AN OVERVIEW OF THE OPEN SCIENCE IN TIMES OF BIG DATA AND INNOVATION TO GLOBAL HEALTH

[1]Jorge Magalhães
[2]Zulmira Hartz
[3]Adelaide Antunes
[4]Maria do Rosário O Martins

ABSTRACT

The last several years have revealed information technology and scientific data to be important allies. However, the most important scientific ally is that which can assist in the complex task of identifying, collecting and treating the exponential amount of data added to the Web every day in the XXI century. Moreover, the volume of this data exceeds 2.5×10^{18} new bytes per day and arrives on the Web at a rapid speed. Thus, it is necessary to assess the veracity of these data and their value to decision-making. Although approximately 43% of the data are related to health and about one million scientific articles published per year are in the health field, it is important to think beyond trivial models to solve problems of local health with a global focus. Thus, this work aims to contribute to the reflection on the use of free tools as well as to discussions of collaborative and effective partnerships for action in the field of Public Health. This study shows opportunities to use open science for global health innovation, mainly, to the countries with difficulty in managing the problems of their ills. Open science in times of Big Data are much more agile than the old model of closed science, i.e., isolated groups that either did not share data or did share but at prices that were unaffordable for developing or underdeveloped countries. Technological development for new chemical entities can be facilitated by using the open science of the Scientific's Big Data.

Keywords: Innovation; Global Health, Technological Trends; Open Science; Collaborative Intelligence; Web 2.0; Big Data

[1] Universidade Nova de Lisboa, UNL, Rua da Junqueira 100, 1349-008 Lisboa, Portugal.
Email: <jorgemagalhaes@far.fiocruz.br>

[2] Universidade Nova de Lisboa, UNL, Rua da Junqueira 100, 1349-008 Lisboa, Portugal.
Email: <zhartz@ihmt.unl.pt>

[3] Nacional Institute of Industry Property –INPI. Federal University of Rio de Janeiro – UFRJ, Brazil.
Email: <adelaide@eq.ufrj.br>

[4] Universidade Nova de Lisboa, UNL, Rua da Junqueira 100, 1349-008 Lisboa, Portugal.
Email: <mrfom@ihmt.unl.pt>

UMA VISÃO GERAL DA CIÊNCIA ABERTA EM TEMPOS DE BIG DATA E INOVAÇÃO PARA SAÚDE GLOBAL

RESUMO

Os últimos anos revelaram que a tecnologia da informação e os dados científicos são importantes aliados. No entanto, a colaboração científica mais importante da ciência da informação, é de poder auxiliar na tarefa complexa de identificar, extrair e tratar a quantidade exponencial de dados adicionados à Web diariamente neste século 21. E volume de dados excede 2,5 x 1018 novos bytes/dia e com rápida velocidade. Assim, é necessário avaliar a veracidade desses dados e seu valor para a tomada de decisões. Com 43% desses dados relacionados à saúde e cerca de um milhão de artigos científicos publicados/ano na área da saúde, é importante pensar além de modelos triviais para resolver problemas de saúde local com foco global. Assim, este trabalho objetiva contribuir na reflexão sobre o uso de ferramentas da Web 2.0, das parcerias colaborativas e efetivas para ações em Saúde Pública. É demonstrado oportunidades para usar a ciência aberta para a inovação global em saúde, principalmente, para os países com dificuldade em gerenciar os problemas de atenção à saúde de sua população. A ciência aberta em tempos de Big Data é muito mais ágil do que o antigo modelo de ciência fechada, ou seja, onde os grupos eram isolados e não compartilhavam seus dados, ou compartilhavam a preços que não eram acessíveis para países em desenvolvimento ou subdesenvolvidos. O desenvolvimento tecnológico para novos compostos químicas pode ser facilitado pela utilização da ciência aberta e do Big Data em dados científicos.

Palavras-chave: Inovação; Saúde Global; Tendências Tecnológicas; Ciência Aberta; Inteligência Colaborativa; Web 2.0; Big Data

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

271

INTRODUCTION

The creation of the World Wide Web is considered one of the greatest recent examples of radical technology. According to O'Sullivan (2008), the web makes it possible to create and offer products, services and distributive processes in most industrial sectors (O'Sullivan, D. & Dooley, L., 2009). The Web has evolved according to progressive socialization: mainly due to the rapid growth of the Internet-using population, the Web has evolved from being business-centered to being user-centered. This change is known as Web 2.0 (Figure 1 from Spivak (Nova Spivack, 2013), a term first used by Tim O'Reilly in 2004 (Lee, In., 2011; O'Sullivan, D. & Dooley, L., 2009). Web 2.0 is defined by the surrounding social and technological environment, participation and positive user interaction (Lee, In., 2011).
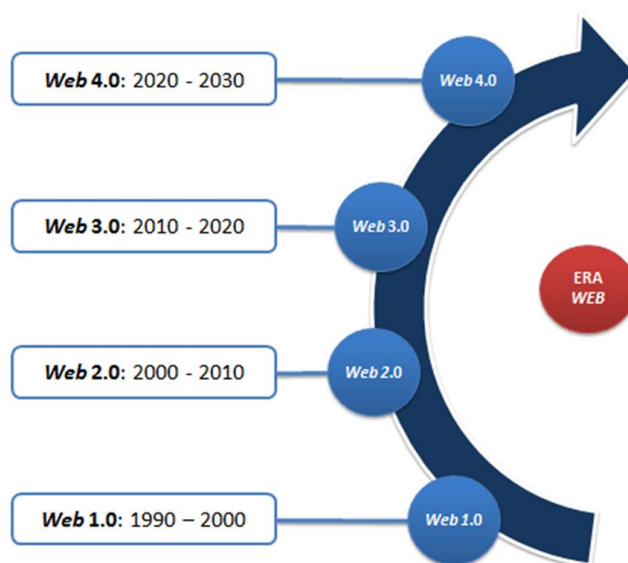

Figure 1: Evolution of the Web over the years.
Source: Created by the authors using Spivak (2013).

The Web provides the links that unite the global economy and enable increasingly global trade. Much of this trade is conducted electronically (currently 50% of services and 12% of product sales). We are approaching a new frontier in the reduction of communication and transaction costs, in addition to customizing consumer profiles. Alongside speed gains, these changes create economic value, thus increasing innovation, competition and productivity ("Big data", 2011).

Thus, the concept of globalization has changed over time. Previously, globalization's image was of nations coming together through the movement of goods, services and finance, while today the flow has a different dynamic, with the focus being on data transmissions that are generated at high levels of speed, volume and variety. The effect of this phenomenon of expanding data – known as Big Data – is accelerated by globalization and reset by data streams that incorporate ideas, information and innovation (McKinsey Global Institute, 2011a).

In the 21st century, the world is becoming fully connected to the Internet, with approximately 40% of the population connected at present (McKinsey Global Institute, 2011c, 2011b). Against this background, O'Reilly (2007) suggested the term Big Data, referring to a huge database updated in real time that easily uses thousands of terabytes of storage in various formats. Traditional systems of relational database management cannot handle these large volumes of data (Magalhães & Quoniam, 2013;

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

272

Quoniam, L, Lucien, A, 2010). Big Data is generating a new generation of methodologies designed to extract economic and strategic value out of a large and varied amount of data (structured and unstructured), thus enabling high-speed capture and analysis ("Gray, J. and Chambers, L. and Bounegru, L., The Data Journalism Handbook, O'Reilly Media, 2012 - InfoVis:Wiki", [s.d.]; O'Reilly, 2007a). The big data revolution is changing the way data is produced, analyzed, and valued. In the environmental sciences, big data has made it onto the agenda through calls to utilize the current data "deluge" more effectively and a desire for more complete measurement (Salmond, Tadaki, & Dickson, 2017).

