

The Datafication of Everything – Even Toilets

Kwok-Chan Lun

Chief Executive Officer, Gateway Consulting Singapore, Singapore

Professor, National University of Singapore and Nanyang Technological University (retired), Singapore

Summary

Health informatics has benefitted from the development of Info-Communications Technology (ICT) over the last fifty years. Advances in ICT in healthcare have now started to spur advances in Data Technology as hospital information systems, electronic health and medical records, mobile devices, social media and Internet Of Things (IOT) are making a substantial impact on the generation of data. It is timely for healthcare institutions to recognize data as a corporate asset and promote a data-driven culture within the institution. It is both strategic and timely for IMIA, as an international organization in health informatics, to take the lead to promote a data-driven culture in healthcare organizations. This can be achieved by expanding the terms of reference of its existing Working Group on Data Mining and Big Data Analysis to include (1) data analytics with special reference to healthcare, (2) big data tools and solutions, (3) bridging information technology and data technology and (4) data quality issues and challenges.

Keywords

Data technology, data-driven culture, data analytics, data governance

Yearb Med Inform 2018:234-6

<http://dx.doi.org/10.1055/s-0038-1641199>

Introduction

In July 2013, FierceBigData [1] published a very interesting paper entitled “The datafication of everything – even toilets”. The attractiveness of the paper does not lie in the catchy title but in the increasing pervasiveness of the Internet of Things (IOT) and how in recent years, IOT has contributed to our increasing appreciation of data as a valued asset and the understanding of data science in shaping the way we live, work, and play. The whole world is collecting data – lots and lots of data. In the example given in the FierceBigData paper, we are beginning to see how even the toilet seat in our daily milieu of living has the potential of becoming a connected device. If you sit on the toilet seat, your weight can be measured. Insert your arms through the cuffs of the armrest and your blood pressures can be taken. The intelligent toilet seat of the future will have the ability to analyse your liquid and solid discharges for various disease markers. Smart toilets are also making debuts all over the world. In these smart toilets, toilet roll holders can notify the cleaners using WiFi technology when toilet papers are about to run out so that toilet users will not be caught in an embarrassing situation of not having toilet papers to clean up after a session. Smart toilets can also count and collect data on the number of flushes to produce a KPI (Key Performance Indicator) on usage. Sensors (thankfully not cameras) can detect non-motion or falls especially for toilets designated for the handicap and the elderly.

Impact of Information Technology

The International Medical Informatics Association (IMIA) and health informatics have come a long way over the past fifty years. At MEDINFO 2017 held in August 21 -25 2017 in Hangzhou, the People’s Republic of China, IMIA celebrated the 50th anniversary of its founding in 1967. The occasion provided a poignant reminder of how information technology has helped to develop and propel the development of health informatics over the last half century. At the time of IMIA’s founding, the healthcare community was just making early attempts to use computers to process large amounts of data collected from hospital billings and health insurance claims. One of these earliest systems was TMIS (Technicon Management Information Systems), a collaboration between Lockheed and the El Camino Hospital in California in the 1960s [2]. The early attempts of using information technology in healthcare to process data were purely administrative. These were followed some years later by attempts to evolve from these administrative, hospital-centric systems, to patient-centric systems integrating clinical components into administrative and financial systems. Some of the early efforts included the work of Larry Weed on the Problem-Oriented Medical Record (POMR) at the University of Vermont in the early 1970s [3] and the work of Homer Warner on the HELP System at the LDS Hospital in Utah at around the same time [4].

Since then, health informatics has made tremendous progress, largely riding on the

wave of advances in information and communications technology (ICT). Progress is evident from the current state-of-the-art in the development of electronic medical and health records, use of standards for messaging, terminology, disease coding and knowledge representation, remote delivery of healthcare through the use of telehealth and telemedicine, development of autonomous and non-autonomous decision support systems in patient care, advances in digital imaging, storage and retrieval within and across the healthcare enterprise, and, more recently, use of mobile devices for patient-centric care.

The Era of Data Technology

The increasingly pervasive use of ICT in healthcare has now started to spur advances in information technology as hospital information systems, electronic health records, mobile devices, social media and IOT are making an impact on the generation of data – big data. The Chinese saying, “长江后浪推前浪” (“the rear waves of the Yangtze River push the waves in front forward”) aptly describes how the success of the ICT era has ushered and propelled the development of the Data Technology (DT) era.

