Xin-She Yang · Simon Sherratt · Nilanjan Dey · Amit Joshi
Editors

Proceedings of Ninth International Congress on Information and Communication Technology

ICICT 2024, London, Volume 8

Springer

Scientific Communication in eScience

Francisco Carlos Paletta

Abstract This paper examines the role of scientific communication and information management in the context of eScience research. It discusses how the evolution of information and communication technologies (ICTs) has shaped the operationalization and fluidity of scientific communication. The paper emphasizes the importance of digital literacy and information literacy in the network society. It highlights the mediating role of libraries and information services in scientific communication through knowledge storage, organization, and access. The concept of eScience is explained as data-driven and computationally intensive research methodology enabled by technological infrastructure like high-performance computing, data storage/management, and networks. The paper describes how eScience involves collaboration between computing researchers and domain experts, as well as creation of computational tools and methods to analyze large, complex data. It also covers the fourth paradigm of data-driven science paradigm facilitated by eScience. The paper concludes by underscoring the importance of managing the life cycle of ICTs in digital libraries and repositories to support efficient scientific communication through open access, data sharing, and collaboration platforms. It advocates for an information technology governance model aligned with the mission of digital libraries to effectively organize and communicate knowledge.

Keywords eScience · Scientific communication · Research · Digital literacy · Data science

F. C. Paletta (✉)
Universidade de São Paulo, São Paulo, Brazil
e-mail: fcpaletta@usp.br

X.-S. Yang et al. (eds.), *Proceedings of Ninth International Congress on Information and Communication Technology*, Lecture Notes in Networks and Systems 1004,
https://doi.org/10.1007/978-981-97-3305-7_33

1 Introduction

All scientific endeavors are shaped by historical and socioeconomic factors as social activities. In this context, the information society necessitates a scientific approach to examine the characteristics of information and the intricacies of its creation, dissemination, and utilization. In the contemporary world, the subject matter of information science diverges from traditional disciplines such as library science and related fields. It no longer revolves around the library, the book, the documentation center, or the physical object; rather, the primary emphasis lies on the concept of information itself [1].

The primary principle of the Web, as proposed by the World Wide Web Consortium Brazil, emphasizes that the core value of the Web is its social utility. It transcends being merely a technological platform, as it facilitates human communication, commercial transactions, and the sharing of knowledge. To achieve universality, the Web must be accessible to everyone, irrespective of the devices and software used, as well as considerations of culture, geographic location, physical and mental abilities, and socioeconomic or educational conditions. Ensuring and enhancing the universality of the Web require a democratic and pluralistic governance model that prioritizes universal access and the ongoing technological advancement of the Web itself [2].

The operationalization and fluidity of scientific communication are structured around the evolution of the ICTs, and for various digital libraries, the ever-growing provision of technologies has shown ambiguity in its own management. From a positive outlook, the introduction of these innovative technologies has contributed to a boost in the efficiency of information professionals, enhanced decision-making processes, and underscored the satisfaction of information users. Nonetheless, there is evidence that the management and support of these heterogeneous and complex environments—full of different personal computers, desktops and portables, mobile devices, printers, networks, and applications—have been found to be difficult and expensive for the information technology departments. In this context, it is important to analyze the main obstacles that the digital libraries will be faced with in relation to the management of the life cycle of their technologies and the consolidation and simplification of their processes within their computational environments in order to increase productivity and build agile environments that allows for the libraries to meet the demands of the digital information management in providing scientific communication with efficiency and quality.

The incorporation of these technologies is regarded as a milestone that should be aligned with the tradition and mission of digital libraries and repositories. The primary strategy should center on assessing the adaptability of computational structures, their appeal, and dynamics, where the user assumes the role of an agent in shaping the environment. This necessitates resources for customization and personalization to establish innovative information services (processes or services) that empower the library to maintain its crucial role in mediating, producing, and communicating new knowledge.

The increasing complexity of the technological devices has motivated IT managers to seek ways to improve the efficiency in operation so that there is a cost reduction, comply with the regulatory aspects of Information Technology Governance [3], and meet the constant demands of digital libraries in order to provide a better response to the demands made by the user, as well as the definition and execution of preservation strategies of digital information.

