© 2014 IMIA and Schattauer GmbH

Challenges and Potential Solutions for Big Data Implementations in Developing Countries

D. R. Luna[1], J. C. Mayan[1], M. J. García[1], A. A. Almerares[1], M. Househ[2]
[1] Hospital Italiano de Buenos Aires, Buenos Aires, Argentina
[2] King Saud Bin Abdulaziz University for Health Science, College of Public Health and Health Informatics, Riyadh, Kingdom of Saudi Arabia

Summary

Background: The volume of data, the velocity with which they are generated, and their variety and lack of structure hinder their use. This creates the need to change the way information is captured, stored, processed, and analyzed, leading to the paradigm shift called Big Data.

Objectives: To describe the challenges and possible solutions for developing countries when implementing Big Data projects in the health sector.

Methods: A non-systematic review of the literature was performed in PubMed and Google Scholar. The following keywords were used: "big data", "developing countries", "data mining", "health information systems", and "computing methodologies". A thematic review of selected articles was performed.

Results: There are challenges when implementing any Big Data program including exponential growth of data, special infrastructure needs, need for a trained workforce, need to agree on interoperability standards, privacy and security issues, and the need to include people, processes, and policies to ensure their adoption. Developing countries have particular characteristics that hinder further development of these projects.

Conclusions: The advent of Big Data promises great opportunities for the healthcare field. In this article, we attempt to describe the challenges developing countries would face and enumerate the options to be used to achieve successful implementations of Big Data programs.

Keywords

Big data, developing countries, computing methodologies, data mining, health information systems

Yearb Med Inform 2014:36-41
http://dx.doi.org/10.15265/IY-2014-0012
Published online August 15, 2014

Introduction

"It's the data, stupid". Jim Gray

Over the last decade, there has been a data explosion, caused by the exponential increase of data generation through the progressive digitization of virtually every aspect of everyday life. As a result, more data have been created in the past 2 years than in the entire history of mankind [1].

Some aspects of this new reality hinder data use, e.g. their volume, the speed with which they are generated and their diversity. This situation creates the need to change the way data are captured, stored, processed, and analyzed [2, 3].

The first organizations to use Big Data (BD) techniques were information technology (IT) companies like Google, which created most of the software infrastructure needed to manage data in a scalable manner [4-8]. Although organizations in the healthcare field have been more cautious to use BD in their processes, this attitude is changing [9]. The adoption of electronic health records has grown considerably in recent years, with a subsequent increase in the amount of information generated [10], either quantitative (e.g. lab results), qualitative (e.g. notes), or transactional (e.g. medication orders) [11]. Furthermore, the emergence of genomics and the shift in focus to personalized medicine will push healthcare organizations towards the adoption of BD techniques [12].

Given the growing importance of BD, and the potential benefits of its use in the healthcare field, the objective of this paper is to describe the challenges that developing countries face when implementing BD projects in healthcare, and ways in which they can solve them.

A non-systematic review of the literature based on the bibliographic references in MEDLINE was performed, broadened by a Google Scholar search. The following keywords and their combinations were used: "big data", "developing countries", "health information systems", "data mining", and "computing methodologies". The search was performed without date or language filters.

A thematic review of the selected articles was performed, in order to provide a conceptual approach to BD in general and its fields of application. Afterwards, the factors that constitute challenges for BD projects, regardless of the application domain, were enumerated. Finally, certain categories were established to gather the challenges faced by developing countries, and actions to overcome them were proposed.

Definition and Conceptual Framework

The term BD lacks a consistent definition, but it can be defined as a new paradigm, with a different outlook on data and how to analyze them, and a set of technologies designed to extract value from data, which could not possibly be gained using previously instituted techniques, given their volume, the speed with which they are acquired, and their diversity [13-15].

