

The Utilization of the Internet of Things for a Learning Health System

Robert Steele

Abstract—The Internet of Things, incorporating new device, sensor and smartphone technologies can provide amongst other things a means for more extensive and more real-time, public health data collection. It can enable a far more sophisticated understanding of public health and how it changes over time than previously possible, and furthermore provide a new mechanism for a learning health system. In this paper we describe how the Internet of Things supports such a learning health system, and also how such a system can help contribute to new techniques for the advancement of healthcare knowledge.

Index Terms—Internet of things, public health, wireless sensor networks.

I. INTRODUCTION

The growth in the capabilities and uptake of Internet of Things-related devices, including biosensors, connected meters, and smartphones capable of acting as health sensor platforms has the potential to advance health data collection and healthcare knowledge creation.

Much of the research and commercial focus to-date has been on smartphones and sensors in their capacity to support individual health data capture, monitoring and feedback, and the full implications for health data capture to advance public health have been less well explored. The implications for a learning health system have been even less well developed.

This paper describes how the Internet of Things can be utilized to support a Learning Health System (LHS), so as to provide new health knowledge and research capabilities and which does so whilst still supporting participant anonymity and privacy.

An Internet of Things-enabled LHS has capabilities to capture population health data that are significantly more extensive and hence the abilities to measure the health of the population and contribute to healthcare knowledge are greater.

The Internet of Things-enabled LHS helps contribute to a new paradigm for health and healthcare knowledge advancement. The LHS provides a platform to learn from data collected from outside of clinical settings. This approach also intersects with the concepts of the quantified self and the citizen scientist, where the actions of members of the community can contribute to knowledge.

The methodologies and predictive analytics approaches supported by an Internet of Things-enabled Learning Health

System are extensible to numerous health application areas and condition areas.

II. RELATED WORK

Systems that involve members of the population in collecting data for some aggregate purpose have been explored in a number of application areas including pollution sensing [1] urban noise levels [2], urban traffic analysis [3] and vehicle fuel efficiency [4], amongst many other applications.

Our previous work has reported a particular form of learning health system, a learning public health system, that works by distributing informational public health interventions to individual's smartphones, measuring the effect of the public health interventions and incrementally making improvements to the public health interventions and thereby making improvements to public health [5]. The underlying architecture of that system, based upon Health Participatory Sensing Networks (HPSNs) [6], emphasizes the protection of privacy. It is the nature of public health data, that complete and precise data for every individual is not essential, but rather aggregate values across population groups and population health trends across time are rather the data of interest. It is this characteristic that allows such a public health information system to protect against identification of individuals and the approach taken also resolves the risk of re-identification based on quasi-identifiers in the form of information known about individuals that could potentially be used to match with and re-identify the submitted data. The platform's anonymity preserving characteristics have been analyzed using the common construct of k-anonymity [7]. K-anonymity is assured for a data set, if any individual is indiscernible from k other records based on quasi-identifiers.

The proposal in [5] is concerned with a learning public health system that incrementally makes improvements to public health interventions, and is a limited sub-case of what is introduced in this paper, a broader form of learning health system concerned with a range of capabilities.

The Patient-Centered Outcomes Research Institute (PCORI) supports various patient-centered research initiatives to consider the health outcomes of various treatments [8]. These are focused on patient outcomes but not focused on sensor-derived patient-generated data per se. The concept of the e-Patient is also relevant as it is a concept indicative of individual patients who are proactive in their healthcare and involved in information exchange [9]. The characteristics for successful adoption of mHealth have also been considered [10].

The concept of a Learning Health System as put forward

Manuscript received July 6, 2017; revised November 15, 2017.

R. J. Steele is with the Florida Polytechnic University, Lakeland, FL, 33805 USA (e-mail: rsteele@flpoly.org).

by the US Institute of Medicine has been published [11], but there is yet to be work published further considering the role of the Internet of Things in achieving such a system and the implications for citizen-based contributions to the advancement of healthcare knowledge that such systems can support. This represents a significant contribution of this current work.

III. AN INTERNET OF THINGS-ENABLED LEARNING HEALTH SYSTEM

Via the ever-increasing range of Internet of Things devices, individuals will be able to better measure and quantify their bodies and health. Where aspects of their health change or they change their behavior or lifestyle affecting their health for example, this can be recorded manually, via consumer-available meters or devices or reflected in changes in the data from sensors. In addition, as has been previously described in relation to HPSNs [6] this can be done in a way that maintains the anonymity of the individual. This new form of data capture can increasingly have an impact on health care knowledge and how the health system can 'learn'. It is also suggestive of new methodologies that can contribute to the advancement of public health and healthcare knowledge.

