Big Data, Smart Homes and Ambient Assisted Living

V. Vimarlund1,2, S. Wass1
1 Jönköping International Business School, Jönköping, Sweden
2 Department of Computer Science, Linköping University, Linköping, Sweden

Summary
Objectives: To discuss how current research in the area of smart homes and ambient assisted living will be influenced by the use of big data.

Methods: A scoping review of literature published in scientific journals and conference proceedings was performed, focusing on smart homes, ambient assisted living and big data over the years 2011-2014.

Results: The health and social care market has lagged behind other markets when it comes to the introduction of innovative IT solutions and the market faces a number of challenges as the use of big data will increase. First, there is a need for a sustainable and trustful information chain where the needed information can be transferred from all producers to all consumers in a structured way. Second, there is a need for big data strategies and policies to manage the new situation where information is handled and transferred independently of the place of the expertise. Finally, there is a possibility to develop new and innovative business models for a market that supports cloud computing, social media, crowdsourcing etc.

Conclusions: The interdisciplinary area of big data, smart homes and ambient assisted living is no longer of interest for IT developers, it is also of interest for decision makers as customers make more informed choices among today’s services. In the future it will be of importance to make information usable for managers and improve decision making, tailor smart home services based on big data, develop new business models, increase competition and identify policies to ensure privacy, security and liability.

Keywords
Big data, smart homes, ambient assisted living

Yearb Med Inform 2014:143-9
http://dx.doi.org/10.15265/IY-2014-0011
Published online August 15, 2014

Introduction

There is today considerable interest in the use of information technology (IT) for collecting big data in networked communities or in domestic environments often called smart homes1 and/or ambient assisted living [1-5] in order to increase volume, variety and velocity of data and information.

Technology vendors and research groups are focused on how to produce, combine, analyze and effective use of big data from heterogeneous and distributed sources in order to provide services at home. Studies, both in the USA and in the EU, have reported that the interest in capturing the benefits of using big data increases due the expectations of its impact on quality and efficiency of health care delivery as in detecting diseases at earlier stages to be treated most successfully [6-13]. There is also an expectation to be able to manage specific health populations and individuals and to detect health care fraud more quickly and efficiently.

Big data is expected to fundamentally transform smart homes and ambient assisted living delivery of services and consequently managerial and economic aspects of health services delivery, business models and governance processes. But also to improve patient centered services, diminish costs and monitoring with real-time analysis as well as preventing potentially adverse events, such as side effects of medications, early development of infections, allergic reactions etc.

In smart homes environments, research projects and R&D ventures try to utilize consumer affordable technology to equip homes with a set of advanced electronics, sensors and automated devices specifically designed for care delivery, remote monitoring, and early detection of problems or emergency cases. Technology is also used as promotion of residential safety, increased quality of life and real-time analysis of the health conditions of individuals [14].

However, at the same time volumes of data are growing and the data used in health and social care are becoming more variable, complex and difficult to manage with the same tools, routines and organizational culture used in the past. In its 2011 report, Big Data: The Next Frontier for Innovation, Competition and Productivity, McKinsey estimates that the potential value from data in US health care could be more than US$300 billion per year [6].

This gives rise to a number of challenges such as:

a) the accuracy of techniques and technologies to enable the capture, storage, distribution, management and analysis of the information sampled
b) the accessibility to large volume of data from distributed sources to identify complex problems and management
c) managerial issues related to the fact that individuals change their roles from passive consumers into active participants
d) economic issues related to reimbursement policies and principles

Objectives

In the near future it will be important to develop a research agenda for smart homes
and ambient assisted living that combines business and organizational initiatives to gain critical insight from the data collected.

In this paper we discuss how current research in the area of smart homes and ambient assisted living will be influenced by the use of big data and the issues that will be of relevance for the area.

