Expert Systems With Applications 152 (2020) 113388

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Review

Healthcare informatics and analytics in big data

Md. Ileas Pramanik[a,*], Raymond Y.K. Lau[b], Md. Abul Kalam Azad[a], Md. Sakir Hossain[c], Md. Kamal Hossain Chowdhury[d], B.K. Karmaker[e]

[a] *Department of Computer Science and Engineering, Begum Rokeya University, Rangpur, Bangladesh*
[b] *Department of Information Systems, City University of Hong Kong, Hong Kong, Hong Kong SAR*
[c] *Department of Computer Science, American International University-Bangladesh, Dhaka, Bangladesh*
[d] *Department of Computer Science, Comilla University, Comilla, Bangladesh*
[e] *Electrical and Electronic Engineering, Jatiya Kabi Kazi Nazrul Islam University, Mymensingh, Bangladesh*

ARTICLE INFO

Article history:
Received 8 January 2019
Revised 5 March 2020
Accepted 13 March 2020
Available online 16 March 2020

Keywords:
Healthcare informatics
Healthcare technology
Data analytics
Big data

ABSTRACT

Healthcare informatics and analytics (HCI&A), also known as healthcare information technology (HIT), healthcare IS (HIS), and so on, has rapidly evolved with the emerge of advanced data analytics technologies applied to the medical domain. Currently, HCI&A has emerged as an important area of study for both practitioners and academic researchers. Accordingly, this emerging field has prompted for an inquiry of the opportunities and challenges related to management of healthcare data, and the application of advanced data analytics to the contemporary healthcare industry. In order to contribute to the literature of healthcare informatics and analytics, this study proposes an HCI&A framework under the context of big data, which covers four important segments such as the underlying technologies, system applications, system evaluations, and emerging research areas. Based on the key features and capabilities of underpinning technologies, the evolution of HCI&A are conceptualized by three stages, namely HCI&A 1.0, HCI&A 2.0, and HCI&A 3.0. By analyzing the technological growth and current research trends, this study outlines the trend map of HCI&A for education and knowledge transfer. We also contributed to conduct a bibliographic study on healthcare informatics and healthcare information systems. To the best of our knowledge, our study is among the very few comprehensive bibliographic studies about HCI&A. We hope that our study can contribute to supplement contemporary thoughts on HCI&A research, and facilitate the related knowledge transfer to the healthcare industry.

Contents

* Corresponding author.
E-mail addresses: mpramanik2-c@my.cityu.edu.hk (Md.I. Pramanik), raylau@cityu.edu.hk (R.Y.K. Lau), akazad@brur.ac.bd (Md.A.K. Azad), sakir.hossain@aiub.edu (Md.S. Hossain), kamal.cou@gmail.com (Md.K.H. Chowdhury), bijoy_ice@jkkniu.edu.bd (B.K. Karmaker).

https://doi.org/10.1016/j.eswa.2020.113388
0957-4174/

1. Introduction

Healthcare informatics and analytics (HCI&A) represents the applications of advanced technologies and data analytics in healthcare. Recently, HCI&A and the almost parallel discipline of big data analytics have become more and more popular for both practitioners and researchers. Over the last 20 years, new inventions in information and communication technologies, artificial intelligence, and advanced data analytics have already shifted the healthcare system toward a more efficient and effective mode than ever before. Therefore, HCI&A has emerged as a dominant field of study presenting the measurement and value of data affiliated problems to be answered in current healthcare industries. Regarding these matters, recently, the President's Council of Advisors on Science and Technology (PCAST) published a report on the application of information technology in US healthcare.[1] This statistic branded a data-centric method to realize the prospects of health technologies. Moreover, according to the National Academy of Engineering, cutting-edge medical care informatics is highlighted as one of the 14 grand challenges of engineering (Yang H. et al., 2014). In business settings, advanced healthcare technologies occupy a superior position to encounter different long-term requirements (Benko and Wilson, 2003). Currently, several research initiatives have also featured the significant development of healthcare technologies. For example, the National Institute of Health (NIH) and US National Science Foundation (NSF) are working together and have issued some joint program solicitations to invite research in data science and medical informatics (i.e., Smart and Connected Health (SCH)[2]and Computational and Data-Enabled Science and Engineering (CDS&E)[3]). Moreover, the collaboration with the NIH Institute of General Medical Sciences, division of mathematical science of the NSF has also released joint initiatives to support bioinformatics research (NSF 2013) .

From the research outcomes, it is verified that information can be an extraordinary resource for healthcare industries and the benefits related to medical data and analysis in different healthcare organizations have helped create appreciable attention to HCI&A. HCI&A attracts researchers from multiple disciplines including computer science, social science, business, physics, biology, and medicine. The multidisciplinary approach to HCI&A creates some taxing challenges and promotes numerous inventive solutions. However, there is no clear and organized study about the advancement of HCI&A to date. To get a clear view of the state-of-the-art technologies, it is imperative to conduct a thorough study of the HCI&A and such study should range from evolution of the HCI concepts to the future direction of research. To fill-up such gap, in this study, we present an overall synopsis of healthcare informatics and analytics, underscoring its numerous challenges and opportunities. The key contribution of the paper is noted below:

- We summarized key concepts of informatics which enables HCI&A to exists. In this regard, the evolution of HCI&A from its beginning to the present days is presented. The roles of the computer networks, social media, ecommerce and web technologies are investigated and summarized.
- From an extensive study of the literature, we figure out the key technologies which enable the state-of-the-art HCI&A. The importance and uses of the key technologies in healthcare are figured out.
- The state-of-the-art research outcomes are summarized. We divide the research efforts into two broad categories: HCI&A research from informatics and analytics perspective, and SCI& research from information systems perspective. Research efforts and results in each category are analyzed and summarized.
- We also carry out bibliographic study of research paper published in different journals and conference to date on HCI&A. From the study, we figure out the trend of research in the field in academia and industries.
- The final contribution is to present the state of the current education systems in academia to address the needs of HCI&A.

The overall structure of the paper is shown in Fig. 1. In Section 2, the evolution of HCI&A from its genesis will be discussed in detail, which will be followed by a thorough description of the state-of-the-art HCI&A technologies in Section 3. Then, the applications of HCI&A will be discussed from the Big data point of view in Section 4. Different domains of current research on HCI&A in academia and industries are described in Section 5. We then report a bibliographic study on health informatics and information system literature in Section 6. Before concluding the paper in Section 8, we present some challenges and prospects of HCI&A education and development in Section 7.

2. HCI&A evaluation: key characteristics and capabilities

Since the birth of modern computing, around the 1950s, the term *informatics* has been used by researchers (Ledley et al., 1959) . Healthcare informatics started to be a well-liked factor for practitioners since the late 1990s, and from then all healthcare en-

[1] www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-health-it-report.pdf.

[2] www.nsf.gov/pubs/2013/nsf13543/nsf13543.htm.

[3] www.nsf.gov/funding/pgm_summ.jsp?pims_id=504813.

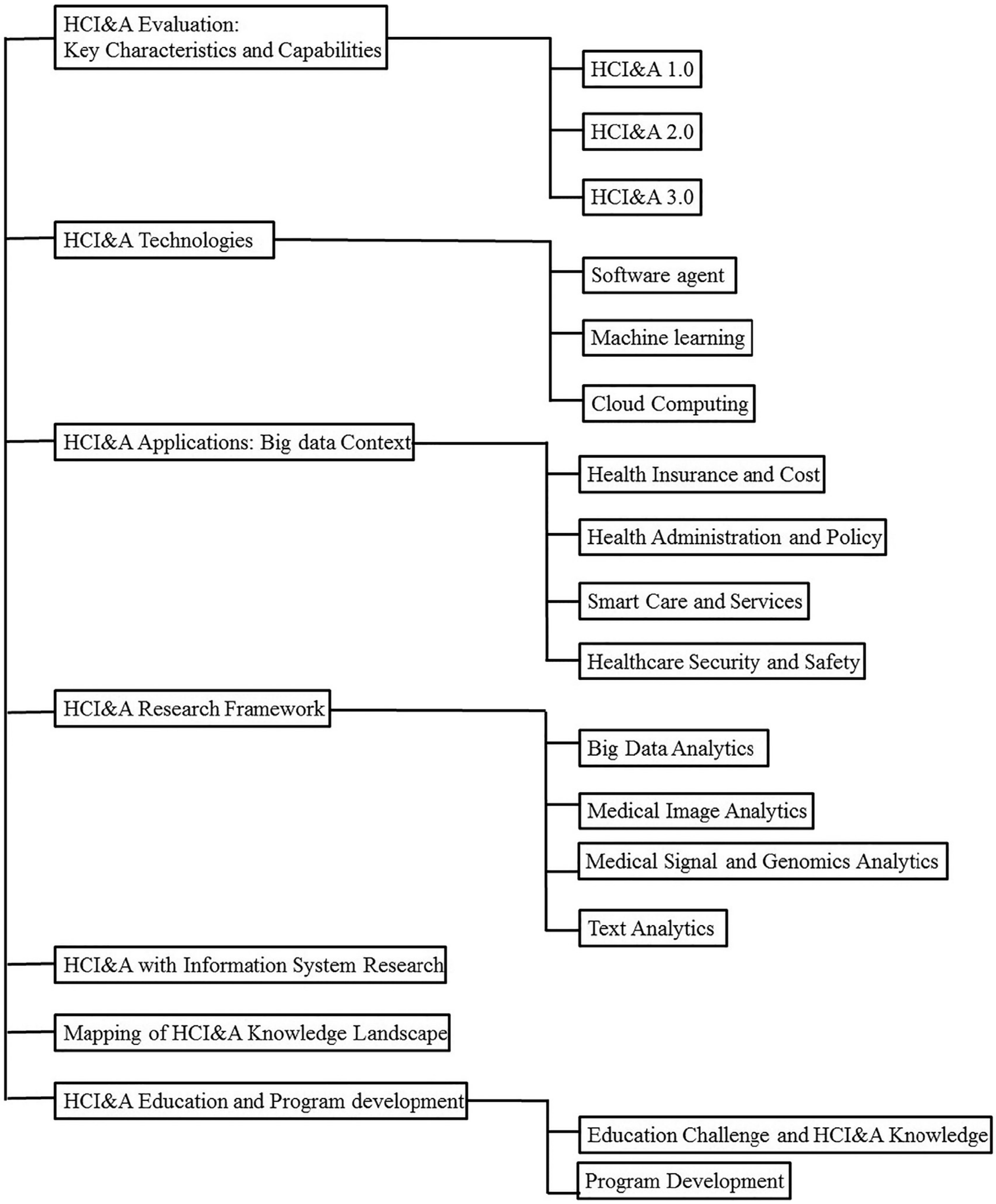


Fig. 1. . HCI&A overview: Evaluation, technologies, applications, and emerging research.

tities such as patients, physicians, and administrators have regularly been enchanted by the capabilities that advanced tools and techniques might have in medicine and medical care services (Acampora et al., 2013). More recently, big data and big data analytic methods have been widely used in healthcare applications. Big data in healthcare also introduces some complicated issues, such as non-uniform data distribution and parallel processing with a large number of variables, which are inefficiently handled by existing analytical methods. Big healthcare data is devastating not only because of its volume but also the heterogeneous nature of data and speed at which it must be managed (Frost, 2015) . In order to extract knowledge from big data, a healthcare system requires unconventional and mature data storage, management, analysis, and visualization tools and techniques. However, in this study, we use HCI&A as an integrated label and manage big data analytics as a new paradigm that enables new guidelines for HCI&A research.