The Internet of Things conceptually includes devices worn by or in some other way associated with the participant, but also includes devices beyond the participant, which could include such things as ambient sensors.

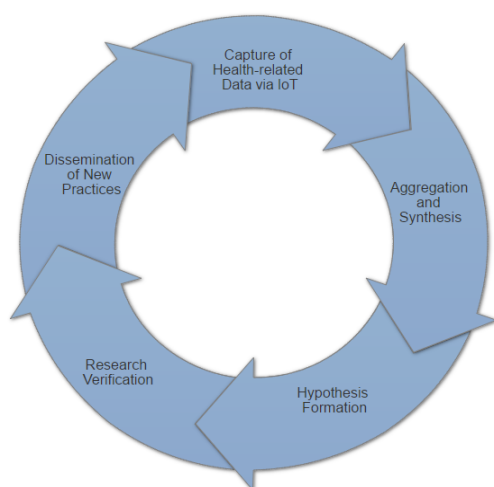


Fig. 1. An internet of things-enabled learning health system.

The Internet of Things can support a Learning Health System via the following cycle (see Fig. 1):

- capture of health-related data via IoT
- aggregation and synthesis
- hypothesis formation
- research verification
- dissemination of new practices
- continue the cycle

A. Capture of Health-Related Data via IoT

There are numerous commercially available sensors for capturing basic vital signs information, physical activity and movement, other physiological measures and increasingly

blood constituent information. For example, physical activity sensors, sleep pattern sensors or nutrition informatics-related sensors can all capture important lifestyle factor information pertinent to health and changes to these factors. There are also various other manually operated, health-related meters and tests available to consumers, which may or may not be digitally connected. With the ongoing advance of the applicable Internet of Things devices, and new types of devices, there will also be new data capture capabilities. An example which can draw upon multiple types of sensors are emerging approaches to automatically capture dietary intake and nutritional intake information [12]. This example also illustrates that the Internet of Things devices do not need to be only worn or physically embedded sensors but can also include ambient, environment or other types of sensors or meters.

Internet of Things capabilities will increasingly provide ways to capture more information about health, and lifestyle and capture this digitally [13]. It is important to note that it is not just information captured automatically, but health-related information is often captured via tests that individuals carry out manually, such as home urine tests, blood glucose tests, breathing tests amongst others.

Where data is captured automatically via sensors this contributes to the data being structured and where it is captured manually the data can also be structured via digitally connected meters or by the approach an individual takes to manually record it.

In this way, far more health-related data than previously available can be captured and this capture can be in a structured and digital form.

B. Aggregation and Synthesis

In this picture of a LHS, there is not a single methodology or LHS functionality implied. The key point is that there is an ability for individuals via Internet of Things devices, to collect healthcare data and submit this in a structured form for aggregation and analysis. This also implies a possible change to the traditional models of health knowledge creation [14].

Via the anonymizing infrastructure of a HPSN [6], it will be possible for various LHS functionalities to be supported while maintaining the privacy and anonymity of all participants. Supported functionalities could include formal research studies where scientists are collecting detailed data in a pre-designed way, about the day-to-day health status of research study participants. It could include public health initiatives where data is collected (anonymously) from a large sub-population to measure public health metrics and fluctuations. It could include public health systems where the effectiveness of public health interventions are studied and improved. Or it could include so-called 'quantified self' initiatives where members of the population capture and document their health data and experiences to contribute to the study of a condition or the understanding of health.

In each of these cases, the specifics of how the data are aggregated and synthesized may vary.

C. Hypothesis Formation

The nature of the hypothesis formation differs between the different possible functionalities supported by the LHS.

In the case of a formal research study there are at least

two options in relation to hypothesis formation. In an approach that follows more directly from the emerging capabilities of an Internet of Things-enabled LHS, the results of the aggregation and synthesis of collected device data could suggest possibly hypotheses for scientists to then further investigate.

Alternatively, for a formal research study of a more traditional nature, the first step in the LHS cycle in this case might be the hypothesis formation stage of the cycle. As per traditional scientific methods, in this case the initial hypothesis might derive from clinical observations, study of the literature, preliminary experiments or via other means.