The process of globalization is driving the evolution of the term Global Health, which presents challenges and opportunities in the public health field. Global Health can be understood as, simultaneously, a condition, an activity, a profession, a philosophy, a discipline and/or a movement. However, it should be recognized that there is no consensus about what Global Health truly is; there is no single definition, and its scope has imprecise limits (Fortes & Ribeiro, 2014). However, it is undeniable that society is pursuing health at a time of globalization (Koplan et al., 2009);

The phenomenon of globalization brings new spatial, temporal and cognitive dimensions. It changes our perception of distance, borders, and barriers to global contacts; it changes our perception of time by connecting everyday life with events taking place in other parts of the world, changing our cognitive perceptions of how we see and understand ourselves and the world around us, allowing engagement with the "other" around the world (Bozorgmehr, 2010).

Thus, it is necessary that we seek to identify, extract and integrate Big Data into health in this globalized world in order to highlight critical information for decision makers of the 21st century. Nevertheless, while the management of Big Data is not trivial, the approach of Open Science also appears imminent.

The expression open science refers to a model of scientific practice that, in line with the development of digital culture, aims at creating an information network that is the opposite of the model of closed research laboratories.

Currently, the term also refers to the generation of research materials that are shared openly, without patents. According to Turner *et al* (2016), Open Science is crucial for the advancement of science (Turner et al., 2015).

In this context, the European Community has shown leadership and maturity in this area with the promotion of Open Science. One can cite, as an example, the Resolution of the Council of Ministers of Portugal: Making science open and accessible to all is a collective challenge – political, cultural, economic and social. See more in Resolução do Conselho de Ministros 20/2016 (Resolução Conselho de Ministros, 2016).

Thus, among organizations that address health research, it is essential to seek better knowledge management of Big Data in Health – a growing area of Open Science – to create a collaborative and constructive intelligence for society.

In this sense, this article seeks to contribute to reflection on the open science in times of Big Data in the global health, where open science and collaborative intelligence can, together, be great allies in the 21st century. Both are ideally suited to help address the flurry of big data and the ephemerality of this new era and thus find solutions to fight and to decimate the ills plaguing our society.

METHODOLOGY

This work is based on mining of literature databases in indexed papers and official organizations databases such as SCOPUS, Web of Science, PubMed, National Institutes of Countries. After identifications of the literature reviews and quantitative data, they were processed in GoPubMed, CarrotLingo4G and Patent2Net open source. The statistic was put in table and maps.

RESULTS

The results of this work concentrate on identifying some existing scenarios in the new millennium on open science and Big Data. In this way, reflect on new ways of dealing with them for global health.

Data science and open science

Data science involves principles, processes, and techniques for understanding phenomena via the analysis of data automated. So, data science with open science improving decision making as this generally is of paramount interest to business. According to business theorist Clayton Christensen, Open Science can be considered like a disruptive innovation and it will become the dominant model for the distribution within the next decade in academic journals (Hossain & Aydin, 2011).

Open Science represents the opening of the scientific process and the strengthening of the concept of scientific social responsibility towards humanity. It is a key tool for expanding and democratizing access to knowledge because its favors several opportunities for science and innovation as it contributes to breaking down territorial, institutional and disciplinary barriers. Considering the social level of Open Science, it provides universal access to scientific knowledge and helps to reduce asymmetries. In this way, it enhances equity and economic development through the creation of value and the return of public investment.

In this sense, it is necessary to change academic culture in order to reduce the sense of ownership that researchers have over their data. Most often these data are the results of research funded with public money and materials donated by others. As an example, researchers funded by the National Institutes of Health (NIH) are obligated to make their results public, under penalty of no longer being funded.

Large corporations also use the concept of open innovation as a way to promote innovation on processes and/or products using the contributions of any researcher, company, etc. (Celadon, 2014; Michelino, Cammarano, Lamberti, & Caputo, 2015). Public health also presents a similar opportunity (Chaifetz, Chokshi, Rajkumar, Scales, & Benkler, 2007).

According to the journal Editorial Science (2016), although there has been progress in recent years, most clinical and genomic data are also collected and studied in isolation – in silos – compartmentalized by disease, institution, country, etc. (Nature News, 2016). Initial efforts in sharing have allowed the development of treatments for rare diseases and some forms of cancer. However, this benefit will not reach the entire population until doctors and researchers can access and compare data on millions of individuals (The Global Alliance for Genomics and Health, 2016).

When there is difficulty in finding a specific book if the content of a dozen different national libraries were gathered in one place, a strategy is necessary to integrate the various ways in which different content is archived, tracked, recorded and made available. It would be much easier to ask each library to store its own books and to share information on how to find them with every other library. Following this line of reasoning, it would be interesting to share science data using the same approach (Nature News, 2016).

Structural Genome Consortium:

The Partnership in the Structural Genome Consortium (www.thesgc.org) aims to generate small molecule inhibitors of protein kinases. These molecules will be available to any research group in Brazil and the world. In this context, new initiatives of publication and peer review, such as the Peerage of Science, the arXiv and PLoS, have confirmed this trend, as has the example of the open encyclopedia (Wikipedia) among others, which was so criticized at its beginning yet has now emerged as one of the best Web 2.0 tools for the dissemination of knowledge [46].

Foundation for Research of the State of São Paulo - Brazil (FAPESP):

In pursuit of this greater interaction and in reference to the "Year of Science", the development agency promoted open data as well as Wulder *et al* (2012). Agency researchers demonstrated that there was a gap of 30 years between the science reflected in research sources and cutting-edge developments being made in this field in other languages (Wulder, Masek, Cohen, Loveland, & Woodcock, 2012). I.e., a case study done by Brazilian initiative in the encyclopedia open, Wikipedia, added thousands of words to Portuguese Wikipedia on topics

related to the mathematical properties of neural dynamics. They expanded the article in Portuguese on Alzheimer's disease and added to articles on more complex issues such as brachial plexus injury. They also created an introduction to biological neuron models and created a video that explains "Spike sorting" – a way to track and measure the electrical properties of cells – that appears in both Portuguese and English Wikipedia (Mena-Chalco, Digiampietri, Lopes, & Cesar-Jr., 2013; Souza et al., 2015). In this model, experts and researchers work to explain concepts to volunteers in the research group in Brazil, who will then write articles based on the involvement of experts, since science is always in motion (Van Noorden, 2012). Its noteworthy that many of the authors are postdoctoral researchers (Hossain & Aydin, 2011; Kvan & Candy, 2000; Wilson, Ko, & Reis, 2011).

Oswaldo Cruz Foundation (FIOCRUZ):

FIOCRUZ is a Brazilian institution that has created and encourages Open Science policy knowledge and dissemination of scientific information (Pinheiro, 2014). This organization aims to help strengthen the preservation of mechanisms of institutional memory and increase access to, and the impact of, the intellectual production of FIOCRUZ, thus creating an important tool to promote, in an organized and united way, the dissemination, accessibility and, consequently, the visibility of knowledge generated by this institution. Your Institutional Repository (ARCA – Brazilian term) is the main instrument for carrying out this policy in light of the objectives to collect, host, preserve, make available and give visibility to the scientific production of the institution. Thus, it is mandatory to deposit in the Ark Institutional Repository both dissertations and theses from the Fiocruz Graduate programs and articles produced under Fiocruz that are published in scientific journals. In addition to texts, the repository can hold images and audio. The ARCA repository should have the ability to integrate with national and international systems, which automatically enables the inclusion and collection of relevant intellectual products, noting in particular the defined protocols and standards of the Open Archives Initiative model (OAI) (Chaifetz et al., 2007; Turner et al., 2015).

