Big data architecture for pervasive healthcare: a literature review

Book or Report Section

Accepted Version

Tan, C., Sun, L. and Liu, K. (2015) Big data architecture for pervasive healthcare: a literature review. In: ECIS 2015 Completed Research Papers. ECIS, Paper 117. ISBN 9783000502842 Available at http://centaur.reading.ac.uk/40392/

Publisher: ECIS

BIG DATA ARCHITECTURE FOR PERVASIVE HEALTHCARE: A LITERATURE REVIEW

Complete Research

Tan, Chekfoung, Informatics Research Centre, Henley Business School, University of Reading, Whiteknights, RG6 6UD, United Kingdom, c.f.tan@pgr.reading.ac.uk

Sun, Lily, School of Systems Engineering, University of Reading, Whiteknights, RG6 6AY, United Kingdom, lily.sun@reading.ac.uk

Liu, Kecheng, Informatics Research Centre, Henley Business School, University of Reading, Whiteknights, RG6 6UD, United Kingdom; School of Information Management and Engineering, Shanghai University of Finance and Economics, Shanghai 200433, China, k.liu@henley.ac.uk

Abstract

Pervasive healthcare aims to deliver deinstitutionalised healthcare services to patients anytime and anywhere. Pervasive healthcare involves remote data collection through mobile devices and sensor network which the data is usually in large volume, varied formats and high frequency. The nature of big data such as volume, variety, velocity and veracity, together with its analytical capabilities complements the delivery of pervasive healthcare. However, there is limited research in intertwining these two domains. Most research focus mainly on the technical context of big data application in the healthcare sector. Little attention has been paid to a strategic role of big data which impacts the quality of healthcare services provision at the organisational level. Therefore, this paper delivers a conceptual view of big data architecture for pervasive healthcare via an intensive literature review to address the aforementioned research problems. This paper provides three major contributions: 1) identifies the research themes of big data and pervasive healthcare, 2) establishes the relationship between research themes, which later composes the big data architecture for pervasive healthcare, and 3) sheds a light on future research, such as semiosis and sense-making, and enables practitioners to implement big data in the pervasive healthcare through the proposed architecture.

Keywords: Pervasive Healthcare, Big Data Approaches, Health Informatics, Semiosis, Sense-making

1 Introduction

Pervasive healthcare, also known as ubiquitous healthcare or mobile healthcare, is a concept to deinstitutionalise healthcare services (Ruotsalainen et al., 2012). This concept emerged because of the accelerating operational costs due to the age related chronic diseases and the growing numbers of medical errors (Varshney, 2003; Shieh et al., 2007; Touati and Tabish, 2013). It is reported that 12.4% of the US population are over the age of 65 and this has cost the US government more than 2 trillion dollars in healthcare due to the chronic diseases (Touati and Tabish, 2013). The delivery of pervasive healthcare relies on the technological foundation that encompasses pervasive computing, sensor networks, mobile and ambience intelligence (Ruotsalainen et al., 2012; Maitland et al., 2011; Weiser, 1991). This includes telemedicine, patient monitoring, location-based medical services, incident detection, emergency response and management, pervasive access to medical data and prevention (Varshney, 2003; Varshney, 2007). Pervasive healthcare improves communication between patients and healthcare professionals by delivering accurate medical information anytime and anywhere (Varshney, 2003). This enables the real time clinical information recording and avoids information duplications, hence leading to a better information sharing and decision making (Varshney, 2003; Drayton, 2012).

Information is pivotal to pervasive healthcare, in providing the right care to the right person at the right time (Bush, 2006). Pervasive healthcare involves remote data collection such as personal health data (i.e. genomic, proteomic, epigenetic data) and environmental data (i.e. geospatial data) through mobile devices and sensor network (Ruotsalainen et al., 2012; Chen et al., 2012a; Akter and Ray, 2010). However, these collected datasets are usually large in volume, varied in formats and high in frequency (Chen et al., 2012a). These new features of data challenges the existing data management technology in ensuring the quality of the data source (Madden, 2012), and facilitating the data acquisition process for healthcare professionals to make their decisions (Feldman et al., 2012).

Given the nature of pervasive healthcare, data from distributed sources are demanded for the healthcare service provision (Heerden et al., 2012). The wide range of mobile devices, sensor and applications produce data in various formats. Big data is therefore enabling the pervasive healthcare by providing an innovative solution in managing the large, varied and high frequency data sets. Big data is commonly defined with 4Vs: Volume, Velocity, Variety and Veracity (Hurwitz et al., 2013; Fernandes et al., 2012; Chen et al., 2012b; Feldman et al., 2012). Volume refers to the rapid rate at which data is growing. Velocity represents the increasing frequency with which data is delivered or accumulated. Variety signifies the many forms in which data exist. There are two types of data defined in the big data context: structured data and unstructured data (Hurwitz et al., 2013). Veracity ensures the data quality especially for the unstructured data. Big data management provides the capabilities in data storage, management, analysis (or known as big data analytics, which is a pivotal element in the big data management) and visualisation for handling large (from terabytes to exabytes) and complex (from sensor to social media) data sets (Chen et al., 2012b).

However, most of the existing research focusses on the technical context of big data (e.g. developing decision algorithm and decision models) in the healthcare sector. There is limited research in intertwining pervasive healthcare with big data. These two domains are seen as complementing each other to better the healthcare services delivery and to reduce the accelerating operational cost. This in turn causes data related challenges. For example, when patients have multiple clinical needs and health conditions at one time, it is often difficult to use and share the data if it is not interoperable and retrievable (Estrin and Sim, 2010; Check, 2013; Heerden et al., 2012). On the other hand, data privacy and security issues are inevitable as these medical data are collected, processed and shared among multiple systems (Coiera and Clarke, 2004). Security issues are increased with the potential and unexpected loss of mobile devices and sensors, theft or inappropriate use of data (Levin, 2011). In addition, most healthcare organisations lose sight of the strategic role of big data which has a direct impact on the quality of healthcare services at the organisational level without considering the big data implementation in a holistic sense. Similar to other sectors, most healthcare organisations tend to dive into the flamboyant hype of big data without truly embracing it from the architectural perspective.

The aim of this study is therefore to propose a big data architecture for pervasive healthcare conceptually through an intensive literature review of pervasive healthcare and big data. The proposed architecture details the nuts and bolts in the domain of big data and pervasive healthcare, which are developed by the identified research themes through the literature review. The proposed architecture provides a holistic picture for healthcare organisations on how to implement big data in the pervasive healthcare setting. In addition, the proposed architecture contains component that addresses the data interoperability, privacy and security issues and radically enhances the pervasive healthcare delivery. This architecture is robust where the relationship of big data provision with pervasive healthcare is clearly shown. It covers the major mechanisms in the big data domain such as big data collection, governance, integration, analytics and visualisation from various users' perspectives. This paper is structured as follows: Section 2 illustrates the review methods adopted for this study, Section 3 discusses the review results in depth, Section 4 proposes the big data architecture for pervasive healthcare, Section 5 elucidates the contributions, limitations and further research avenues in the domain of big data and pervasive healthcare, and Section 6 draws the conclusion of the study in this paper.

2 Review Methods

The review procedure includes the review strategy, inclusion criteria, data abstraction and analysis. The review procedure is adapted from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement developed by Moher et al. (2009) to identify, select, collect and analyse the collected literature. PRISMA is a common review method used in healthcare related research. Descriptive statistic is employed to integrate the results of the review findings.

