

Augmented Personalized Health: How Smart Data with IoTs and AI is about to Change Healthcare

Amit Sheth, Utkarshani Jaimini, Krishnaprasad Thirunarayan, Tanvi Banerjee
Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis)
Wright State University, Dayton OH, USA
Email: {amit,utkarshani, tkprasad, tanvi}knoesis.org

Abstract—Healthcare as we know it is in the process of going through a massive change - from episodic to continuous, from disease focused to wellness and quality of life focused, from clinic centric to anywhere a patient is, from clinician controlled to patient empowered, and from being driven by limited data to 360-degree, multimodal personal-public-population physical-cyber-social big data driven. While ability to create and capture data is already here, the upcoming innovations will be in converting this big data into smart data through contextual and personalized processing such that patients and clinicians can make better decisions and take timely actions for augmented personalized health. This paper outlines current opportunities and challenges, with a focus on key AI approaches to make this a reality. The broader vision is exemplified using three ongoing applications (asthma in children, bariatric surgery, and pain management) as part of the Kno.e.sis kHealth personalized digital health initiative.

Index Terms—Augmented Personalized Health, Smart Data, Internet of Things, Wearable, Sensors, Semantic Computing, Cognitive Computing, Perceptual Computing.

I. INTRODUCTION

The invention of stethoscope in 1816 by Rene Laennec fundamentally changed healthcare [1]. The earliest prototype consisted of a monaural wooden tube that for the very first time, allowed clinicians to investigate a patient's physiology without simply relying on what the patient self-reported. This marked the beginning of data-driven clinical diagnosis that fundamentally changed healthcare. A similar gestalt shift is happening now with the advent of low-cost sensors, wearables, mobile computing, and AI. The days of episodic healthcare where a clinician relies on information collected during a patient visit or what is reported in the lab tests he orders is coming to an end. We now have the ability to continuously monitor a patient not only in the clinical setting but in their homes, capturing physiological and environmental data across personal, public, and population levels.

Augmented Personalized Healthcare (APH) is expected to enhance healthcare by personalizing the use of all relevant physical, cyber, and social data [2] obtained from wearables, sensors and Internet of Things (IoT), mobile applications, Electronic Medical Records (EMRs), web-based information, and social media. The exploitation of all relevant data, relevant medical knowledge, and AI techniques will extend and enhance human health and well-being. The concept of *augmentation* refers to the aggregation and integration of

all the signals at the personal, public and population level obtained by analyzing data and knowledge from sensors and the Web that can affect human health, and then converting these signals and data into actions that improve health-related outcomes. These signals collected both passively (without patient engagement) and actively (with patient or physician engagement) can help make better and more timely decisions. This embodiment of APH is an entirely new approach to human health compared to the current episodic system of periodic care primarily centered around healthcare establishments (such as clinics, hospitals, and labs). APH involves continuous monitoring, engagement, and health management, where rather than treating a patient for a disease, the focus shifts to involving the patient in preventing disease, predicting possible adverse outcomes and preventing them through proactive measures, and keeping them healthy and fit with lifestyle changes. Rather than chronic disease management, it takes a holistic approach to improving the overall quality of life.

Let's review the key enablers of APH and significant progress made in the recent past that centers on the ability to capture, analyze, and exploit big data relevant to an individual [3]:

- 1) Using low-cost sensors we have the ability to continuously collect multimodal data about the patients physiological and psychological condition, activities (including meals, exercise and sleep), and about the surrounding environment (within the living space or outside).
- 2) We have access to patient clinical records (EMRs) as well as a large body of medical knowledge/protocols, and peer-reviewed medical literature online (PubMed) that can be used to obtain up-to-date information relevant to the health of a patient.

There are a lot of challenges faced while analyzing this multimodal data which are discussed in detail in the next section.

II. CHALLENGES

According to IBM, the volume of healthcare data has reached 150 exabytes in 2017 [4]. This is partly because of

our ability to effectively monitor health and the progression of chronic diseases such as asthma, ADHF, dementia and Type 1 diabetes by utilizing data from various sources, including wearable sensors and health information in electronic form. However, there are several challenges to be overcome in making sense of this large amounts of data and deriving actionable insights in a timely manner that can directly benefit the patient.