In the case of a public health functionality, the hypotheses might derive from an analysis of the aggregated device data collected from the Internet of Things. For example, particular demographic groups, or those from particular geographical regions could be suggestive of where public health challenges exist or what might be the causes of these. It should be noted that due to the greater data-capture capabilities of the Internet of Things, there is potential for these initial public health insights and hypotheses to be new or more sophisticated than has occurred traditionally within public health. For example, the ability of the Internet of Things to capture more information (anonymously) in relation to lifestyle practices, could allow for new insights into lifestyle effects on health, in particular in relation to the so-called 'lifestyle diseases'.

In the case of less traditional quantified self approaches the precise methodologies that may occur are still evolving. While quantified self cases are sometimes referred to as experiments where $n=1$, there is nevertheless potential and promise within the quantified self approach to aggregate the experiences and data deriving from many individuals to gain broader insights. The quantified self also supports the concept that the hypotheses are not broad cross-population ones. Rather one of the promises of the quantified self is for it to be an enabler of the personalized medicine or precision medicine goal, with the findings potentially being personalized and having implications for just one individual or a small group of individuals.

In this case, the hypothesis formation stage may involve an individual deriving an hypothesis pertinent to their own health, or for example third-party researchers, if the data is provided to them, deriving a hypothesis just relevant to that one person's health.

D. Research Verification

Again the actions in this stage of the cycle are contingent upon the particular Internet of Things-enabled LHS functionality that is being considered.

In the case of a formal research study, this stage would involve using traditional research methods such as randomized controlled trials (RCTs), bench or wet lab-based research or other methodologies. In this stage the hypotheses derived either from the Internet of Things data or from hypotheses derived from other traditional means, are tested.

In the case of hypotheses derived from the Internet of Things, this allows the data from the public to just be a new source of insight whilst still maintaining traditional research methodologies in determining results.

In the case of a public health intervention-oriented system, this step involves study of the public health hypotheses derived from the Internet of Things data, or data resulting from an implemented public health intervention. This stage will involve the use of traditional research methods or statistical analysis to determine whether an hypothesis is supported.

In the case of quantified self methodologies, again the formalization of the quantified self approach and indeed the fully personalized medicine approach are still a work in progress. For example where data has been collected by an individual about their own health, this stage might involve them working to verify if what they have measured is reproducible.

E. Dissemination of New Practices

In this stage the results of the Internet of Things-enabled investigation are potentially disseminated to achieve the adoption of new practices or use of new discoveries.

In the case of a formal research study, the new practices discovered could be disseminated to clinicians, to patients directly via public health campaigns where appropriate or even via emerging learning health information systems where appropriate [5].

In the case of public health functionalities the new practices can be disseminated via traditional public health campaigns or via emerging smartphone-based informational interventions. In general for the public health case, the result of the preceding stages will have been the development of refined, more effective public health interventions.

In the case of the quantified self functionality, the nature of dissemination is still a work in progress. For example, where a finding is pertinent to just one individual personally, the dissemination would involve that individual being aware of this new insight into their health. For the quantified self functionality, again it corresponds to a means to further develop a personalized medicine approach, and the 'dissemination' may involve each individual learning of different findings that are relevant to them personally. That is, dissemination might not involve many receiving the same input.

F. Continue the cycle

The Internet of Things-enabled LHS process is an iterative or cyclic one. The cycle described above will now continue to occur, so as to iteratively improve healthcare knowledge, public health and personalized medicine.

IV. CHARACTERISTICS OF AN INTERNET OF THINGS-ENABLED LEARNING HEALTH SYSTEM

There are a number of notable characteristics of an Internet of Things-enabled LHS:

Heterogeneity: the implications of the Internet of Things applied to a LHS are heterogeneous and multitudinous in nature. The new data capture and analysis capabilities will affect many health endeavors from research methodologies, public health and public health interventions and personalized medicine to name just a few.

Contributing expertise: depending on the LHS

functionality being considered, the role of the health researcher will be hand-in-hand with the analytics, informatics or data science expert. Computational health approaches [15] will be increasingly intermingled with other methodologies in research, public health and health care.

Latency: the nature of the Internet of Things-enabled LHS is such that data from devices, can be collected almost instantaneously and immediately. This decreases the latency for public health and research purposes.