The term emerged in the 1990s, referring to the growth in volume of data

[16]. In 2001, Doug Laney enriched the conceptualization of the term, stating that BD is characterized by three qualities, the 3 V's: volume, velocity, and variety [17]. Recently, there have been other candidates to integrate this group, the most widespread being veracity [18] (See Figure 1):

'Volume' refers to the exponential increase in the amount of data that are generated and stored. It is estimated that data production will be 44 times greater in 2020 than it was in 2009. BD techniques seek to generate knowledge from these large amounts of data.

'Velocity' represents the increase in frequency with which data are delivered. The growth of integrated sensors in all types of devices, and the increasing adoption rates of mobile phones worldwide contribute to the continuous influx of data. BD projects seek to use these data to enable decision making in real time.

'Variety' describes the different formats data can adopt, such as images, free text, video, and sound, among others. BD tries to harness existing data even if they are not structured or they have non-standard formats.

'Veracity' refers to the confidence level associated with certain types of data. It is impossible to remove the inherent unpredictability of some data, such as weather data, or purchasing decisions of the public, so it is necessary to incorporate this dimension when planning BD projects.

"Business Intelligence" (BI) is a set of techniques used to analyze consolidated information and business processes in order to achieve improvements in competitiveness [19]. BI differs from BD in that BI requires well consolidated data to generate knowledge, whereas in a BD resource this is usually not possible, because data are more diverse, and new data are added constantly.

The adoption of BD is growing, generating benefits across multiple sectors. With the development of these techniques, the acceleration of scientific discovery and innovation is expected, as well as improvements in the understanding of human processes and social interactions, the acceleration of economic growth and improvements in health and quality of life [20].

There are several ways in which BD can help healthcare organizations to improve quality and efficiency [11, 21-24]:

- generating new knowledge, for example allowing the creation of an observational evidence base to answer questions that are very difficult to answer using clinical trials;
- disseminating knowledge, helping physicians through clinical decision support systems (CDSSs), which can provide suggestions and predictions in real time based on the personal data of patients;
- translating personalized medicine initiatives into clinical practice by providing the opportunity to use analytical capabilities, which can integrate genomics with data extracted from medical records;
- empowering patients by providing them with personalized information regarding their health and healthcare;
- improving epidemiological surveillance: tools could be developed and integrated to medical records to predict the occurrence of highly prevalent or deadly diseases in the population.

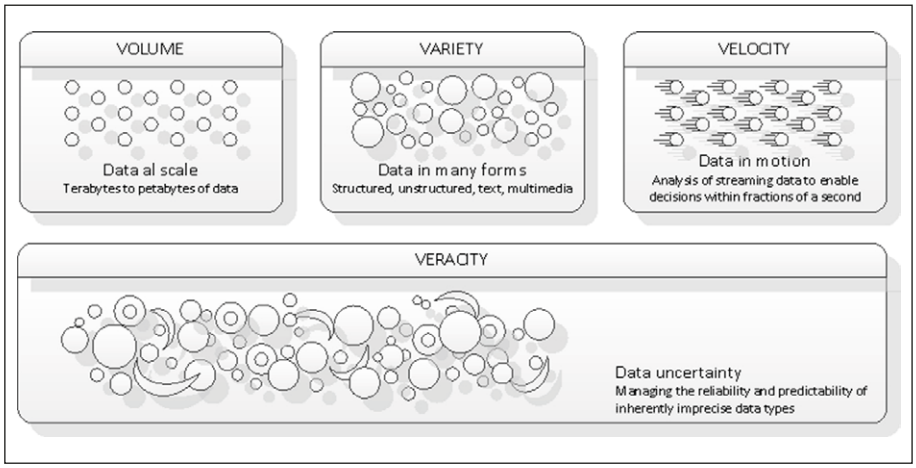

Fig. 1 Adapted from [16] Analytics: the real world use of big data

Implementation Challenges of Big Data Projects

Even though many benefits are expected with the implementation of BD projects in all areas, there are difficulties common to all of them, regardless of whether they are conducted in developing or developed countries. Developing countries and their healthcare industries in particular, have unique characteristics that merit special analysis on the challenges faced by the application of BD and the ways they can be surmounted.

