dean & Ghemawat, 2010). The following table shows some Hadoop projects for data intensive processing and analysis.


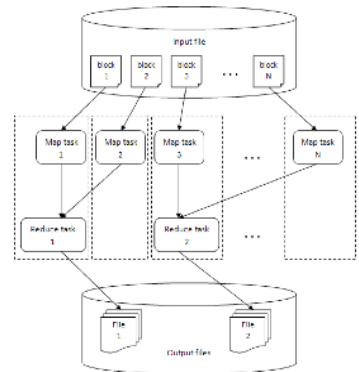


Tool	Description	Illustration
	<p>Hadoop MapReduce supports large distributed datasets on clusters of computers or nodes. It splits the input dataset into independent pieces and distributes them between processing nodes, which are processed by the map tasks in a completely parallel manner. Next the reduce tasks sort the outputs of the map tasks to make the final outputs.</p> <p>Hadoop optimizes the distribution task in a way that data communication overhead between the machines is minimal, and it can handle faulty machines (Prekopcsak, Makrai, Henk & Gaspr-Papanek, 2011).</p> <p>In addition, many bioinformatics researchers are interested in integrating the R environment with Hadoop so that it is possible to code MapReduce algorithms in R (Taylor, 2010).</p>	 <p>(Prekopcsak, Makrai, Henk & Gaspr-Papanek, 2011)</p>
	<p>Hadoop Hive is used for ad-hoc querying with an SQL-type query language, developed at Facebook (Taylor, 2010).</p> <p>It is able to run MapReduce algorithms on an unlimited number of processing nodes to execute SQL. Then any of the analytics tools can be applied to the dataset. However, some of the analytics tools (such as WEKA) have their own SQL analysis engine, and execute the SQL directly in MapReduce (Lei, Kaiping & Bin, 2011). Hadoop Hive can also create a graphical representation of results in the form of diagrams, plots, dashboards, and so forth (Cuzzocrea, Song & Davis, 2011).</p>	See Image 24
	<p>HDFS (Hadoop Distributed File System) stores file system metadata and application data separately and all servers are fully connected and communicate with each other using TCP-based protocols (Shvachko, Hairong, Radia & Chansler, 2010).</p>	See Image 24

Table 24: Data intensive processing and analysis tools

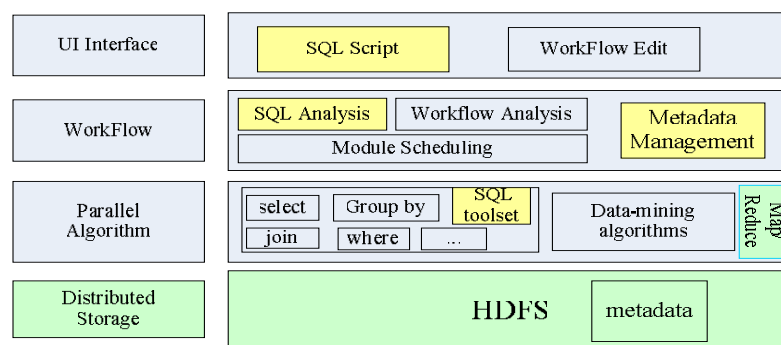


Image 22: Hive data mining system architecture with some enhancements (Lei, Kaiping & Bin, 2011).

Image 24 illustrates the four layers of data analysis in the data mining system:

- The User Interface layer provides an SQL script entrance that can be transformed into an Extraction-Transformation-Loading (ETL) workflow. The ETL function is required for mining the data.
- The WorkFlow layer includes SQL analysis alongside the traditional workflow analysis in data mining systems. The SQL analysis transforms the query into an Abstract Syntax Tree and generates an XML plan for the MapReduce job.
- The Parallel Algorithm layer. Each parallel algorithm is completed by launching one or more MapReduce job. The algorithms contain the modules that SQL needs.
- The Distributed Storage layer: **Hadoop Distributed File System** (HDFS) stores the data and its metadata.

6.3 Online analytics Tools

Another type of data visualisation tool is the web-based application. In this category, users upload their data for analysis. This category is more suitable for self-tracking systems due to its ease of use by laypeople.