Analytics: a combination of human and computationally-derived insights can inform the various stages of the LHS cycle. The structured and digitized data capture will lend itself to the application of numerous supervised and unsupervised analytics approaches. Where machine learning/predictive models are used, there are a plethora of such existing models available: support vector machines, k-means, decision trees, logistic regression, naïve bayes and ensemble approaches to name a few [16]. In general the data collection will support the building of labeled data sets to support supervised learning approaches and will also lend itself to unsupervised approaches such as clustering. It should also be noted that various Online Analytic Processing (OLAP) approaches will be useful tools in analyzing the data collected by the LHS in various of its functionalities.

Big data: clearly such an LHS has the potential to be considered a big data system with it potentially being able to collect very large volumes of data. This implies that contemporary big data technologies, that go beyond the capacity of more traditional data warehouse systems will be applicable. In addition, such an LHS is likely to spur the need for further advances beyond the current crop of big data technologies.

Learning: it should be noted that the LHS provides for ‘learning’ in the broad sense that the iterative and cyclic process is not one limited to the application of ‘automated’ machine learning techniques, but the approach is one to involve an intermingling of human and computational analysis and work.

Extensibility: The LHS approach would be quite resilient to extension via new sensors, sensor systems [13] or consumer-available health measurement devices as they would present just an additional data measure. In this way, the relationship between the Internet of Things and the LHS is a persistent one. As Internet of Things devices improve and become available in more types, the fundamental learning algorithms may stay unchanged, but new analytics and research capabilities will become available due to the new data types available, and the greater accuracy of the data.

Actuators: the role of the Internet of Things for a LHS in this paper has been considered in terms of data capture and aggregation. Fundamentally the Internet of Things is conceived to also consist of actuators, able to affect the physical world to which it is connected or embedded. Even for the case of public health functionalities we have considered informational interventions – that is public health interventions that just involve the sending of some form of information to the end user, which could be textual, multimedia or other, with the goal of positively affecting

their health-related behavior. The addition of the Internet of Things to the LHS does present the theoretical possibility of other forms of intervention where devices or associated actuators are affected. This is not considered within the scope of this work.

Citizen science contributions: the captured or aggregated data can also include manual descriptions of health or actions as is consistent with both quantified self and citizen scientist activities. It is noted that any data would still be captured in an anonymized way. The LHS takes on the role of being an aggregator of these data and a source to support analysis of these aggregated data.

Crowd sourcing: the nature of the Internet of Things-enabled LHS is that it can provide a new form of ‘crowd sourcing’ of health and public health insights, can use the new capabilities of the Internet of Things and the various analytics tools and algorithms available to strengthen the learning capacity, and can also provide for traditional research methodologies to study emerging insights.

This paradigm recognizes that differences in lifestyle, health outcomes and various physiological measures can provide new insights not available from traditional clinical data or traditional research methods. There are a range and distribution of behaviors and health outcomes across the community naturally and this provides a variation which can be input for analysis. Internet of Things devices do have the potential to contribute to the capture of regular and structured data, helping to support aggregate analysis.

LHS vs Healthcare: it should also be mentioned that such LHSs as described blur the lines between public health intervention and “ubiquitous computing”-based telehealth techniques [17]. Whereas the telehealth approach could use Internet of Things devices in the care of an individual patient, the LHS paradigm could involve collecting aggregated data and in some cases providing targeted public health intervention benefits to a group of individuals. As described above, the individual care paradigm and the learning health system paradigm are in part reconciled where the learnings of the LHS are first deployed as guidance to clinicians, who utilize this information within the context of individual care.

Local/ personal architectures: there could be many basis’ for limiting the level of detail of the data flowing to the LHS, but this does not preclude the maintenance of far more detailed health-related data on the individual’s local device or portable personal health record [18]. More complex analysis of this data can also be carried out locally to benefit the healthcare of that individual, without transmitting this more complete data to the LHS server. This is again consistent with the paradigm of the quantified self and the citizen scientist, who may wish to privately work on the improvement of their health, but where desired by the individual, these data can be (anonymously) aggregated as inputs to hypothesis formation and for traditional scientific research approaches. There is also the possibility of the hybrid model where the individual citizen scientist investigates their own health based on more detailed data they have locally, and can contribute in a structured way their findings or conclusions to the LHS to assist in the broader population-wide research and improvement of health knowledge.

V. CONCLUSION

This paper details how the Internet of Things can be utilized to enable a learning health system. This includes consideration of how the cyclic approach of the system can make contributions to healthcare research and knowledge advancement, population health data capture and public health intervention and to personalized medicine. The paper also discusses how an Internet of Things-enabled learning health system can support a more citizen-driven approach to health understanding and improvement.