Another example, the partnership of the Fiocruz with Drugs for Neglected Diseases Initiative (Dnd*i*) that resulted in a new drug for malaria diseases. As part of its quest to develop new and simplified treatments for neglected diseases, the project presented excellent results of the pivotal Phase III clinical studies for a new malaria treatment. Thus, the results shown that the new fixed dose combination of artesunate/mefloquine (AS/MQ) is as effective as the existing separate tablets and is better tolerated with a lower incidence of early vomiting. Because of the reduction of the number of tablets as well as the simplification of the regimen, the new product is more convenient for the patients and the prescribers. This was possible due to the sharing of data and synergy of multidisciplinary actions ("Malaria: A new fixed-dose combination of artesunate and mefloquine is effective and well tolerated. – DNDi", [s.d.]; Santelli et al., 2012; Sirima et al., 2016; Valecha et al., 2013).

European Union:

The European Union is taking a leap forward in opening up knowledge. A study was recently published in the form of book entitled Open innovation, Open Science, open to the world. This study was commissioned to the European Commissioner for Research, Innovation and Science by the President of the European Union. The publication presents the knowledge gap as naturally arising from the potential of new information technologies. It aims to address the knowledge gap to better use public resources invested in science and innovation to produce educational – among other – advantages. It adds further principles for addressing the knowledge gap through open innovation, Open Science and citizen science (EU Bookshop, 2016).

The publication suggests the use of a concept called a "Global Research Area". According to this concept, "researchers and innovators can work with international colleagues so that researchers, scientific knowledge and technology circulate as freely as possible." For Brazil, the desired technologies for teaching science and engineering are those that have the properties of the "Global Research Area", i.e., "scientific knowledge and technology

circulate as freely as possible" (Bhardwaj et al., 2011).

In this context, it is worth noting the seminar held by NOVA University of Lisbon on June 23, 2016. This was a seminar on Big Data, Sustainable Development and Open Science, promoted by the Hygiene and Tropical Medicine Institute (IHMT) and Institut Français, in partnership with FIOCRUZ, Brazil, and the National Innovation Agency in Lisbon. The event was attended by 57 participants who attended the sessions in person and through streaming. At issue were two main themes: the application of big data research and the use of Open Science in technology transfer. In the context of the first theme, the uses of big data and open health information were analyzed. In the second, there was discussion of innovation for social inclusion and low-income contexts, as well as for automated patent analysis using an open source free software, as an example of data processing by Open Science: the crawler Patent2Net (http://patent2netv2.vlab4u.info/).

As a contribution to ways of managing the Big Data in Global Health to decision makers, the figure 4 shows data processed using technological information via Open Science. Using the descriptor "dengue" in the "crawler" Patent2Net it was identified, extracted and correlated the technological information on patents within European Patents Office database (inside the universe of more than 90 million patents).

Several results about dengue disease were obtained and some core information for decision makers are summarized in the infographic (figure 4). In this analysis, 1427 documents of patents were recovered, through which, with the crossing of the data, it was possible to extract essential information such as countries assignees, existing technologies and their interactions, established networks, technology trends, inventors and their locations, networks etc. At the same time, using a private analysis (closed science) in the same infographic, the technological evolution and others information are visualized for the descriptor "chikungunya".
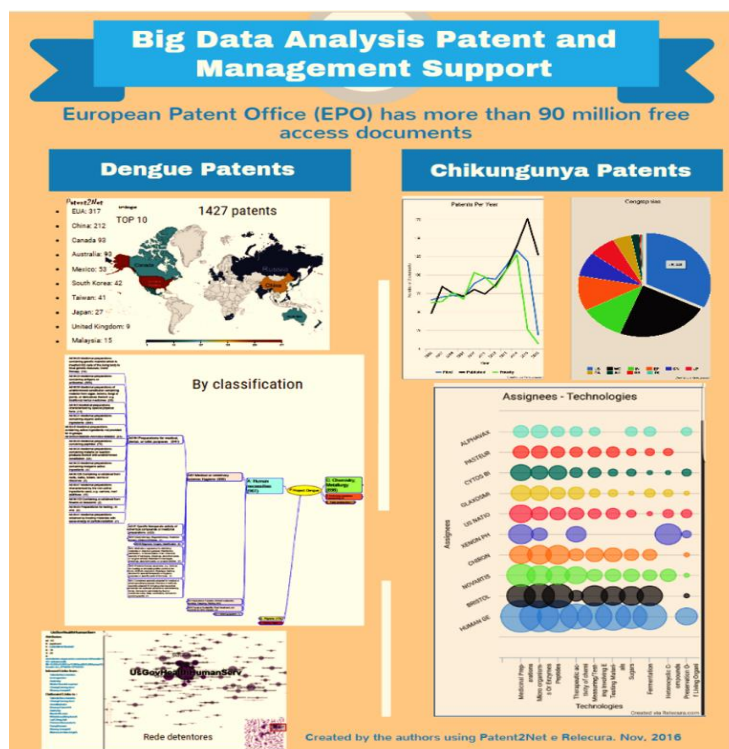

Figure 4: Infographic of Dengue Patent Analysis and Chikungunya

The opportunities of Open Science contribute to all areas of science. In the area of heath, the pharmaceutical industry is the most intense in research, development and innovation

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

276

(R,D&I) (Basil Achilladelis, 2001; Chaves, Oliveira, Hasenclever, & Melo, 2007). Big Pharmas' advances and contributions to public health are surrounded by patent protection. However, the increase of collaborative intelligence among its actors has been observed over the years, whether in papers or in the establishment of research networks for the patenting of their products. Thus, it can be said that even in a scenario monitored by "confidential" research, knowledge management transcends barriers and aligns itself in the context of the research of the new century (Cattell, Chilukuri, & Levy, [s.d.]; Rafols et al., 2014).

Big data in global health and collaborative intelligence

A report delivered to the U.S. Congress in August 2012 defines big data as "a term that describes large volumes of high velocity, complex, and variable data that require advanced techniques and technologies to enable the capture, storage, distribution, management, and analysis of the information" (Chluski & Ziora, 2015). Big Data analysis is already being used successfully in several countries including the United States, which has incorporated the concept in almost all its productive sectors. In 2014, the US government presented the report Big Data: Seizing Opportunities. This report was based on consultations with major US stakeholders such as Apple, IBM, Google, and Bank of America, among others, on issues concerning opportunities and values in the use of Big Data. It evaluated how this new phenomenon will change relations between government, citizens, businesses and consumers (The White House, 2016).

Big Data refers to the third generation of the information age (Magalhaes, JL & Quoniam, L, 2015; Raghupathi & Raghupathi, 2014). Initially, some characteristics of big data go beyond size and volume. This exponential data volume met the criteria of the 3 Vs: Volume, Variety and Velocity (Laney, 2001). Then, an additional 2 Vs were added: Value and, specifically to health care, Veracity. Some authors add a final set of 3 Vs – Verity, Versatility and Viability – such that the combination of all "Vs" generate the "V" of value (ALEIXO & DUARTE, 2015). This configures a great complexity of data. Complexity of the data source can be assessed in four different aspects: data structure, data format, data itself and hierarchies used in the data. The way in which the data is received, read, validated, processed and stored by this depends on the characteristics of the data. Complexity of the data structure means that there can exist various relations between data, including complex keys in tables, that makes it difficult to integrate various data tables (Kim, Kim, & Kim, 2016).