2.1 Review Strategy and Inclusion Criteria

In order to have high quality data for this study, the search is performed in four healthcare databases: MEDLINE, EMBASE, CINAHL and Health Business Elite. The keywords such as 'pervasive healthcare', 'big data', 'big data analytics' and their relevant index terms are used when conducting the search. The period of the publications is omitted in this review since both concepts are relatively new in healthcare (Groves et al., 2013; Levin, 2011). Making sense of big data and addressing its challenges is vital for the pervasive healthcare delivery (Chen et al., 2012a; Bardram, 2008; Touati and Tabish, 2013). Figure 1 shows the review strategy. The large amount of healthcare data should be processed in a right manner to improve patient safety and quality of care (Akter and Ray, 2010). Hence, the inclusion criteria for the review are: 1) the state of art of pervasive healthcare and big data, 2) methods used in implementing the big data solution such as big data analytics in the pervasive healthcare setting, and 3) data related challenges in implementing pervasive healthcare. Full papers in English that fulfil the inclusion criteria were reviewed. The risk of bias is eliminated from this review as the review does not involve any of the randomised control trials or health interventions (Dwan et al., 2008). The technological elements were selected and discussed with no bias to any solution providers in this review.

2.2 Data Abstraction and Analysis

Pervasive healthcare and big data are the emerging fields where the research questions in relation to the approaches and methods are yet to be answered (Bardram, 2008; Groves et al., 2013). Thematic analysis is therefore adopted to gain insight of these two disciplines prior to developing the architecture (Boyatzis, 1998). According to Thomas and Harden (2008), there are three steps in the thematic analysis: 1) coding text, 2) developing descriptive themes, and 3) generating analytical themes. Step one and two are conducted together. Descriptive themes are derived based on the key message that is addressed from the collected literature (e.g. "trend and challenges", "framework", "healthcare delivery", "governance", and "analytics"). The analytical themes are those descriptive themes that contribute to the development of the big data architecture for pervasive healthcare. In this review, architecture is defined as "a set of design artefacts, descriptive representations, structure of components and their inter-relationships, principles or guidelines, that are relevant for describing an object such that it can be used to derive the requirements as well as maintained over a period of time" (TheOpenGroup, 2011; Zachman, 1997). The selected analytical themes are used to construct the components of the architecture. For example, the literature coded in the "framework" theme shows that the elements such as storage, metadata, analytics and visualisation should be included when developing an architecture.

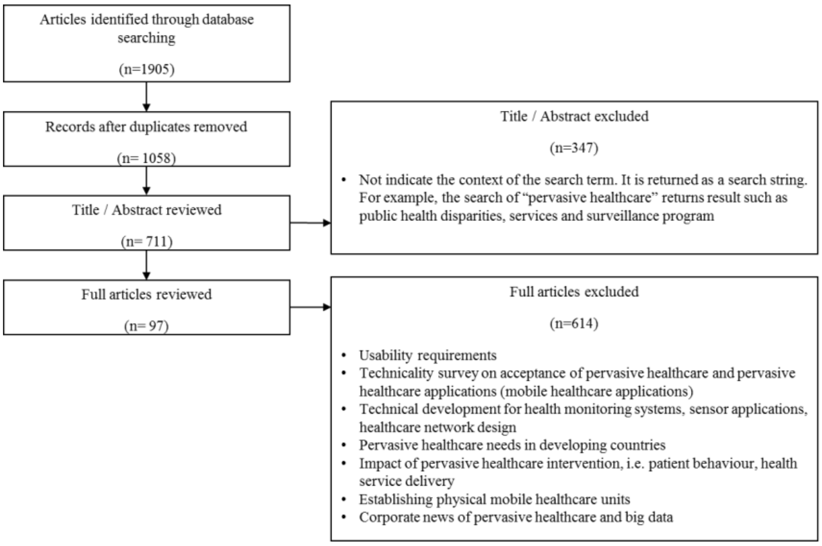

Figure 1. Review Strategy

3 Review Results

3.1 Review Themes and Descriptions

There are 1905 articles sourced for this review. As a result, 97 articles met the inclusion criteria and are reviewed in detail. The analytical themes are derived from these 97 articles and categorised into pervasive healthcare (n=25) and big data (n=72). Table 1 shows the description and the number of reviewed articles for each theme. The subsections provide an overview of pervasive healthcare implementation and big data implementation in the healthcare domain.

Pervasive Healthcare Implementation

Pervasive healthcare is defined as the use of portable devices, sensors and wireless technologies, where data is created, stored, retrieved and transmitted in real time for the purpose of improving patient care (Akter and Ray, 2010; Chao et al., 2005; Levin, 2011). Unlike the conventional delivery of healthcare services, pervasive healthcare extends the provision of healthcare services outside a hospital by heavily utilising the mobile and wireless technologies for enabling real-time data collection, monitoring, and even interactive intervention of individual patient's activities. It leads to a dependency of powerful technical infrastructure which is capable of handling big data in high frequency and delivering context-dependent personalised service. The benefits of pervasive healthcare are (Bardram, 2008; Drayton, 2012; Gaggioli and Riva, 2012): 1) enhancing continuous care provision, where patient's treatment is moved from hospitalisation to home-based treatment, 2) encouraging patient centric care provision, where patients can monitor their health condition with the assistive decision support systems, 3) improve communication among clinicians and patients, 4) avoiding data duplication as clinicians can view and share data of the clinical services involved in patient care, and 5) reducing operational costs as the needs for hospitalisation is decreased.

Review Areas	Themes	Descriptions	Articles (n)
Pervasive Healthcare	Trend and Challenges	Describes the trend, benefits, applications in general and challenges in pervasive healthcare	9
	Healthcare delivery	Illustrates how pervasive healthcare is packaged and delivered technically	2
	Framework	Depicts the components or elements that are used to construct pervasive healthcare framework that addresses data specific issues such as data interoperability and data sharing	7
	Governance	Describes data governance issues such as data privacy, data security and data encryption	7
Big Data	Applications	Shows the application domain of big data in healthcare	51
	Analytics	Explains methods that are used to facilitate data collection, data processing and data access	14
	Challenges	Describes the challenges of big data implementation in healthcare	7

Table 1. Review Themes and Descriptions

Cloud computing is one of the options in delivering pervasive healthcare due to its nature of pervasiveness and on demand service orientation (He et al., 2012). The key success factors of pervasive healthcare implementation are (Lee and Chang, 2012; Heerden et al., 2012; Touati and Tabish, 2013): 1) supporting the evidence-base and patient-centred care, 2) ensuring data interoperability, 3) fulfilling the hardware and communication requirements, 4) having the sustainability feature so it can be implemented elsewhere, and 5) imposing socio-technical design where patient care should be the main focus in the core design.

Data interoperability, privacy, and security are the core challenges in pervasive healthcare (Check, 2013; Kluge and Siegal, 2013; Luxton et al., 2012). The use of data in pervasive healthcare consists of data acquisition, data processing and data visualisation which lead to decision making on providing treatment or prescription for patients (Shieh et al., 2007). The portable devices and sensors in pervasive healthcare usually produce large volumes of high frequency data in various formats (Chen et al., 2012a). High Level 7 (HL7) or other healthcare information exchange (HIE) standards are adopted in an attempt to achieving data interoperability (Yoo et al., 2003; De Toledo et al., 2006; Torrado-Carvajal et al., 2012; Rafe and Hajvali, 2013; Cruz and Garcia, 2010). Pervasive healthcare poses data privacy and security issue as the personal information is collected, processed and shared among a number of heterogeneous devices and systems (Ruotsalainen et al., 2012). The number of the devices and systems is hard to control and it changes regularly (Ruotsalainen et al., 2011). Moreover, the use of unsecured portable devices in pervasive healthcare increases the risk of breaching patient health information (Green, 2013). This should be combated by implementing measures to prevent unauthorised access, data encryption solutions or security components such as MPEG-21 Intellectual Property Management and Protection (IPMP) (Fragopoulos et al., 2008; Silva et al., 2013).