Tool	Description	
	<p>Chartmyself is a web-based application that provides tools for charting multiple aspects of health and lifestyle such as vitality, body measurements, activity, food and drink consumption, menstrual cycle, symptoms and drugs. The data can be illustrated as a log or on a chart. The user needs to enter the data and select how to chart these data.</p> <p>www.chartmyself.com</p>	
Screenshots		
		
A: Enter data	B: Select the chart	C: See the result
Tool	Description	
	<p>TRAQS.me is a web-based application that provides tools for analysing and generating a detailed report of the user's stats. There are four dashboards for presenting data in different parameters. It is able to integrate data from various sources, such as Fitbit and Zeo, into a single interface. It also provides additional tracking options, such as tracking visited locations.</p> <p>http://traqs.me/</p>	
Screenshots		
		
A: Historical stats	B: Daily stats	C: Geo stats
	D: Intraday stats	



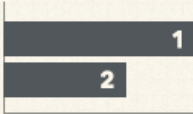
Tool		Description	
		Statwing is a website that enables users to perform basic statistical analysis on any kind of data. The user uploads the data and the results appear in seconds. https://www.statwing.com/	
Screenshots			
 1. Upload your data Paste a spreadsheet or upload a csv A: Enter data		 2. Get results in seconds Slice, dice, and visualize with ease B: Get the result	 3. Communicate your findings C: Communicate the findings

Table 25: Online analytics tools

7 Conclusion

Advances in sensor technology, in conjunction with the proliferation of mobile medical devices, have begun to expand the scope of Self-Quantification systems from solely fitness-oriented systems to more fully-fledged healthcare systems.

Currently, people are able to simultaneously track various aspects of their health and share data instantly with healthcare professionals, subject to the availability of Internet and telecommunications service provision. There are clear implications for the use of high capacity broadband to transmit health data of the varieties represented in Self-Quantification, in the volumes that may be generated as Self-Quantification becomes more widespread, and with the velocity required for timely decision support in healthcare.

However, Self-Quantification systems still need improvement.

Current problems, as discussed in this paper, are as follows:

- Interpreting data is confusing and users usually need help to understand the charts or reports that are generated. To figure out what is causing fluctuations in their readings, users must analyse the data themselves or seek help. There are no guidelines or recommendations for decision support based on health status.
- Data should be presented together. People are tracking different factors either by using a single tool or multiple tools and must consult each corresponding graph separately.
- Sharing data is difficult unless there is cooperation among service providers. People are using different tracking tools from different service providers and wish they could explore their data in a single interface.
- There is no way to find relationships within different collected datasets.
- Data analysing tools are missing. Most Self-Quantification systems do nothing beyond presenting collected data.

On a positive note, standardized protocols for transmitting data are gaining more adoption. As evidence, Self-Quantification systems that are built to use Bluetooth or WiFi technology have a better chance for user adoption than systems that use other, proprietary protocols.