Volume is a reference to the amount of information that is available daily on the Web, while the term "variety" is characterized in different ways by different sources. Accuracy and speed are related to the processing of such data, which aims to make it useful, available in real-time, and reliable (Hilbert & López, 2011). Value is the attribute wherein the effective treatment of Big Data can generate access to essential information and save money for the organization in question.

According to Minelli et al (2013), Big Data is divided into the perfect data storm, the perfect storm of convergence, and the perfect storm of computing, the last of which is the result of four phenomena: Moore's Law, mobile computing, social networking and cloud computing. This data collection should be used to present information in a selectively researched and objective way to increase business intelligence and enable improvements in the decision-making process (Minelli, M., Chambers, M., & Dhiraj, A., 2013).

The velocity and data volume available in the virtual world have demonstrated the power of Big Data, as well as the possibility of using the results of this information in the area of competitive intelligence – that is, to produce more in less time, surpass competitors and save time. The result is that data processing has been a differentiator in decision making.

According to Magalhães et al (2013), out of the daily volume of data added to the web, 47% is related to health and, of this, 43% is related to public health (Magalhaes, JL & Quoniam, L, 2015). Therefore, it is necessary to develop new tools for the identification, extraction and Big Data analysis of health data both globally and locally, or "glocally". According to Humbert (2005), the term glocal refers to the attitude of thinking about problems globally and acting locally, as glocal actions can have a global impact.

This concept reflects the current state of globalization, in which technology has made it possible to reach beyond a single context, thus encouraging the participation of enterprises located in suboptimal locations (Humbert, M., 2003).

The opportunities associated with data and analysis in different organizations have helped generate significant interest in business intelligence and analytics, which is often referred to as the techniques, technologies, systems, practices, methodologies, and applications that analyze critical business data to help an enterprise better understand its business and market and make timely business decisions (Chen, Chiang, & Storey, 2012).

A report on the real-world use of big data conducted by IBM in collaboration with the University of Oxford showed that Big Data analysis makes organizations up to 23 times more likely to overcome their market competitors than those that do not analyze such data. However, some productive sectors, such as healthcare, finance, telecommunications, government and energy, are especially adaptable to Big Data strategies.

Health is considered mostly a global public good: it is not exclusive; no single person or community is excluded from its possession or consumption; and its benefits are available to everyone. There is also the apparent consensus that health is not competitive; that is, the health of one person does not come at the expense of excluding others (Buse & Waxman, 2001; Haines et al., 2009; Hartz, 2012; Vance, Howe, & Dellavalle, 2009).

In the current health scenario, stakeholders are embracing the concept of evidence-based medicine, which means making the best clinical decisions based only on the best scientific evidence available, which Big Data can now provide. In most cases, compiling datasets into Big Data algorithms is the best source for evidence, since distinction in subpopulations (such as the presence of patients with hypertension caused by Pheochromocytoma) could be intermittent enough that individual smaller datasets cannot offer enough evidence to point out which statistical differences are evident. Pioneers in using Big Data are already

achieving positive results, prompting other stakeholders to act (Jee & Kim, 2013).

DISCUSSIONS
Progress

Recent decades were marked by the advent of Web 2.0. The Internet evolved from a purely static platform and assumed a dynamic and interactive role [51], allowing users to exchange large amounts of information instantly. By the end of 2016, the amount of information created and replicated from sensors of all types, posts on social networks, uploaded photos and videos, commercial transaction records, GPS signals, and navigation traces (among other sources) will reach the order of zettabytes - 1 billion Gigabytes (Huyghe, F.B., 2009a; McKinsey Global Institute, 2011c).

The dynamic infographic The Internet in Real Time (http://visual.ly/internet-real-time) shows that approximately 1 million Gigabytes of data are generated every 1 minute on the Internet, producing a profit of US$ 142,000 per minute for the giants of the medium, such as Apple, Google, Microsoft, Facebook, Netflix, Pandora, LinkedIn, etc. From this context – the large amount of data generated at all times – comes the term Big Data.

The web has evolved from version 1.0 to version 4.0. It should be noted that the terms Internet and web are easily confused and are often treated as synonyms by users (Engel et al., 2015). In recent years, technological advances and the development of the Internet have led to the creation of various applications that are used primarily to facilitate interactions among people. Thus, modern society has become dependent on personal computers, email, the web, e-commerce, seekers, wireless network technologies, online music, online videos, smartphones and various social networks. Today, the Internet is a network of people and communities; it is no longer just a computer network (Hossain & Aydin, 2011; Huyghe, F.B., 2009b; Lawrence & Giles, 2000; Lee, In., 2011).

In the Web development context, the globalization we are experiencing today is no longer an exclusive privilege of the largest multinational companies in the world. Today, scanning has eliminated many of the barriers that

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

278

previously prevented small and medium-sized enterprises (SMEs), entrepreneurs and ordinary citizens from making connections with customers and suppliers worldwide. Digital globalization, for example, has significant implications, especially with regards to companies and economies in developing countries. In these countries, companies and individuals can use digital platforms as a way of overcoming restrictions in their local markets. In this way, they can identify and access opportunities, information and ideas anywhere in the world.

Among the many aspects of digital globalization, academic literature clearly stands out; it is now accessible virtually instantly thanks to social media and other Internet platforms through which individuals are connecting. It is estimated that nearly one billion people around the world are direct participants in some form of globalization. Analysis of Facebook, Twitter, LinkedIn and WeChat shows that 914 million people have at least one international connection on a social media platform (Trigo, Gouveia, Quoniam, & Riccio, 2007).

It should be recognized that investigations no longer develop in a linear way. The speed of evolution is rapid. The pace of paradigm change and the emergence of new ideas have been increasing exponentially in our current era and world of knowledge. It is noteworthy that the first changes, although they appeared to be fast, took years to develop. This is exemplified by the sequencing of HIV: this took 15 years while the sequencing of SARS took only 21 days (Bakonyi, Gould, Kolodziejek, Weissenböck, & Nowotny, 2004).

Components

Given the technological advances in all areas of science, coupled with technological developments, it is necessary to treat the large volume of data, especially health data, that accounts for 47% of all information. This is an exhaustive amount of data, and it must be organized and structured to provide possible benefits to decision makers. In this sense, the use of non-trivial search engineering tools is required to treat this volume of data so as to extract essential information for decision makers. Thus,

collaborative tools such as Web 2.0 and free software (O'Reilly, 2007b), (Hossain & Aydin, 2011; Huyghe, F.B., 2009b; Lawrence & Giles, 2000), for example, are some alternatives with which institutions or organizations generally begin their work. The Medline database offer to treat your data in the GoPubMed database. The results are "mined" and correlated within the database and produce results that are very interesting for quick analysis and future studies (Magalhães & Quoniam, 2013).

Management and access to the concentrated data

Considering to the explosive evolution of information technology and computer science because of Big Data, the collaboration among people become still more important in the history of mankind. The several collaboration-based applications are called Collective Intelligence. In order to solve a lot of problems, people can share their experiences with others colleagues and also transfer their knowledge through online networks due to a scientific revolution of the Big Data Age (Jung, 2017).

The data storage principles are based on huge databases in all areas of science. In health areas, for example, the human genome, protein data bank, pharmaceuticals molecules and their correlated assignments makes up an exponential number of stored data of compliance regulations, data policies, and access controls as well as data storage methods which it can be implemented and completed in batch processes or in real time (Janssen, van der Voort, & Wahyudi, 2017; Jung, 2017; A. Paul, Chilamkurti, Daniel, & Rho, 2017; S. M. Paul et al., 2010; Wang & Hajli, 2017).