Big Data Solution in the Healthcare Domain

Big data, together with its analytic technology revolutionises how healthcare works (Hardin, 2013). It has been applied in various healthcare areas such as: 1) optimising the clinical pathways (Schulte, 2012; Song et al., 2012; Kersten, 2013; MacDonald, 2012), 2) using predictive modelling to enhance patient care (Fox, 2012), 3) enhancing personalised care (Blobel, 2012), where medicine is developed through understanding the patterns, commonalities and correlations with patients which leads to embracing the pharmaceutical industry (Murdoch and Detsky, 2013; Serebrov, 2013; Montcheuil, 2012) (e.g. inventing drugs for HIV with big data (Bushman et al., 2013), unravelling new ways to investigate autoimmunity (Dendrou et al., 2013), modelling new drugs for type 2 diabetes (Harrison, 2012), applying large patient clinical data to run amyotrophic lateral sclerosis (ALS) trials (Leitner et al., 2012), and making accurate diagnosis by crowdsourcing the raw data (Mavandadi et al., 2012)), 4) improving healthcare services delivery performance (Sanderson, 2013), 5) enhancing public health surveillance where real time data is collected through mobile devices and sensors to provide insight on how people move and behave which helps in understanding the spread of disease (Hay et al., 2013; Talbot, 2013; Gardner, 2013; Velikic et al., 2012), 6) harnessing the conduct of comparative effectiveness research (CER) where large data sets are analysed for deriving the clinical decisions between alternate treatment strategies (Garrison Jr, 2013), and 7) contributing to genomics and proteomics research (e.g. the ribonucleic acid (RNA) research requires analysis on large scale of genomic data in order to search for a new RNA gene sequence and predict their target (Liu et al., 2012), the Encyclopedia of DNA Element (ENCODE) (Gerstein, 2012; Birney, 2012) identifies all functional elements in the human genome sequence, detects patterns and hypotheses such as the phylomemetic pattern through biological data (Chavalarias and Cointet, 2013), and exploring the genomic data for detecting diseases (Neafsey, 2013).

Big data solution is incorporated with big data analytics where certain algorithms are built for processing the data in order to gain certain insights of a specific application area (Hrickiewicz, 2012; Leventhal, 2013; Müller et al., 2012; Budišić et al., 2012; Rehkopf, 2012; Magallanes et al., 2003). There are other co-products developed on top of the big data analytics process, such as the MOLMeth monitors the process of data capture, data analysis and publications (Klingström et al., 2012), Modular API (MAPI) integrates the large data sets in different formats and protocols prior to the big data processing (Karlsson and Trelles, 2013), and Sagace, a web based search engine that retrieves a wide range of biological data such as gene expression profiles and proteomics data (Morita et al., 2012). On the other hand, data as a service (DaaS), software as a service (SaaS), platform as a service (PaaS) and infrastructure as a service (IaaS) are among the cloud services that are highly adopted in processing bioinformatics data (Dai et al., 2012). The research in cloud-based big data services helps promoting translational bioinformatics research and pervasive healthcare in general (Jalali et al., 2012; Chen et al., 2013). There are economic benefits achieved such as increasing the profitability and operating efficiencies of hospitals and health systems (Iskowitz, 2013a; Fernandes et al., 2012; Harper, 2013; Liveris, 2012; Davenport et al., 2012; Schouten, 2013). Knowledge accumulated through big data solution turns a healthcare organisation into a learning organisation (Glaser and Overhage, 2013). These benefits draw investment for big data solution in pervasive healthcare (Iskowitz, 2013b; Lewis, 2012; Özdemir et al., 2013). However, data privacy and interoperability are the challenges to be addressed (Erdmann, 2013; Hoffman and Podgurski, 2012; Schadt, 2012; Erdman et al., 2013). The Health Insurance Portability and Accountability Act (HIPAA) rule is made to protect individual's privacy where the data is de-identified prior to data sharing so that the patient's identity remained protected (White, 2013; McDavid and Bowen, 2012; Hoffman and Podgurski, 2012). However, HIPAA is not robust enough in regulating the health data collection and distribution. The interoperability of various big data sources is vital to achieve data harmonisation that has a direct impact on retrieving the right information for the right purpose (Hoffman and Podgurski, 2012).

3.2 Relationship among Research Themes

It is essential to establish the relationship between themes prior to developing the architecture. Concepts and the relationship among concepts are the main elements in a typical architectural design (Franke et al., 2009). These relationships indicate the structure of the architecture and one concept impacts on the other, which aligns with the definition of architecture established in this review. Thematic ontology is therefore introduced to portray the relationship among themes (Figure 2). Thematic ontology is a type of lightweight ontology; it is a specification of concepts and the relationship among themes (adapted from Gruber, 2008; Sun et al., 2010). Architectural research as in Tan et al. (2013) and information analytics research as in Sun et al. (2014) in the healthcare domain are adapted to establish the relationship among themes.

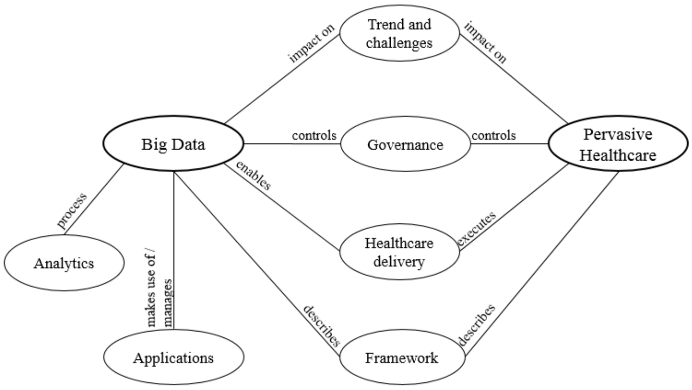

Figure 2. *Relationships between Themes*

The two main themes are *pervasive healthcare* and *big data*. *Pervasive healthcare* is executed by *healthcare delivery* services such as telemedicine, patient monitoring, location-based medical services, incident detection, emergency response and management, pervasive access to medical data and prevention. *Pervasive healthcare* is impacted by the *trend and challenges*. Trend refers to the existing needs or motivations of implementing pervasive healthcare, for example driving the operational costs down, improving patient care quality and reducing medical errors. Challenges indicate the data related issues such as data interoperability, privacy and security. Data collection in the *healthcare delivery* is enabled by *big data*. Big data *applications* make use of or manage *big data*. *Analytics* is required for data processing and analysis. The processed data is provisioned to *big data*, where data is used to enable pervasive *healthcare delivery*. For example, the patient care quality is improved through a better decision. A better decision can be made via the processed sensor data which usually comes in an unstructured format, together with the structured data captured in the electronic health records (EHR). The challenges in *trend and challenges* impact *big data*. *Governance* controls the data quality in *big data* for delivering *pervasive healthcare*. *Governance* considers the requirements from social layer, where information requirements from related processes are collected before they are facilitated in the technical layer at the data level. *Framework* describes *big data provision* for *pervasive healthcare*, with *governance* addresses the identified challenges.

3.3 Literature Gap

There is a lack of literature in the healthcare databases that addresses pervasive healthcare and big data methodologically as a complete solution. The application and motivation in adopting pervasive healthcare (n=7) and big data (n=51) are heavily discussed for achieving certain objectives. The big data analytics (n=14) literature discusses the mathematical algorithms (Budišić et al., 2012; Magallanes et al., 2003; Rehkopf, 2012) or tools (Karlsson and Trelles, 2013; Morita et al., 2012; Klingström et al., 2012) which are developed to process the large volume of data, for instance the genomic and proteomic data. However, the literature does not really illustrate the topic as a whole, for example, what are the processes that are required prior to the development of the algorithms for processing the large data set. In hindsight, the existing literature suggests the ‘what’ and the ‘why’ but does not contribute much to the ‘how’ for this review. The aforementioned gap is rectified by incorporating the big data practice in the industry into the pervasive healthcare delivery. Hadoop is a pioneer solution for big data management (Hurwitz et al., 2013; O’Reilly, 2012). It is designed for reliable and scalable distributed computing (Hadoop, 2012). Its main components are Hadoop Distribute File System (HDFS) and MapReduce (Hurwitz et al., 2013; O'Reilly, 2012; Hadoop, 2012). HDFS is a data storage cluster which facilitates the management of the related files across the machines. MapReduce is a programming model for generating and processing large data sets for structured data and unstructured data such as the large scale images. The delivery of big data solution involves six main components (Russom, 2011; Forrester, 2012; Miller and Mork, 2013; Sherman, 2013): 1) *identification* defines the motivation and requirements for big data implementation, 2) *collection* identifies where data is used and what data is required, 3) *integration* establishes a common data representation, 4) *analytics* adopts advance or discovery analytics techniques to analyse the large integrated data, 5) *visualisation* presents the analytic results in an interactive manner, and 6) *consumption* enables the decision making process.