Scientific communication, facilitated by information technology, demands that managers of digital libraries and repositories possess a comprehensive understanding of collaborative technology resources within digital information environments. Building upon resources identified and gathered from digital libraries and repositories, there is an endeavor to observe the implementation of collaborative technology resources in the Semantic Web context, also recognized as the Web of Data. The integration of these technologies represents a significant breakthrough that should be aligned with the tradition and mission of digital libraries and repositories. This alignment is crucial for effectively managing and organizing knowledge in open systems, leading to the establishment of an information technology governance model for the Digital Library.

The construction of a more efficient Internet leads the way to fomenting a revolution in the digital universe of information and knowledge organization. The models of information search, access, retrieval, appropriation, and utilization must be in line with the new eScience paradigms and play the role of scientific communication inductors. With the use of innovative technologies, the adoption of information and communication technologies is paramount so that the search for information and the generation of new knowledge become more agile. In such a context, a new information user that makes new demands on computational resources and new capacities for producing new knowledge emerges [4].

Over the recent years, information literacy has emerged as a new research domain that captivates the attention of education, information science, and cognitive sciences. The initial challenge lies in acquiring proficiency in basic technological resources, often referred to as digital literacy, followed by the subsequent ability to adeptly apply and leverage these skills to generate new knowledge—known as information literacy.

However, this present overview of both forms of literacy that emerged from the core of the network society is not sufficient to shape their deep ruptures. The challenge of the information field is not only to create intellectual and novel technologies, but also to contribute to the development of awareness, in the network society, concerning the remarkable collective richness at global level that the free access to the public domain of information worldwide represents.

Scientific communication necessitates the oversight of the appropriation process undertaken by digital libraries and repositories concerning collaborative technologies within digital information environments and open systems. Additionally, it involves managing the infrastructure of their computational resources within the technological framework of information services directed toward Web users. This is crucial to ensure open and democratic access to information and knowledge in the digital era.

2 eScience

eScience, as a research methodology, is connected to models involving the search, access, retrieval, appropriation, and utilization of scientific information in the form of data—a domain commonly known as data science. It depends on the use of digital mining tools and data analysis. Moreover, it has provided new opportunities for research in every area of knowledge [5].

The viability of eScience requires the following elements, Fig. 1, in terms of infrastructure:

- Network Infrastructure: global network specifically designed for research purposes with high-capacity links.
- High-Performance Computing (HPC): supercomputing resources.
- Data storage and management, information architecture and security.
- Applications and high-performance services.

The volume of data amassed through empirical science, encompassing experiments and computational simulations, has ushered in the fourth paradigm of science in recent years, commonly referred to as data-driven science. This paradigm has organically evolved from the preceding three paradigms: experiments, theory, and computing/simulation, as illustrated in Fig. 2. With this new data-driven perspective, various fields, including big data and data science, have converged to facilitate the extraction of knowledge from data. Big data is a field of study dedicated to understanding how to manage, analyze, and gather information from exceedingly large datasets that cannot be effectively analyzed using traditional systems due to their size and complexity [6].

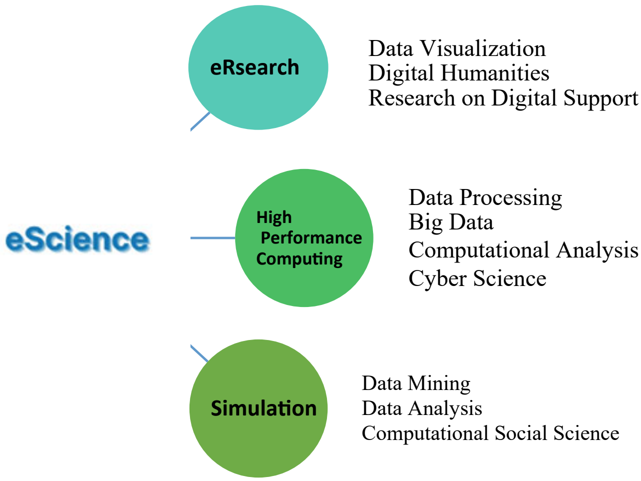

Fig. 1 **eInfrastructure** for eScience. *Source*: Adapted from eScience—"https://vidensportal.deic.dk/en/what-is-eScience"