8 References

- AmyT, 'Testing Out the iBGStar – First Plug-In Glucose Meter for the iPhone', viewed 18 Aug 2012 <http://www.diabetesmine.com/2011/09/testing-out-the-ibgstar-first-plug-in-glucose-meter-for-the-iphone.html>
- Bandyopadhyay, P & Kumar, S 2011, 'Image Hiding in DNA Sequence Using Arithmetic Encoding', *Journal of Global Research in Computer Science*, vol. 2, no. 4, pp. 167-71.
- Carlson, B 2008, 'Can Online Genetic Testing Predict the Future?', *Biotechnology healthcare*, vol. 5, no. 3, p. 11.
- Chang, J 2012, 'Self-Tracking for Distinguishing Evidence-Based Protocols in Optimizing Human Performance and Treating Chronic Illness', paper presented to AAAI Spring Symposium Series, North America, <http://aaai.org/ocs/index.php/SSS/SSS12/paper/view/4323/4670>.
- ChartmySelf, viewed 17 Sep 2012 <https://www.chartmyself.com/>
- Chen, C, Haddad, D, Selsky, J, Hoffman, JE, Kravitz, RL, Estrin, DE & Sim, I 2012, 'Making Sense of Mobile Health Data: An Open Architecture to Improve Individual-and Population-Level Health', *J Med Internet Res*, vol. 14, no. 4, p. e112.
- Choi, JM, Choi, BH, Seo, JW, Sohn, RH, Ryu, MS, Yi, W & Park, KS 2004, 'A System for Ubiquitous Health Monitoring in the Bedroom via a Bluetooth Network and Wireless LAN', in *Engineering in Medicine and Biology Society, 2004. IEMBS '04. 26th Annual International Conference of the IEEE*, vol. 2, pp. 3362-5.
- Cuzzocrea, A, Song, I-Y & Davis, KC 2011, 'Analytics over large-scale multidimensional data: the big data revolution!', paper presented to Proceedings of the ACM 14th international workshop on Data Warehousing and OLAP, Glasgow, Scotland, UK, DOI [10.1145/2064676.2064695](https://doi.org/10.1145/2064676.2064695).
- Dean, J & Ghemawat, S 2010, 'MapReduce: a flexible data processing tool', *Commun. ACM*, 53(1), 72-7.
- Digifit, viewed 7 Aug 2012 <http://new.digifit.com/overview/>
- Estrin, D 2010, 'Participatory sensing: from community data gathering to personal health', viewed 5 Sep 2012 <http://www.cens.ucla.edu/pub/PS-Overview-Nov2010.pdf>
- Estrin, D 2011, 'open mHealth: Reference implementation and Community Building Phase I', viewed 1 Sep 2012 openmhealth.wikispaces.com/file/view/RWJ_CHCF_Final.pdf
- Estrin, D & Sim, I 2010, 'Health care delivery. Open mHealth architecture: an engine for health care innovation', *Science (New York, N.Y.)*, vol. 330, no. 6005, pp. 759-60.
- Estrin, D & Sim, I 2011, 'Open Participatory mHealth: an opportunity for innovation in healthcare, wellness, research', viewed 1 Sep 2012 <http://www.cens.ucla.edu/pub/ParticipatoryOpenMhealth-Apple-DE080711.pdf>
- Fitbit, viewed 7 Aug 2012 <http://www.Fitbit.com/au/product>
- Fitlinxx (2012a), viewed 7 Aug 2012 <http://www.fitlinxx.net/fitlinxx-products.htm>
- Fitlinxx (2012b), viewed 7 Aug 2012 <http://www.fitlinxx.net/bodylan-wireless-protocol.htm>
- Garmin, viewed 7 Aug 2012 <http://www.garmin.com/au/products/fitness-products/>
- Gupta, N & Jilla, S 2011, 'Digital Fitness Connector: Smart Wearable System', in *Informatics and Computational Intelligence (ICI), 2011 First International Conference*, pp. 118-21.
- Happtique 2011, *Mobile health has taken the world by storm*, viewed 6 Feb 2013, http://www.happtique.com/wp-content/uploads/HAPP_booklet010212hi.pdf.
- Healthcare IT news, 'Health Level 7 International (HL7)', viewed 2 Sep 2012 <http://www.healthcareitnews.com/directory/health-level-7-international-hl7?page=3>
- Hicks, J, Ramanathan, N, Kim, D, Monibi, M, Selsky, J, Hansen, M & Estrin, D 2010, 'AndWellness: an open mobile system for activity and experience sampling', paper presented to Wireless Health 2010, San Diego, California, DOI [10.1145/1921081.1921087](https://doi.org/10.1145/1921081.1921087).
- HIMSS 2012, 'Smartphones set to revolutionize the way medicine is practiced', *Medical Device Daily* <http://www.abiresearch.com/press/1484-Mobile+Cloud+Computing+Subscribers+to+Total+Nearly+One+Billion+by+2014>
- HL7 Australia, viewed 2 Sep 2012 <http://www.hl7.org.au/CDA.htm>
- HL7, 'About HL7', viewed 2 Sep 2012 <http://www.hl7.org/about/index.cfm?ref=nav>
- HL7, 'FHIR', viewed 3 Sep 2012 <http://wiki.hl7.org/index.php?title=FHIR>
- HL7, 'HL7 Development Framework', viewed 3 Sep 2012 http://wiki.hl7.org/index.php?title=HL7_Development_Framework
- HL7, 'Resource', viewed 3 Sep 2012 <http://wiki.hl7.org/index.php?title=Resource>