2.1. HCI&A 1.0

In the context of data management and analytics, HCI&A began from the database techniques which are widely used in various healthcare settings from the 1970s (Brook et al., 1976). During that time, IBM designed a Patient Order Management and Communication System (POMCS) which was used by the Coral Gables Variety Children's Hospital in order to achieve three objectives: increase revenue, develop personnel productivity and cost savings, and improve patient care quality (Wiederhold, 1981b). These data management systems rely heavily on health data collection, extraction, and analysis technologies (Wiederhold, 1981a). As a data-centric perspective, HCI&A can be treated as HCI&A 1.0 where data is fully

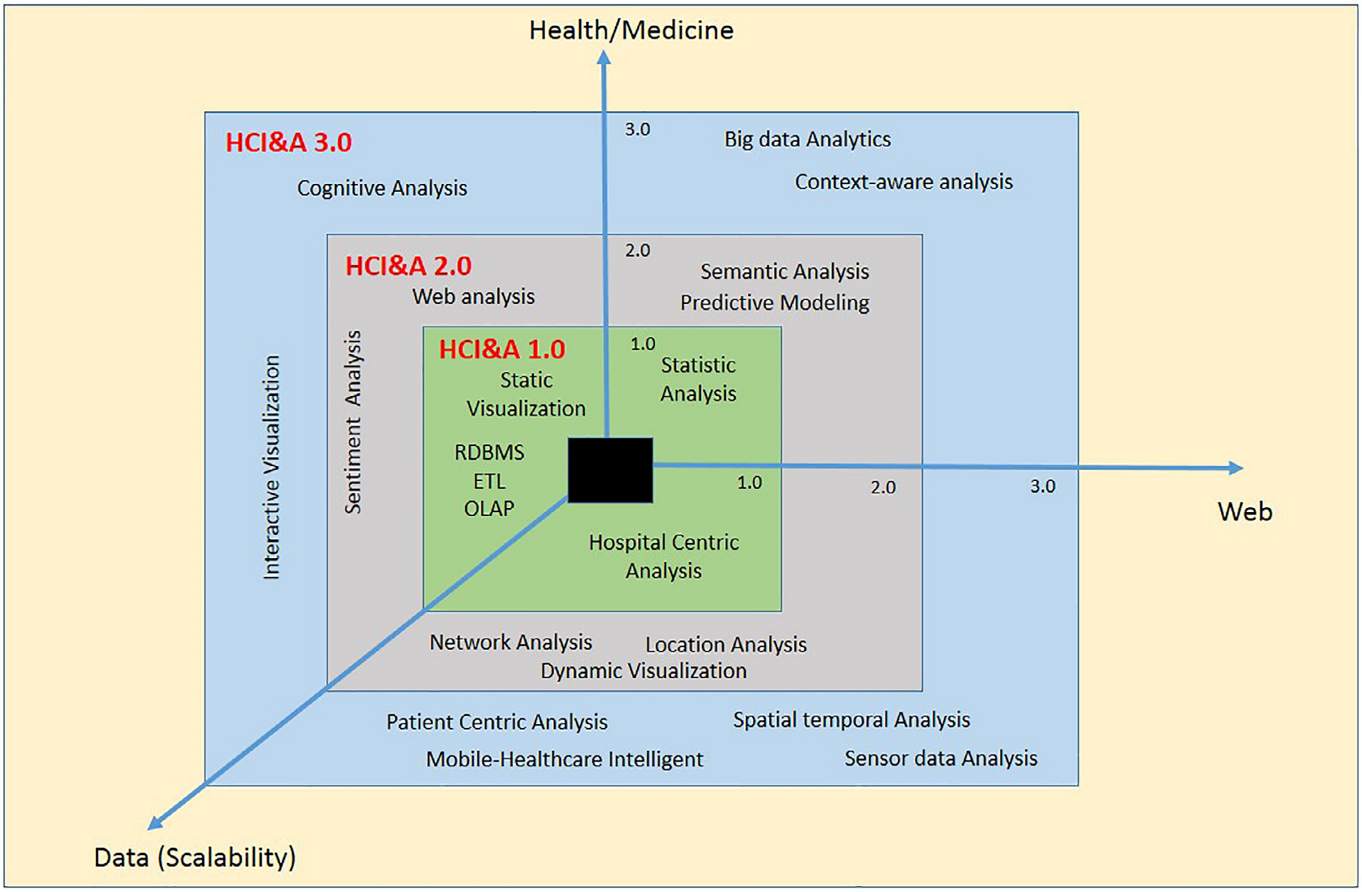


Fig. 2. . Healthcare intelligence and analytics with respect to health, web, and data.

structured, homogeneous and often stored in relational database management systems (RDBMS).

Moreover, the artificial intelligence and data analytics of the medical domain originated from three other important dimensions - health-medicine, web, and data (see Fig. 2). The concept of HCI&A 1.0 is that it is an umbrella term which mostly includes Web 1.0 technologies, Health 1.0 and Medicine 1.0 applications, services and tools. In Web 1.0, hospitals or healthcare organizations produce information without any interaction with patients; it is mainly one way of publishing content. In the healthcare context, the goal of Web 1.0 was to establish an online presence of service providers and make their information available to all customers (mainly patients) at any time. As HCI&A 1.0 includes Web 1.0 tools and techniques, its technologies include all core web protocols (i.e. HTML, HTTP), emerging web protocols (i.e. XML), and hypertext. HCI&A 1.0 technologies have some functional limitations: they are sometimes not able to forward participation between customers and service providers. Medicine 1.0 and Health 1.0 involve broad concepts which include mainly a provider centric approach. However, they define the combination of healthcare data management systems through the use of database technologies such as warehousing. HCI&A 1.0 also includes different statistical analytic tools, datamining tools and techniques for classification, segmentation, clustering, and analysis of health data. Different HCI&A tools have already been consolidated into the top commercial healthcare informatics platforms developed by IBM, Oracle, and Microsoft (Mead, 2006). In the context of software applications, HCI&A 1.0 includes some software packages for extraction, transformation and loading (ETL), online analytical processing (OLAP), database querying and reporting, data mining, and visualization. However, from data collection to knowledge discovery, there are several intelligence and analytical capabilities that must be contained in HCI&A 1.0.

2.2. HCI&A 2.0

In 2004, the term Web 2.0 was introduced to enable users to add information or content to the web, thereby allowing users to collaborate and interact using social media dialog. From that time, the internet has become increasingly popular and now forms a predominant part of our daily life (Oh et al., 2005). Web 1.0 was mostly unidirectional and so it allowed only direct interaction between providers and customers, whereas Web 2.0 allows the managing and assembling of large global crowds with common interests in social interaction (Lau et al., 2012). Moreover, Web 2.0 can deal with unstructured data resources -social media feeds, microblogs, weblogs, wikis, and other things that sometimes aren't stored in a database. Thus, in Web 2.0, patients can share their feelings and emotions on chat logs, online forums, and microblogs (Twitter), and these applications also support the design of cost saving e-healthcare solutions. When Web2.0 technologies are functionalized to the medical domain, the term health 2.0 is used to synthesize health and Web 2.0(Hughes et al., 2008) , and Medicine 2.0 represents the fusion of medicine and Web 2.0 (Van De Belt et al., 2010). However, according to the study in of Hughes B. et al. (2008), there are no potential differences between Health 2.0 and Medicine 2.0, and both of them are new concepts and are still developing (Van De Belt et al., 2010). Web 2.0 technologies, Health

2.0/Medicine 2.0, and advanced datamining technologies have ushered in a novel and thrilling era of HCI&A 2.0 research in the inaugurating period of social media, focused on data and network analytics for unstructured web content. With HCI&A 2.0 patients and professionals are enabled to collaborate with each other for developing collaborative healthcare. They transform their respective role/view using social networks where network content is heterogeneous. Moreover, in the healthcare context, RDBMS-based service information, comprehensive and IP-based exploration and reciprocal records that are accumulated smoothly through cookies and server records have become an innovative gold mine for understanding patients' demands and extracting new prospects of the health industry. HCI&A 2.0 applications can systematically assemble large volumes of timely reactions and feelings from both patients and professionals for different types of healthcare analytics.

According to the Gartner Hype cycle for healthcare (July, 2012), there are four major windows in the priority matrix for healthcare provider applications and systems. Several HCI&A 2.0 systems contain high value in the five to ten-year window (Shaffer, 2012). Moreover, Gartner healthcare platforms emphasize the application of intelligent agents, machine learning, neural networking, and text mining in order to improve quality, control operating costs, better engage patients, improve operational efficiency and meet the growing and challenging health services demand. New computer science and information system courses in artificial intelligence have emerged to address the required methodological exercise and training.

2.3. HCI&A 3.0

Web 3.0, also called the third generation of the worldwide web, promises to be a more mature and emerging web where improved 'trails' for information recovery will be designed, and an advanced scope for cognitive and intelligent processing of information will also be constructed (Lassila and &Hendler, 2007). Web 3.0 is likely to have a big effect on medicine and health (Giustini, 2007). Over the last five years, the new version of the Web Health 3.0/Medicine 3.0 has revolutionized healthcare systems through improving patient–professional communication, presenting context aware services, and offering intelligent and automated healthcare (Nash, 2008). The term HCI&A 3.0 represents the broader concept and is an umbrella term which includes all emerging concepts of Web 3.0, Medicine 3.0/Health 3.0, big data, mobile systems (e.g. m-health), and also smart systems. HCI&A 3.0 contains all mobile HCI&A technologies which are emerging and some are foreseen for the near future. Academic research on mobile HCI&A is still in its nascent phase. However, HCI&A 3.0 technologies can support a wide range of advanced applications and opportunities including mobile telemedicine, sensor-based monitoring, ubiquitous medical services, universal access to healthcare data, intensive monitoring, and lifestyle incentive management. The future HCI&A 3.0 systems will require the integration of mature and scalable techniques in big data analytics, social network analysis (SNA), and spatial-temporal analysis with existing datamining-based HCI&A 2.0 systems.

We have investigated the studies of the Gartner Hype Cycle for healthcare during 2005–2013, and summarized the key characteristics, challenges, and opportunities of HCI&A 1.0, 2.0, and 3.0. in Table 1. The healthcare business community has started to take some important steps to adopt HCI&A for its needs. IS researchers are enjoying unique opportunities for the unprecedented capabilities of HCI&A, though they are also confronting some challenges due to different uncertainties associated with emerging healthcare informatics and analytics. However, as HCI&A is a sociotechnical matter and the IS discipline is also a multidisciplinary area, IS study needs to carefully evaluate research, future plans and guidelines, and curricula from HCI&A 1.0 to HCI&A 3.0.

3. HCI&A technologies and scopes

Recently, many emerging healthcare technologies have reshaped existing HCI&A research capabilities. In the healthcare context, researchers have shown that different intelligent and analytical tools have led to more effective process redesigns of medical services (i.e., diagnosis and medicine), in decision support system (e.g. scheduling), in management (e. g. supply chain), and in communication systems (e.g. global networking systems) (Yeow and Goh, 2015). Moreover, recently, the data deluge from multiple sources has caused a new era, "Big Data," to emerge that has soundlessly descended on every section of healthcare industries, from diagnosis and surgery departments to administration and management sections. In the healthcare context, an overwhelming amount of web-based, mobile, and sensor-generated data is arriving on a terabyte and even exabyte scale (Raghupathi and &Raghupathi, 2014). New drugs, diagnoses, and prognoses discovery, and insight can be attained from the highly detailed, contextualized, and rich content of relevant data (Groves et al., 2013).

However, over the last three decades, the development of healthcare technologies at both academic and industry level has reflected the inception, growth, and maturation of several tools and techniques in different technological streams (see Fig. 2). After investigating the existing literature, we identified three major technical streams of healthcare informatics and analytics. They are software agent, machine learning, and cloud computing. Different tools and techniques under these streams are being developed over time to incorporate different definitions of healthcare problems that relate to information systems and technologies. The remaining part of this section describes each major technology with appropriate examples in the medical domain.

3.1. Software agent

In the healthcare domain, software agents are one of the most exciting research paradigms for developing software applications and they must be able to perceive the physical and virtual world around it using different sensing devices (Wooldridge, 2009). Moreover, agent-based computing systems in healthcare have been addressed as 'the next significant revolution in healthcare service development through software' (Isern et al., 2007, September) and 'the emerging software innovation to shift health paradigm to smart healthcare system' (Yuan et al., 2005). From the 1990s there has been a growing interest in the research and practice on agent-based application in healthcare (Cortés et al., July 2002)(Shankararaman, June 2000), and recently, some significant research works are also beginning to appear which reap big data benefits in the healthcare industry through multi-agent systems (MASs)[4] (Felisberto et al., 2015)(Twardowski and Ryzko, 2014, August).

Moreover, intelligent agents have already been deployed in different tasks such as health information retrieval from large volume data sources (Abasolo and Gomez, 2000, September);(Baujard et al., 1998), decision support systems for diagnosis and care (Baujard et al., 1998), patient, doctor, and nurse scheduling (Rossetti et al., 1999), co-operation among different medical components (e.g., diagnosis and medicine) to manage pervasive healthcare (Su et al., 2011, September);),(Chan et al., 2008) the development of education, training, and services (Zaharakis et al., 1998), and medical information sharing (Mohan et al., 2009). Over

[4] http://www.aamas2015.com/en/WORKSHOP-PROCEEDINGS.html.