Methods

We used scoping methods to examine the extend of the research activity and to sample information about the state of the art in the area. Scoping reviews are relatively a new type of research review that provides a tool for summarizing literature in a topic area. We applied this method to the topic big data and ambient assisted living and smart homes. We summarize the research findings of relevance for our study during the last three years and provide a background that can be further extended in future publications. In doing that we produce a profile of the issues that are not covered but of importance for smart homes and ambient assisted living.

In contrast to a systematic review, a scoping study is not driven by a well-specified, often narrow, research question, but applies an iterative search process [15]. Scoping reviews are used to present the range and extent of research in a specific field, without in-depth analysis or to assess the quality of the literature but to provide a background [16]. This allowed us to extend the number of reviewed papers during the writing of this article with the aim to achieve a broad coverage rapidly.

Search Process

Starting this study, we had no particular research question or interest of studies with any particular research design. Instead we wanted to get a broad overview of how the field has developed during the last years. Nevertheless, we needed to formulate the most important aspects of the expected results. With the assumption that we would find relevant studies in the common health informatics field as well as in applied technical fields and the social science literature, we started up by searching the following databases: PubMed, ScienceDirect, Google Scholar, ACM Digital Library and IEEE Xplore with the keywords “big data” AND “ambient assisted living” OR “smart houses” OR “smart homes”. We searched for papers published during the last years written in English. In addition to the studies that appeared in the searches we also included similar publications, proposed by the database search engine, if they had been published in the same period of time and seemed relevant from the title.

The amount of publications with focus on big data is vast and has expanded rapidly during the last two years and more than 15,000 articles can be found when searching for publications that discuss issues related to big data in general. Almost all the publications (90 %) belong to the area of computer science. Publications that combine big data and healthcare (approx. 15 %) aim to describe how big data can be captured using wearable sensors capable of carrying out continuous monitoring of the elderly, alerting the caregivers when necessary and forwarding pertinent information from big data systems for analysis.

We also found articles (three articles) that argued that data generated at smart houses are sensitive and ownership issues are not always clear. They discuss also the kind of data collected, stored and shared and the importance to link data from different sources. Further, they discuss how sensors can transfer data, how networks should be protected and how confidentiality has to be improved to secure sensitive data to preserve individuals’ integrity and truthfulness of the data. How data is stored and which steps are needed to protect privacy, how to ensure statistical properties and data consistency as well as how access to the system should be ensured through proper authentication and authorization of the users are common issues in current publications [17-19].

Furthermore we found one article published during 2013 that focused on big data challenges to stimulated national and international initiatives in building interconnected data repositories and integrated data resources as well as long-term data management and data stewardship to support cross-disciplinary scientific data discovery and reuse [20]. Another one discussed the need for increasing academic – industry collaboration to maximize the benefits from using big data [21]. Moreover, one publication discussed the importance of big data to generate new knowledge, the costs of answering many clinical questions prospectively and even retrospectively, the importance to create and observational evidence base for clinical questions that would otherwise not be possible and that may be especially helpful to develop new ways to disseminate knowledge [22]. Finally, one article discussed big biomedical data and the importance of linking big data sets to the individual person level [23].

Due to the fact that we, in this paper, discuss how current research in the area of smart homes and ambient assisted living will be influenced by the use of big data we only selected articles that did not focus on how big data can be captured using wearable sensors capable of carrying out continuous monitoring of the elderly, alerting the caregivers when necessary and forwarding pertinent information from big data systems for analysis. We did not either include articles on how big data can be used to customize individual searching for information, how it can be captured, analyzed, stored, transferred and protected or how big data can be used in clinical environments.

Upon the titles of these publications they were sorted into three categories, namely “big data as emerging discipline”, “challenges related to the use of big data” and “the potential of big data in health and social care”. These category names were used for a further iterative search. The knowledge acquired at this first step was used to write the introduction of this article and to identify possible gaps that need to be fulfilled in the future.

