

Internet infodiversity: State of the art and future trends

Jonathan Hernández Pérez

Ph.D. student, Library and Information Studies, National Autonomous University of Mexico (UNAM), Mexico

Abstract. The aim of this paper is to present a state of the art of the Internet infodiversity, through its five aspects; plurality, recovery, conservation, availability and free access to information. It also discusses issues related to censorship and freedom of expression on the Internet.

Keywords. Infodiversity, Internet, Free access to information, information diversity

1. Introduction

Inclusion and respect for political, religious, cultural, and intellectual ideologies and practices, has gained importance in society because only in this way we can reach a social balance among nations. The same applies to the information, if we recognize that we as humans are diverse that would result in recognition of the variety of ideas and thoughts we have, which originate a large amount of information.

The idea of infodiversity was introduced by Morales Campos (1999) and refers to the defense of the existence of multiple and diverse manifestations of information, which implies the convergence of different types of this Morales Campos (1999). Infodiversity has existed since humans began to produce information and encompasses all the information produced by humanity from different social groups in all historical periods. Therefore, the infodiversity can be studied from its manifestations and can be applied to specific geographic space, media or time period. In this way, the subject of study in this research is the Internet.

It is important to mention that the methodological approach in this research was the methodology of content analysis, which is a research technique designed to make-from certain data-, reproducible and valid inferences can be applied to certain contexts. Content analysis is also a research tool focused on

the actual content and internal features of media; the study of recorded human communications, such as books, websites, blogs and laws. Particularly in this study used statistics from international organizations and literature.

2. Cultural Approaches of the Internet

Different aspects regarding the evolution of the Internet have been extensively reported in recent literature. It is known that Internet has revolutionized important aspects of our society, particularly in the fields of communication, identity and information retrieval.

The emergence of Internet as a new medium of communication has generated controversy over the emergence of new patterns of social interaction; this is because one of the most important characteristics is that Internet provides the opportunity for users to be consumers and producers of information at the same time. All of this in a social space where the most diverse expressions converge. The fact of produce and share information along with its open nature, it's the Internet most important quality.

Moreover, the uses of the Internet are closely linked with work, family and everyday life of users, today be on the Internet means having a representation of our self, an identity that is constructed from our own activity on the Web. Is a fact that we generate enormous amounts of online data about our habits: where we go, what we do, and how we feel, consequently we are building a "virtual self". Thus, and according to Giones and Serrat (2010) identity on the Internet can be defined as a digital representation of a set of claims made by one party about itself or another data subject.

Although information retrieval has been developing since the late 1950s, now accounts for a greater role than in previous decades, due in large part to the infodiversity, globalization and the development of Internet.

Information retrieval on the Internet not only has resulted in the development of more sophisticated search engines, it also has been established as a business model, Internet users will never stop looking up information, and the way in which search remains an essential part of Internet use can be attributed to one simple reason: the constant growth of information in all types of formats.

The effects of information retrieval on the Internet have had a greater impact on young people, especially those born with this technology, even several authors have noted a deep change in the way people search for information, establishing a generation gap for these users.

Although there is no standardized terminology to refer this phenomenon, in the literature are attributed names like Millennials, net generation, Generation Y, screenagers, Google generation, and digital native, among others.

3. Internet infodiversity: state of art

The development of the five elements of the infodiversity on the internet (plurality, recovery, preservation, availability and free access to information) is detailed next.

The plurality refers to acknowledge the diversity of Internet users, the variety of languages in which they interact and the different contents they produce. Plurality is the starting point for the construction of the infodiversity because if we have a diversity of users, there will be a variety of content, formats and language in which these contents are produced.

On the Internet, we generalize the term of "user", but also used other terms like Internaut and Cybernaut, these were mostly used when Internet was known as the "Information Superhighway" or "Cyberspace" respectively. However, due to the increasing interactivity as a result of the use of social networks has begun to talk of citizens, emphasizing that they not only use the Internet, but produce and interact in it, calling Netizen, this term is not new, but has gained importance in recent years.

Despite the fact that the Internet is constantly growing, it is noteworthy that there is still much to be done in order to incorporate most of the world's population into this technology because currently only 34% of the world population has access, i.e. some 2, 405,518,376 of people, however these users come from different regions of the world and interact in a variety of languages.

In this sense, on the report of the Internet World Stats (2013) mention that the ten languages that dominate Internet are English, Chinese, Spanish, Japanese, Portuguese, German, Arabic, French, Russian and Korean, having a variety in terms of growth, highlighting the top three, English, Chinese and Spanish, and particularly the Arabic which is the language with the most growth in the last eleven years.

In terms of content, social networks have become a very important factor in the recent years this is because they are now an integral part of daily life for many Internet users and there are now hundreds of millions utilizing these services around the world. These social networks generate a lot of content that is shared and read daily, as an example we mention the production of twitter, in which the six largest countries in terms of information production through this social network according to the Oxford Internet Institute (2013) are the United States, Brazil, Indonesia, the United Kingdom, Mexico, and Malaysia.