Table 1.
HCI&A Evaluation: Key Characteristics, Challenges, and Opportunities.

Healthcare Intelligence and Analytics	Key Characteristics	Opportunities	Challenge
HCI&A1.0	• Structured data and Database Systems RDBMS, OLAP, ETL, Web1.0, Medicine 1.0/Health1.0, Statistical analysis, Hospital-centric analysis, Static visualization.	➢ Advanced service from provider to customer (Runyon et al., 2005). ➢ Develop hospital services (Runyon et al., 2007)	❖ Integrating various system (Detmer et al., 1997) ❖ Standard coordination (Bower, 2005)
HCI&A2.0	• Heterogeneous data and Data mining Web 2.0, Medicine 2.0/Health 2.0, Network analysis, Location analysis, Dynamic visualization, Semantic web analysis, Predictive modeling, Sentiment analysis	➢ Develop participatory service (Runyon et al., 2006). ➢ Include Social network in Healthcare (Shaffer, 2012). ➢ Develop patient centered medical home (Runyon, 2010)	❖ Legal and licensing problems for new technologies (Runyon, 2010). ❖ Technological Adaptation (Runyon et al., 2008). ❖ Developing new protocols for dealing with large volumes of information (Runyon, 2010)
HCI&A 3.0	• Smart Device and Big data Context-aware analysis, Spatial temporal analysis, Web 3.0, Medicine3.0/Health 3.0, Patient-centric analysis, Cognitive analysis, Interactive Visualization	➢ Ubiquitous healthcare service ➢ Include Mobile Social Network in Healthcare (Shaffer V. 2012). ➢ Enable high scalable analysis (Shaffer V., 2013) ➢ Real time healthcare system (RTHS) (Fenn and Raskino, 2011) ➢ Develop integrated clinical/financial business intelligence (BI) system (Shaffer V., 2012).	❖ Privacy and Security (Shaffer, 2012) ❖ Data complexity and heterogeneity (Sensor data, smart device data) (Fenn and Raskino, 2011)(Shaffer, 2013)

the last decade, a number of software agent tools have been included in healthcare research and practices. For instance, the National Electronic Library for Health (NeLH) (Nealon and Moreno, 2003), context-aware hospital information system (CHIS) (Tentori et al., 2006), and Virtual electronic patient record (VEPR) (Cruz-Correia et al., 2005) are the multi-agent tools for healthcare management; CARREL (Cabanillas et al., 2003), CARREL+ (Tolchinsky et al., 2006), Medical information agents (MIA) (Vermeulen et al., 2009;(Braun et al., 2007), and healthcare services (HeCaSe2) (Isern et al., 2007, September) are the multi-agent tools that support the making of appropriate medical decisions; K4Care and Aingeru are the agent tools for pervasive healthcare; and SHARE-IT (Cortés et al., 2007;Walliser et al., 2008), Geriatric Ambient Intelligence (GerAmI) (Corchado et al., 2008) are examples of integrated system agent tools.

3.2. Machine learning

In recent years, a good number of research works have been appearing in the biomedical engineering and artificial intelligent literature, which describe the application of machine learning (ML) techniques to design classifiers for anomalies (diseases and viruses) detection or medical diagnosis. Moreover, there has been a dramatic increase in the application of machine learning methods, tools, and techniques that can help solve diagnostic and prognostic complications in advanced healthcare systems. Machine learning is being used to increase the performance of clinical parameters and their combinations of the variety of medical prognoses such as prediction of disease progression, decision support for therapy or surgery, knowledge extraction from emerging research and practice, and overall healthcare system management (Magoulas and Prentza, 2001). The machine learning approach is also being used to analyze heterogeneous health data (e.g., X-ray reports, ECG reports, tomography reports, temperature, pulse, and blood pressure reports). In existing literature, we see the intellectual growth of machine learning in medical applications via several different research studies, and through investigating literature we also identify four major ML application streams in healthcare: ML in diagnosis (Kononenko, 2001), ML in prognosis and treatment (Frunza et al., 2011; Kukar et al., 1996), ML in drug discovery (Burbidge et al., 2001), and ML in surgery (Lanfranco et al., 2004).

Medical diagnostic reasoning is the most remarkable application area of intelligent systems. Therefore, currently different ML tools and techniques are well suited to analyzing medical diagnosis in specialized diagnostic problems (Kononenko, 2001). Learning-Classification-Propagation (LCP) (Zhou et al., 2008, September), expert system for thyroid disease diagnosis (ESTDD) (Keleş & Keleş, 2008) are examples of ML-based diagnosis tools. Machine learning techniques allow computers to learn from the past examples, to extract hard-to-discern patterns from high volume heterogeneous noisy datasets, and to predict intended anomalies or diseases using different statistical, probabilistic, and optimization techniques. PredictSNP (Bendl et al., 2014), CanPredict (Kaminker et al., 2007) are two well-known tools for prediction disease-related mutation. Deep learning is also used in estimating prognosis and guide therapy in patients with pulmonary hypertension and adult congenital heart diseases (Diller et al., 2019). The estimation accuracy is more than 90%. In case of drug discovery, IBM has introduced its own health ML applications in drug discovery since its early days.[5] Recently, Google has also entered into the drug discovery challenge and has started to work as a host company that raises and makes economic value by working on medicine inventions with the help of ML tools and techniques.[6] GPCRpred is another same type ML-based drug discovery tool that can predict families and subfamilies of GPCRs(G-protein coupled receptors) from the dipeptide composition of proteins (Bhasin and Raghava, 2004).

3.3. Cloud computing

In the big data era, healthcare systems have almost become a data-centric and collaborative endeavor (Almashaqbeh et al., 2014). The application of health data with different text mining or data mining algorithms entails a growing demand for dynamic, scalable resources (Koh and Tan, 2011; Raja et al., 2008). In order

[5] http://watsonhealth.ibm.com/Watson-Drug-Discovery.html?lnk=mpr_buwh.
[6] http://research.googleblog.com/2015/03/large-scale-machine-learning-for-drug.html.

to present the proper communication between different resources, cloud computing is a new, fashionable and fast rising area of development in the healthcare context (Mell and &Grance, 2011). Moreover, in the healthcare framework, permanent infrastructure investments are hard to evaluate, and adjustable on-demand services are required to meet dynamic demands (Griebel et al., 2015). Therefore, these resources are employed temporarily which can produce some viable solutions to fulfill such type of requirements. By adopting the cloud model, all technical processes of a remote organization will be migrated to the vendor's cloud-computing infrastructure where all the manipulations will be made and preserved (Armbrust et al., 2009). Recently some vendor organizations like IBM and Microsoft enable other organizations to access their (vendor) expensive infrastructure (e.g. hardware, software, and experts) through using a new "pay-as-you-go" model (Armbrust et al., 2010). Therefore, adopting cloud health organization can minimize infrastructure setup costs because the cloud computing providers will shoulder it.

In the cloud computing environment, preserving confidential sensitive healthcare data and compliance with key regulations such as the HIPAA is a matter of great concern, which may also cause a regulatory backlash and hinder further organizational innovation (Kuo, 2011). By using different privacy preserving methods and security tools, these security and privacy risks can easily be controlled and then healthcare organizations can certainly take advantage of cloud computing solutions and gain numerous benefits such as helping to improve the quality of care and minimizing overall healthcare expenditures (Muir, 2011; Wang and Tan, 2010, October).

In order to meet the healthcare demands, several third-party vendors have started to design cloud computing tools and different healthcare organizations are embedded in these vendors in customizing these tools with suitable security components to enhance their business. Cloudkick,[7] LogicMonitor (Logic Monitor 2012), and Pandora FMS (Pandora 2011) are examples of cloud computing tools, which are designed for healthcare monitoring. Moreover, by using the cloud, healthcare organizations are collaborating with their partners to create better healthcare services (Kabachinski, 2011). For instance, Microsoft HealthVault[8] is cloud-based platform where Microsoft collaborated with Kaiser Permanente, and Google Health uses a cloud named "gcloud"[9] to obtain health information from Cleveland Clinic's MyChart program (Kabachinski, 2011).

4. HCI&A applications: big data context

Healthcare Informatics and Analytics applications are growing in different dimensions in the health arena; they are now being actively diffused via systems such as electronic medical records (EMR), personal health records (PMR), clinical decision support systems (CDSS), and computerized provider order entry (CPOE) (Romanow et al., 2012). HCI&A is elaborately used and also provides leverage for different prospects generated by the large volume of medical data and multidimensional analytics in the medical domain required in various potential and high-influence application areas (Avison and Young, 2007; Goldschmidt, 2005). Four potential and high-impact applications of HCI&A are - (1) Health insurance and cost (2) Health administration and policy (3) Smart Healthcare and services, and (4) Security and privacy. In the healthcare context, different subdomains use different datasets and analytical approaches, where researchers and practitioners have to adopt or develop appropriate systems to generate the targeted results.

4.1. Health insurance and cost

Currently big data analytics and different health ITs are broadly used by insurers, personal or other healthcare funders to identify cost overruns that might constitute anomalies (Srinivasan and Arunasalam, 2013). Insurance and payment data is collected from customer claims and hospital discharges data including client opinions and behavioral data using the web. Big data enabled technologies (i.e., machine learning) allow us to consider interconnected claims and payments, whereas previously, systems tended to focus on each claim individually, and could not formally make judgments based on clues, which could possibly create abuse, waste, fraud and errors (Groves et al., 2013). Therefore, advanced Health-ITs present not just effective analytics but also relational explanations for actions that help facilitate transparent solutions. Though designing effective analytics is a challenging matter because of data complexities and enormous human factors, big data and its technologies could carry significant financial and healthcare benefits over large populations and timeframes. Big data tools and techniques are enabling private health insurance funds to recover hidden cost overruns, and are ushering in a new era of high-quality patient care at lower cost (Bates et al., 2014). For instance, in order to make effective decisions, predictive models of healthcare insurance in Australia use three levels of analytics- Admission-Level Analytics, Aggregated-level analytics, and Contract-level analytics. In the big data environment, these analytics extract anomalous admissions, and compare the performance of providers with respect to cost effectiveness and quality of care anomalies (Srinivasan and Arunasalam, 2013).

4.2. Health administration and policy

The proliferation of "smart" devices and the growth of electronic communication have generated much excitement for designing new healthcare setups and policies. As the current healthcare standard (Health/Medicine- 2.0,3.0) allows participatory, online, and rich multimedia, there is a great opportunity for adopting HCI&A applications in healthcare management and policy design. Healthcare administration and policy are related to four important factors- healthcare management, planning, resource allocation and decision-making (James et al., 2011). They will be imperative for health administration and policymakers to figure out how to overcome healthcare inconsistencies. In the big data epoch, an enormous amount of health data is being generated to facilitate organizations and persons to innovate and grow. Through converting these vast resources (data) into information the art of healthcare management system is advancing to achieve the objectives. For instance, the National Electronic Library for Health (NeLH) is a data resource that offers a portal to extract evidence-based health data on the internet (Nealon and Moreno, 2003). In order to provide effective and efficient healthcare administration (e.g., automatic resource planning, efficient scheduling among all entities, and accurate real-time medical decisions), the health industry should deploy some big data-enabled intelligent tools and techniques. Moreover, these technologies reduce overall healthcare expenditure, and according to the Mckinsey reports, such technologies can support the unlocking of more than $300 billion a year in additional costs throughout the US healthcare sector (James et al., 2011).

4.3. Smart care and services

Smart health integrates ideas from ubiquitous computing and ambient intelligence applied to predictive, personalized, preventive

[7] http://www.codeproject.com/Articles/157992/High-Availability-Cloud-Environments.

[8] http://msdn.microsoft.com/en-us/healthvault/healthvault-introduction.aspx.

[9] http://cloud.google.com/compute/docs/load-balancing/health-checks.

and participatory healthcare systems (Röcker et al., 2014). Smart health is strongly connected to the concepts of wellness and well-being (Suryadevara and Mukhopadhyay, 2014) and includes a large volume of data, collected by large amounts of biomedical sensors, (e.g., temperature, heart rate, blood pressure, and breathing rate), genomic-driven big data (genotyping, gene expression, and sequencing data), payer–provider big data (electronic health records, insurance records, and pharmacy prescriptions), and social media data (patients' status, feedback, and responses) actuators, to observe and predict patients' physical and mental conditions. Smart health is a nascent but promising field of study at the intersection of medical informatics, public health, and also business, alluding to intelligent healthcare services or enhanced cognitive capabilities through the IoT (internet of things). Big data applications in healthcare organizations can provide significant benefits which include detecting diseases at an early stage when they can be prescribed more easily and effectively. The major initiatives of the National Science Foundation (NSF) related to big health data analytics is the NSF Smart Health and Wellbeing (SHB)[10] program. The main goal of the SHB program is to address ICT issues in the big data context that support a much-needed revolution in healthcare from being reactive and hospital-centered to proactive and patient-centered, and accentuate wellbeing rather than disease control (Chen et al., 2012b).