- 1) **Sensor data reliability and quality:** The sensors available in the consumer market come from a variety of vendors. This brings with it challenges of ensuring the reliability and quality of the data [5] [6]. For instance, there are a number of sensors for activity tracking (Garmin, Fitbit, etc.) that are available in the market, but we need to validate their use to ensure high quality data. There is no gold standard for these health monitoring sensors, so it is difficult to evaluate and determine the fidelity of sensor readings.
- 2) **Sensor data heterogeneity:** Sensor data is diverse and multimodal. To enable the proper interpretation of data and for determining remediation measures, it is essential to convert them into abstractions that ignore inessential differences, and provide an integrated view for taking actions. For instance, even though IBM Watson mines data to discover patterns and synthesize information from vast amounts of literature (such as PubMed), it is not clear how a clinician can utilize that information in the context of sensor data streams we have access to.
- 3) **Contextual interpretation and abstraction:** The knowledge that a patient has taken 10,000 steps in a day does not directly enable us to learn about the true activity level of the patient without knowing the patient's current condition. Just the number is of no value to a patient unless properly interpreted and abstracted, taking into account the physical and physiological characteristics of the patient. For instance, 10,000 steps can signify high activity for a person with a sedentary lifestyle while it may indicate low activity for an athlete. Thus, the same number of steps may have different meaning for different patients/individuals and can get mapped to different abstractions. Quantifying by itself will not succeed in these scenarios.
- 4) **Personalized Health:** Healthcare should be highly personalized. For example, in the case of asthma, the medication prescribed to ensure appropriate control is a function of the severity of the disease and the prevalence of the triggers. A low dosage SABA (Short-Acting Beta Agonists) may help someone to keep asthma symptoms in check in the fall season but it may not work for another patient who might have to resort to a higher dosage because of the higher severity

of the asthma and prevailing intensity of triggers in spring season. Thus, most disease management requires careful treatment personalization in the context of the triggers and vulnerabilities. This is one of the key reasons for active clinician involvement until now.

- 5) **Health objective:** The health objectives are disease specific. For instance, in diseases like asthma we have the objective of preventing asthma attacks, while in bariatric surgery a patient wants to get the intended outcome of the surgery. In the context of pain management, it is important to understand how chronic conditions such as hypertension can affect pain symptoms, which in turn can play a role in predicting the blood pressure levels of a patient. Furthermore, objectives are specific to individual patients (different patients may give different priority to longevity and quality of life) and the guidance may vary according to the clinician managing the patient care (one physician may be quick to prescribe an antibiotic while another may not).

III. SOLUTION OUTLINE: SEMANTIC, COGNITIVE AND PERCEPTUAL COMPUTING

How do we manage a chronic disease given that factors that affect the disease change very fast? In the past few years, the progress has been made in creating technologies for collecting the data and interacting with the patient through telemedicine. However, the clinicians, health workers, and patients cannot keep up with all the data being produced. Patients cannot understand it and clinicians cannot look at it in the available time. The patient-clinician interaction bandwidth is also limited. Instead of providing a clinician or a patient with all the data, we need to provide them with only actionable information. There is a need to convert the raw data into actionable information, interpreting the data in terms of knowledge (context), personalization, and using the appropriate AI techniques. One example of an AI-based reasoning approach we have introduced is of Semantic- Cognitive-Perceptual Computing [7] as in Fig. 1 and is discussed next. Many additional capabilities will also be needed, such as abnormality detection and predictive analytics. For brevity, many details are ignored.

Patient data consists of demographic and medical information from Electronic Medical Records (EMRs) and time series data collected from various environmental sensors, physiological sensors, and public Web resources. This low-level fine-grain data covers various facets ranging from the objective to the contextual and personalized. **Semantic Computing (SC)** deals with determining the type and value of the data, and situates it in relationship to other domain concepts. A large body of existing research on ontologies and Semantic Web techniques and technologies can be leveraged for this purpose [8]. **Cognitive Computing (CC)** deals with the representation and reasoning related to how humans interpret the data. In the healthcare and medical context, this