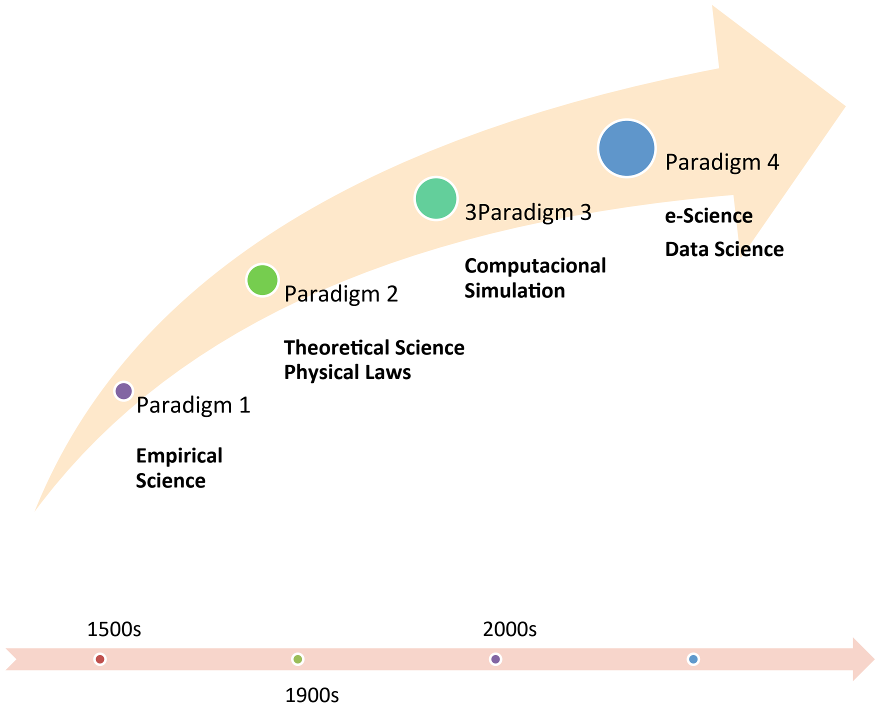

Fig. 2 Paradigms of eScience. *Source* Adapted from GRAY [7]

eScience, where "Information Technology finds the researchers," and the new scientific communication model converge with another for the data collection process using processing, mining, and data analysis tools taking into consideration the fourth paradigm of science [7].

The term e-Science (or eScience), originally emerged in the UK and widely used in the rest of Europe, was coined by John Taylor in 1999 when he was Chief Executive Officer of the Research Councils UK, according to Jankowski (2007). Such a term was translated into Portuguese as e-Ciência, whose meaning represents the power of enhanced science with intensive use of ICTs and its expansion toward a collaborative effort [8].

Founded in the UK in 2001, eScience Institute (2007) aims to "facilitate eScience," promote research, and support interaction of the domains that require the utilization of computational, mathematics, and statistical resources and innovations concerning topics, such as influence and impact of Web, computational development, scientific communication, neuroinformatic, adoption of e-Research technologies, exploration of scientific data sources, and information services for the decision-making process.

In Washington University, the [9] enables researchers and students in all fields to respond to fundamental research questions using a complex analysis of a large quantity of data. It functions as a "hub—portal" for intensive use of data, leading an

Fig. 3 eScience institutes. *Source* https://escience.washington.edu/ and https://web.archive.org/web/20090228073225/ https://www.nesc.ac.uk/esi/index.html

innovative community on the utilization of technologies and better practices of data science considering the fields that depend on it (Fig. 3).

eScience is a globally used term that encompasses research conducted across all fields of knowledge, particularly those involving substantial volumes of data, sophisticated computational methods, and high-performance computing. [10]. The eScience research can be described as follows:

- It involves the collaboration between computing researchers and researchers from other areas of knowledge.
- It requires the creation of sophisticated computational methods to deal with massive quantities of data (known as big data), and/or run simulations and programs that demand complex computational systems.

The formal difficulty of defining the term eScience due to its premature life and lack of common ground in the scientific community is important to highlight. The most effective means of comprehending the definition were to identify its origins. When the term was mentioned for the first time by John Taylor, it was linked to the acknowledgement of the ever-growing importance of the role that information technology plays in the twenty-first century for the data-intensive, multidisciplinary, and collaborative scientific research.

The eScience research [10] addresses every stage of the research process, starting from developing computational tools that assist scientists in formulating research problems and collecting/analyzing data, to modeling, simulation, promotion, and reuse of research outcomes.

3 Scientific Communication and Open Science

The scientific communication and digital identity of the researcher must be coherent in terms of his analogical identity. In that sense, areas and lines of research must be established in tandem with aspects related to professional activities and scientific communication that one seeks to potentialize.