In smart systems, healthcare workers are enabled to review and update a patient's medical data from every positional setting using handheld devices (Arnrich et al., 2010 a; Arnrich et al., 2010 b). Besides, HCI&A under integrated systems is used in the synchronization of actions that have to be performed to provide smart (autonomous, interactive, and intelligent) healthcare to citizens. These intelligent systems integrate different AI techniques with a specific purpose among different purposes (i.e. diagnoses, treatments, therapy, and surgery) under the umbrella of e-health.

4.4. Healthcare security and safety

Prognosticating threatening events and measuring healthcare security risk in real-time are highly needed and also very challenging issues in the burgeoning healthcare industry. Due to the inclusion of different advanced technological approaches, the healthcare industry is being revolutionized through presenting smart healthcare services (Agrawal et al., 2007), but otherwise industry is encountering a deluge of avant-garde attacks ranging from Distributed Denial of Service (DDoS) to secret disrupting software (Patil and &Seshadri, 2014, June; Pramanik et al., 2017a). Moreover, according to the Institute of Medicine report, security and safety issues have received a lot of attention in the US since 2000 (Institute of Medicine 2000). That report presented a real panorama of health security and safety problems in details; it also provides some guidelines on how an organization should be reformed to improve the security system in regard to healthcare. Since that report, US healthcare industries have been uninterruptedly struggling to develop coherent programs to refine security and safety systems, and also these programs have been examined substantially. In recent years, using heterogeneous data (i.e. financial and spatial data) from the advanced healthcare monitoring systems, HCI&A will steer new directions in real-time security intelligence, and using the analytical outcomes, healthcare providers can take preemptive measures before affecting the healthcare systems (Bates et al., 2014; Meingast et al., 2006, August). Table 2 summarizes the rising HCI&A applications, the nature of data, analytical approaches, and contributions in different sections of healthcare implementations.

[10] "Smart Health and Wellbeing (SBH)," Program Solicitation NSF 12-512 (http://www.nsf.gov/pubs/2012/nsf12512/nsf12512.htm; accessed August 2, 2012).

5. HCI&A research

In this section, the research related to HCI&A are analyzed and summarized. We divide the research efforts into two broad categories: research from informatics and analytics perspective and from the information systems perspective. Each perspective will be discussed in detail next.

5.1. HCI&A research framework: characteristics, tools, techniques, and emerging research in analytics

The high-impact applications of HCI&A have generated a great deal of excitement in both healthcare research and practice. In the healthcare context, research and practice are complementing each other, where researchers mainly develop analytical technologies and practitioners use them to present scalable and integrated healthcare services. In this section, we mainly describe the emerging research opportunities that we have categorized into four major technical fields - Big data analytics, Medical image analytics, Medical Signal and Genomic Analytics, and Text analytics. These analytical approaches can contribute to HCI&A1.0, HCI&A2.0, and HCI&A 3.0. These four approaches have individual characteristics, and a set of different tools and techniques, though few of these fields may leverage analogous fundamental technologies. In Table 3 we summarize the tools, technologies, and emerging research areas in each analytic technique.

5.1.1. Big data analytics

Big data analytics refers to those advanced HCI&A technologies that are involved in analyzing large-scale heterogeneous datasets, big data mining, and statistical analysis. Most of the traditional data analytic techniques can be operated with healthcare informatics tools installed in standalone systems, whereas in big data environments, datasets are broken down and executed across multiple nodes in parallel (Raghupathi and &Raghupathi, 2014). In the computer science area (i.e. artificial intelligence and datamining), over the last five years several big data analytic algorithms have been developed. From 2014 to 2016, only the IEEE has arranged 20 international conferences, workshops, and symposia on big data.[11] Big data research communities are trying to redesign most data mining algorithms (such as K-means, SVM, Naïve Bayes) in a scalable fashion where they use distributed large-scale data storage and computing mechanisms (Wu et al., 2014). In the big data environment, data mining methods are expanded in many ways including the increasing performance of single-source knowledge discovery methods (Chang et al., 2009, October), developing large-scale data mining techniques where data sources are multiple and distributed (Wu and Zhang, 2003; Wu et al., 2005). Like data mining approaches, these algorithms and methods mainly cover classification, clustering, regression, association analysis, and network analysis, the key difference lying in how they control big data challenges- volume, velocity, and variety- in data execution and manipulation processes (Chen et al., 2012a).

In the medical domain, big data sources are classified into two main classes: one is internal sources (e.g., EHR, clinical decision support systems, computerized provider order entry (CPOE), etc.), and the other is external sources (government records, insurance records, diagnoses data, pharmacies, etc.). Moreover, big data in healthcare also comes from web and social media data such as clickstream and interaction data from Facebook, Twitter, LinkedIn, blogs, and the like.

Healthcare research communities are using different big data analytic approaches on the above datasets to improve care and

[11] http://bigdata.ieee.org/conferences/past-conferences.

Table 2.
HCI&A applications in different sections of healthcare implementations.

	Health Insurance and Cost	Health Administration and Policy	Smart Care and Services	Healthcare Security and Safety
Data	Transaction recordsMayment records Provider finantial statement, Customer generated content, Medical claims data	Official rules and regulations, Resource information Feedback and comments from different entities (i.e. doctors, nurches, patients, and other employees)	Electronic health records(EHR), Treatment records Patients' feedback and comments, Genomic data DNA records Medical data (i.e. blood pressure, X-ray, ECG etc.)	Fraud records Data anomalies Finantial data Spatial data Social media data
Analytics	Anomaly detection Sentiment analysis Social network analysis Web analytics Information integration, Segmentation and Clustering	Information integration Administrative data analytics and ontologies Text analytics Performance analysis Association rule mining	Association mining and clustering, Health ontologies Social network analysis Data integration Health monitoring and analysis, Network analysis text mining Visualization	Multilingual text analytics, Financial data analytics GPS data analytics Social network analytics, Sentiment analytics Annomaly Detection Criminal network analysis Visualization
Applications	Funding and Donation systems, Recommender systems, Transparent dispertion system	Resource management Policy design Patients engagement and participation	Healthcare management Healthcare decision support, Healthcare service evaluation, Knowledge discovery Patient awarness Ubiquitous healthcare	Criminal detection Healthcare security Patient safety and smart care, Recommender system Security management
Contributions	Increase clients' satisfaction, Develop trasparency and increase healthcare support, Assure accountability	Improving adminitrative mobility, Assure appropriate actions in right place at right time, Remove congestion and recommend for effective and efficient policy	Improve healthcare (i.e. diagnosis, treatment, therapy), Patient centric healthcare Develop uninterrupted health monitoring system.	Improve healthcare security and safety.

reduce costs through discovering associations and extracting patterns, classifications, and predictions within collected information. Therefore, big data analytic applications in the medical domain have gained the benefit of the explosion of data to generate knowledge to develop well-informed decisions (jStart 2012; Knowledgent 2013). As a research category, these data-centric managements and decisions are referred to as big data analytics in healthcare systems (IBM 2012; Intel 2012). Big data analytics provides better solutions in healthcare applications across multiple scenarios, for instance: patient profiling using segmentation and predictive modeling, fraud detection by using unsupervised machine learning methods, producing new revenue streams by integrating and synthesizing patients' health data and claims datasets to deliver data and services to third parties, and so on. Big data analytic platforms are usually open-source such as Hadoop, which was developed by Apache. Hadoop enables the processing of extremely large volumes of heterogeneous data where data is petitioned into different sub-sections and then distributed to different servers to solve different portions of a large problem (Borkar et al., 2012; Frank, 2012; Zikopoulos et al., 2012). The solutions are then synthesized to present final outcomes.

5.1.2. Medical image analytics

In the healthcare domain, medical images are a key source of data commonly applied for prognosis, diagnosis, and therapy assessment (Ritter et al., 2011); therefore, image analytics has become an active field of research within HCI&A. Different HCI&A techniques such as segmentation, clustering, information extraction, and integration methods are used in medical image analytics. In clinical settings, these analytical techniques improve decision support systems through analyzing large volumes of medical images, which are increasing exponentially. The recent development of diagnostic systems as well as the increasing number of patients are the main causes of the increase of the volume of medical images. For example, 66,000 images were stored in ImageCLEF medical image dataset during 2005–2007, while in 2013 the number of images increased to 300,000 (Widmer et al., 2014, April), and this number will be 1 million in 2017. Though the growing volume of medical images introduces some new challenges for integration and mining, it effectively helps to improve diagnostic accuracy and also reduces the time and cost taken for diagnosis (Dougherty, 2009).

In the data dimension context, medical images might have two, three and even four dimensions. For instance, computed tomography is treated as four-dimensional and 3D ultrasound is three-dimensional medical data. Some medical images also have excessive resolutions such as microscopic scans of human brains, which are high-resolution images and require 66 TB of storage space (Scholl et al., 2011). Its multidimensional and high-resolution character make medical data more complex, complicated, and voluminous, which are also the characteristics of big data.

Recently some research works have developed some methods and frameworks to improve medical image processing in the big data environment. A MapReduce framework has been proposed by (Markonis et al., 2015), which is used for large-scale image processing based on parallel computing and algorithm optimization. In (Chen et al., 2010, December), the researchers developed a computer-aided decision support system to assist doctors to formulate accurate treatment planning for patients suffering from traumatic brain injury (TBI). An automatic data collection technique using a patient information processing software from military onboard ships for TBI classification is proposed in (Rodger, 2016). Then, capturing data from the images are done and inputted to database. Then, multiple ship databases are used for analyzing data and reporting results. The software used for data collection is hybrid Hadoop Hive. The proposed system shows that the mortality, survival and morbidity rates can be extracted from the medical operation data. Such information can be used for future decision making and planning. A hybrid digital-optical correlator (HDOC) was developed in (Zheng et al., 2014) in which molecular images are used to predict advanced ovarian cancer. Though a number of advanced image processing algorithms and meth-

Table 3.
Tools, technologies, and emerging research areas within HCI&A.

	Key Characteristics /Challenges of Datasets	Emerging Research	Technologies	Tools/System
Big data Analytics	➢ Volume (Raghupathi and &Raghupathi, 2014) ➢ Velocity (Raghupathi and &Raghupathi, 2014) ➢ Variety (Raghupathi and &Raghupathi, 2014) ➢ Veracity (Feldman et al., 2012), and ➢ Value (LaValle et al., 2011)	❖ Cloud computing ❖ Social network Analysis ❖ Statistical predictive model ❖ Privacy and Security analytics ❖ Real-time healthcare system ❖ Machine learning, Neural network, and Text mining. ❖ Hadoop, Mapreduce based healthcare data analytics ❖ Smart healthcare ❖ Sentiment Analytics	❖ Distributed computing ❖ Clustering, classification, segmentttion, and data integration ❖ Genetic algorithms ❖ Annomaly detection ❖ Association rule mining ❖ Optimization	• Hadoop • MapReduce • Cassandra • Zookeeper • Mahout
Medical Image Analytics	➢ Volume (Von Landesberger et al., 2013) ➢ Value (Belle et al., 2015) ➢ Variety (Seibert, 2009)	❖ Image processing ❖ Hybride Medical Image Processing ❖ Medical diagnoses ❖ Statistical predictive model ❖ Optical signal analytics ❖ Data correlation and dependencies ❖ Computer graphics ❖ Mapreduce in medical image analytics	❖ Cognitive radiology ❖ Image segmentation ❖ Data Modeling ❖ Pattern recognation ❖ SVM ❖ Optimal control ❖ Filtering	• Hadoop • MapReduce • Hybrid HDOC (T. Zheng, Cao, He & Jin, 2014)
Medical Signal and Genomic data analytics	➢ Volume (Pramanik et al., 2017b) ➢ Level of scrutiny for privacy(Berndt et al., 2001), and ➢ Complexity (for spatiotemporal nature) (Belle et al., 2015).	❖ Time series data analytics ❖ Wave analytics ❖ Time series data analytics ❖ Machine learning ❖ Genomic Structure ❖ MapReduce ❖ Cloud computing ❖ Visualization ❖ Mobile web analytics ❖ Sensor data analytics	❖ Pattern recognation ❖ Frequency analysis ❖ Fuzzy logic ❖ Component analysis ❖ Genomic signal processing ❖ Radiology ❖ Fourier transformation ❖ Wave filtering	• Hadoop • MapReduce • iPOP (H. Chen et al., 2012) • ClueGo • GoMiner • GESA • Onto-Express
Text analytics	➢ Volume (Raghupathi and &Raghupathi, 2014) ➢ Unstructured (Chen et al., 2005), ➢ and ➢ Veracity (Connolly et al., 2013)(++)	❖ Information extraction ❖ Web analytics ❖ Text visualization ❖ Statistical NLP ❖ Knowledge mining ❖ Multi-lingual analysis ❖ Hadoop ❖ Mapreduce ❖ Sentiment Analysis	❖ Query processing ❖ Information retrieval ❖ Artificial intelligence ❖ Computational linguistics ❖ Content analysis ❖ Semantic processing ❖ Lexical and syntactic processing	• MapReduce • MedScan • TXTGate • MedLEE

ods have been proposed in the healthcare domain, their acceptable accuracy and efficiency are still critical (Dougherty, G.2009). Moreover, when these methods are dealing with big data, these challenges seem to be more serious. According to (Belle et al., 2015), there are eight major challenges- preprocessing, compression, parallelization, mapping, security, segmentation, integration, and validation- of medical image analytics in big data environments. In order to address these challenges, new HCI&A technologies need to be deployed in clinical settings. These methods can help to improve the accuracy of diagnosis, and they also enable the use of disparate sets of data to improve healthcare quality at minimal cost.