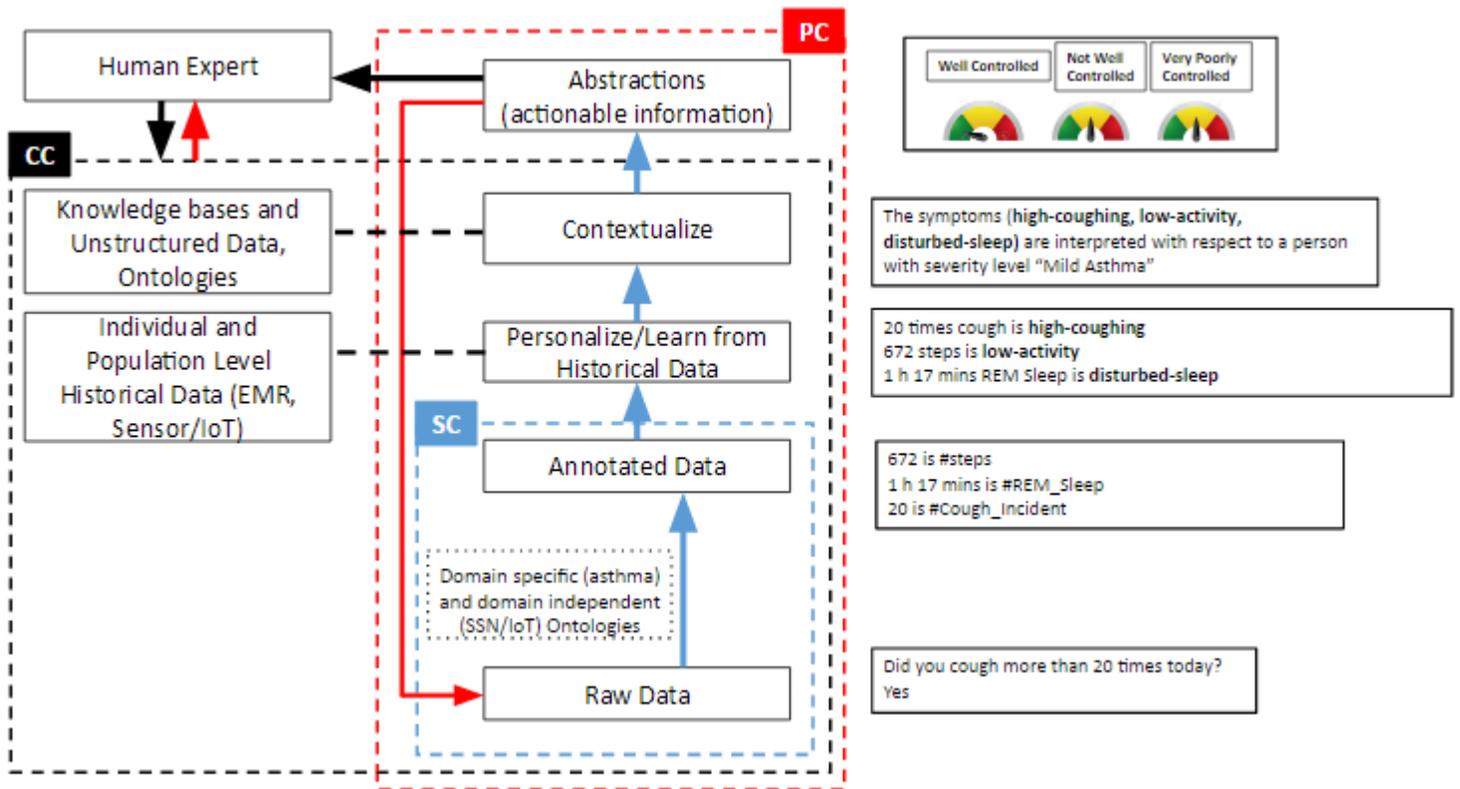


Fig. 1. Semantic, Cognitive, and Perceptual computing in an asthma-management scenario. The Semantic Cognitive (SC) system gives meaning to the raw data. The Cognitive Computing (CC) system gleans statistics and abstractions from patient observations and asthma literature to enable human interpretation, and the Perceptual Computing (PC) system iteratively interprets and explores them to learn personalized and context-specific normalcy for the patient and from it anomalies. REM: Rapid eye movement [7])