As an example of using the scientific databases, the medical field (MEDLINE - PubMed), which contain over 25 million citations, a search found 8,059 documents containing the term Big Data. The top 10 countries that have published on the subject of big data in the medical field, from 1970 to August 2016, are the US, China, the United Kingdom, Germany, Canada, Japan, Australia, Italy, France and Spain (see table 1 and figure 2).

Country	Number of papers
United States	1.110
China	377
United Kingdom	224
Germany	207
Canada	145
Japan	142
Australia	135
Italy	117
France	111
Spain	104

Table 1: Number of papers published with term Big data by countries.
Source: Created by the authors using GoPubMed database, March 2017.

It should be noted that the number of publications was stable until the beginning of the 21st century, specifically the year 2011. That year was the start of a 700% increase in publications, from 111 in 2011 to 775 in the first half of 2016.

Figure 2: World map of countries publishing on the term big data
Source: Created by the authors using GoPubMed database, March 2017.

Although the United States is the country that publishes most in this area, it does not have the top-ranked city. China's Beijing is the city with the most publications on the term Big Data in the medical field. Table 2 shows the top 10 cities.

Cities	Numbers of papers
Beijing	106
New York City	68
London	61
Boston	56
Sydney	48
Shanghai	45
Chicago	43
Seoul	42
Tokyo	38
Los Angeles	37

Table 2: Number of papers published with term Big data by cities.
Source: Created by the authors using GoPubMed database, March 2017.

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

280

Using the same search term and the free platform Carrot Search Lingo4G®, one can obtain other forms of extraction and plots of the data, thereby greatly aiding the processing of big data in Global Health. Similar to GoPubMed, one can estimate the number of available informational documents in general. Searching with Carrot Lingo3G® for documents on the web containing "big data", a total of 349 000 000 results were extracted. Next, mining revealed the terms that are more "obvious" (more often repeated) and indexed thereafter only in PubMed and then segregated those that used the term "big data" in scientific papers in the medical field.

Thus, it is observed that in the scientific world, as expected, the documents are more selective. At this stage, there were 100 clusters total of 8,059 essential documents, such as those in PubMed; these were classified according to their major implications and plotted for analysis according to the most relevant sub-topics within Big Data. The larger the subject's segment on the foam tree, the more relevant the papers published become, as noted in Figure 3 (created using CarrotLingo4G).
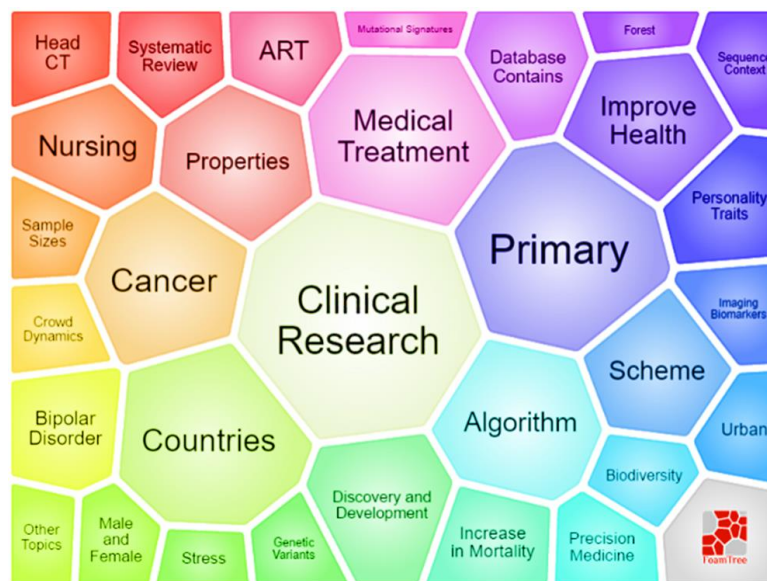

Figure 3: Key terms correlated to the term Big Data in the PubMed database.
Source: Created by the authors using CarrotLingo3G®, March 2017.

Due to the increasing amount of information added to databases daily, analyzing the state of the scientific and technological art and extracting essential information for decision making becomes an almost impossible task when employing traditional means. In this context, Information Technology has contributed tools that can assist this process in any area of science. Nevertheless, in addition to paid software (text mining, datamining etc.), Web 2.0 offers free tools that provide options for handling this volume of data to organizations that cannot easily acquire and maintain private software and search engines (e.g., by using licenses).

Thus, when using free tools, it is possible to obtain essential information to streamline the decision-making process by refining data, ordering chronologically, and grouping themes, thus giving the decision-maker the advantages and convenience of having essential and strategic information.

Relevance to global health goals

Global Health is an area of study, research and practice that has as a priority the attainment of equity in the health of the world population and includes, in addition to the priorities of the Sustainable Development Objectives, other areas related to fundamental aspects of health such as diseases Non-communicable chronic diseases (cardiovascular diseases, diabetes, etc.) that represent high disease loads worldwide; The social determinants of health; Climate change and maldistribution of human resources in health

(Gagnon, 2011; Koplan et al., 2009; Univesity of California San Francisco - UCSF, [s.d.]).

These aspects are particularly relevant in 21st century for several reasons: we live in a world of global connectivity (media), with high mobility (global citizen), with epidemics (HIV, Ebola) and pandemics (flu). In this context, it is essential to have an approach based not only on the disease, but also on social, economic and cultural aspects. Thus, the global health area is distinct from public health, international health and tropical medicine, with broader goals for achieving health equity (Koplan et al., 2009).

The contribution of Big Data to Global Health through Open Science is aligned with Global Health 2035 – The Grand Convergence in Public Health is an effort to address and eliminate global disparities in infectious, child and maternal mortality rates (Cesario, 2016). According to Kaslow *et al* (2016), the effort reaches its aims; by 2035 the mortality rates of lower-income countries (LICs) and rural areas of middle-income countries (MICs) will converge with those of higher-income countries (HICs) and the best-performing middle-income countries (MICs). Therefore, investment in health and developing a new investment framework to achieve dramatic health gains by 2035 creates opportunities for action for national governments of low-income and middle-income countries and by the international community (Kaslow et al., [s.d.]).

In this sense, to use Open Science in collaborative research and to identify, extract and correlate the Big data to Global Health presents itself as a path to the Global Health goals for 2035. This is an opportunity to anticipate a scenario of stagnant investments and no improvements in science, development and technology (Cesario, 2016; Jamison et al., 2013a; Terry et al., 2012; Univesity of California San Francisco - UCSF, [s.d.]; Wassan, 2001).

Many analyses of the relationship between globalization and health have considered health as a byproduct, as a spontaneous result – sometimes positive, sometimes negative – of strange globalizing forces, perhaps only motivated by other interests. Global Health is a desirable social goal today regardless of whether it is distorted by the influence of monetary fundamentalism. It deserves the best evidence,

whether because of its intrinsic value or as a symbol of the dominance of human values over other interests (S. M. Paul et al., 2010).

Regarding this process, a cooperative network is necessary. According to Pierre Lévy (1994), "collective intelligence is an intelligence distributed throughout, constantly valued, coordinated in real time, resulting in effective mobilization of skills," which seeks the recognition and enrichment of people. The concept of collective intelligence was created based on the insights of Pierre Lévy (1994), and is related to intelligence technologies. It is characterized by a new form of sustainable thinking using the social connections that make feasible the use of the open networks of Internet computing. These intelligence technologies are represented in particular by the languages, sign systems, logical resources and instruments that we use. All our intellectual functioning is induced by these representations. Humans are unable to think alone and without the aid of any tools (A. Paul et al., 2017).