HL7, 'Standards', viewed 3 Sep 2012 <http://www.hl7.org/implement/standards/index.cfm>

iBGStar, viewed 18 Aug 2012 <http://www.bgstar.com/web/ibgstar>

iBGStar, viewed 18 Aug 2012 <http://www.bgstar.com/web/ibgstar/app/app>

InfoQ 2007, 'A Brief Introduction to REST', viewed 6 Sep 2012 <http://www.infoq.com/articles/rest-introduction>

Imai, K, Kricka, LJ & Fortina, P 2011, 'Concordance study of 3 direct-to-consumer genetic-testing services', *Clinical Chemistry*, vol. 57, no. 3, pp. 518-21.

Inmon, W.H 2003, 'Information Architecture and Budget', viewed 1 Sep 2012 http://inmongif.com/_fileCabinet/gifbudarch.pdf

johnfass 2012, *BodyTrack/Fluxstream*, Word Press, viewed 12 Feb 2013, <http://johnfass.wordpress.com/2012/09/06/bodytrackfluxstream/>

Kim Y, Moon J, Lee H-J & Bae C-S, 2012, 'Knowledge Digest Engine for Personal Bigdata Analysis' in *Human Centric Technology and Service in Smart Space, Lecture Notes in Electrical Engineering* vol. 182, pp. 261-7.

KNIME, viewed 18 Sep 2012 <http://www.knime.org/>

Lei, Z, Kaiping, L & Bin, W 2011, 'The research and design of SQL processing in a data-mining system based on MapReduce', in *Cloud Computing and Intelligence Systems (CCIS), 2011 IEEE International Conference on*, pp. 301-5.

Li, I, Dey, AK & Forlizzi, J 2011, 'Understanding my data, myself: supporting self-reflection with ubicomp technologies', paper presented to Proceedings of the 13th international conference on Ubiquitous computing, Beijing, China, DOI 10.1145/2030112.2030166.

Lifescan, viewed 18 Aug 2012 http://www.onetouch.com/software_kit

Madelin, R 2012, *The European Directory of Health Apps*, PatientView, England, http://www.patient-view.com/uploads/6/5/7/9/6579846/pv_appdirectory_final_web_300812.pdf.

Matlab, viewed 18 Sep 2012 http://www.mathworks.com.au/products/matlab/?s_cid=wiki_matlab_2

Mehta, R 2011, 'The Self-Quantification Movement—Implications For Health Care Professionals', *SelfCare Journal*, vol. 2, no. 3, pp. 87-92.

Moodpanda, viewed 9 Aug 2012 <http://moodpanda.com/>

Mun, M, Hao, S, Mishra, N, Shilton, K, Burke, J, Estrin, D, Hansen, M & Govindan, R 2010, 'Personal data vaults: a locus of control for personal data streams', paper presented to Proceedings of the 6th International Conference, Philadelphia, Pennsylvania, DOI 10.1145/1921168.1921191.

Ng, PC, Murray, SS, Levy, S & Venter, JC 2009, 'An agenda for personalized medicine', *Nature*, vol. 461, no. 7265, pp. 724-6.

Nike+, viewed 7 Aug 2012 <http://nikeplus.nike.com/plus/>

Payne, T 2012, 'A phenomenal legacy for London 2012', viewed 15 Oct 2012 <http://mediacentre.dh.gov.uk/2012/08/01/a-phenomenal-legacy-for-london-2012/>

Prekopcsak, Z, Makrai, G, Henk, T & Gaspar-Papanek, C 2011, 'Radoop: Analysing Big Data with RapidMiner and Hadoop', in *RCOMM 2011: RapidMiner Community Meeting And Conference*.

R Project for Statistical Computing, viewed 15 Sep 2012 <http://www.r-project.org/>

Ramanathan, N, Alquaddoomi, F, Falaki, H, George, D, Hsieh, C, Jenkins, J, Ketcham, C, Longstaff, B, Ooms, J, Selsky, J, Tangmunarunkit, H & Estrin, D 2012, 'ohmage: An open mobile system for activity and experience sampling', in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on*, pp. 203-4.