Most of the literature in the area of e-health, in which issues related to ambient assisted living and smart homes are included, has focused on the importance of developing a sustainable information chain, and on the need to develop innovative business models in order to stimulate the market of e-health as well on the managerial implications of collaboration between industries and health-care organizations [21, 24]. Relevant studies were selected by reading the abstracts of all papers that were pre-sc-
lected by their title. The inclusion criteria used was whether the paper was in the intersection between the use of big data and the expectations and potential of big data, this resulted in ten articles. To our best knowledge, there are no scientific publications that report empirical findings of the use of big data in smart homes and ambient assisted living. The authors, after finishing the review, discussed the findings and the criteria to follow. After five interaction meetings the discussion ended in an iterative draft, in which three issues considered as innovative and of relevance for the ambient assisted living and smart homes area, because the absence of publications that handle these issues, were established. They are:

a) The development of a coherent and sustainable information chain
b) Managerial implications
c) Innovative business models

The goal when identifying the issues has been, to create an initial framework that would guide stakeholders to identify issues of relevance for ambient assisted living and smart homes when discussing the future and potential of big data in the area. Further, the identification of future research issues in this area has to be of relevance for several stakeholders as well as for the research community. Reading the literature produced by the EU, it is possible to understand that one of the mayor reasons why e-health, including the use of big data, has not succeeded in its totality is because of the absence of empirical findings in the areas selected as well as the absence of evidence of the value added that big data can bring to managers, providers and receivers of services [24-26].

Results

What Is Big Data?
The mobile revolution which we are experimenting right now has completely changed how we capture the data and build intelligent systems. Complex organizations as health and social care organizations are indeed facing challenges to keep all the data on a platform which give them a single consistent view of their data. This unique challenge to make sense of all the data coming in from different sources and deriving the useful actionable information out of is what we today call as Big Data. “Big Data” refers to datasets whose size is beyond the ability of typical database tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered big data—i.e., we don’t define big data in terms of being larger than a certain number of terabytes (thousands of gigabytes). We assume that, as technology advances over time, the size of datasets that qualify as big data will also increase. Also note that the definition can vary by sector, depending on what kinds of software tools are commonly available and what sizes of datasets are common in a particular industry. With those caveats, big data in many sectors today will range from a few dozen terabytes to multiple petabytes (thousands of terabytes) [6, page 1].

In health care this is characterized by the increasing volume of data through digitalizing of already existing data and from new forms of data. Variety of data can be described by data being either structured, unstructured or semi-structured. Traditionally, point of care activities has generated unstructured data while financial and some clinical data can be more structured. Velocity of data is related to static or real-time data, something that is becoming more common in today’s health care [27]. Big data require, however, advanced techniques and technologies to enable the capture, storage, distribution, management and analysis of trustworthy information. Specifically information that has, until today, been hidden and embedded in work-routines and/or not considered when building hypothesis to analyze a phenomena and to deliver personalized health and social care as well as for specific populations.

Grove et al, 2012 suggest that big data comprises further categories of streams of information. They are [28]:
1. Web and social media data: Clickstream and interaction data from social media such as Facebook, Twitter, LinkedIn, and blogs. It can also include health plan websites, smartphone apps etc.
3. Big transaction data: Health care claims and other billing records increasingly available in semi-structured and unstructured formats.
4. Biometric data: Fingerprints, genetics, handwriting, retinal scans, and similar types of data. This would also include X-rays and other medical images, blood pressure, pulse and pulse-oximetry readings, and other similar types of data.
5. Human-generated data: Unstructured and semi-structured data such as electronic medical records (EMRs), physicians’ notes, email, and paper documents.

It seems interesting to note that all these categories are considered as key issues for the delivery of health and social care services at distance, through the use of telemedicine applications or in smart homes and ambient assisted living contexts.

Smart Homes and Ambient Assisted Living
The term smart home or ambient assisted living has been used to refer to and study the applications of state-of-the-art IT-based products that are emerging for use by individuals of private household [29]. It covers not only items of hardware like home computers and new consumer electronic goods but also software and services, like online information systems, that may be used with the hardware and software.