In this section we develop six broad categories to organize the content; with each domain we describe at first the difficulties that are common to all BD projects, and finally the challenges and opportunities to overcome them that are specific to developing countries and the healthcare sector.

Data Capture

As we stated previously, the exponential growth in the amounts of data generated is one of the most important factors to understand the BD phenomenon. Data sets are becoming larger and more difficult to manage using traditional database tools. As a result, organizations are faced with difficulties to capture, store, manage, and analyze data in a timely manner [15].

Consequently, this situation creates new infrastructure needs, and significant economic costs. Fortunately, storage costs are also decreasing. This allows for the capture of useful data, such as location data, which permit the mapping of real-time events for epidemiological surveillance.

Perhaps the greatest potential regarding the capture of data for developing countries originates from the progress mobile

networks are making [25], and the growing market penetration of cell phones [26], which hold more sensors at lower prices with every passing year.

The growing adoption of mobile phones, 80% of which are located in developing countries [27], offers the possibility to use the data they provide to improve development programs. For example, SMS for Life uses a combination of mobile phones, SMS messages, the Internet, and electronic mapping technology to track weekly stock levels of malaria drugs at public health facilities. This program improved the distribution of malaria drugs in rural Tanzania, reducing facilities without stock from 78% to 26% [28]. In 2013, this initiative encompassed several countries in sub-Saharan Africa from Ghana to Kenya, with plans to increase the number of countries reached [29].

Infrastructure

A robust physical infrastructure is a key point for the operation and scalability of a BD project. It is based on a distributed model, where data can be physically stored in different places and integrated through networks. The fundamental condition to take advantage of this capacity lies in the quality of telecommunications, which offer a gateway to BD projects [30].

To perform the intensive analytical tasks required by BD projects, a special server architecture is necessary, comprising thousands of nodes with multiple processors and disks connected by a high speed network working in a distributed way [31]. Large Internet companies like Google, Microsoft, Yahoo, and Amazon use this architecture with centers distributed throughout the world offering their services.

All these changes in infrastructure involve substantial costs, generating economies of scale that favor large Internet companies [32], which take advantage of these barriers to provide infrastructure as a service (IaaS) to organizations who cannot afford them [33].

In addition, apart from the hardware infrastructure, an additional component is required: the software used to implement BD. The production, adoption, and adaptation of this software are key ingredients for BD, and require a properly trained workforce [30].

Many developing countries lack the storage and communications infrastructure needed to organize and integrate the amount of information that is generated in a BD project. Not only do these countries lack these resources, but they don't have the computing capacity to analyze them. Deficiencies in electrical grids and telecommunication networks are common to many of them, and technological improvement projects have to compete with more pressing needs such as alimentation [34-36].

The vast majority of the necessary hardware resides in developed countries, and access to information and resources is skewed by a very unequal distribution of telecommunication capabilities to access them [30].

As an inexpensive alternative, it is possible to use clusters of video game consoles to substitute the supercomputers used in the analysis of BD [37,38]. Another possibility is the use of resources present in developed countries, through private companies, which provide the necessary IaaS, such as Amazon Web Services and Google Compute Engine [39,40], usually at a competitive price, through the cloud. These resources are useful because the efficiency these companies achieved allows them to offer safe and high quality services at affordable prices.

One example of the use of IaaS in developing countries is DHIS 2, a tool for the collection, validation, analysis, and presentation of aggregate statistical data for health information management activities. DHIS 2 has been implemented in more than 30 countries in Africa, Asia, Latin America, and the South Pacific, and countries that have adopted DHIS 2 as their nation-wide health information system software include Kenya, Tanzania, Uganda, Rwanda, Ghana, Liberia, and Bangladesh [41].

Regarding software used for organizing, integrating, and analyzing data, production is limited by the lack of a trained workforce, and the possibility to purchase or license the necessary systems is often not an option for developing countries. However, there are open source options with strong communities that provide the necessary functionalities for free. The most outstanding example is Apache Hadoop [42], a platform for processing large amounts of data distributed on computer clusters used by companies like Yahoo and Facebook.