RapidMiner, viewed 16 Sep 2012 <http://rapid-i.com/content/view/4/77/lang,en/>

Rossouw, L 2012, 'Big Data—Big Opportunities', *Risk Insights*, vol. 16, no. 2.

RunKeeper, viewed 7 Aug 2012 <http://itunes.apple.com/au/app/runkeeper-gps-running-walking/id300235330?mt=8>

Sensaris Senspod, viewed 20 Aug 2012 <http://www.sensaris.com/products/senspod/>

Shvachko, K, Hairong, K, Radia, S & Chansler, R 2010, 'The Hadoop Distributed File System', in *Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on*, pp. 1-10.

Sinkovits, RS, Cicotti, P, Strande, S, Tatineni, M, Rodriguez, P, Wolter, N & Balac, N 2011, 'Data intensive analysis on the gordon high performance data and compute system', paper presented to Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, San Diego, California, USA, DOI 10.1145/2020408.2020526.

Statwing, viewed 1 Oct 2012 <https://www.statwing.com/>

Taylor, RC 2010, 'An overview of the Hadoop/MapReduce/HBase framework and its current applications in bioinformatics', *BMC bioinformatics*, vol. 11, no. Suppl 12, p. S1.

- Turisco, F & Garzone, M 2011, *Harnessing the Value of mHealth for Your Organisation*, Computer Sciences Corporation (CSC), Chicago, viewed 6 Jan 2013, <http://www.csc.com/health_services/insights/69713-harnessing_the_value_of_mhealth_for_your_organisation>.
- uBiome, viewed 17 Feb 2013 <http://www.indiegogo.com/ubiome>
- Vashist, SK 2012, 'Non-invasive glucose monitoring technology in diabetes management: A review', *Analytica Chimica Acta*, no. 0.
- Weka, viewed 16 Sep 2012 <http://www.cs.waikato.ac.nz/ml/weka/>
- Wikipedia, 'ANT+', viewed 2 Sep 2012 <http://en.wikipedia.org/wiki/ANT%2B>
- Wikipedia, 'Health Level 7', viewed 2 Sep 2012 http://en.wikipedia.org/wiki/Health_Level_7
- Wikipedia, 'Stovepipe system', viewed 1 Sep 2012 http://en.wikipedia.org/wiki/Stovepipe_system
- Wizard for Mac, viewed 1 Oct 2012 <http://wizard.evanmiller.org/>
- Wright, A, Kemmler, C & Gibson, R 2012, *BodyTrack: Open Source Tools for Health Empowerment through Self-Tracking*, Open Source Convention (OSCON), viewed 12 Feb 2013, <http://www.oscon.com/oscon2012/public/schedule/detail/24733>
- Young, M 2012, 'Olympic anti-doping lab will become medical and genetic research centre', viewed 15 Oct 2012 http://www.bionews.org.uk/page_165941.asp
- Zeo, viewed 7 Aug 2012 <http://www.myzeo.com/sleep/shop/featured-products/zeo-sleep-manager-bedside.html> and <http://www.myzeo.com/sleep/shop/zeo-sleep-manager-mobile.html>

8.1 Image Credits

Every effort has been made to contact the copyright holders of images reproduced in this paper. If any parties are aware of any issues relating to images please contact contact@broadband.unimelb.edu.au

1. Credit: MyZeo. Source: www.myzeo.com/sleep/shop/media/catalog/custom/zeo-bedside-homepage.jpg
2. Credit: MyZeo. Source: <http://myzeo.co.uk/products/zeo-sleep-manager-mobile>
3. Credit: Roobix. Source: <http://www.roobix.co.uk/fitbit-ultra-wireless-fitness-activity-tracker-p-54.html#UWVjLxmkIkw>
4. Credit: 23andMe. Source: <https://www.23andme.com/howitworks>
5. Credit: uBiome. Source: <http://www.indiegogo.com/projects/ubiome-sequencing-your-microbiome>
6. Credit: MoodPanda. Source: <https://itunes.apple.com/us/app/moodpanda/id447452124?mt=8>
7. Credit: Sanofi. Source: www.bgstar.com/web/ibgstar
8. Credit: Sensaris. Source: www.sensaris.com/products/senspod
9. Credit: WikiLife. Source: <http://wikilife.org/blog/2012/09/14/usage-apps-and-devices-integrated-wikilife/>
10. Credit: DigiFit. Source: www.digifit.com/overview
11. Credit: Happtique. Source: www.happtique.com/wp-content/uploads/HAPP_booklet010212hi.pdf
- 12 - 14. Credit: W. Inmon. Source: Unpublished (via email).
15. Credit: D. Estrin. Source: <http://www.cens.ucla.edu/pub/PS-Overview-Nov2010.pdf>
16. Credit: D. Estrin. Source: www.cens.ucla.edu/pub/ParticipatoryOpenMhealth-Apple-DE080711.pdf
17. Credit: N. Ramanathan. Source: <http://chipts.ucla.edu/wp-content/uploads/2011/11/Ramanathan-11-8-11.pdf>
18. Credit: HL7. Source: www.hl7.org/implement/standards/fhir/v0.05/resources.htm
- 19 - 21. Credit: Stefan Tilkov. Source: www.infoq.com/articles/rest-introduction