It has often been assumed that individuals have at least a computer or a smartphone in every home, and that information structures are used at all levels to move from public spheres (e.g. health or social care organizations) to more private ones (homes). Presently, many forms of information carriers are used in the home from simple handwritten notes, through printed documents to full audio-video information in the form of television broadcast, smartphones, mobile applications or telemedicine [30-36].

The research reports in this area have often focused on the private home life and its “electronics” rather than being concentrated on how to innovate management or how to take advantage of big data to deliver and plan
health and social care services. The discussion has focused on what kind of state-of-the-art technology that can be implemented and used in homes of the future to support and facilitate everyday tasks, which kind of platforms that can be used to improve monitoring of people within the home [37-41], the behavioral aspects of individuals when interacting with the technology in use, the number of devices that a system can handle at once and its relationship with time aspects. Studies have also focused on the selection of technological applications and sensors to measure physical activity as well as usefulness and acceptance of the technology in use [42-63]. Thereby, more often, current research in the area of smart homes and ambient assisted living has mainly studied one side of the equation, the market space perspective.

The use of IT-based applications as for instance telemedicine, to provide distant services from one site to another using information and communication technologies, is generally postulated to hold the capability to improve the accessibility of healthcare to underserved and rural communities [57-59]. No differences are made regarding the synchronous or asynchronous alternatives to deliver services or the importance to sample big data to identify personal needs of individuals with chronic diseases when interacting with them in real time (synchronous telemedicine) or to use clinical data, such as diagnostic images, or teleradiology or remote monitoring (asynchronous telemedicine). Most of the studies related to the advantages of telemedicine are focused on the possibilities to improve the quality of the services, reduce hospitalization and emergency visits and not on how to innovate managerial issues when offering health and social care at smart homes or ambient assisted living environments.

Another interesting trend in the area has been the need to empower citizens and patients and stimulate them to use technological innovations as support and assist living in smart homes. Many of the studies in this area has been performed as field studies involving mobile health care teams and IT support for supervision of status and symptoms in the home, where security, trust and accountability constitute important issues. An interesting trend in the area is the potential for integration of consumer products for communication and information processing with medical equipment.

Systems failures and disconnection caused by e.g., missed alarms, interference are still today considered a major problem in smart homes and ambient assisted living environments, because the consequences are often loss of confidence, anxiety, suffering and thus diminishing of life quality for the patients who choose to receive health care services at home. Three major issues requiring improvements have been identified: systems failures and disconnection, complicated interfaces, lack of specific knowledge for using central functions when IT innovations are implemented and environmental issues. Especially complicated interfaces threaten to cause confusion and stress both for healthcare personnel and patients. By developing prototype systems to field test with respect to reliability, security, privacy and usability, it is possible to reduce unexpected negative effects and thus costs for both households and health and social care organizations.

In the private area, volume of data has always been collected and used to improve supply chain management, logistic issues, development of business-to-business processes, or for quantification of risk (financial or civil) but also for fraud detection and to identify trends among specific target groups and thus to understand consumer behaviour. This is now starting to show in the area of smart homes as it has been assumed that the new generation of homes will be equipped with telematic gateways, allowing, in principle, online monitoring of every single individual to improve i.e. elderly care. Further, the digital presence is becoming part of everyday life in forms of smart phones, sensors, apps multimedia systems etc. Technology is expected to be embedded and connected with homes, cars, and will result in unprecedented amount of data allowing health and social care to combine information from several different sources and to create services limited by the constrains that exists in the manufacturing area. The ongoing digitalization of smart homes and ambient assisted living is consequently assumed to create an innovation platform for innovators and stimulate technological and economic growth.

Smart Homes, Ambient Assisted Living and Big Data: A Given Equation?

The health and social care market has, however, lagged behind other markets in the use and implementation of innovative technologies, especially in smart homes and for ambient assisted living due to problem stems from resistance to change and acceptance of new work processes (own clinical evidence rather than relying on protocols based on big data). Structural obstacles can also be related to the uncertain returns of IT or due to lack of procedures to integrate data, communicate findings and manage the rising demand for insight and evaluate outcomes.