Organizational Changes - Workforce

According to Villars et al, BD deployments require new IT administration and application developer skill sets. Additionally, the people who possess these skills are a scarce resource given the high market demand [15]. Hal Varian, Google's chief economist, contends that statisticians will have the job most in demand in the next decade [43].

To take advantage of the opportunity created by BD, trained human resources are needed, with the ability to manage and analyze data, with knowledge in computer science, statistics, and mathematics. These resources are not only scarce in developing countries: it is estimated that the United States will have a shortage of 160,000 professionals with these skills by 2018 [44]. Most of these scientists are recruited by major technology companies in the core countries, with the consequence that countries with fewer resources will suffer this deficit even more.

Some developing countries are better positioned in this regard, including Brazil, Russia, India and China (the BRIC countries). In 2008, 40% of the specialized resources were trained in these countries [30].

As Internet and technological advances allow the outsourcing of infrastructures, there also exists the possibility to recruit the human resources needed for a BD project over the web. As an example, the Kaggle platform allows any organization to set a prize, and specialists from around the world can compete to solve BD problems [45]. Ultimately, this possibility depends on the economic resources that can be offered. One important example of a nonprofit organization is Datakind, a group of data scientists that work with high impact social organizations to improve their decision making processes [46].

Integration and Interoperability

One of the greatest challenges BD faces is to integrate data from many different sources. The use of standards to achieve interopera-

bility between systems is a core requirement to effectively integrate information [47].

The major difficulty for achieving interoperability among multiple repositories of BD lies in the differences in the metadata used in one repository with respect to other repositories. Without standards for these metadata, the integration of data generated in BD projects will be even more challenging [48].

Health information systems are often fragmented and isolated in information silos hindering analysis and improvements in healthcare assistance [49]. This problem requires a political rather than a technological solution. In most cases, the required standards for systems to interoperate already exist, and they are the same in developing countries than in developed countries [50]. It is necessary to achieve consensus between government organizations, businesses, and stakeholders in order to advance in the development of digital agendas.

Developed countries have made progress in spreading digital agendas in the last decade, and are now better positioned than developing countries, although lately this gap is narrowing. According to the World Health Organization (WHO), since 2008 more than 20 developing countries are in the process of implementing strategic plans for eHealth [51].

The WHO and the International Telecommunications Union (ITU) published a document in order to help countries in the process of generating a national eHealth vision and an action plan (National eHealth Strategy Toolkit) [52]. These resources are especially useful for governments in developing countries.

Privacy and Security

Some characteristics of BD, such as the relative lack of structure and the informal nature of some data, can be a problem if they are sensitive, with potential privacy, safety or legal issues. Traditional database management systems support granular security policies that protect data at various levels. The software used in BD projects does not usually have these safety measures [15].

Another important challenge includes the security infrastructure and privacy policies. It is crucial to apply not only legal but also ethical considerations on the security of the data as soon as possible. The development of strategies to report on how data are collected, how they are protected, and how they will be used should be considered and recognized as a necessity [53].

Likewise, an action plan should be contemplated in case of possible data losses or security breaches. Sharing information in a clear and careful way will help reduce concerns related to security and privacy [54].

It is essential to ensure the privacy and confidentiality of personal data, especially with regard to the use of BD in healthcare. These factors should be considered part of the structure of a BD project from the beginning.

In the U.S., the Health Insurance Portability and Accountability Act's (HIPAA) privacy and security rules govern how personal information should be protected, and define what safety standards should be applied [55,56]. A regulatory framework is essential to build trust between all parties involved. This is why developing countries are advised to advance policies and regulatory frameworks in order to ensure the privacy and security of sensitive data.

Whatever the data, when they are related to humans, safety concerns will inevitably arise. If the goal is to share data, those who provide them have to be able to trust those who assume the responsibility of caring for their information [57,58]. This will only be achieved with an appropriate regulatory framework.