5.1.3. Medical signal and genomics analytics

Medical signals mainly mean physiological signals that come from continuous monitoring systems. These signals also include volume and velocity obstacles like medical images; additionally, physiological signals are of a spatiotemporal nature that makes medical signals more complex and complicated. Moreover, medical signals are interconnected and interdependent. Therefore, before analysis, simplification of data complexity is of foremost importance. Analysis of these signals is more essential if the healthcare system is smart and context aware (Solanas et al., 2014). Currently physiological waveform data comes from multiple sources of information and so existing healthcare systems are not able to generate analytical results because they are designed to analyze only singular physiological signals (Belle et al., 2015). To analyze these clinical time series data, there is a need to deploy advanced HCI&A technologies which are able to extract interactions and complex relations among multisource physiological signals.

A big data-based clinical decision support system is proposed in Bloom, B. S. (2002), where advanced HITs are used to improve the quality of care. Similarly, in order to develop patient care management systems, a scalable infrastructure was designed in (Han et al., 2006, June), where static and monitoring data is analyzed. For the critical events in ICUs, this system mainly issues real-time alerts for medical employees. By using physiological signals, another similar decision support system was designed in (Bressan et al., 2012, January). Time series data (i.e. blood pressure and cardiac) was used by the study of (Lee and Mark, 2010, September) to improve therapeutic intervention in regard to hypotensive episodes. To develop neurological care, a physiological monitoring system was developed in (Le Roux et al., 2014); (Tisdall and Smith, 2007), and (Hemphill et al., 2011).

Genomics mainly enables DNA sequencing to empower medical diagnostics and other features of healthcare, including disease identification, health risk, therapeutic credentials, and prenatal development and analysis. As human genome encompasses 30 to 35 thousand genes, analytics of advanced-throughput sequencing techniques in genomics is an inherently big data problem (Drmanac et al., 2010; Lander et al., 2001). Though genomic data can contribute to accuracy and reduce the time taken for a diagnosis, analyzing it to develop real-time applicable recommendations is an important challenge in the area of computational biology. The combination of physiological data and high-throughput "-omics" methods to convey recommendations is the challenging factor in the discipline of computational biology. In order to address these challenges, advanced research on genome network analysis is required. Therefore, current healthcare systems demand advanced HCI&A approaches and analytics to deliver recommendations from genomic data in clinical settings.

5.1.4. Text analytics

Healthcare information systems comprise a large volume of textual and numeric data about patients, treatments, administrative notes, visits, physician notes, etc. The electronic health records (EHR) encapsulate information that could lead to: development of the quality of healthcare, promotion of healthcare research initiatives, reduction in erroneous medical diagnoses and prognoses, and elimination of unnecessary healthcare costs (Koh and Tan, 2011). However, the unstructured textual documents that contain the health information are wide ranging in complexity, length, and use of technical terminology, making knowledge discovery more challenging. In the medical domain, a number of text analytical tools can facilitate a unique opportunity to extract critical knowledge from textual data archives (Chen et al., 2005).

The applications of text analytics in healthcare are two-fold: medical research and medical services. In the research context, due to the limited capacity of the human brain, researchers have only specialized in a few sub-domains (e.g. several particular genes), while text mining tools and techniques offer appropriate links among many subdomains for discovering new knowledge or formulating hypotheses (Yandell and Majoros, 2002). MedScan (Novichkova et al., 2003), TXTGate (Glenisson et al., 2004)), PubMatrix (Becker et al., 2003), and BioRAT(Corney et al., 2004) are examples of text analytical tools which are used in medical research. In the medical service aspect, text analytic techniques have been applied to patient records (e.g. diagnosis reports, prognosis reports, and treatments) and other medical documents to facilitate cost-effective and real-time services in healthcare systems. MedLEE is an example of a text analytical tool (Friedman and Hripcsak, 1998). A similar text mining approach was proposed by (Chapman et al., 2004), which is an automated fever detection system (AFDS) from clinical records. In the big data context, in order to address different medical research and service challenges, advanced and scalable text analytic solutions are required. Therefore, current healthcare systems demand advanced HCI&A approaches and analytics to discover knowledge from large-scale textual datasets.

5.2. HCI&A with information system research

HCI&A technologies are already identified as a potential way of improving clinical outcomes (Garg et al., 2005), and an effective mechanism for reducing healthcare costs (Hillestad et al., 2005) where information system research can contribute to the constructive growth and utilization of advanced ITs to manage and synchronize different health services(Chiasson and Davidson, 2004). However, only IT applications are not enough to develop effective and efficient healthcare systems (Romanow et al., 2012); therefore, recently the healthcare industry poses intriguing research possibilities for researchers interested in the development of healthcare information systems. When effective information systems and efficient healthcare technologies work together in the healthcare domain, their implications are likely to be so pervasive, and their primary, secondary, tertiary, and subsequent-order benefits so penetrating, that they will connect to everyone's life and affect almost every corner of our society (Goldschmidt, 2005).

Thus, from 2004, significant changes have appeared in research on the healthcare industry. IS researchers have changed their attention and so, within the last decade they have published a good number of articles in leading IS journals (Romanow et al., 2012). A brief study on IS in healthcare was conducted by Romanow et al. who classified healthcare research trends into four categories - a) IS research about IS theories without regard to the healthcare context, b) IS research about IS theories with some regard to the healthcare context, c) IS research with rigorous empirical studies to develop IS theories or concepts, and d) IS research in the healthcare context without any IS theories. The inclusion of IS research in the healthcare industry has changed the perception of healthcare informatics. Now healthcare information technologies (HIT) are also variously known as health informatics, health IS, and so on. Over all the permutations of these terms have grown up with the effective propagation of IT in healthcare. In the following section, we conduct a bibliographic study on IS and Health informatics journals to extract integrated research trends in HCI&A. Through the bibliographic study of these journals, we identify the current state of HCI&A research and future foundations of insights into healthcare areas.

6. Mapping of HCI & A knowledge landscape: a bibliographic study on is and healthcare informatics research

In this section, we conduct a bibliographic study analyzing relevant literature on both information systems and health informatic disciplines. To conduct this study, we followed a collection, transformation and analytic process that is like the typical informatics and analytics process adopted in other business applications such as e-commerce and e-government. In order to determine the research trends in HCI&A, the top 20 related journals were selected with ten journals taken from information systems and ten from health informatics journals. We explored HCI&A publications in three bibliographic databases (ABI Inform collection, Business Source Complete, and Web of science) through our university library.[12] These databases contain high-quality bibliographic information including authors' short bios, article titles, article abstracts, article keywords, and journal information such as names, indexing, and citation. To ensure consistency and quality of appropriateness within our collection process, we extracted only those publications that contained the keywords hospital, medical/medicine, healthcare (and health care) for the IS discipline, and the keywords big data and technology (machine learning, software agent, and cloud computing) for health informatics journals. These keywords were relevant to our research interest and they also bore significant weight to meet our research goal. To conduct this study, we used metalanguage, XML, and SQL servers to analyze the collected datasets. Moreover, our collected data was validated by the respective journal authority.

Before going on to discuss the results of our bibliographic study, we would like to recognize some boundaries of our study. Elec-

[12] http://www.cityu.edu.hk/lib/.

Table 4.
Summary of ten major health informatics journals in the HCI&A context.

SL.No.	Journal Name	Big Data	Technology	Indexing Year	Journal launching year	Issues per year	Total Issues	Est.Total article	Est. (%) HCI&A
1	Journal of the American Medical Informatics Association, (JAMIA)	17	439	1994	1994	6	144	2381	19.95
2	Medical Image Analysis	1	15	1996	1996	4 to 8	114	1250	60.57
3	International Journal of Medical Informatics	15	684	1996	1970	4 to 12	409	3368	25.98
4	Journal of Clinical Bioinformatics	3	13	2012	2011	1	6	151	21.85
5	Computer Methods and Programs in Biomedicine	10	252	1985	1971	3 to 15	415	4259	7.18
6	Internet Interventions	1	14	2014	2014	4	12	126	11.90
7	Journal of Biomedical Semantics	2	23	2010	2010	1	7	327	12.23
8	Applied Clinical Informatics	11	15	2009	2009	4	29	395	10.63
9	Computers in Biology and Medicine	5	133	1970	1970	2 to 16	320	3325	10.17
10	Computerized Medical Imaging and Graphics	0	77	1988	1977	4 to 16	264	2790	5.66

tronic indices restricted our choice of journals to those indexed by typical commercial settings in our university library. Therefore, our outline of the literature study may not equally reflect all areas (e.g. North America and South Asia) of HCI&A research. Moreover, we entirely ignored book chapters and conference papers which also contained significant information relevant to this research. We also excluded articles that used other HCI&A terms (e.g. artificial intelligence and data warehousing) but not the given keywords (e.g. big data, technology, and healthcare) in the article titles, keywords, or abstracts. Though these boundaries could affect our final outcomes, such confines are common in bibliographic studies.

From the empirical analysis, we retrieved papers which have been published in health informatics journals. In order to quantify the whole research contribution, we considered all articles in our selected journals. Table 4 summarizes the absolute number and relative percentage of research work on HCI&A in health informatics publications. More than half of these journals dedicated less than 15% of their publications to HCI&A, although Medical Image Analysis has the highest contribution to HCI&A research. Although two journals - JAMIA and IJMI (SL.Nos.1 and 3) - have published the highest absolute number of articles on big data and technologies in healthcare, they have around 20%−25% impact on HCI&A. Like healthcare informatics, these two journals also emphasize other issues such as policy issues, laws, finance, and economics. Thus, their total number of publications is high but contribution to HCI&A is low. On the other hand, almost 800 papers of the Medical Image Analysis Journal relate to the computational approach of healthcare areas, though our desired keywords (big data and technology) do not frequently appear in article titles, keywords, or abstracts. Hence, even though it has published only 16 journals on big data and technologies in relation to healthcare, its overall impact on HCI&A is maximum, at more than 60%.

