

Fall 2015

A Brief Essay on the Organization of Knowledge outside Libraries

Samuel, Noah Oluwafemi

Graduate School of Library and Information Science

University of Illinois at Urbana Champaign

IL, USA

Email: nosamue2@illinois.edu

Abstract

It is a given, today, that the internet as a platform has been used to index more documents than the world's libraries put together. Additionally, majority of the documents that were previously stored on the shelves by libraries in the past are now part of these online collections. In this paper, a brief review of early attempts at the organization of knowledge outside of the libraries based on Rayward (1994) was carried out with a view to linking the past to the present. What exactly was the motivation for the establishment of the *office of publicke addressse* by Samuel Hartlib and John Dury, and the *information bureau* suggested by Leibniz? Is there any commercial motivation behind Paul Otlet's Mundaneum? With the understanding that the most capitalized internet service organization (Google) today is involved in reference services similar to those proposed by Hartlib, Dury, and Leibniz, is the end in sight for knowledge organization within the library?

Introduction

Since the inception of the written form, humans have been preoccupied with the need to preserve what is written and to make it available for reuse. This functionality has been performed in various forms even before the first set of the libraries of the ancient world. While several things have changed over the years regarding the way we live and the activities we perform as humans, the basic purpose of the library has remained intact – provision of access to the record of the past. This reference functionality is still the bedrock of the library till date. In this paper, a consideration of the knowledge organization function of the library vis-à-vis the beginning and current status of knowledge organization outside of the library is discussed with a focus on exploring the rationale behind early efforts such as those linked with Samuel Hartlib, John Dury, and the likes of Paul Otlet, and those of the 21st century commercial-driven knowledge organization enhanced by the Internet. For clarity, except when the usage is otherwise explained to mean something different, the term *knowledge organization* is used in this paper to denote a process of organizing knowledge for use rather than referring to an establishment concerned with what may be referred to as *knowledge work*.

Organization of knowledge in libraries

It is appropriate to include albeit a short background of the organization of knowledge in the library in an article that has its focus on the organization of knowledge outside of the library. The library, as one of the leading memory institutions according to Hjerppe (1994), has been the main institution concerned with knowledge organization over the centuries. The development of the library as an institution that it is today has been linked with human civilization efforts. According to Tolzmann (2001, pp1-2), fundamental activities involving knowledge organization in the ancient times included the collection, organization, and preservation of documents recorded on cave walls and on clay tablets both of which were common media through which the experiences of the time were documented. Tolzmann wrote about a number of ancient collections discovered by archaeologists with most of them having significant number of documents that are yet to be fully transcribed.

The Royal Library of Nineveh was reported to be one of the greatest ancient Libraries followed by the Library of Alexandria in Egypt with vast collection of ancient knowledge (Tolzmann, 2001 pp3-5). Fast forward to the early modern period and then to the 21st century, the library as an institution has witnessed tremendous growth in terms of the number of established libraries as well as developments in the way collections and other activities surrounding the institution are managed. It is also interesting to note that while much development in knowledge organization in the ancient world was in the part of the world now known as the Global South (for example Nineveh; an ancient city located in modern day Iraq, and Egypt located in Africa), the development that took place regarding knowledge organization over the Middle Ages and up till the present time have been championed by most countries in the Global North, mainly Europe and Northern America.

While the primary focus of this paper is not on the organization of knowledge in the library as stated above, it is expedient to note some of the works of early scholars who have immensely contributed to making the library the institution that it is today. The research literature is replete with the efforts of scholars such as Samuel Swett Green who founded the American Library Association and who is often regarded as the father of reference, Melvil Dewey, profiled by Miksa (1986) as a prominent figure in the history and the development of the library classification scheme known as Dewey Decimal Classification (DDC), and S. R Ranganathan who worked extensively in the development of the library profession and developed what is today known as the five laws of library science which still remain relevant in the practice of today's librarianship. These early scholars put in place theory that helped and is still helping in the daily activities perform by hundreds of thousands of librarians around the world.

The Colon Classification (CC) scheme developed by Ranganathan, for example, has been said to be the very first form of faceted classification and the majority of website today make use of faceted classification in their online item classification (Broughton, 2001; Uddin and Janecek, 2007). The techniques developed by these early librarians are still as effective as ever and their relevance in the future seems sure. The Dewey Decimal Classification of 1876, for example, is still the world's most widely used library classification system, a signal to the enduring efforts of these early knowledge organizers. In addition to the efforts of Dewey and Ranganathan, Paul Otlet, whose early works in information organization has been described as the foundation for the modern day internet, and Henri La Fontaine, also developed a classification system known as the Universal Decimal Classification (UDC). The UDC, like Dewey's DDC and Ranganathan's CC, is extensively used today in the classification of knowledge across various works of life including many electronic platforms like the modern library platform known as the Online Public Access Catalog (Slavic, 2006).