Adoption

Data should be managed as a strategic asset within organizations. Existing barriers to the adoption of BD are usually cultural. Many organizations do not implement BD programs because they cannot appreciate the way in which data analysis can enhance their businesses [15].

Defining objectives and expected outcomes are critical in order to establish a governance capable to sustain projects of this magnitude. A BD program should include the people, processes, and policies needed [59].

The difficulties that were previously reviewed: economic issues, poor infrastructure, and lack of trained personnel, are common to most developing countries, and generate a gap in the adoption of BD as compared to developed countries that is equivalent to the digital divide [30].

Some ways to accelerate the adoption of BD techniques in developing countries are simple, such as sharing experiences and lessons learned [36]. Currently, developing countries have more access to sources of scientific information, due to the increased penetration of Internet, the emergence of the Open Access movement, which allows to access to scientific articles of prestigious publications for free, and the advent of new tools for searching scientific literature, like Google Scholar. A recent paper shows that Google Scholar provides greater access to free full-text articles than PubMed [60].

Spearheading a movement of BD for development, the United Nations launched Global Pulse in 2009, an innovation lab that aims to raise awareness of BD opportunities and catalyze the adoption of BD tools to help policymakers understand human wellbeing and emerging vulnerabilities in real time [61]. Other international institutions, like the World Economic Forum and the OECD are working to use BD for development goals [62,63].

The implementation of strategic partnerships allowing for regional integrations will help developing countries reduce costs and lower the difficulty of implementing BD projects. If these countries want to adopt BD, it is essential for their governments to have a strategic vision on information technology, and implement regulatory mechanisms as well as incentives for its development.

Conclusions

The advent of BD provides great opportunities for the healthcare field. There are challenges to the implementation of any BD program, consisting of the exponential growth of data, the special infrastructure needed to analyze them, the need for a specially trained workforce, the need to agree on interoperability standards to integrate

data, the privacy and security risks involved in these projects; and the requirement of a strategic vision that contemplates the need to include people, processes, and policies to ensure their adoption.

Developing countries have particular characteristics that hinder further development of these projects. In this article we attempted to describe these challenges focusing on the options these countries could use to achieve successful implementations of BD programs. We hope this article will be useful for those in charge of making policies or leading BD projects in developing countries.

Table 1 Recommendations for developing countries

Data capture	Take advantage of the high penetration rates of mobile phones to collect usage-associated data and sensor data for innovative BD projects.
Infrastructure	Circumvent infrastructure and economic deficits using IaaS and open source software.
Organizational changes - Workforce	Increase the number of data scientists trained. Make partnerships with nonprofit organizations like Datakind when trained resources are needed.
Integration and interoperability	Advance in the creation and adoption of digital agendas.
Privacy and security	Institute policies and regulatory frameworks to ensure the privacy and security of sensitive data.
Adoption	Implement strategic partnerships with private and public institutions with expertise in BD tools and techniques.

Acknowledgements
The authors would like to thank Mariana Licia Napoli for her contribution to the translation of the original manuscript.