According to Bonabeau (2009), collective intelligence contributes strongly to the transfer of knowledge and power from the individual to the collective (Wang & Hajli, 2017). Open sources of collective intelligence will eventually generate results superior to the knowledge generated by the proprietary software developed within corporations. Education, and how people learn to participate in cultures outside of formal learning contexts, is crucial in this new global context. Learning through the means of collective intelligence is crucial; it is important for the democratization of science because it is linked to a culture based on knowledge and is sustained by sharing collective ideas, thus contributing to a better understanding of our diverse society (Gagnon, 2011; Jamison et al., 2013b; Janssen et al., 2017). Collective ideas may come from or flow into the science of institutions (Cesario, 2016; Kaslow et al., [s.d.]; Terry et al., 2012; Wassan, 2001).

In this sense, the creation of the collective knowledge environment for Global Health is justified by the development of innovative activities, such as these Collaborative Knowledge Networks (Cesario, 2016). This collective "open" learning environment produces positive effects

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

282

on health outcomes if certain educational and institutional strategies are implemented. These reforms span the exploitation of the potential of Information and Communication Technologies (ICT) - as an example, Web 2.0 tools for the treatment of Big Data in health.

CONCLUSIONS

The 21st century has brought new challenges and opportunities thanks to the increasing volume of new data added to the Web every day. Science and technological development are no different, especially in the health field. Thus, it is important to quickly develop new methodologies for the identification, extraction and processing of data to obtain essential information. Therefore, mining Big Data in health is an urgent and emerging issue, as it provides agility in the processes of decision making. One alternative to assist organizations in this process is the availability of Web 2.0 tools.

Open Science Policy has been widespread in this new era of the twenty-first century as a form of responding to technological advances and the desire ofsociety to seek solutions to its ailments faster and more dynamically. Europe and the US are in the process of developing cutting edge solutions.

The "treatment" (knowledge management) of big data contained in patent documents shows itself as an important source of innovation for scientific technological development. Thus, data mining and processing to obtain essential information for decision makers are available and for use after the end of a patent domain or in countries where there is no patent protection by the holder of a patent.

Individual research work has lost some of its academic territory in recent times. Instead, collective work – collaborative intelligence – has succeeded not only in private institutions but also in public institutions in response to the expectations of society. There is a development in health where scientific and technological advances have been achieved in mere days of research rather than years exclusively through the cross (data mining) of open data, compared to the traditional investigations of the previous century.

REFERENCES

ALEIXO, J. A., & DUARTE, P. (2015). Big Data Opportunities in Healthcare. How Can Medical Affairs Contribute? *Revista Portuguesa de Farmacoterapia*, *7*, 230–236.

Bakonyi, T., Gould, E. A., Kolodziejek, J., Weissenböck, H., & Nowotny, N. (2004). Complete genome analysis and molecular characterization of Usutu virus that emerged in Austria in 2001. *Virology*, *328*(2), 301–310. doi:10.1016/j.virol.2004.08.005

Basil Achilladelis, N. A. (2001). 'The Dynamics of Technological Innovation: The Case of the Pharmaceutical Industry'. *Research Policy*, *30*(4), 535–588. doi:10.1016/S0048-7333(00)00093-7

Bhardwaj, A., Scaria, V., Raghava, G. P. S., Lynn, A. M., Chandra, N., Banerjee, S., … Brahmachari, S. K. (2011). Open source drug discovery– A new paradigm of collaborative research in tuberculosis drug development. *Tuberculosis*, *91*(5), 479–486. doi:10.1016/j.tube.2011.06.004

Big data: The next frontier for innovation, competition, and productivity | McKinsey Global Institute | Technology & Innovation | McKinsey & Company. (2011). Recuperado 7 de março de 2013, de http://www.mckinsey.com/insights/mgi/researc h/technology_and_innovation/big_data_the_ne xt_frontier_for_innovation

Bozorgmehr, K. (2010). Rethinking the "global" in global health: a dialectic approach. *Globalization and Health*, *6*, 19. doi:10.1186/1744-8603-6-19

Buse, K., & Waxman, A. (2001). Public-private health partnerships: a strategy for WHO. *Bulletin of the World Health Organization*, *79*(8), 748–754. doi:10.1590/S0042-96862001000800011

Cattell, J., Chilukuri, S., & Levy, M. ([s.d.]). How big data can revolutionize pharmaceutical

R&D | McKinsey & Company. Recuperado 24 de agosto de 2016, de http://www.mckinsey.com/industries/pharmaceuticals-and-medical-products/our-insights/how-big-data-can-revolutionize-pharmaceutical-r-and-d

Celadon, K. L. (2014). Knowledge Integration and Open Innovation in the Brazilian Cosmetics Industry. *Journal of Technology Management & Innovation*, *9*(3), 34–50. doi:10.4067/S0718-27242014000300003

Cesario, S. K. (2016). Sustainable Development Goals for Monitoring Action to Improve Global Health. *Nursing for Women's Health*, *20*(4), 427–431. doi:10.1016/j.nwh.2016.06.001

Chaifetz, S., Chokshi, D. A., Rajkumar, R., Scales, D., & Benkler, Y. (2007). Closing the access gap for health innovations: an open licensing proposal for universities. *Globalization and Health*, *3*(1), 1. doi:10.1186/1744-8603-3-1

Chaves, G. C., Oliveira, M. A., Hasenclever, L., & Melo, L. M. de. (2007). Evolution of the international intellectual property rights system: patent protection for the pharmaceutical industry and access to medicines. *Cadernos de Saúde Pública*, *23*(2), 257–267. doi:10.1590/S0102-311X2007000200002

Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business Intelligence and Analytics: From Big Data to Big Impact. *Business Intelligence Research*, *36*(4), 1165–1188.

Chluski, A., & Ziora, L. (2015). The application of big data in the management of healthcare organizations. A review of selected practical solutions. *Informatyka Ekonomiczna*, (1). doi:10.15611/ie.2015.1.01

Collective intelligence. (2016, agosto 2). In *Wikipedia, the free encyclopedia*. Recuperado de https://en.wikipedia.org/w/index.php?title=Collective_intelligence&oldid=732642532

Engel, D., Woolley, A. W., Aggarwal, I., Chabris, C. F., Takahashi, M., Nemoto, K., … Malone, T. W. (2015). Collective Intelligence in Computer-Mediated Collaboration Emerges in Different Contexts and Cultures (p. 3769–3778). ACM Press. doi:10.1145/2702123.2702259

EU Bookshop. (2016). *Open innovation, open science, open to the world - Research policy and organisation - EU Bookshop*. Recuperado de http://bookshop.europa.eu/en/open-innovation-open-science-open-to-the-world-pbKI0416263/

Fortes, P. A. de C., & Ribeiro, H. (2014). Saúde Global em tempos de globalização. *Saúde e Sociedade*, *23*(2), 366–375. doi:10.1590/S0104-12902014000200002

Gagnon, M. L. (2011). Moving knowledge to action through dissemination and exchange. *Journal of Clinical Epidemiology*, *64*(1), 25–31. doi:10.1016/j.jclinepi.2009.08.013

Gray, J. and Chambers, L. and Bounegru, L., The Data Journalism Handbook, O'Reilly Media, 2012 - InfoVis:Wiki. ([s.d.]). Recuperado 6 de março de 2014, de http://www.infovis-wiki.net/index.php?title=Gray,_J._and_Chambers,_L._and_Bounegru,_L.,_The_Data_Journalism_Handbook,_O%27Reilly_Media,_2012