reflects non-trivial experience and domain expertise exhibited by doctors in abstracting and integrating multimodal data to enable actionable insights, taking into account contextual factors such as patient health history, physical characteristics, environmental factors, activity and lifestyle, to personalize the future course of action and treatment plans. To mechanize this we need to develop techniques to map raw sensor values to action-related abstractions, taking into account personal details (e.g., high activity translates to different amounts of workout based on age, weight, current health, weather, sport, etc. a low risk of heart problems depends on demographic and ancestry information, food habits, etc.). In general, we need to develop hybrid techniques that combines probabilistic as well as declarative models to formalize normalcy, and thereby detect anomaly. The anomaly itself may later be correlated and explained using patient-volunteered answers to health and symptoms-related questions. Anomaly detection is non-trivial because the notion of normalcy itself is intrinsically dynamic, based on spatio-temporal context, and requires personalization. It also requires uncovering various correlations among multi-modal data streams and discovering medically-relevant abstract interpretations and the factors that influence them. If sufficient patient data can be obtained through large-scale clinical studies or is personally volunteered, we can also explore the use of deep learning

techniques to uncover correlations and abstractions with predictive power.

Perceptual Computing (PC), in its simplest incarnation, is founded on rich domain knowledge that connects causes with effects and on reasoning strategies that can predict the effects of causes and explain the effects using causes. In a more general setting, perceptual cycle refers to interpreting current sensed data, attempting to build a model of the current situation, determining incomplete or ambiguous information and automatically seeking additional data in a targeted fashion, to minimize uncertainty. This knowledge can take several forms, and learning for creating relevant knowledge applicable to a range of data abstractions, from fine-grained data to coarse-level groupings, is challenging. The knowledge can be deterministic or probabilistic, transcending abstraction levels. In general, the symptoms manifested by a patient are a function of patient characteristics, how vulnerable/susceptible a patient is, what preventive measures/medications/treatments the patient takes, and how intense are the triggers. A key open problem is how to synthesize the *vulnerability score* associated with a patient with respect to a relevant health management objective to better capture the influences of aforementioned issues, and a control level to quantify and express the effectiveness of remedial measures in a manner

that is readily accessible to end users (whether a patient or clinician). Orthogonal to these issues is the development of efficient and effective strategies to perform requisite perceptual computations (such as the interleaved use of abductive and deductive reasoning steps) on a resource-constrained mobile platform that hosts all the sensors and to query necessary sensors for additional data needed to synthesize actionable information and alerts [9]. This is very challenging due to the multi-factorial nature of chronic diseases and usability issues for lay end users (patients). The Perceptual computing builds on Semantic and Cognitive computing to recommend timely and highly personalized actions understandable to humans (patient and clinician) which can lead to improved health outcomes.

Let us exemplify the evidence-based approach we advocate with the help of three concrete examples from our kHealth project – asthma, bariatrics and pain management. In each of the cases, we will demonstrate the kind of data captured, and their relevance. It will help in understanding what is the control, severity, and vulnerability score of a patient, which is the crux of everything. Specifically, we are able to collect data about the patient and relate to it, what is happening outside the clinical system. For example, in the case of asthma what leads to wheezing is a causal relation. Similarly, the relation between chronic conditions such as hypertension and physiological symptoms such as blood pressure is causal in nature (example described further below).

IV. KHEALTH APPROACH TO PERSONALIZED DIGITAL HEALTH

Usually, doctors see patients afflicted with chronic diseases infrequently but at well-defined time intervals such as monthly, every three months, semi-annually or annually (depending on the chronic disease and the patient's condition). The doctor's understanding of a patient's condition comes from what the patient tells the doctor (self-reporting) and what a doctor observes in the patient in person. Unfortunately, this may not provide a doctor with all the relevant information because the patient may not recall all the events that have happened in the intervening time, or may relate only recent events that may or may not relate to the current symptoms. We have developed a framework called kHealth for continuous monitoring of patients by collecting large quantity of physical-cyber-social and medical (EMRs) data with the intention of converting data into actionable information to make timely medical decision.

To demonstrate augmented health we discuss some of the action points and how augmentation assists with diagnosis, mitigation, and treatment associated with three applications of kHealth.