In this regard, it is important to mention the importance of the Internet domain names because these are a reflection of the Web pages diversity, which can be academic, popular, specialized, with content specific, etc. On this subject, in 2012 some businesses and other organizations applied for a wide array of domains in a list of nearly 2,000 top-level domain names at the Internet Corporation for Assigned Names and Numbers (ICANN) and indicated that they would like to operate the registry as -closed- which means that they may not allow other organizations to buy second-level domains. The new domains would vastly expand the pool of suffixes beyond ".com" and ".net" and perhaps add such new Web address endings as ".lol," ".google" or ".free." This has brought a combination of trademark disputes and arguments about which companies or organizations will be appropriate owners of TLDs. Unfortunately, there were no requests for .library or .biblioteca.

Moreover, the World Wide Web Foundation launched the project the Web Index (2012) that is the world's first multi-dimensional measure of the Web's

growth, utility and impact on people and nations. This can give an approximate measure of what the Web in content and political influence. The Web Index covers a large number of countries, some of which data have serious deficiencies or were not covered by the data providers.

It highlights Sweden; of all 61 countries, Sweden takes top place in this year's ranking, with high marks across the three sub-indexes, nevertheless some of its scores are surprising: Sweden tops the list for overall impact of the web, taking first place for political, second place for social and third place for Economic Impact. And it is second highest on the global list in terms of Readiness, scoring third for Communications Infrastructure and fifth for Institutional Infrastructure.

Mexico takes fourth place in the Americas and is ranked 22nd globally. Its highest ranks are in the areas of Web Content and use, as well as Political Impact, this is noteworthy because there have been significant results due to sociopolitical activity on social networks. Brazil, in contrast, ranks higher than Mexico in Readiness, but lower in terms of overall Political Impact, taking 5th place in the regional ranking.

The cyclical nature of the information, i.e. the fact that it is based on prior information, only serves to highlight the need to rescue information from the past in its various representations. In this sense, the Recovery is the second element for the development of the infodiversity, in general it can be said that the recovery is the rescue of the documentary heritage for use and dissemination through ICT. The potential of Internet can be used to recover a large part of the infodiversity that is not on the internet, like printed books, magazines, newspapers, artwork, comics, etc.

The rescue of Web pages that have disappeared it is an important issue because the Internet is a reflection of today's society, the rescue provides great benefits for research and reference, for example if some researcher wants to know at what point particular social group / country / person began to have presence and visibility on the Internet, or if some specialist want to document online political activity of each country over time or know how it has changed the way Web sites are structured, etc.

The preservation, the third aspect of the infodiversity is closely related with previous: the rescue. Every day the information turns into a digital form and is a fact that much of this is generated only in digital form or is convert it from analog or physical material, integrating into global infodiversity, so that the disappearance or mutilation of any information that has been produced in any language and support restrict the development of the infodiversity, by this reason the preservation of this is essential for current and future generations can consult.

The information availability aims to provide a user all available information required to satisfy their needs regardless of where they are. This is technically possible, but often the political and social factors do not allow. Particularly in the context of Internet infodiversity, availability is the characteristic, attribute or condition of the information being available to those who need access to it,

whether people, processes or applications. In this sense, digital libraries, repositories, databases play a key role in this aspect of the infodiversity.

Free access to information is a key factor for the existence of a balanced infodiversity, in this case on the Internet. Free access to information is a fundamental right for human beings as defined in Article 19 of the United Nations Universal Declaration of Human Rights: "Everyone has the right to freedom of opinion and expression; this right includes freedom to hold opinions without interference and to seek, receive and impart information and ideas through any media and regardless of frontiers".

Also, in 2011 a report from the Human Rights Council of the United Nations General Assembly declares access to the Internet a basic human right which enables individuals to "exercise their right to freedom of opinion and expression".

As stated by the Human Rights Council (2011) all people should be allowed to connect to and express themselves freely on the Internet, and that disconnecting people from the internet is a human rights violation and against international law.

Even the IFLA Internet Manifesto (2002) declares that the Internet provides a medium through which everyone can enjoy this right. Consequently, access should neither be subject to any form of ideological, political, religious, or economic barriers.

However, we must not forget that the Internet is a technology which is not free and costless at all. According to Hamilton (2004) the Internet mixes freedom and control at different levels. The physical level is fundamentally controlled – wires and computers around the world are generally the property of governments, businesses or individuals. And at the content layer, much is controlled. Not everything is free for the taking, and this is the area that protected by copyright law.

The wide diffusion of the Internet in developed and developing nations has made it a key medium for political debate and activism. More access to information entails, among other things, a greater participation in the democratic processes of each country. Access to information via the Internet is especially valuable in special political moments such as elections, times of social unrest or historical anniversaries. And social media tools can give users reach that was previously only available to much larger organizations.

In light of this, some authorities restrict and regulate access to the Internet and the content of it. Those services are mostly focused on the publication, dissemination and retrieval of information such as blogs, video sharing sites, podcasts, wikis and search engines, among others, or even prevent access to the Internet through firewalls, high costs for service with limited speed, criminal sanctions, and so on.