REFERENCES

- [1] B. Predic, K. Aberer *et al.*, "Exposuresense: Integrating daily activities with air quality using mobile participatory sensing," *IEEE in Proc. International Conference on Pervasive Computing And Communications Workshops*, 2013, pp. 303-305.
- [2] M. Wisniewski, G. Demartini, A. Malatras, and P. Cudré-Mauroux, "Noizcrowd: A crowd-based data gathering and management system for noise level data," in *Mobile Web And Information Systems*, F. Daniel, G. Papadopoulos, and P. Thiran, Eds. Berlin: Springer, 2013, pp. 172-186.
- [3] R. Ganti, I. Mohamed, R. Raghavendra, and A. Ranganathan, "Analysis of data from a taxi cab participatory sensor network," in *Mobile And Ubiquitous Systems: Computing, Networking, and Services*, A. Puiatti and T. Gu, Eds. Springer, 2012, pp. 197-208.
- [4] R. K. Ganti, N. Pham, *et al.*, "Greengps: A participatory sensing fuel-efficient maps application," in *Proc. The 8th International Conference On Mobile Systems, Applications, And Services*, 2010, pp. 151-164.
- [5] R. Steele and A. Clarke, "A sensor-based learning public health system," in *Proc. The 50th Hawaii International Conference on Systems Sciences (Hicss)*, IEEE, 2017.
- [6] A. Clarke and R. Steele, "Health participatory sensing networks," *Mobile Information Systems*, vol. 10, no. 3, 2014, pp. 229-242.
- [7] P. Kalnis and G. Ghinita, "Spatial k-anonymity," in *Encyclopedia Of Database Systems*, L. Liu and M. T. Özsu, Eds. Us: Springer, 2009, pp. 2714-2714.
- [8] S. Daugherty *et al.*, "Patient-powered research networks: Building capacity for conducting patient-centered clinical outcomes research," *Journal of The American Medical Informatics Association*, vol. 21, no. 4, 2014, pp. 583-586.
- [9] S. Okun and C. Caligian, "The engaged epatient," *Health Informatics: An Interprofessional Approach*, vol. 1, p. 204, 2017.
- [10] T. Davis, R. Diclemente, M. Prietula, "Taking mhealth forward: Examining the core characteristics," *Jmir Mhealth And Uhealth*, vol. 4, no. 3, 2016.
- [11] C. P. Friedman, A. K. Wong, and D. Blumenthal, "Achieving a nationwide learning health system," *Science Translational Medicine*, vol. 2, no. 57, 2010.
- [12] R. Steele, "An overview of the state of the art of automated capture of dietary intake information," *Critical Reviews in Food Science And Nutrition*, vol. 55, no. 13, pp. 1929-1938, 2015.
- [13] M. Swan, "Sensor Mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0," *Journal of Sensor and Actuator Networks*, vol. 1, no. 3, pp. 217-253, 2012.
- [14] C. Storni, "Report on the 'reassembling health workshop: Exploring the role of the internet of things,'" *J. Participat Med*, 2010.
- [15] R. Steele and A. Lo, "Future personal health records as a foundation for computational health," *Computational Science and Its Applications-Iccsa*, Berlin, Heidelberg: Springer, pp. 719-733, 2009.
- [16] X. Wu, V. Kumar *et al.*, "Top 10 algorithms in data mining," *Knowledge And Information Systems*, vol. 14, no. 1, 1-37, 2008.
- [17] R. Steele and A. Lo, "Telehealth and ubiquitous computing for bandwidth-constrained rural and remote areas," *Personal and Ubiquitous Computing*, vol. 17, no. 3, pp. 533-543.
- [18] R. Steele, K. Min, and A. Lo, "Personal health record architectures: Technology infrastructure implications and dependencies," *Journal of the American Society for Information Science and Technology*, vol. 63, no. 6, pp. 1079-1091.



Robert J. Steele holds a Ph.D. in computer science from Flinders University, SA, Australia awarded in 2000.

He currently serves as a full professor and institute director of the Health Informatics Institute at Florida Polytechnic University, Florida, USA. He has previously served as a division director of the Division of Health Informatics at the Medical University of South Carolina, full professor and chair of health informatics at The University of Sydney, and Associate Professor and Senior Lecturer of Computer Systems at the University of Technology, Sydney. His research interests include ubiquitous information systems, data analytics and health informatics.

Prof. Steele has previously served as the Vice Chair of ACM SIGMOBILE and has authored over 120 refereed publications and his research has been patented and successfully commercialized.