The construction of a digital identity, reputation, and brand requires that the researcher focuses on identification on the institutional webpage, areas of research, teaching activities, experience in management, and profile ID: ORCID, ResearchID, Scopus author ID.

Digital repositories, such as e-Archive, Zenodo, Github, Figshare, LinkedIn, Slideshare; Academic platforms, such as ResearchGate, Academia.edu, Microsoft Academic, Google Scholar, ImpactStory; Social Media, such as Twitter, Facebook, LinkedIn, Instagram; and Bibliography managers, such as Mendeley, Zotero, and Records of peer review on Publons, must be incorporated as part of the scientific communication strategy of the researcher.

The success of a scientific communication strategy is related to the digital identity of the researcher, the community where he practices, the impact of the research, and his reputation and personal brand.

Science emerged as a byproduct of sociocultural evolution, and the globalization of science in terms of accumulation and exchange of scientific knowledge has become a determining factor for modern society [11].

According to Calzada-Prado [12], the viability of the scientific communication through time and space depends on the following elements:

- Codes: Latin was the primary language of science from the scientific revolution until the contemporary age, whereas English has gradually been considered a universal language. The use of a language seen as a reference facilitates the access and the scientific communication.
- Communication Media: The cost reduction of documentary support during the Renaissance period and the possibility of document reproduction (print) favor knowledge communication. Digital technologies can supersede the analog media through new possibilities of information reproduction and dissemination in a virtual manner.
- Channels: The era of conquests, exploration, and empires has allowed for the development of communication channels that preceded the worldwide digital communication networks. The administrative, cultural, and scientific structures, as well as the techniques and technologies, developed throughout the years have been fundamental to store, organize, control, and disseminate knowledge (from archives to libraries with databases and engines of scientific search).
- Sociocultural Context: Factors, such as religion, politics, economy, and cultural boundaries, have been, in different moments of history, restrictive or favorable to scientific globalization.

The role of the university is to generate, disseminate, and receive knowledge, and the Library through Information Services becomes aware of its mediating role in terms of scientific communication, taking care of the documentary processes and preserving information to subsequently transform it into knowledge in a spiral of scientific and technological evolution. In that sense, the university focuses on amplifying the production of technology, innovation, processes, reflection, public policy, and professional training of superior quality, and the library plays the role of an agent that promotes the socialization of knowledge and the production of new knowledge. The foundational functions of the library stem from this social dynamic, creating a circular movement that contributes to its own continuity. The main functions considered are as follows:

- Knowledge Storage: Collection development, archival of scientific and technological output, preservation, and conservation;
- Knowledge Organization: Quality of thematic and descriptive treatment that can favor the record exchange among the libraries and its retrieval;
- Knowledge Access: The need for information surpasses its value, location, and form, necessitating access. Consequently, it is essential not only to supply information but also to enable simultaneous access to all of it.

In the context of the information and knowledge society, the digital era presents as the main challenge for professionals of information science the scientific data curation that must remain accessible throughout the information life cycle on the Web of Data and, at the same time, this is a strategic opportunity for digital libraries to amplify the aggregated value to the user of the information through innovations in their processes and services.

In Jim Gray's presentation to the Computer Science and Telecommunications Board [7], he outlined his perspective on the Fourth Paradigm of scientific research [...] data-intensive science encompasses three fundamental activities: capture, curation, and analysis. The development of a set of universal tools that encompass the entire spectrum of activities is imperative—from data capture and validation to data curation, analysis, and permanent archiving.

In the UC3M Ticket to Open Science training activity organized by the Vicerrectorship for Research Policy as part of the HRS4R actions regarding training in Open Science for Early Career Researchers, we could consider the following presented by Prof. Dra. Eva Méndez. Deputy Vice-president for Research Policy. Open Science; Prof. Dr. Tony Hernández. Vice Dean of Information and Digital Content management. Library and Information Science Department.

Open sciences are defined as an inclusive framework that amalgamates diverse movements and practices with the goal of making scientific knowledge openly available, accessible, and reusable for all. This initiative aims to enhance scientific collaborations, facilitate information sharing for the betterment of science and society, and open the processes of scientific knowledge creation, evaluation, and communication to individuals beyond the conventional scientific community.