Issues related to the lack of specific knowledge on how to manage IT can, however, be expected to be solved through improving the level of technological knowledge of the personnel and the end-users, but also by giving clear and exact instructions about how to use the technique and to whom to ask specialized questions. The issues related to these areas seem also to be relatively easy to solve through an appropriate test of the technique in use, or through the involvement of the end-users when developing IT-based solutions. Experiences from research on the effects of end-users participation in the development of health information systems has shown that involving end-users during the whole IT development process stimulate learning by doing process and thus induce to a better acceptation of the final system implemented [64,65].

More difficult to solve are however risks factors related to virtualized cloud environments and scalable, robust and secure storage of solutions that support huge amounts of data in smart houses. This is due to the fact that to process big data demand computational frameworks and models presupposes interactivity, local processing of data as well as a well-working infrastructure and integration of the social space with the technological space.

The future of big data for smart homes and ambient assisted living is apparently dependent on the well-functioning interaction between technology and a number of organizational and social dimensions, such as individuals’ preferences, organizational
culture, level of sophistication in the use of technology to collect data at different levels and special organizational and managerial factors that have to be managed to ensure an efficient future for the area.

**Discussion: The Challenges**

The challenges of the use of big data in smart homes and ambient assisted living can be divided into three different areas: the development of a coherent and sustainable information chain, managerial implications and the development of innovative business models.

**The Development of a Coherent and Sustainable Information Chain**

Health and social care in smart homes is today delivered by a diverse range of actors. Both private companies and non-profit organizations as well as public entities are actively involved in the care and support of elderly or individuals with specific needs. The devices used in smart homes i.e. sensors connected to smartphones, or mobile applications can transfer information of relevance either for the organizations or for the individual herself, that can be entered into electronic patient records and being used to produce personalized services. The increased amount of information demands however, that the information is sorted, analyzed and processed before it is used or interpreted by practitioners or specialists. This also demands the creation of a sustainable and trustfully information chain where all relevant information can be transferred from all producers to all consumers in a structured, well defined, complete, and undistorted manner to produce and consume information of relevance. Applications that are involved in a person’s health care and social services have to both produce and consume information of relevance. A trustfully information chain demands that the following prerequisites are fulfilled:

- **Semantic Interoperability**: the ability to exchange and use information between participating organizations, processes and systems. This includes all health care providers such as county councils, municipalities, private providers and even the patient and, if necessary, the patient’s relatives.
- **Aggregation**: the ability to combine and use information from different sources.
- **Intelligent views**: the ability to view the information whether it is text, graphics or photos, structured or unstructured in a way that is intuitive and easy to use.
- **Ethics and information privacy**: the patient information is anonymous where possible; a correct level of integrity is applied to all levels of the care and research, depending on the needs of each point of care. The information exchange is absolutely secure in terms of both technical redundancy and encryption.
- **Reusable, exploitable and scalable**: the coherent information chain allows the integration of a wide variety of information sources and systems, to be used in many existing and future system, so that the data does not needs to be manually transformed to fit into future applications.

It is also interesting to stress the possibilities of sharing open data and the implications that it might have on future development of services as there is a possibility for private actors, municipalities, county councils and citizens to collaborate and create services together. In practice this will probably mean that managers will have to collaborate across organizational boarders in order to identify the need for new services and the possibilities for sustainable collaboration among several different actors.

**Managerial Implications**

The tools available to handle the volume, velocity, and variety of big data demand new skills to integrate all the relevant internal and external sources of data. This is a new situation to most IT-units that support smart houses and ambient assisted living services. Although attention to technology is not sufficient any longer, there is a need to be able to develop a big data strategy and innovative policies and routines to handle a new phenomenon. Namely, that the information is created and transferred independent of the place of the expertise. Until today, information and relevant decision rights have been placed in the same location.

Managers of the future will need to be able to change the way smart houses and ambient assisted living leaders and organizations make many of their decisions. The managerial and cultural changes are enormous, data driven decisions will not only be a better decision, it will become significant for the leaders because of the possibility of being replaced by other who are better trained to confront the challenges. Apart from this there is also a need to set aside resources and time in order to analyze the data and make use of the future possibilities that big data gives.