A. Asthma

Around 6.3 million children in United States suffer from asthma [10]. Asthma remains one of the leading reasons



Fig. 2. The kHealth kit for Asthma Management. The personal level data includes forced exhaled volume in 1 second (FEV1) measured using the spirometer, activity measurement using Fitbit, contextually-relevant health questions using the android application, indoor air quality parameters such as carbon dioxide, volatile organic compounds, indoor temperature and humidity measured using Foobot. The population level real-time IoT data includes pollen level, air quality index, outdoor temperature, and humidity. Finally, public level data includes asthma-related social media tweets. We will utilize the tweets from nearby locations with reference to the patient location to explain anomalous dynamics between personal level, population level, and public level signals.

for pediatric admissions in children's hospitals, and has a prevalence rate of approximately 10% in children, which leads to missed days from school and other societal costs. The kHealth kit for asthma management (Fig. 2) uses low-cost consumer-grade sensor devices such as Foobot, Fitbit, Spirometer, and Android Tablet [11]. The kit provides a platform for continuous monitoring of the patient's personal, public, and population-based health signals and sends alerts to the patient and/or to the clinician based on the patient's condition. These augmentations assist a clinician in determining the precise triggers and the patient susceptibilities, and in deciding the future course of action for prevention and treatment of the disease. More importantly, it can also enable a patient to take better control of their health and well-being by taking more timely actions on their own (e.g., in case of asthma, using an inhaler to ward off an asthmatic attack, or remaining indoors to minimize exposure to triggers such as weed pollens). For instance, we have anecdotal evidence of pediatric patients forced to go to hospital emergency rooms in the wee hours to deal with wheezing attacks and the family incurring significant financial costs that could have been avoided if only we had environmental data and timely alerts.

Our kHealth approach goes significantly beyond the current efforts that are focused on collecting data but are incapable of contextual and personalized processing of diverse, multimodal data to lead to action that can help control asthma [12]. In other words, our approach can help detect changes in conditions that can trigger an adverse event specific to an individual patient, and enables proactive measures for timely intervention. This effort is in close collaboration with a clini-



Fig. 3. kHealth kit for post-surgery bariatrics patients. The kit contains a water bottle sensor to remind the patient to drink water, a pill bottle sensor for sending reminders to the patient to take their vitamins and minerals, a Fitbit which measures patient’s daily level of activity, and a weighing scale which measures a patient’s body mass index, heart rate, and weight. All these sensors communicates via Bluetooth with an android application that asks relevant questions twice a day

cian at the Dayton Children’s Hospital and is being evaluated with 200 children with asthma.

B. Bariatrics

More than 36.5% of US adults were reported clinically obese during 2011-2014 according to a CDC report [13]. The estimated annual medical cost of obesity in the U.S. was \$147 billion US dollars in 2008; and the medical cost for a person who is obese was \$1,429 higher than those of normal weight [14]. By 2018 more than 40% of the US adults will be obese and the healthcare cost will rise to \$344 billion [15]. It is well established that weight loss surgery can play a significant role in reducing, or even eliminating medical problems associated with obesity. Unfortunately, weight regain is one of the biggest challenges, and more than 50% of patients regain weight within two years or more following their surgery. A lifetime commitment to diet and behavior modifications after surgery is essential for success after undergoing surgery. The main issue in a post-bariatric patient is the lack of follow up years after of their surgery and not checking on their diet plan adherence to other guidance.

The kHealth framework for bariatrics (Fig. 3) for managing and monitoring the post-surgery bariatrics patients uses the multimodal data from devices such as the Fitbit, pill bottle sensor, water bottle sensor, weighing scale, and android application. The framework monitors patient’s compliance with post-surgery progress, recommends post-surgical guidelines and motivates patients to have proper follow-ups. It aids bariatric surgeons in identifying noncompliance with direction by providing aggregated data of all the primary parameters to be monitored. For example, the kHealth application detects when a person does not take his/her vitamins, does drink 64