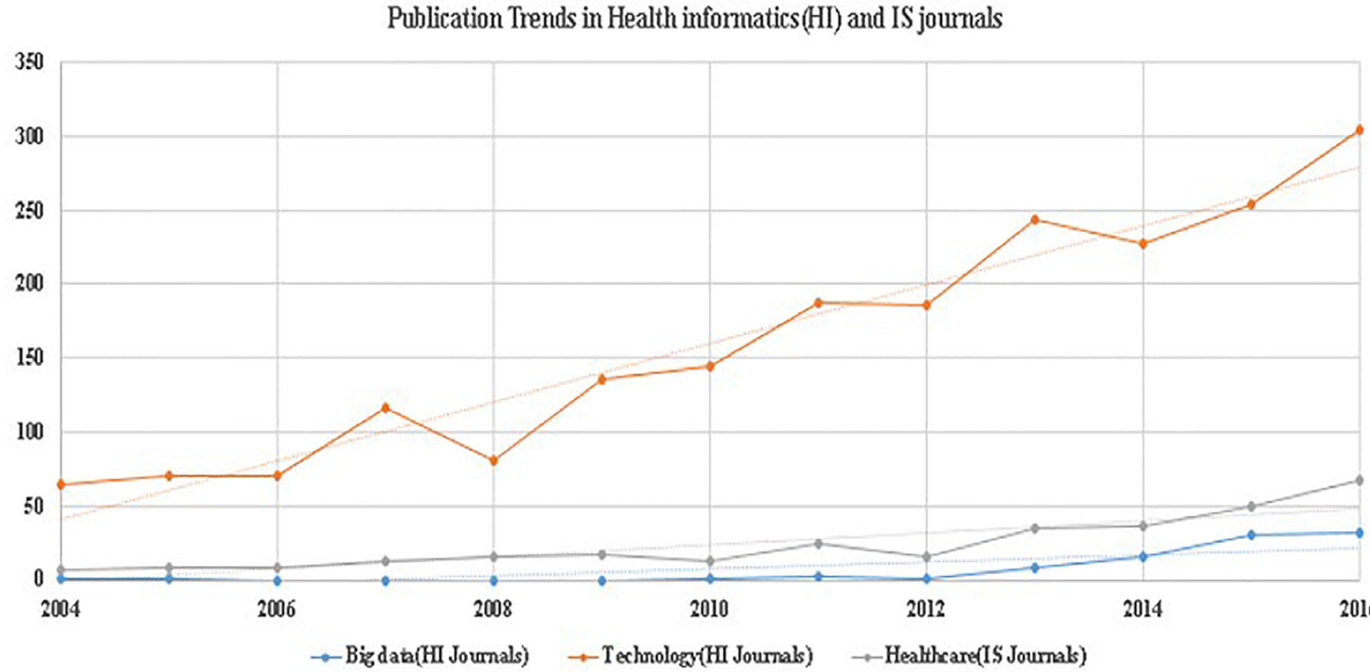
In order to measure the statistics and growth trends of publications on HCI&A, we considered related literature from health informatics and also from IS journals during the past decade (2004–2016). In 2004, different social media were introduced and from then the data deluge has also produced some opportunities in healthcare like other areas. Moreover, in the literature from 2004 there appeared different advanced technological research to gain big data opportunities which has received much more attention in the medical domain after about 2010. Fig. 3 shows the statistics and growing trends of health informatics and IS publications. Here, Health Informatics (HI) articles are related to big data and the technological context of the healthcare industry, and IS articles are just related to healthcare. As four journals (SL. No. 4, 6, 7, and 8) in Table 4 started publication after 2004, we excluded them for measuring the trends of health informatics publications. However, we also included another four health informatics publications - Journal of Biomedical Informatics, Value in Health, The American Journal of Medicine, and Artificial Intelligence in Medicine - which launched before 2004 and they all are also top-tier publishers in the healthcare field. The quantities of healthcare publications are summarized in Tables 5 and 6 for IS and health informatics journals respectively.

Table 5.
Major IS journals with healthcare publications.

Name of IS journal	Quantity of Healthcare publications (2004–2016)
MISQ	23
ISR	12
Organization Science	6
Decision Support System	18
Management Science	21
Information & Management	13
Information Science	35
Journal of Operations Management	44
Computers & Operations Research	49
Technological Forecasting and Social Change	91

Table 6.
Major health informatics journals with HCI&A publications.

Name of IS journal	Quantity of Healthcare informatics and Analytics publications (2004–2016)
Journal of the American Medical Informatics Association(JAMIA)	342
Medical Image Analysis	16
International Journal of Medical Informatics	514
Computer Methods and Programs in Biomedicine	163
Computers in Biology and Medicine	106
Computerized Medical Imaging and Graphics	195
Journal of Biomedical Informatics	713
Value in health	57
The American Journal of medicine	42
Artificial intelligence in Medicine	30



Key Words	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
Big data (HI Journals)	1	1	0	0	0	0	1	3	1	8	15	30	32	92
Technology (HI Journals)	65	70	71	117	81	136	144	187	186	243	228	254	304	2086
Healthcare (IS Journals)	7	8	9	13	16	17	13	24	16	35	37	49	68	312

Fig. 3. . Publication trends on HCI&A in health informatics (HI) and IS journals from 2004 to 2016.

A few IS publications contain HCI&A-related research. To conduct the bibliographic study, we selected ten reputed journals from the bibliographic database. Though our collected IS journals have published 312 articles within the last decade, we found only 220 journal articles related to big data or technology issues in the healthcare context. In health informatics journals, we found more papers on HCI&A because these journals are dedicated to the healthcare domain, whereas IS journals cover multiple domains such as e-governance and e-commerce. For instance, we could consider all big data papers in health informatics journals for this study, in contrast to IS journals where only a few big data research projects focused on the medical domain. Some research works in IS journals refer to IT in “name only” without substantive attention in the paper content, though we considered them in our study. For instance, though research work in (Kohli et al., 2012) is not a fully technical work, it describes the impact of IT investment and also inspires academics to undertake further research on IT’s influence over different sectors of the healthcare industry. However, from Fig. 3 we can perceive that research trends in both publications have been continuously rising since 2010. When we investigated the publication years of 312 papers in IS journals, we found that around 229 journals were published during 2010 to 2016. Similarly, in health informatics journals we found that 75% of technical articles within the last decade have been published since 2010. According to our bibliographic data, almost every big data article on healthcare has been published from 2010 to 2016. Though social media started to be introduced from 2004, they became popular within the next six years. Thus, from 2010 different social media data generated some opportunities and challenges for the healthcare industry. Through existing technologies, processing the data is challenging not only due to its volume but also the heterogeneous nature and speed at which it must be managed. Therefore, recently, both health informatics and IS publications have paid special attention to the computational view (i.e. processing efficiency, retrieval and manipulation of information, and technological capabilities) of healthcare technologies.

7. HCI&A education and program development

HCI&A has made significant contributions to developing the existing educational activities and also supporting every initiative regarding education in biomedical and health informatics (Mantas J. et al., 2010). HCI&A bridges several professional disciplines including medicine and allied health professions, information communication technology (ICT), information systems (IS), computer science (CS), law, business, social science, and project management. In the big data era, the National Institute of Health announced $50 million support for developing healthcare informatics education (United States, 2014). Moreover, the US government has announced $56 billion in support of modifying the existing U.S. healthcare system with advanced HCI&A technologies.[13] These initiatives indicate that thousands of new jobs will be created in healthcare informatics over the next ten years. Moreover, according to the US Bureau of Labor Statistics (BLS), medical and healthcare service managers are expected to see a 17% increase in jobs from 2014 to 2024. Currently healthcare informatics mostly rely on data science for presenting better outcomes. And so, existing healthcare professionals are not sufficiently qualified, and lack the required knowledge to manage the 3 V challenges - volume, velocity, and variety - of big data (Feldman et al., 2012). In order to find the solutions to these problems, in 2013 the White House arranged an event, named - "Data to Knowledge to Action" that featured many public and private initiatives.[14] They also presented some guidelines on how academic institutions (e.g. universities and colleges) should prepare students to become data scientists who will be actively engaged in solving the difficult problems in the healthcare industry. In the subsequent part of this section we identify some challenges for IS education for healthcare informatics and also describe program development on HCI&A education.

7.1. Education challenges and HCI&A knowledge

HCI&A mainly focuses on caring systems such as management, decision support, and strategizing within the healthcare industry. In order to achieve the goals, several disciplines run together under HCI&A where the IS discipline makes a more significant contribution than others (Fichman et al., 2011). In the IS discipline, a number of advanced programs such as data management, statistical analysis, computer programs and algorithms, management science techniques, economics model, and so on are used to train a new generation of academicians and practitioners. Though information systems (IS) is a young discipline, within the last four decades it has become a major stream to bridge business and technology. In order to achieve the new vision of healthcare systems, the IS curriculum should address the following three requirements:

- IS programs should emphasize courses on big data analytics, data mining, text mining, and social network analytics.
- IS programs should include courses on management of information systems in healthcare, characteristics and functionalities of IS in healthcare, and the construction of health and medical coding systems.
- IS departments should assure more applied courses related to science and engineering disciplines, and so should admit students with a strong math and statistical background.

HCI&A knowledge should be multidisciplinary covering data analytics, IT skills, business, and domain knowledge in healthcare environments. According to John Mantas et al. IT skills in healthcare have two dimensions - one is healthcare computing which includes artificial intelligence, human computer interaction, simulation and modeling, and cognitive aspects of information processing; the other is healthcare services including healthcare management, decision support systems, and statistical analytics. Coverage of both these dimensions ranges from HCI&A 1.0 to HCI&A 3.0. In order to produce qualified and competent professionals on HCI&A, academic courses must cover these technological and research streams that are included in Table 3.

7.2. Program development

HCI&A generates an unparalleled opportunity for the IS community in healthcare education to develop new programs, diploma curricula, and also degree programs configured to procure the future generation of systematic philosophers. In the healthcare industry, there are many courses of action to deliver HCI&A education, where graduation programs are a more suitable choice. Some reputed universities in the world have already included some new programs at bachelor level. HCI&A education programs mainly present integrated applied programs where students can learn from experiments such as hands-on projects, lab work, and industrial internships. In the big data environment, analytical techniques require experimentation with various methods of doing something until one finds the most successful solution. A strong relationship between industry and academia can boost the practical and realistic study of the HCI&A curriculum.

8. Conclusion

The first generation of healthcare informatics and analytics, namely HCI&A 1.0, leverages traditional relational database technologies and relational database management systems (RDBMs) for supporting data collection, data transformation, data extraction and loading, and data analysis over disparate healthcare databases. With the popularity of Web 2.0, HCI&A 2.0 starts to emerge and leads to the revolutionization of the whole healthcare industry. HCI&A 2.0 supports the notion of "collaborative healthcare" by connecting the crowds sharing common healthcare interests and needs around the globe by means of public or corporate social networks. HCI&A 2.0 offers advanced analytics tools and techniques to analyze large amount of unstructured data collected from social networks or corporate data stores and transform them to generate useful healthcare insights such as the sentiments of patients. In the past decade, a large number of people, devices, and sensors are connected via digital networks and the cross-plays among these entities generate enormous volume of data. The big healthcare data enable healthcare organizations to innovate and grow. Meanwhile, the healthcare big data deluge also gives rise to serious concerns about the retrieval, storage, analysis, and security of patent or healthcare professionals' information. HCI&A 3.0 offers the promise of proving large-scale data management and data analytical tools to cope with the big data challenge in healthcare. In this study, we conduct a critical analysis of numerous data analytics tools and techniques under the proposed HCI&A framework that encompasses the technological evolution from HCI&A 1,0 to HCI&A 3.0. In particular, the discussions of these emerging data analytics technologies are given in the context of many real-world healthcare applications such as health insurance, healthcare administration and policy, healthcare services, healthcare security, and healthcare personnel privacy, and so on. Our study contributes to a critical analysis of the state-of-the-art technologies in healthcare informatics and analytics, and it provides a clear road map to foster future research and knowledge transfer of the next generation of HCI&A.

Declaration of Competing Interest

On behalf of all authors, I, Dr. Ileas Pramanik certify that we have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus; membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

[13] http://healthit.cns.utexas.edu/.

[14] http://www.cfr.org/technology-and-science/white-house-big-data-seizing-opportunities-preserving-values/p32916.

Acknowledgements

This project was partly supported by grants from the Research Grants Council of the Hong Kong Special Administrative Region, China (Projects: CityU 11507219 and CityU 11525716), the NSFC Basic Research Program (Project: 71671155), BRUR_Research Project/2019–20/40, and the CityU Shenzhen Research Institute.