More importantly, the reviewed literature did not seem to provide sufficient emphasis on how to achieve the "Value" of big data, which is a potential add on to the "V" series when describing the characteristics of big data. Making sense of the health data is imperative for delivering the value of pervasive healthcare such as to increase patient safety and to reduce the operational costs. Forming a hypothesis is a departure point for any sense-making process and it articulates the process of how data is collected, processed, analysed and disseminated in the pervasive healthcare setting. In hindsight, sense-making of health data is hypothesis-driven. Despite of the existing technological platforms that enable the features such as finding patterns, trends and relationship through the collected health data, it is still short of literature in postulating the methodology of making sense of these data. This imparts a new research direction which is further discussed in sub-section 6.1.

4 Big Data Architecture for Pervasive Healthcare

The big data architecture for pervasive healthcare in Figure 3 is developed from the thematic ontology (cf. Figure 2) and the review of big data practice in the industry. It demonstrates holistically the role of big data in delivering pervasive healthcare. The thematic ontology informs the fundamental constructs of the architecture. In this research context, an architectural framework from the organisational perspective should portrays the six dimensions (adapted from Sowa and Zachman, 1992): *Who* are the users of the architecture, *what* are the components in the architecture and *why* these components are required, the mechanism of big data provision in supporting pervasive healthcare (*how* and *when*), and *where* big data is stored. The architecture hence consists of four main domains that reflect the six dimensions: *stakeholders* (*who* dimension), *pervasive healthcare applications* (*what* and *why* dimensions), *pervasive healthcare data provision* (*how* and *when* dimensions), and *big data source* (*where* dimension).The architecture is described with Hadoop big data solution.

The *identification* component suggested by Russom (2011), Forrester (2012), Miller and Mork (2013), and Sherman (2013) is not included in the architecture. The assumption made is that the motivations and requirements (software and hardware) are deliberated prior to adopting the architecture. Healthcare organisations should identify the potential value derived by big data in pervasive healthcare delivery, and then decide the solution provider for implementation. *Data integration* and *data governance* component addresses the imminent data related challenges such as data interoperability, privacy and security in the pervasive healthcare delivery (Check, 2013; Kluge and Siegal, 2013; Luxton et al., 2012; Erdmann, 2013; Hoffman and Podgurski, 2012; Schadt, 2012; Erdman et al., 2013).

Stakeholders

The *stakeholders* involved in the pervasive healthcare setting are (Feldman et al., 2012; DepartmentOfHealth, 2012): 1) patients, 2) healthcare professionals (doctors, nurses and care workers), 3) government or commissioners of healthcare services (those who choose and buy services from the service providers), 4) pervasive healthcare service providers, 5) data scientists (people who make sure that information can be accessed through computer), 6) information intermediaries (organisations that take information available centrally and present it in different ways such as in interactive formats or applications), and 7) technology solution providers (people who provide the pervasive healthcare technology which enables the process of collecting, analysing and disseminating information). These *stakeholders* can be either data providers or users, depending on what context they provide and use data.

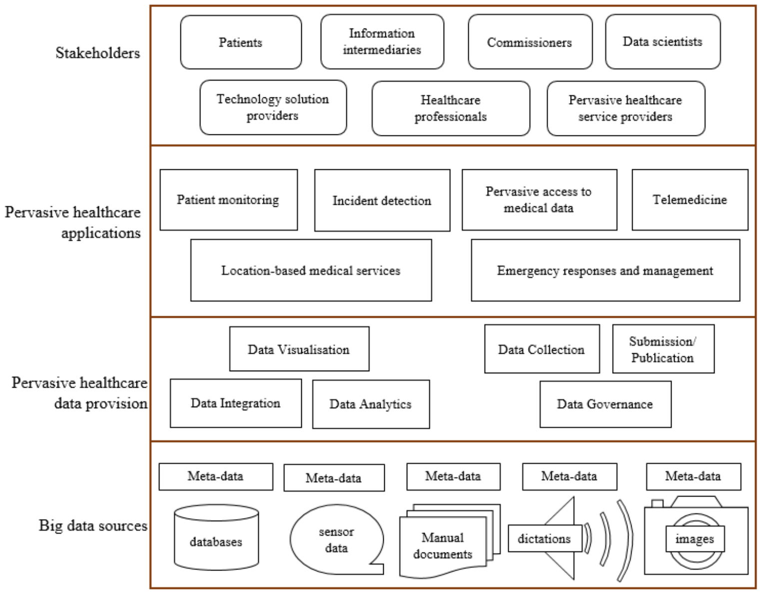

Figure 3. *Big Data Architecture for Pervasive Healthcare*

Pervasive Healthcare Applications

Pervasive healthcare applications delivers six main services aiming to improve patient safety and quality of care by optimising the value of the collected large data set (Akter and Ray, 2010; Chao et al., 2005; Levin, 2011; Drayton, 2012; Bardram, 2008; Gaggioli and Riva, 2012). These services are *telemedicine*, *patient monitoring*, *location-based medical services*, *incident detection*, *emergency responses and management* and *pervasive access to medical data* (Varshney, 2003; Varshney, 2007). Telemedicine enables healthcare professionals such as doctors and specialists to diagnose and recommend treatment to patient from a distance. *Patient monitoring* enables healthcare professionals to remotely monitor patients' health condition. *Location-based medical services* enable pervasive healthcare service provider to track the location of the elderly or the mental health patients who are ambulatory but restricted to certain areas and to help patients locating the nearest healthcare facility. *Incident detection* notifies healthcare professionals if there is a critical change on patients' health condition. *Emergency responses and management* efficiently facilitates the emergency services by dispatching the nearest ambulatory vehicles to where the emergency event happens and sending patients to the nearest hospital. *Pervasive access to medical data* allows healthcare professionals or patients to view and update the clinical records by using the mobile devices.

Pervasive Healthcare Data Provision

Pervasive healthcare data provision consists of six components: *data collection, submission / publication, data governance, data integration, data analytics* and *data visualisation*. *Data collection* indicates how data is collected, where data is used and what data is required in which context. The *submission / publication* is an acknowledgement of stakeholders before the data is offered to the *big data sources*. It is usually a set of terms and conditions that is agreed by stakeholders that the data collected from them for serving certain purposes. *Data governance* deals with the privacy and security issue by establishing a set of rules; it establishes who in the defined context has the rights in determining the data standards and managing the indicators for data quality. Data quality is assessed in terms of completeness, validity, consistency, timeliness and accuracy (Khatri and Brown, 2010). It also sets up access control rules to protect data privacy and security. It identifies the syntax, structure and semantics for the *big data sources.* The access control rule makes *providers* aware about the *data collection* process (Miller and Mork, 2013), and determines who has access to the collected data. The collection and consumption of the data should comply with standards such as HIPAA. The de-identification rule ensures patient's identity is protected (White, 2013). The security rule incorporates a data encryption solution across the sensor devices within the pervasive healthcare network. *Meta-data* standards are defined in this component. It acts as an index to the content of the big data sources and establishes the semantic of the data so that it is interpretable by the users who query the data. Hadoop Distributed File System (HDFS) and Hbase (column oriented database which can be scaled up to billions of rows) are adopted for managing *big data sources*. Hive (the data warehouse solution with SQL access) enables *big data sources* in the pervasive healthcare network.