Censorship and other mechanisms that hinder the free access to information also hinder scientific and technological innovations, therefore the infodiversity. In this respect Castells (2008) said that "although society does not determine technology, it can stifle their development, especially through the state. Or, alternatively, and especially through state intervention, can embark on an

accelerated process of technological modernization, able to change the fate of the economy, military power and social welfare in a few years.”. Thus, in a medium like the Internet, the damage may be higher since the vast majority of Internet innovations have been made thanks to its main feature: its openness.

In addition of the censorship that each government can practice, access to information on the Internet can be controlled and their uses can be biased or even monopolized by commercial, ideological and political interests.

As an example of this are the various bills that have wanted to stop the free flow of online information.

- Anti-Counterfeit Trade Agreement (ACTA)
- Stop Online Piracy Act (SOPA)
- The PROTECT IP Act (Preventing Real Online Threats to Economic Creativity and Theft of Intellectual Property Act of 2011 PIPA)
- Research Works Act (RWA)
- Cyber Intelligence Sharing and Protection Act (CISPA)

In Latin America we have other examples

- Colombia - Proyecto de ley 241 (Ley Lleras)
- Argentina - Régimen para proveedores del servicio de Internet (Ley Pinedo)
- México - Ley Döring

As has been noted, the panorama of the Internet infodiversity has been immersed in permanent changes. These changes have resulted from technological progress and social needs in an increasingly globalized world. Based on the development of this work, it provides the prospective of the Internet infodiversity.

According to the study in the Plurality section, the number of Internet users will grow considerably, integrating users from around the world, which will affect the content, format and intensity of information flows. Also, more and more languages will be represented on the Internet significantly, thereby promoting linguistic diversity and the rescue of endangered languages. Evidently, search engines will have a key role in this aspect because with the increasing production of information on the Internet these will have to provide better mechanisms for information retrieval.

Regarding the aspects of Rescue and Preservation, these areas will benefit greatly by the user who will have the tools to carry out processes to retain the information they produce. Moreover, it will be necessary that national libraries begin to store all the information that occurs on the Web, considering cover all Web production that is done in their own countries.

The availability and free access to information will continue to be the most sensitive aspect, since the socio-political events that have happened so far are just the beginning of major challenges that society faces as a result of censorship, privacy violations and legislations to control the citizen's right to access to the Internet.

4. Conclusions

Internet is a medium that enhances infodiversity by its diverse and plural information offered. However, it is necessary to implement local strategies in order to achieve a greater production online which have to be accompanied by policies that integrate people to this kind of technology, such policies should be consistent with the recognition of the infodiversity, avoiding any risk exclusion and domination of information and promoting respect for difference.

Also, the library will be a fundamental part in order to keep the balance of infodiversity, as an institution that either physically or digitally promote and defend the development of the Internet infodiversity, promoting innovation in information production and dissemination, contributing to the rescue and conservation and ensuring availability and free access to information.

References

- Castells, M. (2008). *La era de la información* Vol. 1. Madrid: Alianza Editorial.
- CNN Mexico. (2011, 25 December). La ONU declara el acceso a Internet como un derecho humano. Retrieved July 16, 2012, from <http://mexico.cnn.com/tecnologia/2011/06/08/la-onu-declara-el-acceso-a-internet-como-un-derecho-humano>
- Giones, A., & Serrat, M. (2010). La gestión de la identidad digital: una nueva habilidad informacional y digital. En: BiD: textos universitaris de biblioteconomia i documentació. BiD: textos universitaris de biblioteconomia i documentació. Retrieved January 25, 2013, from <http://www.ub.edu/bid/24/>
- Hamilton, S. (2004). *To what extent can libraries ensure free, equal and unhampered access to Internet-accessible information resources from a global perspective?* Copenhagen: Department of Library and Information Management, Royal School of Library and Information Science/FAIFE.
- Human Rights Council. (2011, 16 May). *United Nations Human Rights*. Retrieved January 20, 2013, from http://www2.ohchr.org/english/bodies/hrcouncil/docs/17session/A.HRC.17.27_en.pdf
- IFLA. (2002, 26 March). *The IFLA Internet Manifesto*. Retrieved January 22, 2013, from <http://www.ifla.org/files/assets/faife/publications/policy-documents/internet-manifesto-es.pdf>
- Internet World Stats. (2010, 30 June). *Top Ten Languages Used in the Web*. Retrieved March 12, 2013, from <http://www.internetworldstats.com/stats7.htm>
- Morales, E. (2006). *Infodiversidad y cibercultura : globalización e información en América Latina*. Buenos Aires. Buenos Aires: Alfagrama.
- Organización de las Naciones Unidas. (2012, 02 February). *Declaración Universal de los Derechos Humanos*. Retrieved March 12, 2013 <http://www.un.org/es/documents/udhr>
- Oxford Internet Institute (OII) . (2012, 19 June). *Visualizing data, a Geography of Twitter*. Retrieved March 12, 2013 From <http://www.oii.ox.ac.uk/vis/?id=4fe09570>
- World Wide Web Foundation. (2012). *The Web Index*. Retrieved March 18, 2013 <http://thewebindex.org/2012/10/2012-Web-Index-Key-Findings.pdf>