References

Acampora, G., Cook, D. J., Rashidi, P., & Vasilakos, A. V. (2013). A survey on ambient intelligence in healthcare. *Proceedings of the IEEE, 101*(12), 2470–2494.
Agrawal, R., Grandison, T., Johnson, C., & Kiernan, J. (2007). Enabling the 21st century health care information technology revolution. *Communications of the ACM, 50*(2), 34–42.
Almashaqbeh, G., Hayajneh, T., Vasilakos, A. V., & Mohd, B. J. (2014). QoS-aware health monitoring system using cloud-based WBANs. *Journal of Medical Systems, 38*(10), 1–20.
Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R.H., Konwinski, A. et al. (2009). Above the clouds: A berkeley view of cloud computing.
Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., et al. (2010). A view of cloud computing. *Communications of the ACM, 53*(4), 50–58.
Arnrich, B., Mayora, O., & Bardram, J. (2010,). Pervasive or ubiquitous healthcare?.
Arnrich, B., Mayora, O., Bardram, J., & Tröster, G. (2010,). Pervasive healthcare. *Methods of Information in Medicine, 49*(1), 67–73.
Avison, D., & Young, T. (2007). Time to rethink health care and ICT? *Communications of the ACM, 50*(6), 69–74.
Bates, D. W., Saria, S., Ohno-Machado, L., Shah, A., & Escobar, G. (2014). Big data in health care: Using analytics to identify and manage high-risk and high-cost patients. *Health Affairs, 33*(7), 1123–1131.
Baujard, O., Baujard, V., Aurel, S., Boyer, C., & Appel, R. D. (1998). Trends in medical information retrieval on internet. *Computers in Biology and Medicine, 28*(5), 589–601.
Becker, K. G., Hosack, D. A., Dennis, G., Lempicki, R. A., Bright, T. J., Cheadle, C., et al. (2003). PubMatrix: A tool for multiplex literature mining. *BMC Bioinformatics, 4*(1), 1.
Belle, A., Thiagarajan, R., Soroushmehr, S.M., Navidi, F., Beard, D.A., & Najarian, K. (2015). Big data analytics in healthcare. BioMed research international, 2015.
Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E. D., Zendulka, J., et al. (2014). PredictSNP: Robust and accurate consensus classifier for prediction of disease-related mutations. *PLoSComputBiol, 10*(1), e1003440.
Benko, A., & Wilson, B. (2003). Online decision support gives plans an edge. *Managed Healthcare Executive, 13*(5), 20.
Berndt, D. J., Fisher, J. W., Hevner, A. R., & Studnicki, J. (2001). Healthcare data warehousing and quality assurance. *Computer, 34*(12), 56–65.
Bhasin, M., & Raghava, G. P. S. (2004). GPCRpred: An SVM-based method for prediction of families and subfamilies of G-protein coupled receptors. *Nucleic Acids Research, 32*(suppl 2), W383–W389.
Borkar, V. R., Carey, M. J., & Li, C. (2012). Big data platforms: What's next?. XRDS: Crossroads. *The ACM Magazine for Students, 19*(1), 44–49.
Bower, A.G. (.2005). The diffusion and value of healthcare information technology (Vol. 272). Rand Corporation.
Braun, L. M., Wiesman, F., van den Herik, H. J., Hasman, A., & Korsten, E. (2007). Towards patient-related information needs. *International Journal of Medical Informatics, 76*(2), 246–251.
Bressan, N., James, A., & McGregor, C. (2012, January). Trends and opportunities for integrated real time neonatal clinical decision support. In Biomedical and Health Informatics (BHI), 2012 IEEE-EMBS International Conference on (pp. 687-690). IEEE.
Brook, R. H., Williams, K. N., & Avery, A. D. (1976). Quality assurance today and tomorrow: Forecast for the future. *Annals of Internal Medicine, 85*(6), 809–817.
Burbidge, R., Trotter, M., Buxton, B., & Holden, S. (2001). Drug design by machine learning: Support vector machines for pharmaceutical data analysis. *Computers & Chemistry, 26*(1), 5–14.
Cabanillas, D., Vàzquez-Salceda, J., Willmott, S., Cortés, U., & Bouissou, D. (2003). Carrel: Secure deployment of an agent—Based health care application. In Applications of Software Agent Technology in the Health Care Domain (pp. 161-179). Birkhäuser Basel.
Chan, V., Ray, P., & Parameswaran, N. (2008). Mobile e-Health monitoring: An agent-based approach. *IET Communications, 2*(2), 223–230.
Chang, E.Y., Bai, H., & Zhu, K. (2009, October). Parallel algorithms for mining large-scale rich-media data. In Proceedings of the 17th ACM international conference on Multimedia (pp. 917-918). ACM.
Chapman, W. W., Dowling, J. N., & Wagner, M. M. (2004). Fever detection from free-text clinical records for bio-surveillance. *Journal of Biomedical Informatics, 37*(2), 120–127.
Chen, H., Chiang, R. H., & Storey, V. C. (2012a). Business intelligence and analytics: From big data to big impact. *MIS Quarterly, 36*(4), 1165–1188.
Chen, H., Fuller, S.S., Friedman, C., & Hersh, W. (2005). Knowledge management, data mining, and text mining in medical informatics. In Medical Informatics (pp. 3-33). Springer US.
Chen, R., Mias, G. I., Li-Pook-Than, J., Jiang, L., Lam, H. Y., Chen, R., et al. (2012b). Personal omics profiling reveals dynamic molecular and medical phenotypes. *Cell, 148*(6), 1293–1307.
Chen, W., Cockrell, C., Ward, K.R., & Najarian, K. (2010, December). Intracranial pressure level prediction in traumatic brain injury by extracting features from multiple sources and using machine learning methods. In Bioinformatics and Biomedicine (BIBM), 2010 IEEE International Conference on (pp. 510–515). IEEE.
Chiasson, M. W., & Davidson, E. (2004). Pushing the contextual envelope: Developing and diffusing is theory for health information systems research. *Information and Organization, 14*(3), 155–188.
Connolly, S., Wooledge, S., & Aster, T. (2013). Harnessing the value of big data analytics.
Corchado, J. M., Bajo, J., & Abraham, A. (2008). GerAmi: Improving healthcare delivery in geriatric residences. *IEEE Intelligent Systems, 23*(2), 19–25.
Corney, D. P., Buxton, B. F., Langdon, W. B., & Jones, D. T. (2004). BioRAT: Extracting biological information from full-length papers. *Bioinformatics (Oxford, England), 20*(17), 3206–3213.
Corteś, U., Fox, J., & Moreno, A. (2002, July).Eds.. In: Proceedings of the Workshop on Agents Applied in Health Care, in the 15th European Conference on Artificial Intelligence, ECAI-02. Lyon, France: IOS Press;
Cortés, U., Urdiales, C., Annichiarico, R., Barrué, C., Martínez, A., & Caltagirone, C. (2007). Advanced computational intelligence paradigms in healthcare-1, Vol. 48 of Studies in Computational Intelligence.
Cruz-Correia, R., Vieira-Marques, P., Costa, P., Ferreira, A., Oliveira-Palhares, E., Araújo, F., et al. (2005). Integration of hospital data using agent technologies–a case study. *Ai Communications, 18*(3), 191–200.
Detmer, D. E., Steen, E. B., & Dick, R. S. (Eds.). (1997). *The computer-based patient record: An essential technology for health care*. National Academies Press.
Diller, G. P., Kempny, A., Narayan, S. B., Henrichs, M., Brida, M., … Gatzoulis, M. A. (2019). Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: Data from a single tertiary centre including 10 019 patients. *European Heart Journal, 40*(13), 1069–1077.
Dougherty, G. (2009). *Digital image processing for medical applications*. Cambridge University Press.
Drmanac, R., Sparks, A. B., Callow, M. J., Halpern, A. L., Burns, N. L., Kermani, B. G., et al. (2010). Human genome sequencing using unchained base reads on self--assembling DNA nanoarrays. *Science (New York, N.Y.), 327*(5961), 78–81.
Feldman, B., Martin, E.M., & Skotnes, T. (2012). Big data in healthcare hype and hope. October 2012. Dr. Bonnie, 360.
Felisberto, F., Laza, R., Fdez-Riverola, F., & Pereira, A. (2015). A distributed multiagent system architecture for body area networks applied to healthcare monitoring. BioMed research international, 2015.
Fenn, J., & Raskino, M. (2011). *Gartner's hype cycle special report for 2011*. Stamford, CT: Gartner.
Fichman, R. G., Kohli, R., & Krishnan, R. ((2011).Eds.).Editorial overview—The role of information systems in healthcare: Current research and future trends. Information Systems Research, 22(3), 419–428.
Frank, J.O. (.2012). Big data analytics: Turning big data into big money.
Friedman, C., & Hripcsak, G. (1998). Evaluating natural language processors in the clinical domain. *Development (Cambridge, England), 22*, 24.
Frost, S. (2015). Drowning in big data? Reducing information technology complexities and costs for healthcare organizations.
Frunza, O., Inkpen, D., & Tran, T. (2011). A machine learning approach for identifying disease-treatment relations in short texts. *IEEE Transactions on Knowledge and Data Engineering, 23*(6), 801–814.
Garg, A. X., Adhikari, N. K., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., et al. (2005). Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *JAMA, 293*(10), 1223–1238.
Giustini, D. (2007). Web 3.0 and medicine. *BMJ (Clinical Research Ed.), 335*(7633), 1273–1274.
Glenisson, P., Coessens, B., Van Vooren, S., Mathys, J., Moreau, Y., & De Moor, B. (2004). TXTGate: Profiling gene groups with text-based information. *Genome Biology, 5*(6), 1.
Goldschmidt, P. G. (2005). HIT and MIS: Implications of health information technology and medical information systems. *Communications of the ACM, 48*(10), 68–74.
Griebel, L., Prokosch, H. U., Köpcke, F., Toddenroth, D., Christoph, J., Leb, I., et al. (2015). A scoping review of cloud computing in healthcare. *BMC Medical Informatics and Decision Making, 15*(1), 1.
Groves, P., Kayyali, B., Knott, D., & Van Kuiken, S. (2013). The 'big data'revolution in healthcare. McKinsey Quarterly, 2.
Han, H., Ryoo, H.C., & Patrick, H. (2006, June). An infrastructure of stream data mining, fusion and management for monitored patients. In Computer-Based Medical Systems, 2006. CBMS 2006. 19th IEEE International Symposium on (pp. 461–468). IEEE.
Hemphill, J. C., Andrews, P., & De Georgia, M. (2011). Multimodal monitoring and neurocritical care bioinformatics. *Nature Reviews Neurology, 7*(8), 451–460.
Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., et al. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs, 24*(5), 1103–1117.
Hughes, B., Joshi, I., & Wareham, J. (2008). Health 2.0 and Medicine 2.0: Tensions and controversies in the field. *Journal of Medical Internet Research, 10*(3), e23.
IBM: IBM big data platform for healthcare." Solutions brief; (2012).
Institute of Medicine. (2000). To err is human: Building a safer health system.
Intel: Leveraging big data and analytics in healthcare and life sciences: Enabling personalized medicine for high-quality care, better outcomes; (2012).
Isern, D., Sánchez, D., & Moreno, A. (2007, September). HeCaSe2: A multi-agent ontology-driven guideline enactment engine. In International Central and East-