Data integration deals with the structured and unstructured data to enable further data processing (Hurwitz et al., 2013; Sherman, 2013). It constitutes the common data representation in relation to the *big data sources*, tracks the established meta-data to facilitate further data analysis and ensures that the right data is retrieved for the right purpose. *Data integration* maintains the data provenance and involves the extract, transform and load (ETL) which processes structured data. The extract process reads data from the *big data source*. The transform process converts the extracted data into the agreed format so that they meet the data requirements and then loaded to the targeted database. The unstructured data is transformed into structured data, and then only applied to business intelligence solution to gain further insights (Hurwitz et al., 2013). *Data analytics* refers to analytics techniques to analyse the large integrated datasets in order to discover the patterned behaviour (Russom, 2011). Predictive analytics, data mining, statistical analytics, complex SQL, fact clustering, natural language processing and text analytics are the recognised techniques for big data analytics (Russom, 2011). MapReduce is the component in Hadoop which processes large data sets (Russom, 2011; Miller and Mork, 2013). *Data visualisation* presents the analytic result to the healthcare professionals in their desired format and views for effective decision making.

Big Data Sources

Big data sources consist of structured and unstructured data. Structured data is data that has predefined format such as the clinical and sensor data and can be processed by the computing equipment (Baars and Kemper, 2008). Unstructured data is data that usually does not have a specific format such as the hand written note, clinician's dictations, CT and MRI images (Zikopoulos et al., 2012).

5 Discussions

The big data architecture for pervasive healthcare is developed from the studies selected for this review and industrial practice. Thematic analysis is adopted to codify the themes. Thematic ontology, a type of light weight ontology is employed to find the relationship among themes. The themes formulate the components in the architecture. The methodological gap in establishing the architecture is complemented with the empirical big data practices from the industry. The architecture aims to provide a holistic view of how big data solution can be adopted in the pervasive healthcare delivery for healthcare organisations.

5.1 Contributions

The first contribution of this study is to identify the research themes for big data and pervasive healthcare. These research themes are beneficial for the academics and practitioners for suggesting the future research insights and key considerations prior to big data implementation in the pervasive healthcare setting. The review sets the scene by studying the literature in the four healthcare databases (MEDLINE, EMBASE, CINAHL and Health Business Elite) in order to gain an insight on the current research in the healthcare domain. These healthcare databases are suggested by the health informatics personnel from one of a National Health Services (NHS) hospital in the UK for this review. The review themes are derived through thematic analysis following the aim of this study which is to propose a conceptual view of the big data architecture for pervasive healthcare.

The second contribution is to identify the relationship between research themes via thematic ontology. A typical architectural design consists of three main elements: concepts, relationships and viewpoints (Franke et al., 2009; Lankhorst, 2009). Concept refers to the design artefacts, descriptive representations or a structure of components. Relationship reflects the behaviour between concepts which enables the process of developing the architecture. Viewpoint concerns with stakeholders' perception towards the architecture which later helps with deriving the architectural components. The relationship between the research themes is therefore established prior to the design of the proposed architecture.

The third contribution is to establish the conceptual view of the big data architecture for pervasive healthcare (cf. Figure 3) which shed a light on future research opportunities. The delivery of pervasive healthcare services requires large volume, varied format and high frequency data as opposed to the delivery of conventional healthcare services which is enabled by more controlled and structured data. The architecture clearly demonstrates how pervasive healthcare applications can be delivered, including the changes made to the delivery of pervasive healthcare services in an efficient manner by integrating the big data solutions. The architecture fulfils the five key success factors for pervasive healthcare implementation (adapted from Chen et al., 2012b; Gaggioli and Riva, 2012; He et al., 2012): 1) supports the evidence base and patient care model where healthcare interventions are made by healthcare professionals based on the analysis of the patients' data in the *data collection* component, 2) equips with the data interoperability capability which is executed by the *data integration* component, 3) addresses the hardware and the hardware and communication requirements through the *identification* component, 4) possesses the generic characteristic where each component is customisable to the adopter's requirements, and 5) reflect the socio-technical acumen where different *stakeholders'* concern are considered for designing the data pervasiveness which is enabled by big data solution.

5.2 Limitations

Limitations have been identified after the review and the consultation with the health informatics experts from one of the UK hospitals. The limitations are the literature gap and the technical issues affecting the proposed architecture. From the literature perspective, there is limited published work about big data in pervasive healthcare as an integrated solution. From the technical perspective, the informatics experts find that it remains challenging to manage and process the unstructured data in the pervasive healthcare context. They emphasise explicitly that the data provision process should focus on the semantics of the data, i.e. the right content in a right context. In addition, the *data governance* component should include the data quality concerns from the commissioners. There should be controls in each component in the architecture in order to maintain the data quality for right processing. For example, data such as dates, patients and procedure details will be kept in the staging database and only sent to the *data analytics* component once a discharge letter is issued. Furthermore, the architecture is subjected to the feedback loop from the users to revise the meta-data standards in order to improve the health outcome. They suggest that the architecture to be developed as a distributed model. For example, in the UK healthcare setting, the whole data collection process is facilitated by the General Practitioners, and the hospital takes on the responsibility for data integration and processing. This practice will reduce the implementation costs and at the same time the stakeholders involved will hold the same accountability. Then again, when the architecture is a live system, it should deliver a timeliness and accurate data for increasing the pervasiveness of the health services delivery.

6 Future Research Directions

This review suggests seven key research themes for big data implementation in the pervasive healthcare setting based on the review finding. The research themes together with the big data practice in the industry articulate the components of the big data architecture for pervasive healthcare. The review sheds a light on future research insights for the academic researchers and practitioners.

6.1 Semiosis and Sense-making of Big Data

Sense-making is vital for unleashing the full potential of big data. Big data only possess a value if it is being interpreted by the *stakeholders* with a purpose in the pervasive healthcare setting. Big data can be seen as a collection of signs according to the semiotics study (adapted from Peirce, 1935). Semiosis is a sign mediation process, or also known as a sense-making process which involves three aspects (Figure 4): 1) a sign as a signifier, vehicle or representamen (firstness), 2) an object to be signified, described or represented (secondness), and 3) an interpretant or the effect of the sign on someone who reads and interprets the sign (thirdness).

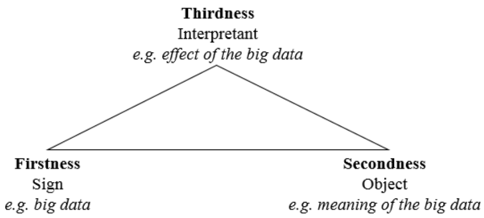

Figure 4. Semiosis and Sense-making (adapted from Liu, 2000)

In line with the semiosis perspective, the sense-making process dictates the components in *pervasive healthcare data provision.* For example, the *interpretant* which is seen as the hypothesis before the signs have been analysed determines what data should be collected for a specific context, what purpose the data should be processed for, and what way results should be visualised to support decision making. In Piercean logical system, the sense-making process is pertinent to the three reasoning approaches: abduction, deduction and deduction, which are corresponding with the firstness, secondness and thirdness (Peirce, 1935). Abduction is a form of semiotics interpretation which guides the process of forming a hypothesis through the big data collection and leads to introducing new ideas (Moriarty, 1996). Deduction is the process of explaining the hypothesis and induction is the process of evaluating the hypothesis. Future avenues of research should pay attention in the human reasoning process in order to ensure the big data is efficiently delivering the pervasive healthcare services. For instance, *stakeholders'* intention of big data usage could be abstracted in delivering or consuming (i.e. the sign users' expectations) the *pervasive healthcare applications*, the algorithm (i.e. *data analytics*) could be developed to serve a purpose and the outcome of sense-making could be reflected through *data visualisation* (i.e. what *stakeholders* expects to see and how the graphic displays guides them through reasoning process). Sun et al. (2014) demonstrates an exemplary work of the sense-making process in big data.

6.2 Social and Technical Considerations

The social and technical aspects should be considered prior to implement big data in the pervasive healthcare setting. This echoes with IOM's (2011) claim that healthcare organisation is a large socio-technical systems. In this research context, the social system is concerned with how information is comprehended and used (e.g. the business processes which are associated with the pervasive healthcare service delivery), and how it impacts on the human actions in the pervasive healthcare setting. The technical system concerns with the big data solution used to process input (e.g. data) to outputs (e.g. information) (adapted from Bostrom and Heinen, 1977; Mumford, 2006).