References

1. Jee K, Kim G-H. Potentiality of Big Data in the Medical Sector: Focus on How to Reshape the Healthcare System. Healthc Inform Res 2013 Jun;19(2):79–85.
2. Ohlhorst FJ. Big Data Analytics: Turning Big Data Into Big Money. John Wiley & Sons; 2012.
3. Ouellette, J. Scientific Data Has Become So Complex, We Have to Invent New Math to Deal With It [Internet]. Wired Science. [cited 2013 Dec 10]. Available from: http://www.wired.com/wiredscience/2013/10/topology-data-sets/
4. Ghemawat S, Gobioff H, Leung S-T. The Google file system. SOSP '03 Proceedings of the nineteenth ACM symposium on Operating systems principles; 2003. p. 29–43.
5. Melnik S, Gubarev A, Long JJ, Romer G, Shivakumar S, Tolton M, et al. Dremel: interactive analysis of web-scale datasets. Proc VLDB Endow 2010;3(1-2):330–9.
6. Corbett JC, Dean J, Epstein M, Fikes A, Frost C, Furman JJ, et al. Spanner: Google's globally-distributed database. Proceedings of OSDI. California, United Sates: USENIX Association; 2012. p. 251–64.
7. Chang F, Dean J, Ghemawat S, Hsieh WC, Wallach DA, Burrows M, et al. Bigtable: A distributed storage system for structured data. ACM Trans Comput Syst TOCS 2008;26(2):4.
8. Dean J, Ghemawat S. MapReduce: simplified data processing on large clusters. Commun ACM 2008;51(1):107–13.
9. Burghard C. Big Data and Analytics Key to Accountable Care Success [Internet]. IDC Health Insights; 2012. Available from: http://www-01.ibm.com/common/ssi/cgi-bin/ssialias?infotype=SA&subtype=WH&htmlfid=IML14338USEN
10. Charles D, Furukawa M, Hufstader M. Electronic Health Record Systems and Intent to Attest to Meaningful Use Among Non-federal Acute Care Hospitals in the United States: 2008-2011. ONC Data Brief 2012;1:1–7.
11. Murdoch TB, Detsky AS. The Inevitable Application of Big Data to Health Care. JAMA 2013;309(13):1351–2.
12. Chawla NV, Davis DA. Bringing Big Data to Personalized Healthcare: A Patient-Centered Framework. J Gen Intern Med 2013 Jun 25;28(S3):660–5.
13. Bourne PE. What Big Data means to me. J Am Med Inform Assoc 2014 Mar 1;21(2):194.
14. Ward JS, Barker A. Undefined By Data: A Survey of Big Data Definitions. ArXiv Prepr ArXiv13095821 [Internet]. 2013 [cited 2014 Mar 28]; Available at: http://arxiv.org/abs/1309.5821
15. Villars RL, Olofson CW, Eastwood M. Big data: What it is and why you should care [Internet]. IDC; 2011 [cited 2013 Dec 10]. Available at: http://sites.amd.com/us/Documents/IDC_AMD_Big_Data_Whitepaper.pdf
16. Diebold F. On the Origin (s) and Development of the Term'Big Data' [Internet]. Penn Institute for Economic Research; 2012 [cited 2013 Dec 19]. Available at: http://economics.sas.upenn.edu/pier/working-paper/2012/origins-and-development-term-%E2%80%9Cbig-data
17. Laney D. 3D Data Management: Controlling Data Volume, Velocity, and Variety [Internet]. META Group; 2001 Feb. Available at: http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf
18. Schroeck, M, Shockley, R, Smart, J, Romero-Morales, D, Tufano, P. Analytics The real-world use of big data [Internet]. IBM Institute for Business Value; 2012. Available at: http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-big-data-at-work.html
19. Evelson, B, Nicolson, N. Topic Overview: Business Intelligence – An Information Workplace Report [Internet]. Forrester Research. 2008 [cited 2013 Dec 16]. Available at: http://www.forrester.com/Topic+Overview+Business+Intelligence/-/E-RES39218?objectid=RES39218
20. Core Techniques and Technologies for Advancing Big Data Science & Engineering (BIGDATA) [Internet]. National Science Foundation; 2012. Available at: http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.pdf
21. MD Anderson Taps IBM Watson to Power "Moon Shots" Mission [Internet]. MD Anderson Cancer Center. 2013 [cited 2013 Dec 17]. Available at: http://www.mdanderson.org/newsroom/news-releases/2013/ibm-watson-to-power-moon-shots-.html
22. Okun, S, McGraw, D, Stang, P, Larson, E, Goldmann, D, Kupersmith, J. Making the Case for Continuous Learning from Routinely Collected Data [Internet]. IOM; 2013. Available at: http://www.iom.edu/~/media/Files/Perspectives-Files/2013/Discussion-Papers/VSRT-MakingtheCase.pdf
23. Davis DA, Chawla NV, Blumm N, Christakis N, Barabasi A-L. Predicting individual disease risk based on medical history. Proceedings of the 17th ACM conference on Information and knowledge management. ACM; 2008. p. 769–78.
24. Davis DA, Chawla NV, Christakis NA, Barabási A-L. Time to CARE: a collaborative engine for practical disease prediction. Data Min Knowl Discov 2010;20(3):388–415.
25. Asangansi I, Braa K. The emergence of mobile-supported national health information systems in developing countries. Stud Health Technol Inf 2010;160(Pt 1):540–4.
26. Lewis T, Synowiec C, Lagomarsino G, Schweitzer J. E-health in low- and middle-income countries: Findings from the center for health market innovations. Bull World Health Organ 2012;90(5):332–40.
27. Big Data for Development: Challenges & Opportunities [Internet]. UN Global Pulse; 2012. Available at: http://www.unglobalpulse.org/sites/default/files/BigDataforDevelopment-UNGlobalPulseJune2012.pdf
28. Barrington J, Wereko-Brobby O, Ward P, Mwafongo W, Kungulwe S. SMS for Life: a pilot project to improve anti-malarial drug supply management in rural Tanzania using standard technology. Malar J 2010 Oct 27;9(1):298.
29. Novartis Malaria Initiative: SMS for Life [Internet]. [cited 2014 Mar 27]. Available at: http://www.malaria.novartis.com/innovation/sms-for-life/
30. Hilbert M. Big Data for Development: From Information-to Knowledge Societies. Univ South