According to [13], scientific communication is essential for the sharing of research, and the information can be a real instrument for scientific and technological development, especially the publications in widely known journals where the peer review has become an important process to evaluate the quality of the results of scientific work. In the context of the network society, the scientific journals have not lost their relevance and remain the main channel for science communication.

4 Conclusions

Scientific communication leverages the Internet as a potent and efficient technological tool, potentially the most sophisticated information and communication technology currently accessible to society. This is attributed to its organizational structure and its profound impacts on technological, social, economic, and political domains. Additionally, it serves as the essential infrastructure for one of the most extensive and well-known applications, the World Wide Web, which has played a pivotal role in popularizing the Internet to the extent that the two are often conflated in today's world. It is important to note that Internet and Web are distinct concepts.

Knowledge organization links the three processes involved in the strategic use of information—meaning creation, knowledge construction, and decision-making in a continuous cycle of learning and adaptation known as the knowledge cycle. Among the most influential factors, shaping information use is an individual's behaviors toward information and their search for it. These behaviors are shaped by education, training, past experiences, and personal preferences.

It becomes precarious when the information user is perceived either as someone seeking specific and conclusive information or as someone willing to exert effort in searching and exploring databases. The reality is that individuals constantly fluctuate between the desire to extract precise information and the inclination to explore. The utilization of information is a disorderly and chaotic process, susceptible to the whims of human nature, much like any other activity [14].

The Web can be characterized as a segment of the Internet accessible through Web browsers. The repercussions of Internet and Web utilization on society, individuals, and organizations have become subjects of research, extending beyond the specialized domain of applied computing to encompass areas within organizational and sociological studies. Due to its inherently dynamic and boundless nature, both physically and virtually, meticulous attention must be given to details to ensure its unhindered evolution and to enable its availability, reliability, and accessibility for everyone [15].

The Web is a network with contents that are intertwined with one another through hypertext documents. It can be mapped by analyzing and collecting successive pages of content from a set of locations of previously known documents. These searches are conducted automatically through computer programs widely referred to as crawlers, bots, or Web spiders. Even though the mapping is achieved in an automatic manner, the set of locations previously considered to carry the research influences on the result obtained. Furthermore, the Web is not fully intertwined, since there are many subsets of documents with varied sizes, and without any connection to the rest of the network, which are intertwined with each other; that is, they form information islands [16].

Web 1.0 links information, with users primarily in the role of viewers and limited interactivity in the content. Web 2.0, on the other hand, connects people by emphasizing the collaborative construction of knowledge. The Social Web aims to transform users from mere viewers into active collaborators. The Semantic Web, also known as the Web of Data, connects knowledge by utilizing algorithms capable of interpreting user preferences and guiding browsing. This is associated with a set of more efficient technologies designed to assist computers in organizing and analyzing information available on the network. This evolution is anticipated to lead to open-source Web-based applications, fostering interactivity across various areas of the Web.

The digital information universe is expanding within the framework of what the contemporary world terms as big data, where the quest for information becomes increasingly challenging. This is particularly noteworthy considering that the digital universe comprises unstructured data, necessitating organization, indexing, access, appropriation, and utilization to generate new knowledge. The Web of Data aims to streamline the organization of all information, making it more accessible for users. The Semantic search within the Web of Data categorizes information based on specific topics, fostering the connection of knowledge. Additionally, it classifies results into categories, enhancing the speed, organization, and effectiveness of information searches.

Within the context of Library 2.0, it can be asserted that much of the innovation seen in the services offered by libraries during the initial wave of the Web revolution was associated with static services. For instance, online public access catalogs necessitate users to actively search for information, even though many have begun integrating Web 2.0 techniques related to research data. Similarly, the first online library was established using static tutorial texts that failed to meet user needs or facilitate interaction. However, libraries have evolved, adopting a more interactive structure by leveraging social communication media replete with tutorials and more advanced databases.

In Libraries, Web 2.0 can serve as a tool for fostering collective intelligence and knowledge management, enabling interactive discovery. The concept of a library for the user has evolved into a library with the user. In merging and complementing these two paradigms, it is essential to acknowledge, understand, explore, and evaluate the new communication, organization, participation, and collective knowledge construction tools available on the Web. These tools are mostly intuitive, free of charge, and address emerging needs in terms of information and community user engagement. Thus, a noticeable transformation emerges—an innovative approach to the relationship between information and knowledge within the realm of digital libraries and repositories compared to traditional environments that solely facilitated user interactions with physically stored content [17].