In addition, emphasis should be given to the infrastructure in setting up for pervasive healthcare services delivery such as coverage of wireless and mobile network, reliability of the wireless infrastructure, design of the handheld devices and sensor, and addressing the data governance and management issues (Varshney, 2007).

6.3 Architectural Techniques

The components proposed in each domain of the architecture (cf. Figure 3) lead to further development of the techniques in enabling the big data implementation in the pervasive healthcare setting. For example, the future research should focus on how to accurately elicit the information requirement from the stakeholders which are later facilitated by the big data solution. Theories such as norms analysis from organisational semiotics as in Liu (2000) can be adopted for this purpose.

In addition, focus should be given to increase the capability of delivering pervasive healthcare services with the right big data solution by incorporating the service-oriented architecture and enterprise architecture research. Service oriented architecture offers a set of design principles which enable the units of functionality to be provided and consumed as services (Lankhorst, 2009), whereas enterprise architecture ensures the technology investments is delivering the value of pervasive healthcare (Henderson and Venkatraman, 1993).

6.4 Information Management Framework

The features of *pervasive healthcare data provision* can be further developed into an information management framework for a healthcare organisation. For instance, Tan and Liu (2014) has proposed a preliminary architecture for information provision in a pervasive healthcare setting. Information management framework is vital for a healthcare organisation (Berg, 2004), especially in the pervasive healthcare setting as information is collected from various sources in different format (Reddy and Jansen, 2008). Health information management will only help in reducing operational cost when healthcare resources are utilised efficiently (adapted from Choo, 2008) for providing high quality care to patients (DepartmentOfHealth, 2010).

Future research should therefore emphasise on a suite of techniques which provide capabilities to enable *data collection, submission / publication, data governance, data integration, data analytics* and *data visualisation* in a robust manner. These techniques should clearly demonstrate information pathway as information come from heterogeneous sources. Moreover, these techniques should evolve accordingly to the changes happened in the pervasive healthcare deliver.

7 Conclusions

This study reviewed the topical literature on pervasive healthcare and big data and the result was coded in the themes. Thematic ontology is introduced to establish the relationship between the research themes. The research themes lead to developing the big data architecture for pervasive healthcare. The literature gap in terms of the methodology is identified and bridged with the empirical big data practices from the industry. This study also intended to address the data related challenges in implementing big data. The proposed architecture aimed to fulfil the key success factors as recognised in the literature. This study offers three contributions: 1) identifying the research themes of big data and pervasive healthcare, 2) establishing the relationship between research themes, and 3) proposing a conceptual view of big data architecture for pervasive healthcare. The limitations are illustrated from both literature and technical perspectives which lead to the future research opportunities such as semiosis and sense-making. The review findings benefit the academic researchers and practitioners who are engaged in the emerging research field of pervasive healthcare and big data.

References

Akter, S. & Ray, P. 2010. mHealth-an ultimate platform to serve the unserved. Faculty of Commerce-Papers, 94-100.

Baars, H. & Kemper, H.-G. 2008. Management support with structured and unstructured data—an integrated business intelligence framework. Information Systems Management, 25, 132-148.

Bardram, J. E. 2008. Pervasive healthcare as a scientific discipline. Methods of information in medicine, 47, 178-185.

Berg, M. 2004. Health information management: integrating information technology in health care work, London, Routledge.

Birney, E. 2012. The making of ENCODE: lessons for big-data projects. Nature, 489, 49-51.

Blobel, B. Co-production of Health Enabled by Next Generation Personal Health Systems. PHealth 2012: Proceedings of the 9th International Conference on Wearable Micro and Nano Technologies for Personalized Health June 26-28, 2012, Porto, Portugal, 2012. IOS Press, 52.

Bostrom, R. P. & Heinen, J. S. 1977. MIS problems and failures: A socio-technical perspective, Part II: The application of socio-technical theory. MIS quarterly, 11-28.

Boyatzis, R. E. 1998. Transforming qualitative information: Thematic analysis and code development, SAGE Publications, Incorporated.

Budišić, M., Mohr, R. & Mezić, I. 2012. Applied Koopmanism. Chaos: An Interdisciplinary Journal of Nonlinear Science, 22, 047510-047510-33.

Bush, G. W. 2006. Reforming Health Care for the 21st Century. URL: http://georgewbush-whitehouse.archives.gov/stateoftheunion/2006/healthcare/ (visited on 21-May-2013).

Bushman, F. D., Barton, S., Bailey, A., Greig, C., Malani, N., Bandyopadhyay, S., Young, J., Chanda, S. & Krogan, N. 2013. Bringing it all together: big data and HIV research. AIDS, 27, 835-838.

Chao, C. C., Jen, W. Y., Li, Y.-C., Chi, Y., Chen, C.-I. & Feng, C. C. 2005. Using mobile technology to improve healthcare service quality. Studies in health technology and informatics, 116, 352.

Chavalarias, D. & Cointet, J.-P. 2013. Phylomemetic Patterns in Science Evolution—The Rise and Fall of Scientific Fields. PloS one, 8, e54847.

Check, R. 2013. A Reality Checkpoint for Mobile Health: Three Challenges to Overcome.

Chen, C., Haddad, D., Selsky, J., Hoffman, J. E., Kravitz, R. L., Estrin, D. E. & Sim, I. 2012a. Making Sense of Mobile Health Data: An Open Architecture to Improve Individual-and Population-Level Health. Journal of medical Internet research, 14.

Chen, H., Chiang, R. H. & Storey, V. C. 2012b. Business intelligence and analytics: from big data to big impact. MIS Quarterly, 36, 1165-1188.

Chen, J., Qian, F., Yan, W. & Shen, B. 2013. Translational Biomedical Informatics in the Cloud: Present and Future. Pattern recognition, 2, 4.

Choo, C. W. 2008. FAQs on Information Management. URL: http://choo.fis.utoronto.ca/Imfaq/ (visited on 09-September-2014).

Coiera, E. & Clarke, R. 2004. e-Consent: The design and implementation of consumer consent mechanisms in an electronic environment. Journal of the American Medical Informatics Association, 11, 129-140.

Cruz, W. A. & Garcia, R. Modeling of ubiquitous technology integration process in health services. Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, 2010. IEEE, 446-449.

Dai, L., Gao, X., Guo, Y., Xiao, J. & Zhang, Z. 2012. Bioinformatics clouds for big data manipulation. Biology direct, 7, 43.

Davenport, T. H., Barth, P. & Bean, R. 2012. How 'Big Data'is Different. MIT Sloan Management Review.

De Toledo, P., Lalinde, W., Del Pozo, F., Thurber, D. & Jimenez-Fernandez, S. Interoperability of a mobile health care solution with electronic healthcare record systems. Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE, 2006. IEEE, 5214-5217.

Dendrou, C. A., Bell, J. I. & Fugger, L. 2013. Weighing in on autoimmune disease: Big data tip the scale. Nature medicine, 19, 138-139.

Departmentofhealth 2010. Liberating the NHS: An Information Revolution. London: Department of Health.

Departmentofhealth 2012. The power of information: Putting all of us in control of the health and care information we need. Department of Health.

Drayton, K. 2012. How mobile technology can improve healthcare. Nursing times, 109, 16, 18-16, 18.

Dwan, K., Altman, D. G., Arnaiz, J. A., Bloom, J., Chan, A.-W., Cronin, E., Decullier, E., Easterbrook, P. J., Von Elm, E. & Gamble, C. 2008. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. PLoS One, 3, e3081.

Erdman, A., Keefe, D. & Schiestl, R. 2013. Grand Challenge: Applying Regulatory Science and Big Data to Improve the Innovation and Approval Processes for Medical Devices.

Erdmann, J. 2013. As personal genomes join big data will privacy and access shrink? Chemistry & biology, 20, 1.