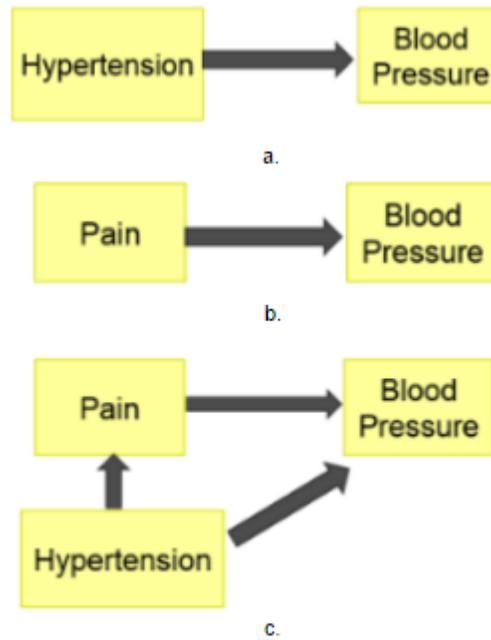


Fig. 4. Causal relations between a) hypertension and blood pressure, b) pain and blood pressure, and c) between the three components together.

ounces of water, eats at least 60 grams of protein in their diet, or performs some type of physical activity. It also determines the probability that a patient is likely to regain weight in a couple of months? The patient gets actionable information (personalized alerts) that is expressed as: drink more water, exercise a little, take a lab report, meet your doctor, or declare that you continuing to make the progress up to 18 months after the surgery, which is the expected, continuous timeline to determine you are keeping up with the timeline or there is something which is keeping you away from the timeline for the optimum results on the surgery, etc.

C. Pain Management

Our third example provides insights in the use of AI for augmented health. With age, pain is a common problem – studies report a prevalence ranging from 45%-80%, depending on age, the cohort under investigation, and the type of residence (independent vs. assisted living) [16]. The physical pain that a patient experiences is both subjective and difficult to quantify, which is what makes the problem both challenging and interesting. It is well known that hypertension causes an increase in blood pressure [17] and that an increase in pain levels (as experienced by patients suffering from chronic healthcare conditions such as sickle cell disease [18]) can cause an increase in their blood pressure levels [19]. However, there are a specific group of individuals who experience hypalgesia and exhibit symptoms of hypertension but do not feel pain as intensely as others [20]. In order to illustrate the nature of probabilistic modeling, we provide a concrete example of using Bayesian reasoning to model pain

and its potential effects using known medical knowledge. Fig. 4 indicates the simple one-to-one relationships between hypertension and pain as factors affecting blood pressure.

Using the Bayesian learning, we can compute the likelihood of the blood pressure being a certain value, say 120/80 (systolic/diastolic) (Fig. 4a) as a function of purely hypertension as the prior of the person exhibiting hypertension symptoms, as well as the conditional probability of blood pressure given the specified symptom of hypertension. In a similar manner, we can represent the causal relationship between pain and blood pressure. In the third case, however, i.e., with the fully connected, directed, acyclic graph (Fig. 4c), we compute the probability of the blood pressure as a function of both pain and hypertension, with pain also being affected by the hypertension symptoms. This indicates that given the conditional probabilities (of blood pressure given pain, and blood pressure given hypertension) and the priors of the probability of a person having a certain pain level and a certain level of hypertension, we can compute the joint distribution of these parameters co-occurring at these specific values.

Through the iterative use of the Bayes Rule, we can compute the probability of blood pressure as a conditional dependency with respect to pain and hypertension, with pain further dependent on hypertension. Using only two factors to measure physiological change in a person, we can see the need to formalize these conditional and causal relationships. If we add other physiological factors such as heart rate, pulse rate, galvanic skin response, and activity into this system, we can get as high as $n*(n-1)/2$ edges (where n is the number of variables used in the model) for the fully connected system. This highlights the need for using a *causal model* for inferencing behavior changes in corroboration with *background knowledge* to reduce the number of connected edges, which in turn reduces the amount of training data required to learn the model. We are currently exploring this hypothesis using AI techniques on EMR data collected from patients in collaboration with the Duke School of Medicine.

V. CONCLUSION

In as much as big data is an opportunity, it presents the challenges of deriving smart data [3]–actionable information that a patient can make sense of with or without a clinician, but typically within the context of relevant medical knowledge, that can allow them to take control of their health. This era of augmented personalized digital health is all about taking the data of an individual in its entirety, including relevant demographic information (e.g., ethnic group, or persons with similar lifestyles and health conditions), along with relevant medical knowledge, and converting them into both short-term actions and long-term (lifestyle, behavioral) changes to improve a person's quality of life.