No other industry probably fits the definition of Big Data better than the healthcare industry in terms of volume, variety, and velocity [5]. Every day, the world generates about 2.5 quintillion (10^{18}) bytes of data of which healthcare accounts for a sizeable share. In terms of variety, the code set in the base classification of ICD-10 allows for more than 14,400 different codes. ICD-10 CM ramps up the code set to more than 14,400 different codes. SNOMED CT contains more than 311,000 active concepts and provides the core general terminology for the electronic health record. There are an estimated 30,000 human diseases known to medicine. The number of investigation procedures, laboratory investigations, and drug-use in healthcare are just mind-boggling. Anyone, even lacking an intimate knowledge of healthcare, will not fail to realize that the velocity of data in the field is phenomenal. Hospitals and clinics all over the world have to attend to millions of patients requiring care one way or another.

Jack Ma, the founder of the Ali Baba Group, and an “evangelist” for DT, gave a very interesting opening address at CeBit in Hannover, Germany, in March 2015. In his landmark speech, Jack Ma declared that “the era of Data Technology is here and in the society of tomorrow, data will be the most important means of production, innovation, social development to improve lives, productivity and even to increase profit” [6].

Today, we have no shortage of examples to show how very successful businesses have leveraged on data to succeed. Facebook has the world’s biggest and most up-to-date registration office (and this is done without legal regulation). Pixar has won 16 Oscars (and has never employed an actor/actress). Airbnb is the world’s biggest hotel chain (and does not own a single room). Uber is the world’s largest cab operator (and it does not operate a single car). Amazon is the world’s biggest book store (and does not have a single bookshelf). All of these companies have one success factor in common – data [7].

The strategic importance of a data-driven culture has been clearly expounded in a paper in the *Economist* entitled “Fostering a data-driven culture” [8]. The paper was based on the results of a survey done by the Economic Intelligence Unit (EIU) in 2012 on 530 senior executives from around the world. The results of the survey clearly demonstrated a clear link between the use of data and financial performance of the surveyed companies. Financially successful companies have found that the promotion of data-sharing across their rank-and-file has helped to generate a data-driven culture and recognize the importance of data as a corporate asset.

It is therefore timely for healthcare organizations and institutions to value the data that they collect and leverage on their data to gain business competitive advantage, improve operational workflow and efficiency, and provide better quality and safer patient care. For example, Geisinger Health System [9], one of the largest health care systems in Pennsylvania, USA, has, in recent years, moved beyond one-off projects to recognize and adopt data-driven care as an enterprise-wide strategy.

In order to leverage on data to achieve business competitiveness, organizations and institutions must have access to good

quality data and not just big data. Checks must be in place to ensure that data quality is good. There have been numerous cases in which the analyses of poor quality big data have resulted in spurious results. Indeed, papers have been published questioning the robustness and quality of medical data collected from IT systems. The paper by Wagner and Hogan [10] questioned the accuracy of medication data from the electronic medical record (EMR) system of a geriatric centre in the USA. Analyses of just the medication records showed significant levels of data error.

Not long ago, *Scientific American* published an article on “Why we must keep track of errors in Electronic Medical Records” [11]. It noted that “for all their promise, however, electronic medical records have their own flaws. In dozens of known cases, caregivers have entered information into the wrong chart or listed important details — such as drug dosages — incorrectly. Data sometimes disappear. In one case, a patient’s allergy to penicillin was improperly entered into an electronic record. The patient later received ampicillin and nearly died of shock”. The *Scientific American* article advocated the setting up of a national database for electronic record errors as learning from mistakes can help to save lives. Proper tracking can only help. Robust data, especially in healthcare, are obviously more important than big data because it is the value and not the size of the data that counts.

Demand for Data Analytic Skills

With the increasing availability of data, it is important that healthcare organizations and institutions know how to leverage on the data collected. The adoption of a strategic data-driven culture within the organization or institution calls for, among other things, the possession of in-house data analytic skill sets for the rank-and-file – from the C-Suite (CEO, CIO, CMIO, CFO). Everyone within the organization or institution must be mindful of the importance of data – the instruments for data collection, the kind of data collected, whether the collected data are properly defined and documented, guidelines

for inventory, classification, storage, and utilization etc. These are part and parcel of good data governance within the organization which should identify, define, and enforce processes, policies, standards, and technologies that are required to manage and ensure the availability, accessibility, quality, consistency, auditability, and security of data within the organization.