Estrin, D. & Sim, I. 2010. Open mHealth architecture: an engine for health care innovation. Science(Washington), 330, 759-760.

Feldman, B., Martin, E. M. & Skotnes, T. 2012. Big Data in Healthcare Hype and Hope. US: Dr. Bonnie 3600.

Fernandes, L., O'connor, M. & Weaver, V. 2012. Big data, bigger outcomes. Journal of AHIMA/American Health Information Management Association, 83, 38-43; quiz 44.

Forrester 2012. The Forrester Wave™: Advanced Data Visualization (ADV) Platforms, Q3 2012. Forrester Research, Inc.

Fox, B. 2012. Using big data for big impact. Health Management Technology, 32.

Fragopoulos, A., Gialelis, J. & Serpanos, D. 2008. Security framework for pervasive healthcare architectures utilizing MPEG-21 IPMP components. International journal of telemedicine and applications, 2009.

Franke, U., Hook, D., Konig, J., Lagerstrom, R., Narman, P., Ullberg, J., Gustafsson, P. & Ekstedt, M. EAF2-a framework for categorizing enterprise architecture frameworks. Software Engineering, Artificial Intelligences, Networking and Parallel/Distributed Computing, 2009. SNPD'09. 10th ACIS International Conference on, 2009. IEEE, 327-332.

Gaggioli, A. & Riva, G. 2012. From mobile mental health to mobile wellbeing: opportunities and challenges. Studies in health technology and informatics, 184, 141-147.

Gardner, E. 2013. THE HIT APPROACH TO BIG DATA. Health Data Management, 21, 34.

Garrison Jr, L. P. 2013. Universal Health Coverage-Big Thinking versus Big Data. Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research, 16, S1.

Gerstein, M. 2012. Genomics: ENCODE leads the way on big data. Nature, 489, 208-208.

Glaser, J. & Overhage, J. 2013. The role of healthcare IT: becoming a learning organization. Healthcare financial management: journal of the Healthcare Financial Management Association, 67, 56-62, 64.

Green, H. 2013. Strategies for safeguarding security of mobile computing. Healthcare financial management: journal of the Healthcare Financial Management Association, 67, 88.

Groves, P., Kayyali, B., Knott, D. & Kuiken, S. V. 2013. The 'big data' revolution in healthcare: Accelerating value and innovation. Center of US Health System Reform, Business Technology Office.

Gruber, T. 2008. Ontology. In: Liu, L. & Özsu, M. T. (eds.) Encyclopedia of Database Systems. Springer-Verlag.

Hadoop 2012. Welcome to Apache™ Hadoop®! : The Apache Software Foundation.

Hardin, S. 2013. ASIS&T annual meeting plenary speaker: Edward chang: Mobile opportunities. Bulletin of the American Society for Information Science and Technology, 39, 46-48.

Harper, E. M. 2013. The Economic Value of Health Care Data. Nursing Administration Quarterly, 37, 105-108.

Harrison, C. 2012. Deal watch:'Big data'deal for diabetes clinical trial modelling. Nature Reviews Drug Discovery, 11, 822-822.

Hay, S. I., George, D. B., Moyes, C. L. & Brownstein, J. S. 2013. Big Data Opportunities for Global Infectious Disease Surveillance. PLoS medicine, 10, e1001413.

He, C., Fan, X. & Li, Y. 2012. Toward ubiquitous healthcare services with a novel efficient cloud platform.

Heerden, A. V., Tomlinson, M. & Swartz, L. 2012. Point of care in your pocket: a research agenda for the field of m-health. Bulletin of the World Health Organization, 90, 393-394.

Henderson, J. C. & Venkatraman, N. 1993. Strategic alignment: Leveraging information technology for transforming organizations. IBM systems journal, 32, 4-16.

Hoffman, S. & Podgurski, A. 2012. Big Bad Data: Law, Public Health, and Biomedical Databases. Public Health, and Biomedical Databases (October 30, 2012). Journal of Law, Medicine and Ethics, Forthcoming.

Hrickiewicz, M. 2012. Thinking big about data. Health Facilities Management, 25, 3.

Hurwitz, J., Kaufman, M., Halper, F. & Kirsch, D. 2013. Big Data For Dummies, For Dummies.

Iom 2011. Health IT and Patient Safety: Building Safer Systems for Better Care. Washington, DC Institute of Medicine.

Iskowitz, M. 2013a. Can Big Data justify big budgets? Medical Marketing & Media, 48, 20.

Iskowitz, M. 2013b. NYU launches degree in data science. Medical Marketing & Media, 48, 26.

Jalali, A., Olabode, O. A. & Bell, C. M. 2012. Leveraging Cloud Computing to Address Public Health Disparities: An Analysis of the SPHPS. Online journal of public health informatics, 4.

Karlsson, J. & Trelles, O. 2013. MAPI: a software framework for distributed biomedical applications. Journal of biomedical semantics, 4, 1-12.

Kersten, S. K. 2013. What Is Big Data, and How Will It Affect Your Healthcare? Accurate Data Excellent Care. AHIMA Advantage, 17, 5.

Khatri, V. & Brown, C. V. 2010. Designing data governance. Communications of the ACM, 53, 148-152.

Klingström, T., Soldatova, L., Stevens, R., Roos, T. E., Swertz, M. A., Müller, K. M., Kalaš, M., Lambrix, P., Taussig, M. J. & Litton, J.-E. 2012. Workshop on laboratory protocol standards for the molecular methods database. New Biotechnology.

Kluge, E.-H. & Siegal, G. 2013. Will the rapid proliferation of mobile health, or mHealth, technology pose a threat to patient confidentiality? Ocular Surgery News, 31, 11.

Lankhorst, M. 2009. Enterprise architecture at work: Modelling, communication and analysis, Berlin Heidelberg, Springer-Verlag.

Lee, Y. & Chang, H. 2012. Ubiquitous Health in Korea: Progress, Barriers, and Prospects. Healthcare informatics research, 18, 242-251.

Leitner, M., Sherman, A., Atassi, N., Berry, J., Shui, A., Zach, N., Walker, J., Sinani, E., Katsovskiy, I., Schoenfeld, D. & Cudkowicz, M. 2012. ALS Enters the World of Big Data: Initial Description of and Results from the PRO-ACT Platform. Neurology.

Leventhal, R. 2013. Trend: big data. Big data analytics: from volume to value. Healthcare informatics: the business magazine for information and communication systems, 30, 12-14.

Levin, D. 2011. MHealth: promise and pitfalls. Frontiers of health services management, 29, 33-9; discussion 40-4.

Lewis, N. 2012. Pittsburgh Medical Center Invests$100 M In Big Data Analytics. Information Week.

Liu, C. H., Wu, D.-Y. & Pollock, J. D. 2012. Bioinformatic challenges of big data in non-coding RNA research. Frontiers in Genetics, 3.

Liu, K. 2000. Semiotics in information systems engineering, Cambridge, UK, Cambridge University Press.

Liveris, A. N. 2012. The Rise of "Big Data". Chemical and Engineering News, 90.