ACKNOWLEDGMENT

This work was supported in part by National Institutes of Health under the Grant Number: 1 R01 HD087132-01 and NIH 1 K01 LM012439. The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

- [1] (2017, May) Stethoscope. [Online]. Available: <https://en.wikipedia.org/wiki/Stethoscope>
- [2] A. Sheth and P. Anantharam, "Physical cyber social computing for human experience," pp. 1:1–1:7, 2013. [Online]. Available: <http://doi.acm.org/10.1145/2479787.2479865>
- [3] A. Sheth, "Smart data how you and i will exploit big data for personalized digital health and many other activities," pp. 2–3, 2014.
- [4] (2017, Apr) Is Watson the best medicine? how big data analysis impacts healthcare. [Online]. Available: <https://www.ibm.com/blogs/internet-of-things/iot-and-healthcare/>
- [5] T. Banerjee and A. Sheth, "IoT quality control for data and application needs," *IEEE Intelligent Systems*, vol. 32, no. 2, pp. 68–73, 2017.
- [6] U. Jaimini, T. Banerjee, W. Romine, K. Thirunarayan, A. Sheth, and M. Kalra, "Investigation of an indoor air quality sensor for asthma management in children," *IEEE Sensors Letters*, 2017.
- [7] A. Sheth, P. Anantharam, and C. Henson, "Semantic, cognitive, and perceptual computing: Paradigms that shape human experience," *Computer*, vol. 49, no. 3, pp. 64–72, 2016.
- [8] A. Sheth and K. Thirunarayan, "Semantics empowered web 3.0: managing enterprise, social, sensor, and cloud-based data and services for advanced applications," *Synthesis Lectures on Data Management*, vol. 4, no. 6, pp. 1–175, 2012.
- [9] H. Patni, C. A. Henson, M. Cooney, A. P. Sheth, and K. Thirunarayan, "Demonstration: real-time semantic analysis of sensor streams," 2011.
- [10] (2017, Jan) Asthma. [Online]. Available: <https://www.cdc.gov/nchs/fastats/asthma.htm>
- [11] Khealth: Semantic multisensory mobile approach to personalized asthma care. [Online]. Available: http://wiki.knoesis.org/index.php/KHealth:_Semantic_Multisensory_Mobile_Approach_to_Personalized_Asthma_Care
- [12] [Online]. Available: https://www.med.umich.edu/1info/FHP/practiceguides/asthma/EPR-3_pocket_guide.pdf
- [13] [Online]. Available: <https://www.cdc.gov/nchs/data/databriefs/db219.pdf>
- [14] (2016, Sep) Adult obesity facts. [Online]. Available: <https://www.cdc.gov/obesity/data/adult.html>
- [15] CBSNews. (2009, Nov) Study: 40% of us may be obese by 2018. Available: <http://www.cbsnews.com/news/study-40-of-us-may-be-obese-by-2018/>
- [16] F. Landi, G. Onder, M. Cesari, G. Gambassi, K. Steel, A. Russo, F. Lattanzio, and R. Bernabei, "Pain management in frail, community-living elderly patients," *Archives of Internal Medicine*, vol. 161, no. 22, pp. 2721–2724, 2001.
- [17] High blood pressure or hypertension. [Online]. Available: https://www.heart.org/HEARTORG/Conditions/HighBloodPressure/High-Blood-Pressure-or-Hypertension_UCM_002020_SubHomePage.jsp
- [18] O. S. Platt, B. D. Thorington, D. J. Brambilla, P. F. Milner, W. F. Rosse, E. Vichinsky, and T. R. Kinney, "Pain in sickle cell disease: rates and risk factors," *New England Journal of Medicine*, vol. 325, no. 1, pp. 11–16, 1991.
- [19] P. S. Chawla and M. S. Kochar, "Effect of pain and nonsteroidal analgesics on blood pressure." *WMJ: official publication of the State Medical Society of Wisconsin*, vol. 98, no. 6, pp. 22–5, 1998.
- [20] Medscape log in. [Online]. Available: http://www.medscape.com/viewarticle/465355_3