Calif - Annenberg Sch Commun [Internet]. 2013; Available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2205145
31. Barroso LA, Hölzle U. The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines. Synth Lect Comput Archit 2009 Jan;4(1):1–108.
32. Shapiro C, Varian HR. Information rules: a strategic guide to the network economy. Boston, Mass.: Harvard Business School Press; 1999.
33. Infrastructure as a Service (IaaS) [Internet]. Gartner IT Glossary. [cited 2013 Dec 10]. Available at: http://www.gartner.com/it-glossary/infrastructure-as-a-service-iaas
34. Latourette MT, Siebert JE, Barto Jr. RJ, Marable KL, Muyepa A, Hammond CA, et al. Magnetic resonance imaging research in sub-Saharan Africa: Challenges and satellite-based networking implementation. J Digit Imaging 2011;24(4):729–38.
35. Shiferaw F, Zolfo M. The role of information communication technology (ICT) towards universal health coverage: The first steps of a telemedicine project in Ethiopia. Glob Health Action 2012;5(1):15.
36. Simba DO. Application of ICT in strengthening health information systems in developing countries in the wake of globalisation. Afr Health Sci 2004 Dec;4(3):194–8.
37. Gardiner, B. Astrophysicist Replaces Supercomputer with a Cluster of Eight PlayStation 3s [Internet]. WIRED. 2007 [cited 2013 Dec 10]. Available at: http://www.wired.com/techbiz/it/news/2007/10/ps3_supercomputer
38. Zyga, L. US Air Force connects 1,760 PlayStation 3's to build supercomputer [Internet]. PhysOrg. 2010 [cited 2013 Dec 10]. Available at: http://phys.org/news/2010-12-air-playstation-3s-supercomputer.html
39. Amazon Web Services [Internet]. Amazon. [cited 2013 Dec 10]. Available at: http://aws.amazon.com/
40. Google Compute Engine [Internet]. Google Cloud Platform. [cited 2013 Dec 10]. Available at: https://cloud.google.com/products/compute-engine/
41. Purkayastha S, Braa J. Big Data Analytics for developing countries–Using the Cloud for Operational BI in Health. Electron J Inf Syst Dev Ctries [Internet]. 2013 [cited 2014 Mar 25];59. Available at: https://ejisdc.org/ojs2/index.php/ejisdc/article/view/1220
42. Apache Hadoop [Internet]. Hadoop. [cited 2013 Dec 10]. Available at: http://hadoop.apache.org/
43. Lohr S. For Today's Graduate, Just One Word: Statistics. The New York Times [Internet]. 2009 Aug 6 [cited 2013 Dec 10]; Available at: http://www.nytimes.com/2009/08/06/technology/06stats.html?_r=3&
44. Manyika J, Chui M, Brown B, Bughin J, Dobbs R, Roxburgh C, et al. Big data: The next frontier for innovation, competition, and productivity [Internet]. McKinsey Global Institute; 2011. Available at: http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation
45. Competitions | Kaggle [Internet]. [cited 2014 Mar 27]. Available at: https://www.kaggle.com/solutions/competitions
46. DataKind | DataKind [Internet]. [cited 2014 Mar 27]. Available at: http://www.datakind.org/
47. Hammond WE, Bailey C, Boucher P, Spohr M, Whitaker P. Connecting Information To Improve Health. Health Aff (Millwood) 2010 Feb 1;29(2):284–8.