Haines, A., McMichael, A. J., Smith, K. R., Roberts, I., Woodcock, J., Markandya, A., … Wilkinson, P. (2009). Public health benefits of strategies to reduce greenhouse-gas emissions: overview and implications for policy makers. *The Lancet*, *374*(9707), 2104–2114. doi:10.1016/S0140-6736(09)61759-1

Hartz, Z. M. A. (2012). Meta-evaluation of health management: Challenges for "new public health". *Ciencia e Saude Coletiva*, *17*(4), 832–834. doi:10.1590/S1413-81232012000400004

Hilbert, M., & López, P. (2011). The world's technological capacity to store, communicate, and compute information. *Science (New York, N.Y.)*, *332*(6025), 60–65. doi:10.1126/science.1200970

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

284

Hossain, M. M., & Aydin, H. (2011). A Web 2.0-based collaborative model for multicultural education. *Multicultural Education & Technology Journal*, *5*(2), 116–128. doi:10.1108/17504971111142655

Humbert, M. (2003). *Globalization and glocalization: problems for developing countries and policy (supranational, national and subnational) implications. In Cassiolato, J. E., Lastres, H.M.M. and Maciel, M.L. (eds). Systems of Innovation and Development*. Cheltenham, UK: EE Elgar. Recuperado de https://www.e-elgar.com/shop/systems-of-innovation-and-development?___website=uk_warehouse

Huyghe, F.B. (2009a). Web 2.0: Influence, outils et réseaux. *Revue Internationale d'Intelligence Economique*, 11.

Huyghe, F.B. (2009b). Web 2.0: Influence, outils et réseaux. *Revue Internationale d'Intelligence Economique*, 11.

Jamison, D. T., Summers, L. H., Alleyne, G., Arrow, K. J., Berkley, S., Binagwaho, A., … Yamey, G. (2013a). Global health 2035: a world converging within a generation. *The Lancet*, *382*(9908), 1898–1955. doi:10.1016/S0140-6736(13)62105-4

Jamison, D. T., Summers, L. H., Alleyne, G., Arrow, K. J., Berkley, S., Binagwaho, A., … Yamey, G. (2013b). Global health 2035: a world converging within a generation. *The Lancet*, *382*(9908), 1898–1955. doi:10.1016/S0140-6736(13)62105-4

Janssen, M., van der Voort, H., & Wahyudi, A. (2017). Factors influencing big data decision-making quality. *Journal of Business Research*, *70*, 338–345. doi:10.1016/j.jbusres.2016.08.007

Jee, K., & Kim, G.-H. (2013). Potentiality of Big Data in the Medical Sector: Focus on How to Reshape the Healthcare System. *Healthcare Informatics Research*, *19*(2), 79. doi:10.4258/hir.2013.19.2.79

Jung, J. J. (2017). Computational Collective Intelligence with Big Data: Challenges and Opportunities. *Future Generation Computer Systems*, *66*, 87–88. doi:10.1016/j.future.2016.08.021

Kaslow, D. C., Kalil, J., Bloom, D., Breghi, G., Colucci, A. M., De Gregorio, E., … Xu, X. ([s.d.]). The role of vaccines and vaccine decision-making to achieve the goals of the Grand Convergence in public health. *Vaccine*. doi:10.1016/j.vaccine.2016.10.088

Kim, J.-S., Kim, E.-S., & Kim, J.-H. (2016). Towards Conceptual Predictive Modeling for Big Data Framework. *International Journal of Software Engineering and Its Applications*, *10*(1), 35–42. doi:dx.doi.org/10.14257/ijseia.2016.10.1.0 4

Koplan, J. P., Bond, T. C., Merson, M. H., Reddy, K. S., Rodriguez, M. H., Sewankambo, N. K., … Consortium of Universities for Global Health Executive Board. (2009). Towards a common definition of global health. *Lancet (London, England)*, *373*(9679), 1993–1995. doi:10.1016/S0140-6736(09)60332-9

Kvan, T., & Candy, L. (2000). Designing collaborative environments for strategic knowledge in design. *Knowledge-Based Systems*, *13*(6), 429–438. doi:10.1016/S0950-7051(00)00083-6

Laney, D. (2001). *3D Data Management: Controlling Data Volume, Velocity, and Variety* (Application Delivery Strategies) (p. 4). Recuperado de http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf

Lawrence, S., & Giles, C. L. (2000). Accessibility of information on the Web. *Intelligence*, *11*(1), 32–39. doi:10.1145/333175.333181

Lee, In. (2011). Overview of Emerging Web 2.0-Based Business Models and Web 2.0 Applications in Businesses: An Ecological Perspective: Business IS&T Journal Article | IGI Global. *International Journal of E-Business*

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

285

Research (IJEBR), *7*(4), 1–16. doi:10.4018/jebr.2011100101

Magalhães, J. L., & Quoniam, L. (2013). Perception of the Information Value for Public Health: A Case Study for Neglected Diseases: Library and Information Science Book Chapter | IGI Global. In *Rethinkin the Conceptual Base for New Practical Applications in Information Value and Quality* (p. 345). IGI Global. Recuperado de http://www.igi-global.com/chapter/perception-information-value-public-health/84218

Magalhaes, JL, & Quoniam, L. (2015). Percepção do valor da informação por meio da inteligência competitiva 2.0 e do Big Data na saúde. In *Análise da Informação para Tomada de Decisão: desafios e soluções* (Vol. 1, p. 365). Brasil: Kira Tarapanoff (Org.).

Malaria: A new fixed-dose combination of artesunate and mefloquine is effective and well tolerated. – DNDi. ([s.d.]). Recuperado 8 de abril de 2017, de https://www.dndi.org/2005/media-centre/press-releases/malaria-a-new-fixed-dose-combination-of-artesunate-and-mefloquine-is-effective-and-well-tolerated/

McKinsey Global Institute. (2011a). *Big Data: The Management Revolution - Harvard Business Review*. Recuperado de http://hbr.org/2012/10/big-data-the-management-revolution/ar/1

McKinsey Global Institute. (2011b). *Big Data: The Management Revolution - Harvard Business Review*. Recuperado de http://hbr.org/2012/10/big-data-the-management-revolution/ar/1

McKinsey Global Institute. (2011c). Big data: The next frontier for innovation, competition, and productivity | McKinsey Global Institute | Technology & Innovation | McKinsey & Company. Recuperado 7 de março de 2013, de http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation

Mena-Chalco, J. P., Digiampietri, L. A., Lopes, F. M., & Cesar-Jr., R. M. (2013). Brazilian bibliometric co-authorship networks, *in press*.

Michelino, F., Cammarano, A., Lamberti, E., & Caputo, M. (2015). Knowledge Domains, Technological Strategies and Open Innovation. *Journal of Technology Management & Innovation*, *10*(2), 50–78.