ern European Conference on Multi-Agent Systems (pp. 322–324). Springer Berlin Heidelberg.
James, M., Michael, C., Brad, B., Jacques, B., Richard, D., Charles, R., et al. (2011). Big data: The next frontier for innovation, competition, and productivity. *The McKinsey Global Institute.*.
jStart: "How big data analytics reduced medicaid re-admissions." A jStart Case Study; (2012). http://www-01.ibm.com/software/ebusiness/jstart/portfolio/uncMedicaidCaseStudy.pdf
Kabachinski, J. (2011). What's the forecast for cloud computing in healthcare? *Biomedical Instrumentation & Technology, 45*(2), 146–150.
Kaminker, J. S., Zhang, Y., Watanabe, C., & Zhang, Z. (2007). CanPredict: A computational tool for predicting cancer-associated missense mutations. *Nucleic Acids Research, 35*(suppl 2), W595–W598.
Keleş, A., & Keleş, A. (2008). ESTDD: Expert system for thyroid diseases diagnosis. *Expert Systems with Applications, 34*(1), 242–246.
Knowledgent: Big data and healthcare payers; (2013). http://knowledgent.com/mediapage/insights/whitepaper/482
Koh, H. C., & Tan, G. (2011). Data mining applications in healthcare. *Journal of Healthcare Information Management, 19*(2), 65.
Kohli, R., Devaraj, S., & Ow, T. T. (2012). Does information technology investment influence a firm's market value? A case of non-publicly traded healthcare firms. *MIS Quarterly, 36*(4), 1145–1163.
Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine, 23*(1), 89–109.
Kukar, M., Kononenko, I., & Silvester, T. (1996). Machine learning in prognosis of the femoral neck fracture recovery. *Artificial Intelligence in Medicine, 8*(5), 431–451.
Kuo, M. H. (2011). Opportunities and challenges of cloud computing to improve health care services. *Journal of Medical Internet Research, 13*(3), e67.
Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature, 409*(6822), 860–921.
Lanfranco, A. R., Castellanos, A. E., Desai, J. P., & Meyers, W. C. (2004). Robotic surgery: A current perspective. *Annals of Surgery, 239*(1), 14–21.
Lassila, O., & Hendler, J. (2007). Embracing" web 3.0. *IEEE Internet Computing, 11*(3), 90–93.
Lau, R. Y., Liao, S. S., Wong, K. F., & Chiu, D. K. (2012). Web 2.0 environmental scanning and adaptive decision support for business mergers and acquisitions. *MIS Quarterly, 36*(4).
LaValle, S., Lesser, E., Shockley, R., Hopkins, M. S., & Kruschwitz, N. (2011). Big data, analytics and the path from insights to value. *MIT Sloan Management Review, 52*(2), 21.
Le Roux, P., Menon, D. K., Citerio, G., Vespa, P., Bader, M. K., Brophy, G. M., et al. (2014). Consensus summary statement of the international multidisciplinary consensus conference on multimodality monitoring in neurocritical care. *Neurocritical Care, 21*(2), 1–26.
Ledley, S., Lusted, L. B., & Ledley, R. S. (1959). Reasoning foundations of medical diagnosis. *In Science.*.
Lee, J., & Mark, R.G. (.2010, September). A hypotensive episode predictor for intensive care based on heart rate and blood pressure time series. In Computing in Cardiology, 2010 (pp. 81–84). IEEE.
Logic Monitor. (2012). LogicMonitor: Architecture white paper. Retrieved from http://www.logicmonitor.com/downloads/Architecture.pdf?84cd58
Magoulas, G.D., & Prentza, A. (2001). Machine learning in medical applications. In Machine Learning and its applications (pp. 300–307). Springer Berlin Heidelberg.
Markonis, D., Schaer, R., Eggel, I., Müller, H., & Depeursinge, A. (2015). Using MapReduce for large-scale medical image analysis. arXiv:1510.06937.
Mead, C. N. (2006). Data interchange standards in healthcare it-computable semantic interoperability: Now possible but still difficult. do we really need a better mousetrap? *Journal of Healthcare Information Management, 20*(1), 71.
Meingast, M., Roosta, T., & Sastry, S. (2006, August). Security and privacy issues with health care information technology. In Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE (pp. 5453–5458). IEEE.
Mell, P., & Grance, T. (2011). The NIST definition of cloud computing (draft). *NIST Special Publication, 800*(145), 7.
Mohan, A., Bauer, D., Blough, D.M., Ahamad, M., Bamba, B., Krishnan, R. et al. (2009). A patient-centric, attribute-based, source-verifiable framework for health record sharing.
Muir, E. (2011). Challenges of cloud computing in healthcare integration. *Special to ZDNet.*.
Nash, D. B. (2008). Health 3.0. *Pharmacy and Therapeutics, 33*(2), 69.
Nealon, J., & Moreno, A. (2003). Agent-based applications in health care. In Applications of software agent technology in the health care domain (pp. 3–18). Birkhäuser Basel.
Novichkova, S., Egorov, S., & Daraselia, N. (2003). MedScan, a natural language processing engine for MEDLINE abstracts. *Bioinformatics (Oxford, England), 19*(13), 1699–1706.
NSF, Joint DMS/NIGMS initiative to support research at the interface of the biological and mathematical sciences, program announcement, (2013).; www.nsf.gov/pubs/2013/nsf13570/nsf13570.htm
Oh, H., Rizo, C., Enkin, M., & Jadad, A. (2005). What is eHealth (3): A systematic review of published definitions. *Journal of Medical Internet Research, 7*(1), e1.
Abasolo, J.M., & Gomez, M. (2000, September). MELISA: An ontology-based agent for information retrieval in medicine. In Proceedings of the first international workshop on the semantic web (SemWeb2000) (pp. 73–82).
Pandora. (2011). FMS-Virtualization and cloud computing monitoring. Retrieved from
Patil, H.K., & Seshadri, R. (2014, June). Big data security and privacy issues in healthcare. In 2014 IEEE international congress on big data (pp. 762–765). IEEE.
Pramanik, M.I., Lau, R.Y., Demirkan, H., & Azad, M.A.K. (2017). Smart health: Big data enabled health paradigm within smart cities. Expert Systems with Applications.
Pramanik, M. I., Lau, R. Y., Yue, W. T., Ye, Y., & Li, C. (2017b). Big data analytics for security and criminal investigations. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 7*(4).
Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: Promise and potential. *Health Information Science and Systems, 2*(1), 1.
Raja, U., Mitchell, T., Day, T., & Hardin, J. M. (2008). Text mining in healthcare. *Applications and Opportunities. J HealthcInfManag, 22*(3), 52–56.
Ritter, F., Boskamp, T., Homeyer, A., Laue, H., Schwier, M., Link, F., et al. (2011). Medical image analysis. *IEEE Pulse, 2*(6), 60–70.
Röcker, C., Ziefle, M., & Holzinger, A. (2014). From computer innovation to human integration: Current trends and challenges for pervasive HealthTechnologies. In Pervasive Health (pp. 1–17). Springer London.
Rodger, J. A. (2016). Discovery of medical big data analytics: Improving the prediction of traumatic brain injury survival rates by data mining patient informatics processing software hybrid hadoop hive. *Informatics in Medicine Unlocked, 1*(2015), 17–26.
Romanow, D., Cho, S., & Straub, D. (2012). Editor's comments: Riding the wave: Past trends and future directions for health IT research. *MIS Quarterly, 36*(3) III-A18.
Rossetti, M.D., Trzcinski, G.F., & Syverud, S.A. (.1999). Emergency department simulation and determination of optimal attending physician staffing schedules. In Simulation Conference Proceedings, 1999 Winter (Vol. 2, pp. 1532–1540). IEEE.
Runyon, B. (2010). Hype cycle for healthcare provider technologies and standards, 2010.
Runyon, B., Gabler, J., Handler, T., Heib, B., Lovelock, J. D., Rishel, W., et al. (2005). *Hype cycle for healthcare provider technologies*. Gartner Inc Research.
Runyon, B., Handler, T.J., Shaffer, V., Hieb, B.R., Edwards, J., Rishel, W. et al. (2008). Hype cycle for healthcare provider applications and systems, 2008.
Runyon, B., Handler, T., Hieb, B., Shaffer, V., Jonathan, E., & Lovelock, J.D. (.2007). Hype cycle for healthcare provider applications and systems. Gartner Industry Report ID, (G00148329).
Runyo, B., Edwards, J., Thomas, J., Barry, R., Kelly, B. A., Lovelock, J. D., et al. (2006). *Hype cycle for healthcare provider applications and systems*. Gartner Inc Research.
Scholl, I., Aach, T., Deserno, T. M., & Kuhlen, T. (2011). Challenges of medical image processing. *Computer Science-Research and Development, 26*(1–2), 5–13.
Seibert, J.A. (.2009). Modalities and data acquisition. In Practical Imaging Informatics (pp. 49–66). Springer New York.
Shaffer, V. (2012). *Hype cycle for healthcare provider applications and systems*. Stamford, CT: Gartner Research.
Shaffer, V. (2013). Hype cycle for healthcare provider applications, analytics and systems, 2013. In.: Gartner.
Shankararaman, V. (2000, June).editor. In: Proceedings of the Workshop on Agents in Health Care, in the 4th International Conference on Autonomous Agents, Agents-00. Barcelona, Spain: ACM Press;
Solanas, A., Patsakis, C., Conti, M., Vlachos, I. S., Ramos, V., Falcone, F., et al. (2014). Smart health: A context-aware health paradigm within smart cities. *IEEE Communications Magazine, 52*(8), 74–81.
Srinivasan, U., & Arunasalam, B. (2013). Leveraging big data analytics to reduce healthcare costs. *IT Professional, 15*(6), 21–28.
Su, K., Li, J., & Fu, H. (2011, September). Smart city and the applications. In Electronics, Communications and Control (ICECC), 2011 International Conference on (pp. 1028–1031). IEEE.
Suryadevara, N. K., & Mukhopadhyay, S. C. (2014). Determining wellness through an ambient assisted living environment. *IEEE Intelligent Systems, 29*(3), 30–37.
Tentori, M., Favela, J., & Rodriguez, M. D. (2006). Privacy-aware autonomous agents for pervasive healthcare. *IEEE Intelligent Systems, 21*(6), 55–62.
Tisdall, M. M., & Smith, M. (2007). Multimodal monitoring in traumatic brain injury: Current status and future directions. *British Journal of Anaesthesia, 99*(1), 61–67.
Tolchinsky, P., Cortes, U., Modgil, S., Caballero, F., & Lopez-Navidad, A. (2006). Increasing human-organ transplant availability: Argumentation-based agent deliberation. *IEEE Intelligent Systems, 21*(6), 30–37.
Twardowski, B., & Ryzko, D. (2014, August). Multi-agent architecture for real-time big data processing. In WI-IAT (3) (pp. 333–337).
Van De Belt, T. H., Engelen, L. J., Berben, S. A., & Schoonhoven, L. (2010). Definition of Health 2.0 and Medicine 2.0: A systematic review. *Journal of Medical Internet Research, 12*(2), 1–14.
Vermeulen, I. B., Bohte, S. M., Elkhuizen, S. G., Lameris, H., Bakker, P. J., & La Poutré, H. (2009). Adaptive resource allocation for efficient patient scheduling. *Artificial Intelligence in Medicine, 46*(1), 67–80.
Von Landesberger, T., Bremm, S., Kirschner, M., Wesarg, S., & Kuijper, A. (2013). Visual analytics for model-based medical image segmentation: Opportunities and challenges. *Expert Systems with Applications, 40*(12), 4934–4943.
Walliser, M., Brantschen, S., Calisti, M., & Schinkinger, S. (2008). Whitestein series in software agent technologies and autonomic computing.
Wang, X., & Tan, Y. (2010, October). Application of cloud computing in the health information system. In *2010 International Conference on Computer Application and System Modeling (ICCASM 2010): 1* (pp. V1–179). IEEE.
Widmer, A., Schaer, R., Markonis, D., & Müller, H. (2014,). Gesture interaction for content–based medical image retrieval. In *Proceedings of International Conference on Multimedia Retrieval* (p. 503). ACM.

Wiederhold, G. (1981a). Database technology in health care. *Journal of Medical Systems, 5*(3), 175–196.
Wiederhold, G. (1981b). Databases in health care. In *Databases for Health Care* (pp. 1–16). Springer Berlin Heidelberg.
Wooldridge, M. (2009). *An introduction to multiagent systems*. John Wiley & Sons.
Wu, X., Zhang, C., & Zhang, S. (2005). Database classification for multi-database mining. *Information Systems, 30*(1), 71–88.
Wu, X., & Zhang, S. (2003). Synthesizing high-frequency rules from different data sources. *IEEE Transactions on Knowledge and Data Engineering, 15*(2), 353–367.
Wu, X., Zhu, X., Wu, G. Q., & Ding, W. (2014). Data mining with big data. *IEEE Transactions on Knowledge and Data Engineering, 26*(1), 97–107.
Yandell, M. D., & Majoros, W. H. (2002). Genomics and natural language processing. *Nature Reviews Genetics, 3*(8), 601–610.
Yeow, A., & Goh, K. H. (2015). Work harder or work smarter? information technology and resource allocation in healthcare processes. *Mis Quarterly, 39*(4) 12135_RA_YeowGoh.
Yuan, S., Lai, X., Zhao, X., Xu, X., & Zhang, L. (2005). Distributed structural health monitoring system based on smart wireless sensor and multi-agent technology. *Smart Materials and Structures, 15*(1), 1.
Zaharakis, I. D., Kameas, A. D., & Nikiforidis, G. C. (1998). A multi-agent architecture for teaching dermatology. *Med Inform, 23*(4), 289–307.
Zheng, T., Cao, L., He, Q., & Jin, G. (2014). Full-range in-plane rotation measurement for image recognition with hybrid digital-optical correlator. *Optical Engineering, 53*(1) 011003-011003.
Zhou, J., Xiong, W., Tian, Q., Qi, Y., Liu, J., Leow, W. K., et al. (2008, September). Semi-automatic segmentation of 3D liver tumors from CT scans using voxel classification and propagational learning. *In MICCAI Workshop (Vol, 41*, 43.
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.*.