**Innovative Business Models**

Chen 2012 suggests that big data analytics “enables companies to create new products and services, enhance existing ones, and invent entirely new business models” [7]. The ability to collect fine grained, location-specific, context-aware, highly personalized content through smart devices has opened new possibilities for advanced and innovative business models. Many different revenue models have begun to emerge for mobile apps, from paid or free to ad-supported apps or mobile gamification, which incentivizes participants (e.g., users or practitioners) by giving rewards for contributions. For mobile smart homes devices, companies are considering enterprise apps, industry-specific apps, e-commerce apps, and social apps.

A significant market transformation of smart houses and ambient assisted living contexts are supported by the accessibility to big data, the development of cloud computing and social media platforms as well as the emergence of customer generated forums, social platforms and crowd-sourcing systems. This will offer new opportunities to develop business models that follow the signals from the market but also support co-creation and co-finance between businesses and customers instead of the traditional business-to-customer one-way marketing approach used until today [65].
Future Research

Smart homes and ambient assisted living contexts can be understood as a multidisciplinary research area concerned with systems and services delivered in the home (in a wide sense, including also other everyday localities and whereabouts, also mobile, inhabited by individuals and their relatives) and as such, it has to be studied from both the supplier (services providers) and the demand (services consumers) perspective. The use of IT-tools to collect and use big data is therefore not any longer only of interest for HCI designers or IT developers but also for managers and decision makers.

In the society of today, consumers are more and more often making deliberate choices rather than being passive actors following the dictates of producer’s marketing efforts. To involve end-users to influence the design of services and user interfaces before industrial production is of crucial importance. Furthermore, issues related to consumer behavior and its implications for the supply side have been studied from rather diverse perspectives in several sub-disciplines of economic theory e.g., managerial economics contributing the results to important strategies for marketing, product development and information policies. To use big data to identify groups of consumers that bear some similarities e.g., preferences regarding latent consumption patterns, is crucial for the further development of smart homes area.

This is due to the fact that if consumers have positive attitudes to some IT-based products and services, they are more likely to accept, use and buy the products when they are finally produced and available at the market. However, it is important to differentiate between early and late adopters even in this area. Almost all research today is directed to study how early adopters, who are usually more educated people and thus with more economic resources, use and adopt innovations at home. If home informatics products and services will sell, it is crucial to concentrate on later adopters, their willingness to accept and pay for IT for home communication.

In the near future it will be important to develop a research agenda for smart homes and ambient assisted living that combine business and organizational initiatives to gain critical insight from the data collected.

Big data is today considered a new discipline. It is however, not obvious that this is correct. Big data is, still, without any doubt an inter-disciplinary phenomenon that span from cloud computing to methods for controlling false discovery rates when testing hypothesis and that demand cross-disciplinary collaboration to interpret its outcomes. Big data is still maturing and many questions remain without answers. To identify the real benefits of big data it is therefore important to discuss:

1. How to make information transparent and usable for managers and decision makers?
2. How to use big data to allow segmentation of customers and consequently more precisely tailored smart home services?
3. How to improve decision making and the link between evidence base delivery of services and big data analysis?
4. How to develop innovative business models in smart homes?
5. How to increase competition, productivity growth and consumer surplus?
6. Which policies will need to be developed to ensure privacy, security and liability?

References

2. Venkatesh V, Davis F D. A theoretical extension of the technology acceptance model: Four longitudinal field studies, Manage Sci 2000;46(2):186-204.

25. Communication from the commission to the European Parliament, the Council, The European Economic and Social Committee and the Committee of the Regions. A Digital Agenda for Europe.


Correspondence to: Vivian Vimarlund Jönköping International Business School PO Box 1026 551 11 Jönköping Sweden Tel: +46 (0)36 107175 Fax: +46 (0)36 165069 E-mail: vivian.vimarlund@iibs.hj.se