As not everyone is comfortable with looking at data, there has been rapid development of interactive data visualization techniques in recent years. Commercial dashboard fabrication to aid decision-making has become a profitable industry. Actually, the great Chinese sage Confucius saw the importance of data visualization as early as 504 B.C. “A picture is worth a thousand words” is an adage accredited to Confucius. Today, a spatial map can readily be visualized by rendering several million data points. One such study is demonstrated by the project to visualize a million geocodes from taxicab pickup and drop-off locations in New York [12]. The study was intended as a stress test on the R package ggplot2 to render one million data points. The analytic value of rendering and analysing the taxicab traffic data is to be able to identify busy hours of traffic in the city and make necessary plans for maintenance schedule, traffic engineering, and urban traffic planning.

The classic work of Dr. John Snow in describing the massive cholera outbreak that ravaged the Soho area of London in the 1850s amply demonstrated the value of data visualization in healthcare. By mapping the outbreak sites (data points), Snow was able to trace the source of the outbreak to the Broad Street pump. On ordering the removal of the pump handle so that Londoners could no longer draw water contaminated with *Vibrio cholerae* from the River Thames, the outbreak was promptly arrested [13].

The Time is Now

So, as the era of data technology dawns on us, there will be, increasingly, a whole lot of data out there in the healthcare ecosystem but there is not a whole lot of people who knows how to deal with them! There is already a worldwide shortfall of data scientists, leading to the proclamation by some that “the data scientist is the sexiest job of the 21st century” [14].

It is therefore timely to recognize the strategic importance of leveraging on data to advance healthcare and the time is now. And the time is now for IMIA as an international organization in health informatics to take that lead. Already, IMIA has a Working Group on Data Mining and Big Data Analysis. The time is now for IMIA to broaden the terms of reference of this Working Group to incorporate other aspects of data-driven healthcare and rename it as IMIA Working Group on Data-Driven Healthcare. The re-organized Working Group should focus on:

- Guidelines on data governance in healthcare;
- Data analytics with special reference to healthcare;
- Big data tools and solutions (the old Big Data and Data-Mining),
- Bridging information technology and data technology;
- Data quality issues and challenges.

With a timely revamp of the Working Group on Data Mining and Big Data Analysis to a Working Group on Data-Driven Healthcare, IMIA should be well poised to provide the international leadership in the era of data technology in the field of healthcare.

References

1. Pam Baker. The datafication of everything—even toilets. FierceBigData, July 29, 2013
2. Technicon Medical Information System (TMIS) – in clinfowiki http://clinfowiki.org/wiki/index.php/Technicon_Medical_Information_System
3. Problem Oriented Medical Record. Stritch School of Medicine, USA <http://www.stritch.luc.edu/lumen/MedEd/MEDICINE/subint/POMR.pdf>
4. Gardner RM, Pryor TA, Warner RH. The HELP Hospital Information System: update 1998. *Int J Med Inform* 1999;54(3):169-82.
5. Raghupathi W, Raghupathi V. Big data analytics in healthcare: promise and potential. *Health Inf Sci Syst*.2014;2:3.
6. Jack Ma's Opening Address at CeBit, Hannover, Germany 2015. <https://www.cebit.de/en/news-trends/news/jack-mas-opening-address-1609>
7. Uber, FB, Airbnb What do they have in common? Post by Scott Wang in E-Commerce. <https://scottwang333.wordpress.com/2015/11/25/uber-fb-airbnb-what-do-they-have-in-common/>
8. Fostering a data driven culture. A report from the Economist Intelligence Unit, *The Economist*; 2013
9. Hall SD. Geisinger moves from one-off projects to data use in everyday practice. *FierceHealthcare*, May 9, 2016
10. Wagner MM, Hogan WR. The accuracy of medication data in an outpatient electronic medical record. *J Am Med Inform Assoc* 1996 May-Jun;3(3):234-44.
11. Why We Must Keep Track of Errors in Electronic Medical Records. *Scientific American*. <https://www.scientificamerican.com/article/why-we-must-keep-track-of-errors-in-electronic-medical-records/>
12. Giday A. A Million Geocode Data Visualization With ggplot2. https://rstudio-pubs-static.s3.amazonaws.com/239436_cc5ab5d271ca4c229c01e33d0788b49d.html#
13. The John Snow Archive and Research Companion <http://johnsnow.matrix.msu.edu/index.php>
14. Davenport TH, Patil DJ. Data Scientist: The Sexiest Job of the 21st Century. *Harvard Business Review*, October 2012. <https://hbr.org/2012/10/data-scientist-the-sexiest-job-of-the-21st-century>

Correspondence to:

Dr KC Lun
Gateway Consulting Pte Ltd
#02-02 Crown Centre
557 Bukit Timah Road
Singapore 269694
E-mail: lunkc@gatewaypl.com