Minelli, M., Chambers, M., & Dhiraj, A. (2013). *Big Data, Big Analytics*. EUA: John Wiey & Sons, Inc. Recuperado de https://books.google.com.br/books/about/Big_Data_Big_Analytics.html?hl=pt-BR&id=Mg3WvT8uHV4C

Nakayama, K., Ito, M., Hara, T., & Nishio, S. (2008). Wikipedia Mining for Huge Scale Japanese Association Thesaurus Construction (p. 1150–1155). IEEE. doi:10.1109/WAINA.2008.37

Nature News. (2016). The ups and downs of data sharing in science. *Nature*, *534*(7608), 435–436. doi:10.1038/534435b

Nova Spivack. (2013). Articles | Nova Spivack - Minding the Planet. Recuperado 28 de janeiro de 2013, de http://www.novaspivack.com/articles

O'Reilly, T. (2007a). *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software* (SSRN Scholarly Paper No. ID 1008839). Rochester, NY: Social Science Research Network. Recuperado de http://papers.ssrn.com/abstract=1008839

O'Reilly, T. (2007b). *What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software* (SSRN Scholarly Paper No. ID 1008839). Rochester, NY: Social Science Research Network. Recuperado de http://papers.ssrn.com/abstract=1008839

O'Sullivan, D., & Dooley, L. (2009). *Applying Innovation*. London: SAGE PUblications, Inc. Recuperado de https://books.google.com.br/books/about/Applying_Innovation.html?hl=pt-BR&id=azgXBAAAQBAJ

Paul, A., Chilamkurti, N., Daniel, A., & Rho, S. (2017). Chapter 8 - Big Data collision analysis framework. In *Intelligent Vehicular Networks and Communications* (p. 177–184). Elsevier. Recuperado de http://www.sciencedirect.com/science/article/pii/B9780128092668000089

Paul, S. M., Mytelka, D. S., Dunwiddie, C. T., Persinger, C. C., Munos, B. H., Lindborg, S. R., & Schacht, A. L. (2010). How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature Reviews Drug Discovery*, *9*(3), 203–214. doi:10.1038/nrd3078

Pinheiro, L. V. R. (2014). Do acesso livre à ciência aberta: conceitos e implicações na comunicação científica. *Revista Eletrônica de Comunicação, Informação & Inovação em Saúde*, *8*(2). doi:10.3395/reciis.v8i2.629

Quoniam, L, Lucien, A. (2010). *Intelligence compétitive 2.0 : organisation, innovation et territoire*. France: Librairie Lavoisier. Recuperado de http://www.lavoisier.fr/livre/notice.asp?ouvrage=2139418&pos=8

Rafols, I., Hopkins, M. M., Hoekman, J., Siepel, J., O'Hare, A., Perianes-Rodríguez, A., & Nightingale, P. (2014). Big Pharma, little science? *Technological Forecasting and Social Change*, *81*, 22–38. doi:10.1016/j.techfore.2012.06.007

Raghupathi, W., & Raghupathi, V. (2014). Big data analytics in healthcare: promise and potential. *Health Information Science and Systems*, *2*. doi:10.1186/2047-2501-2-3

Resolução Conselho de Ministros. Resolução do Conselho de Ministros n.º 20/2016 – Diário da República n.º 70/2016, Série I de 2016-04-11, Pub. L. No. 70/2016, Série I de 2016-04-11 (2016). Recuperado de http://legislacaoportuguesa.com/resolucao-do-conselho-de-ministros-n-o-202016-diario-da-republica-n-o-702016-serie-i-de-2016-04-11/

Salmond, J. A., Tadaki, M., & Dickson, M. (2017). Can big data tame a "naughty" world?: Environmental big data. *The Canadian Geographer / Le Géographe Canadien*, *61*(1), 52–63. doi:10.1111/cag.12338

Santelli, A. C., Ribeiro, I., Daher, A., Boulos, M., Marchesini, P. B., dos Santos, R. L. C., … Ladislau, J. L. (2012). Effect of artesunate-mefloquine fixed-dose combination in malaria transmission in amazon basin communities. *Malaria Journal*, *11*, 286. doi:10.1186/1475-2875-11-286

Sirima, S. B., Ogutu, B., Lusingu, J. P. A., Mtoro, A., Mrango, Z., Ouedraogo, A., … Carn, G. (2016). Comparison of artesunate–mefloquine and artemether–lumefantrine fixed-dose combinations for treatment of uncomplicated Plasmodium falciparum malaria in children younger than 5 years in sub-Saharan Africa: a randomised, multicentre, phase 4 trial. *The Lancet Infectious Diseases*, *16*(10), 1123–1133. doi:10.1016/S1473-3099(16)30020-2

Souza, A. L. P. de, Pitassi, C., Bouzada, M. A. C., Gonçalves, A. A., Souza, A. L. P. de, Pitassi, C., … Gonçalves, A. A. (2015). The Brazilian Network for Public Production of drugs in the perspective of supply chain management: the role of ICTs. *Revista de Administração Pública*, *49*(3), 615–641. doi:10.1590/0034-7612132157

Terry, R. F., Allen, L., Gardner, C. A., Guzman, J., Moran, M., & Viergever, R. F. (2012). Mapping global health research investments, time for new thinking - A Babel Fish for research data. *Health Research Policy and Systems*, *10*(1), 28. doi:10.1186/1478-4505-10-28

The Global Alliance for Genomics and Health. (2016). A federated ecosystem for sharing genomic, clinical data. *Science*, *352*(6291), 1278–1280. doi:10.1126/science.aaf6162

The White House. (2016). *Big Data: A Report on Algorithmic Systems, Opportunity, and Civil Rights* (p. 29). EUA. Recuperado de https://www.whitehouse.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf

Trigo, M. R., Gouveia, L. B., Quoniam, L., & Riccio, E. L. (2007). *Using competitive intelligence*

as a strategic tool in a higher education context. (D. Remenyi, Org.). Nr Reading: Academic Conferences Ltd.

Turner, W., Rondinini, C., Pettorelli, N., Mora, B., Leidner, A. K., Szantoi, Z., … Woodcock, C. (2015). Free and open-access satellite data are key to biodiversity conservation. *Biological Conservation*, *182*, 173–176. doi:10.1016/j.biocon.2014.11.048

Univesity of California San Francisco - UCSF. ([s.d.]). The Report | Global Health 2035. Recuperado 18 de julho de 2016, de http://globalhealth2035.org/report

Valecha, N., Srivastava, B., Dubhashi, N. G., Rao, B. H. K., Kumar, A., Ghosh, S. K., … Anvikar, A. R. (2013). Safety, efficacy and population pharmacokinetics of fixed-dose combination of artesunate-mefloquine in the treatment of acute uncomplicated Plasmodium falciparum malaria in India. *Journal of Vector Borne Diseases*, *50*(4), 258–264.

Van Noorden, R. (2012). Global mobility: Science on the move. *Nature*, *490*(7420), 326–329. doi:10.1038/490326a

Vance, K., Howe, W., & Dellavalle, R. P. (2009). Social Internet Sites as a Source of Public Health Information. *Dermatologic Clinics*, *27*(2), 133–136. doi:10.1016/j.det.2008.11.010

Wang, Y., & Hajli, N. (2017). Exploring the path to big data analytics success in healthcare. *Journal of Business Research*, *70*, 287–299. doi:10.1016/j.jbusres.2016.08.002

Wassan, J. T. (2001, janeiro 1). Big Data Paradigm for Healthcare Sector [chapter]. Recuperado 26 de fevereiro de 2016, de http://www.igi-global.com/chapter/big-data-paradigm-for-healthcare-sector/141111

WikiEdu. (2016). The Year of Science spreads to Brazil – Wiki Education Foundation. Recuperado 25 de agosto de 2016, de https://wikiedu.org/blog/2016/08/02/brazil-year-of-science/

Wilson, M. E., Ko, A. I., & Reis, M. G. (2011). Collaborative Teaching and Learning: A Model for Building Capacity and Partnerships to Address NTDs. *PLoS Neglected Tropical Diseases*, *5*(3), e939. doi:10.1371/journal.pntd.0000939

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. doi:10.1016/j.rse.2012.01.010

Int. J. Innov., São Paulo, v. 5, n. 3, pp. 270-288, Sep/Dec. 2017.

288