Early efforts at making knowledge accessible outside libraries

It is important to note that until the advent of computing, some of the major functions associated with the library have been the identification and provision of access to knowledge in the form of physical books. For centuries libraries have been involved with the provision of physical materials that provide content about the knowledge sought. Also, early attempts at making knowledge available outside of libraries predate the Internet age that began in very late 20th century. For example, Rayward (1994) documented much of such attempts including the efforts of Samuel Hartlib and John Dury in England during the Mid-Seventeenth-Century. The duo had proposed the establishment of what was called *the office of publicke addresse* to help with the management of the enormous information produced as a result of massive increase in publication at the time. The establishment of the office was not in the absence of library since Thomas Bodley already established the Bodleian Library as the first public library in England on November 8, 1602 (Norman, 2015). As explained by Rayward, Dury's proposition for *the office of publicke addresse* was a special kind of information bureau with a purpose to achieve a *well-ordered society* by providing information for the *relief of human necessities*.

Apart from Hartlib and Dury, Rayward (1994) also profiled other prominent figures and their attempts at developing various schemes that would help with access to knowledge. Leibniz like Hartlib and Dury in the writing of Rayward had also suggested *a special information bureau to alert people to the existence of*

goods and services for sale or for rent, of educational opportunities, and of things to see and do throughout the state. These proposed functions by Leibniz are obviously familiar with us today as the internet-based companies involved with knowledge organization are either performing all of these functions or a combination of some of them. Leibniz who is known as a polymath was himself a librarian along with being a practitioner in fields like mathematics, logic and others. Rayward also explained that as a librarian, much of Leibniz's concerns were regarding the organization, simplification, and codification of existing knowledge as represented in books. While the explicit motive behind establishing both the office of the *Publicke Adresse* as well as Leibniz's *information bureau* was for the purpose of making knowledge easily accessible to the people, it is unclear if any commercial reason was imagined by any of these early efforts. By inference however, Rayward's description of the function that was to be performed by the *Adresse of Accommodations*; one of the two divisions of Hartlib and Dury's *office of publicke adresse* showed that a number of the services proposed would attract money from the patrons in exchange for the desired service.

Another prominent figure in the writing of Rayward was Paul Otlet whose several efforts regarding documentation and information management have been researched by a couple of historians. Paul Otlet has been reported to be an early driver of the paperless idea having believed that books are nothing more than containers for ideas that could be replaced by such media as graphics and diagrams through which he himself expressed much of his initiatives (Laaff, 2011). In obscurity until 1968 when W. Boyd Rayward discovered parts of his collection, Paul Otlet has since been profiled as the father of the modern internet with his extensive work in information organization over his life time that culminated in the building of what is regarded as the Mundaneum – a center for the world knowledge. Several authors have made a number of remarks about Otlet's early attempt at knowledge organization and his ideas being a springboard for what is today known as the internet (see Laaff, 2011; Popova, n.d; Rayward, 1994; Wright, 2014). Indeed the enormous work done by Otlet over his life time can pass for the efforts of today's multinational organizations like Google at building platforms that ensure that all the digital needs of everyone is met regardless. In appraising the work done by Otlet over his career, Laaff (2011) wrote: *He dreamed of a "mechanical, collective brain" and his complex system for indexing information could be considered an analog version of Google.*

The central reason for most of the early efforts at making knowledge available outside of the libraries like the efforts of Otlet and others have been to achieve world peace, or so it has been reported (Rayward, 1994). In essence, it is believed that if knowledge is organized in such a way that access to it is without any form of restriction, then individuals can learn what they really want to learn and orderliness can be created in the society. Apart from creating orderliness and enhancing access to knowledge, it is unclear if for example someone like Otlet and others with similar initiatives were not driven by commercial reasons while putting together their ideas for knowledge organization. For instance Laaff (2011) wrote that Otlet's Mundaneum received support from the Belgian government until 1934 when the support stopped due to a lost of interest in the project. It will be of interest to know if the Mundaneum was operated as a social institution while it was actively operated by Otlet with support from the Belgian government or its activities all through the early period of its existence were commercial in nature. For example, as reported by Laaff (2011), there were over 1,500 enquiries received in a year by Otlet and his team and the answers to these enquiries were found by hands by sorting through the catalogs which often took weeks for each enquiry. Were these enquiries paid for by the enquirer(s), or was the Mundaneum functioning at the time as a social institution of some sort as the library does? This question is necessary to understand the motivation for knowledge organization outside of the library in the past in relation to the current model used by major institutions involved in knowledge organization outside of the library.