Luxton, D. D., Kayl, R. A. & Mishkind, M. C. 2012. mHealth data security: The need for HIPAA-compliant standardization. Telemedicine and e-Health, 18, 284-288.
Macdonald, C. 2012. Using Big Data to Improve Health. E–The Environmental Magazine, 15-16.
Madden, S. 2012. From databases to big data. Internet Computing, IEEE, 16, 4-6.
Magallanes, J. F., Zupan, J., Gomez, D., Reich, S., Dawidowski, L. & Groselj, N. 2003. The mean angular distance among objects and its relationships with Kohonen artificial neural networks. Journal of chemical information and computer sciences, 43, 1403-1411.
Maitland, J., Mcgee-Lennon, M. & Mulvenna, M. 2011. Pervasive healthcare: from orange alerts to mindcare. SIGHIT Record, 1, 38-40.
Mavandadi, S., Dimitrov, S., Feng, S., Yu, F., Yu, R., Sikora, U. & Ozcan, A. 2012. Crowd-sourced BioGames: managing the big data problem for next-generation lab-on-a-chip platforms. Lab on a Chip, 12, 4102-4106.
Mcdavid, J. & Bowen, R. 2012. Everday risk. Protecting against breach in release of information. Journal of AHIMA/American Health Information Management Association, 83, 26.
Miller, H. G. & Mork, P. 2013. From Data to Decisions: A Value Chain for Big Data. IT Professional, 15, 57-59.
Moher, D., Liberati, A., Tetzlaff, J. & Altman, D. G. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine, 151, 264-269.
Montcheuil, Y. D. 2012. Drug Development. Pharmaceutical Technology, November 2012, 22.
Moriarty, S. E. 1996. Abduction: A theory of visual interpretation. Communication Theory, 6, 167-187.
Morita, M., Igarashi, Y., Ito, M., Chen, Y.-A., Nagao, C., Sakaguchi, Y., Sakate, R., Masui, T. & Mizuguchi, K. 2012. Sagace: A web-based search engine for biomedical databases in Japan. BMC research notes, 5, 604.
Müller, H., Hanbury, A. & Al Shorbaji, N. 2012. Health Information Search to Deal with the Exploding Amount of Health Information Produced. Methods of information in medicine, 51, 516.
Mumford, E. 2006. The story of socio technical design: reflections on its successes, failures and potential. Information Systems Journal, 16, 317-342.
Murdoch, T. B. & Detsky, A. S. 2013. The Inevitable Application of Big Data to Health Care. JAMA, 309, 1351-1352.
Neafsey, D. E. 2013. ‘Big data’from shrinking pathogen populations. Molecular ecology, 22, 271-272.
O’reilly 2012. Big Data Now: 2012 Edition, O’Reilly Media.
Özdemir, V., Badr, K., Dove, E., Endrenyi, L., Geraci, C., Hotez, P., Milius, D., Neves-Pereira, M., Pang, T. & Rotimi, C. 2013. Crowd-Funded Micro-Grants for Genomics and “Big Data”: An Actionable Idea Connecting Small (Artisan) Science, Infrastructure Science, and Citizen Philanthropy. Omics: a journal of integrative biology, 17, 161-172.
Peirce, C. S. 1935. Collected Papers of Charles Sanders Peirce: Pragmaticisms and Pragnoaticism, Scientific Metaphysics, MA, US, Belknap Press.
Rafe, V. & Hajvali, M. 2013. Designing an Architectural Style for Pervasive Healthcare Systems. Journal of medical systems, 37, 1-13.
Reddy, M. C. & Jansen, B. J. 2008. A model for understanding collaborative information behavior in context: A study of two healthcare teams. Information Processing & Management, 44, 256-273.
Rehkopf, D. H. 2012. Quantile Regression for Hypothesis Testing and Hypothesis Screening at the Dawn of Big Data. Epidemiology, 23, 665-667.
Ruotsalainen, P., Blobel, B., Nykänen, P., Seppälä, A. & Sorvari, H. 2011. Framework model and principles for trusted information sharing in pervasive health. Studies in health technology and informatics, 169, 497.
Ruotsalainen, P. S., Blobel, B. G., Seppälä, A. V., Sorvari, H. O. & Nykänen, P. A. 2012. A conceptual framework and principles for trusted pervasive health. Journal of medical Internet research, 14.
Russom, P. 2011. big data analytics. TDWI Best Practices Report, Fourth Quarter.

Sanderson, M. 2013. Maximize performance with BI and big data. Comparative analytics enables organizations to benchmark performance against their peers. Health management technology, 34, 18.

Schadt, E. E. 2012. The changing privacy landscape in the era of big data. Molecular Systems Biology, 8.

Schouten, P. 2013. Big data in health care: solving provider revenue leakage with advanced analytics. Healthcare Financial Management, 67, 40.

Schulte, D. 2012. Using Big Data Analytics to Optimize Clinical Care Pathways. The Forum 12. Atlanta, GA: Care Continuum Alliance.

Serebrov, M. 2013. House Vote Has Makers of MCMs Breathing a Little Easier. BioWorld Today, 24, 1.

Sherman, R. 2013. Taking Advantages of Big Data Analytics. SearchBusinessAnalytics.com (e-publication): TechTarget Inc.

Shieh, Y. Y., Tsai, F. Y., Wang, M. & Lin, C. Mobile Healthcare: Opportunities and Challenges. Management of Mobile Business, 2007. ICMB 2007. International Conference on the, 2007. IEEE, 50-50.

Silva, B. M., Rodrigues, J. J., Canelo, F., Lopes, I. C. & Zhou, L. 2013. A Data Encryption Solution for Mobile Health Apps in Cooperation Environments. Journal of medical Internet research, 15.

Song, M. H., Park, D. K. & Lee, Y. H. 2012. Medical informatics methods for the clinical evidence extraction. Journal of the Korean Medical Association, 55, 741-747.

Sowa, J. F. & Zachman, J. A. 1992. Extending and formalizing the framework for information systems architecture. IBM systems journal, 31, 590-616.

Sun, L., Ousmanou, K. & Cross, M. 2010. An ontological modelling of user requirements for personalised information provision. Information Systems Frontiers, 12, 337-356.

Sun, L., Yamin, M., Mushi, C., Liu, K., Alsaigh, M. & Chen, F. 2014. Information analytics for healthcare service discovery. Journal of Healthcare Engineering, 5.

Talbot, D. 2013. Big Data from Cheap Phones. MIT Technology Review, 116, 51-54.

Tan, C., Liu, K. & White, E. 2013. Information Architecture for Healthcare Organizations: the case of a NHS Hospital in UK Thirty Fourth International Conference on Information Systems (ICIS2013). Milan, Italy.

Tan, C. & Liu, S. 2014. Information Architecture for Pervasive Healthcare Information Provision with Technological Implementation. In: Michell, V., Rosenorn-Lanng, D. J., Gulliver, S. R. & Currie, W. (eds.) Handbook of Research on Patient Safety and Quality Care through Health Informatics. US: IGI Global.

Theopengroup. 2011. TOGAF Version 9.1. The Open Group. URL: http://pubs.opengroup.org/architecture/togaf9-doc/arch/ (visited on 03-September-12).

Thomas, J. & Harden, A. 2008. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC medical research methodology, 8, 45.

Torrado-Carvajal, A., Rodriguez-Sanchez, M. C., Rodriguez-Moreno, A., Borromeo, S., Garro-Gomez, C., Hernandez-Tamames, J. A. & Luaces, M. Changing communications within hospital and home health care. Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE, 2012. IEEE, 6074-6077.

Touati, F. & Tabish, R. 2013. U-Healthcare System: State-of-the-Art Review and Challenges. Journal of medical systems, 37, 1-20.

Varshney, U. 2003. Pervasive healthcare. Computer, 36, 138-140.

Varshney, U. 2007. Pervasive healthcare and wireless health monitoring. Mobile Networks and Applications, 12, 113-127.

Velikic, G., Sukic, E., Jevtovic-Stoimenov, T., Bocko, M. F., Stoimenov, L. & Pentland, A. 2012. Ongoing diagnostics mapped: from an individual to the community health index. HEALTHMED, 6, 3152-3157.

Weiser, M. 1991. The computer for the 21st century. Scientific American, 265, 94-104.

White, S. E. 2013. De-identification and the Sharing of Big Data. Journal of AHIMA 84, April 2013, 44-47.

Yoo, S., Kim, B., Park, H., Choi, J. & Chun, J. Realization of real-time clinical data integration using advanced database technology. AMIA Annual Symposium Proceedings, 2003. American Medical Informatics Association, 738.

Zachman, J. A. 1997. Enterprise architecture: The issue of the century. Database Programming and Design, 10, 44-53.

Zikopoulos, P., Parasuraman, K., Deutsch, T., Giles, J. & Corrigan, D. 2012. Harness the Power of Big Data The IBM Big Data Platform, McGraw Hill Professional.