Library 3.0 pertains to the library leveraging resources such as the Semantic Web, cloud computing, mobile devices, and information technologies to enhance the dissemination of user-generated content and engagement within collaborative networks. The evolution to Library 3.0 has led to the realization of a "Library without borders," making collections accessible to library users irrespective of their physical location. It serves as a virtual complement to the traditional physical library space [18].

The role of the university is to generate, disseminate, and receive knowledge, and the Library through Information Services becomes aware of its mediating role in terms of scientific communication, taking care of the documentary processes and preserving information to subsequently transform it into knowledge in a spiral of scientific and technological evolution.

In this context, the institution strives to enhance the production of technology, innovation, processes, reflection, public policy, and high-quality professional training. The library serves as an agent that fosters the socialization of knowledge and the generation of new knowledge by developing effective strategies for scientific communication.

References

1. Le Coadic, Y-F (1996) A ciência da informação. Briquet de Lemos, Brasília, DF
2. Comite (2010) Gestor da Internet no Brasil – CGI.br. Dimensões e características da WEB brasileira: um estudo do.gov.br. Access: https://www.ufmg.br/proex/cpinfo/educacao/docs/01b.pdf
3. Weill P, Ross JW (2006) Governança de TI, Tecnologia da Informação. M. Books do Brasil Editora, São Paulo
4. Castells M, Cardoso G (2005) The network society: from knowledge to policy. Washington, DC: Johns Hopkins Center for Transatlantic Relations
5. Danish Roadmap for Research Infrastructure. Danish Agency for Science, Technology, and Innovation. 2011. ISBN: 978-87-92776-20-4. Access: https://ufm.dk/en/publications/2011/files-2011/danish-roadmap-for-research-infrastructure-2011.pdf
6. Gabriel RS et al (2019) J Phys: Mater 2(3). https://doi.org/10.1088/2515-7639/ab084b
7. Gray JJ (2009) Gray on eScience: a transformed scientific method. In: HEY T, Tansley S, Tolle K (ed) The fourth paradigm: data-intensive scientific discovery. Washington: Microsoft Research. Retrieved from: http://languagelog.ldc.upenn.edu/myl/JimGrayOnE-Science.pdf
8. Ferreira VB (2018) E-science. In: E-science e políticas públicas para ciência, tecnologia e inovação no Brasil. Salvador: EDUFBA. p 13–30. ISBN: 978-85-232-1865-2. https://doi.org/10.7476/9788523218652.0003
9. eScience Institute (2020). Retrieved June 27, 2021, from https://escience.washington.edu/
10. Fapesp (2021) Fundação de Amparo à Pesquisa do Estado de São Paulo. Programa FAPESP de Pesquisa em eScience. https://fapesp.br/publicacoes/2015/folder_escience.pdf
11. Renn J, y Hyman M (2021) Survey: the globalization of modern science. The globalization of knowledge in history [en línea]. Max Planck Institute for the History of Science, Berlin, pp 561–604. ISBN 978-3-945561-23-2. Retrieved from: http://edition-open-access.de/studies/1/28/index.html
12. Calzada-Prado F (2021) Identidad digital, difusión y reputación online en LIS [PDF]. Formación transversal Programa de Doctorado Departamento de Biblioteconomía y Documentación Universidad Carlos III de Madrid
13. Freire GHA (2017) A busca da qualidade na comunicação científica. Revista Conhecimento em Ação 2(1):1–1. Retrieved from: http://hdl.handle.net/20.500.11959/brapci/71142

14. Canclini NG (2009) Diferentes, desiguais e desconectados. Editora UFRJ, Rio de Janeiro
15. Russel M (2013) Mining the social web, 2nd edn. O'Reilly, Sebastopol, CA
16. Paletta FC (2019) Biblioteca digital: gestão do ciclo de vida da tecnologia da informação e comunicação TICs. 2019, Anais. Florianópolis: Escola de Comunicações e Artes, Universidade de São Paulo, 2019. Retrieved from: https://www.eca.usp.br/acervo/producao-academica/002976594.pdf
17. Kuhlthau C (2004) Seeking meaning: a process approach to library and information services, 2nd ed. Libraries Unlimited, Westport, Conn
18. Belling A (2010) Exploring library 3.0 and beyond. Access: https://dokumen.tips/documents/exploring-library-30-and-beyond-pdf-investigate-collaborative-purchasing.html?page=1