Images in Table 23:

- Credit: R Foundation. Source: www.r-project.org/screenshots/icon-RAqua-scrshot1.jpg
- Credit: Weka. Source: <http://mloss.org/software/view/16/>
- Credit: Rapid-I. Source: <http://rapid-i.com/content/view/181/190/>
- Credit: Knime. Source: <http://www.knime.org/screenshots>
- Credit: NASA. Source: <http://www.hec.nasa.gov/news/features/2008/matlab.072508.html>
- Credit: Evan Miller. Source: http://wizard.evanmiller.org/#why_wizard

Image in Table 24 – Credit: Gabor Makrai. Source: <http://prekopcsak.hu/papers/preko-2011-rcomm.pdf>

22. Credit: Lei, Z, Kaiping, L & Bin, W. Source: <http://tinyurl.com/bbcztvf>

Images in Table 25:

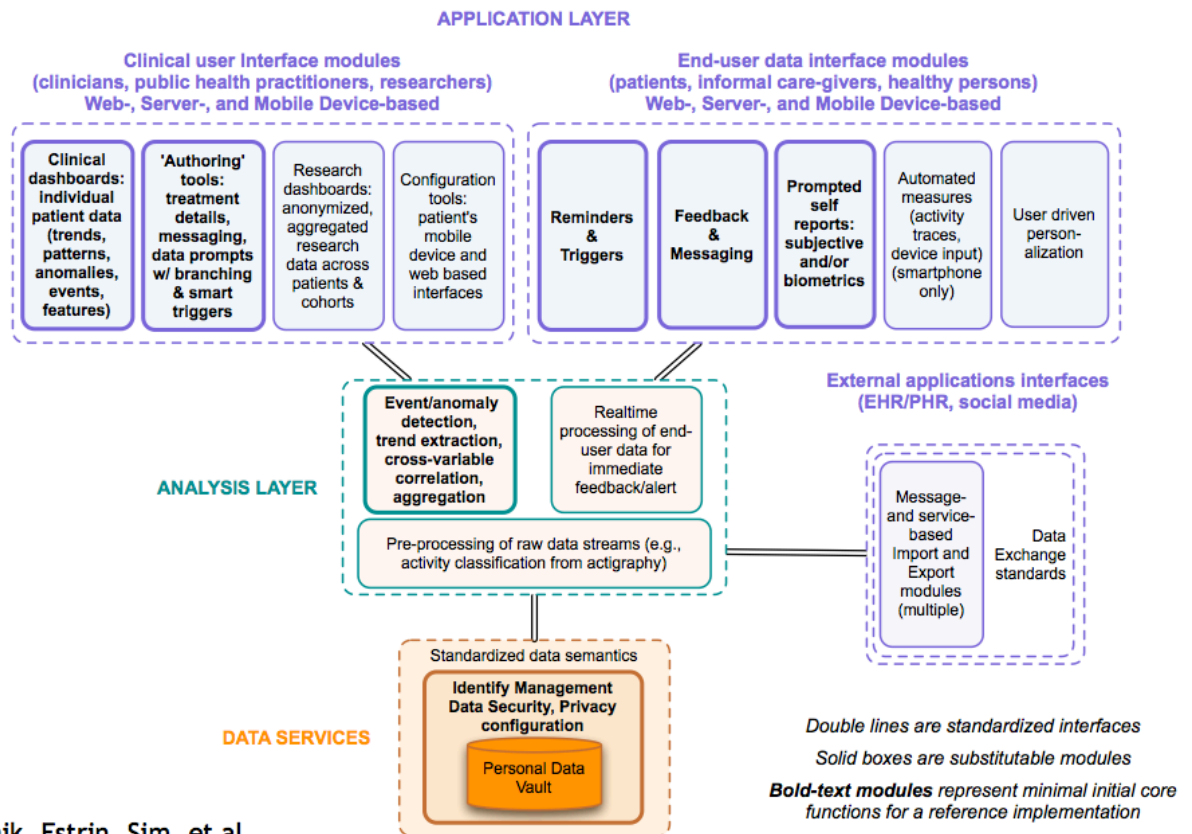
- Credit: ChartMyself. Source: www.health2apps.com/2011/08/01/chartmyself
- Credit: Eric Blue. Source: www.slideshare.net/ericblue76/traqsme-presentation
- Credit: StatWing. Source: www.statwing.com/