48. Searching for standards in big data [Internet]. FCW; 2012 [cited 2013 Dec 17]. Available at: http://fcw.com/microsites/2012/snapshot-managing-big-data/05-establishing-big-data-standards.aspx
49. Glaser J. Interoperability: the key to breaking down information silos in health care. Healthc Financ Manage 2011 Nov;65(11):44–6, 48, 50.
50. Luna, D, García, M, Nishioka, A, Franco, M. OPS - Revisión de estándares de interoperabilidad para la e-salud en latinoamérica y el caribe. In Press. 2013;
51. Country health information systems: a review of the current situation and trends [Internet]. Geneva: World Health Organization; 2011 [cited 2013 Nov 1]. Available at: http://www.who.int/healthmetrics/news/chis_report.pdf
52. National eHealth strategy toolkit. [Internet]. World Health Organization and International Telecommunication Union; 2012. Available at: http://www.itu.int/pub/D-STR-E_HEALTH.05-2012/
53. Committee on the Role of Institutional Review Boards in Health Services Research Data Privacy Protection. I of M. Protecting data privacy in health services research [Internet]. National Academies Press.; 2000. Available at: http://www.nap.edu/openbook.php?isbn=0309071879
54. Meslin EM. Shifting Paradigms in Health Services Research Ethics. J Gen Intern Med 2006 Mar;21(3):279–80.
55. Summary of the HIPAA Security Rule [Internet]. HHS. [cited 2013 Dec 17]. Available at: http://www.hhs.gov/ocr/privacy/hipaa/understanding/srsummary.html
56. Summary of the HIPAA Privacy Rule [Internet]. HHS. [cited 2013 Dec 17]. Available at: http://www.hhs.gov/ocr/privacy/hipaa/understanding/summary/index.html
57. Campbell AV. The Ethical Challenges of Genetic Databases: Safeguarding Altruism and Trust. Kings Law J 2007 Jan 1;18(2):227–45.
58. Chalmers D, Nicol D. Commercialisation of biotechnology: public trust and research. Int J Biotechnol 2004 Jan 1;6(2):116–33.
59. Michele O, Fernandes, L, Weaver, V. Big Data, Bigger Outcomes. J AHIMA 2012;83(10):38–43.
60. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, et al. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res 2013;15(8):e164.
61. Big Data for Development: a primer. Harnessing Big Data For Real-Time Awareness [Internet]. UN Global Pulse; 2013. Available at: http://www.unglobalpulse.org/sites/default/files/Primer%202013_FINAL%20FOR%20PRINT.pdf
62. Vital Wave Consulting. Big Data, Big Impact: New Possibilities for International Development [Internet]. World Economic Forum; 2012. Available at: http://www3.weforum.org/docs/WEF_TC_MFS_BigDataBigImpact_Briefing_2012.pdf
63. New Data for Understanding the Human Condition: International Perspectives [Internet]. OECD; 2013. Available at: http://www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.pdf

Correspondence to:
Daniel Luna, MD, MSc, PhD (Cand.)
Chief Information Officer
Hospital Italiano de Buenos Aires
Pte. Perón 4190 (C1181ACH)
Capital Federal, Argentina
Tel/Fax: +54 11 4959 0507
E-mail: daniel.luna@hospitalitaliano.org.ar