Images in Appendix A and Appendix B. Credit: D. Estrin. Source: www.cens.ucla.edu/pub/PS-Overview-Nov2010.pdf

Image in Appendix C. Credit: M. Mun. Source: <http://remap.ucla.edu/burke/publications/Mun-et-al-2010-Personal-Data-Vaults.pdf>

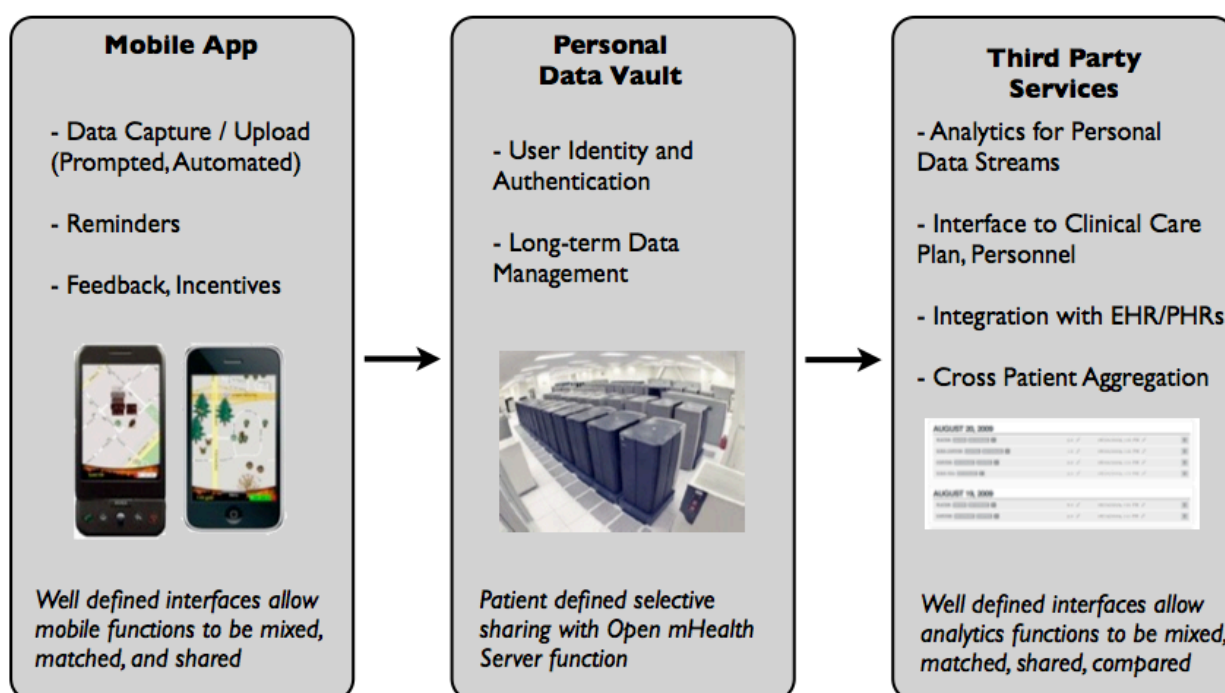
9 Appendix

Open mHealth System: detailed functional components

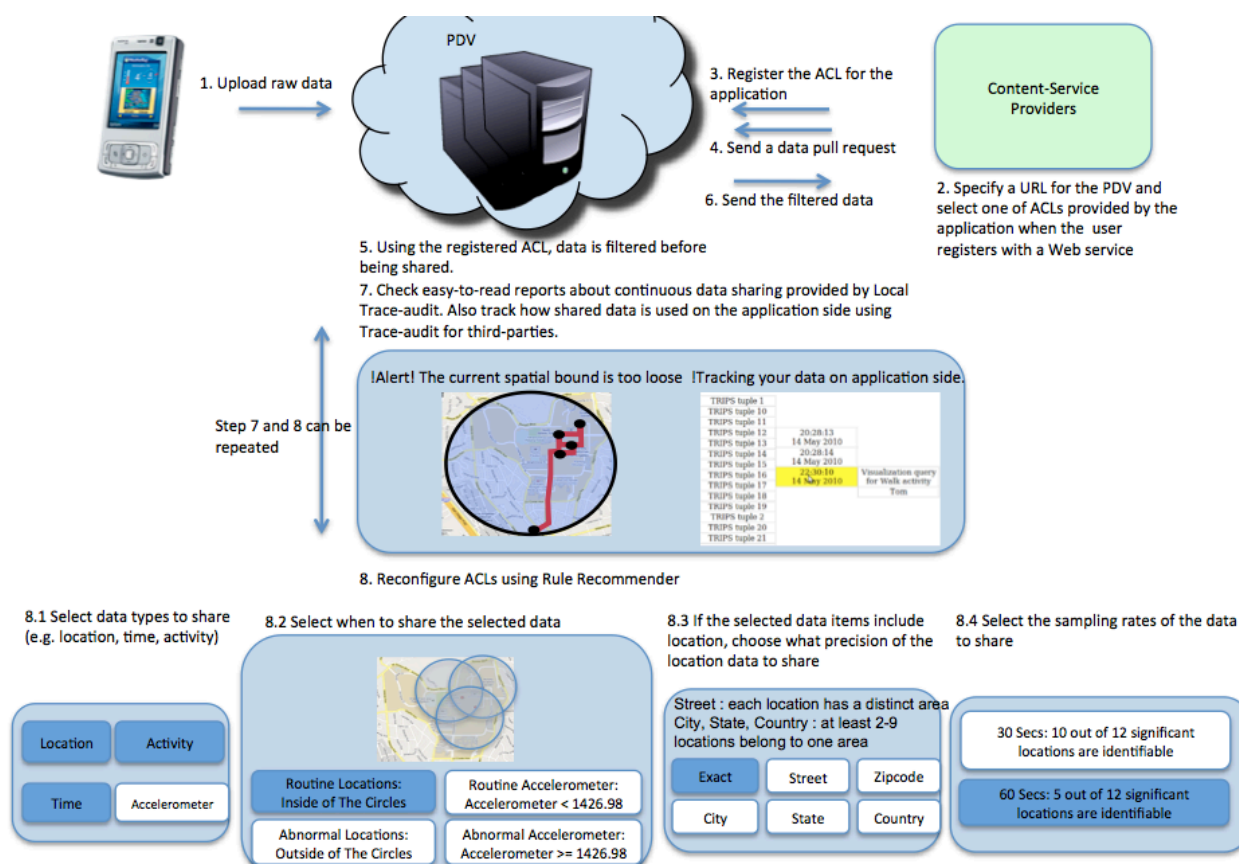


Swiernik, Estrin, Sim, et al

Appendix A: Open mHealth Architecture (Estrin, 2010)



Appendix B: PDV for privacy architecture (Estrin, 2011)



Appendix C: PDV allows participants to retain control over their raw data (Mun, Hao, Mishra, Shilton, Burke, Estrin, Hansen & Govindan, 2010)





IBES
Institute for a
Broadband-Enabled Society

Institute for a Broadband-Enabled Society

Level 4, Building 193

The University of Melbourne, Victoria 3010

e: contact@broadband.unimelb.edu.au

www.broadband.unimelb.edu.au