Rethinking the Internet and activities around current Knowledge Organization

Early knowledge organization outside of the libraries indeed started with the objective of making knowledge available to the majority within a given country with scarce account of such efforts envisioning providing global access to knowledge apart from Otlet's Mundaneum that was intended to be

a global knowledge house of some sort. Although the world began to support global trade after the World War II with advancement in air travel and communications that made shipping cost to be considerably low, the truly global nature of international commerce was heralded by the coming of the World Wide Web in the early 1990s (Teichmann, 2004 pp4). While it is true that the Internet has been in existence since the late 1960s as a result of the research conducted at the Defense Advanced Research Projects Agency (DARPA) by Licklider, Roberts and their colleagues, the World Wide Web invention by Tim Berners-Lee and Robert Cailliau eventually established the Internet as a platform for globalization other than just pair to pair communication of the earlier period (Cailliau, 1995; Leiner et al., n.d.). Hence there have been exponential growth in Internet usage heralded by the invention of the World Wide Web with people now being able to connect and use the internet in multiple locations around the world. Just as this development lead to improvements in ways things are done in different spheres of life, it does more improve the traditional process of knowledge organization known in the library world.

The internet started as a decentralized platform with users having control over their online presence and being able to manage their resources locally – for example, early description of the internet shows that every node usually has its own server and its data are stored on the local computer unlike what is currently the case where for example, users' email data are hosted somewhere on a server they do not have control over (Ali, 2015; Kopstein, 2013). Obviously, much of such early individuality has given way to a centralized system as more and more powerful players took over the control of things or rather as more convenient platforms are developed to entice people to the benefits of using centralized systems. Today, it can be said that more than 80 percent of the entire internet traffic is generated by just about 20 percent of established content providers some of which are even bigger than some governments in terms of revenue (Fox, 2013). Google, Facebook, Amazon, Yahoo, Netflix and a number of other digital giants have built massive businesses around online knowledge organization and the trends are spiraling. These digital giants perform a lot more functionality associated with libraries in several advanced way. Their duties just like those of librarians are to ensure that *patrons* as they are called in the library environment or account holders or users as most appropriate in the digital environment, have access to what they need per time – a gesture that will always win the loyalty of such customers. Google for example have developed diverse tools that enhance individual and organizational functionality most of which come at no charge to the users. Facebook, Twitter, Instagram, Flickr and other platforms alike allow users to share memories with family and friends regardless of distance. These platforms are engaging and they have taken an important aspect of the life of most individual today that it may seem an excommunication from the world of the living if an individual is stripped of any or some of these tools.

All these companies are indeed engaged in knowledge organization using the capability of Information and Communication Technologies. Amazon just like the name has built a forest of online platform, offering convenient reference services to customers. Apart from its provision of platform for businesses dealing in, for example, products from electronics to clothing, Amazon has been involved in books not just in relation to sale but also providing platform for writers to self-publish and market their works on its platform. Indeed convenience is the appropriate word that can be used to qualify the efforts of these digital giants – convenient for buyers to purchase products on the platform and convenient for writers to also publish their work for example on Amazon's self-publishing platform. Google is perhaps the number one digital reference provider in cyberspace. From search engine that provides basic search which everyone is familiar with to Google Scholar built for those in the academic and research community to Google Map that helps with location identification to YouTube that houses millions of videos, etc, this digital giant is currently leading the pack in information organization including its digitization efforts and similar book projects like that of Amazon's kindle. While the majority of the reference services provided by a company like Google (except for those provided through play books – a similar platform to Amazon's kindle) come at no apparent cost to the users, for example, users do not have to pay to use the email service, the search engine, watch videos on YouTube (except for premium service), use the map application and so on, the motivation for the development of these services were actually based on the exploitation of users' data. This is where the library and the companies involved in knowledge

organization outside of the library differ. While one of the entrenched attribute of the library is to protect users' data and without any commercial reference for collecting such data, today's digital institution's business model are based on the usage of such data. These data are used to profile individuals in order to know what type of advertisement to push to them as these organizations are driven primarily by advertising revenue in addition to other sources of income (Grunes, 2015; Rosenberg, 2015).

With the development and improvement in technologies, there is every likelihood that knowledge organization outside of the library is going to continue on the rise with projection that the majority of the established companies will continue to build platforms and applications that will help them provide more and more convenience for users of their platforms and also an increase in new companies wanting to provide similar services. As more of these companies begin to discover the power in providing free access to platforms and applications – that is having access to users' data that can in turn be used to win advertising dollars, it is predicted that privacy issues will be on the increase as users willingly give these providers the authority to use their data often without reading most of the privacy agreements in detail. Therefore, while the advancement in technology that lead to the improvement in the way information is organized outside of the library is something to be proud of at this time because of the various conveniences that come with it, it is also expedient to begin to examine the flip side as more and more personal data are turned in to these digital entities. Apart from concerns attached to individuals and how their data are used by these companies, another concern will be the future of library itself as more and more traditional functions of the library are now being performed on a number of these digital platforms. For example, books are being digitized for easy online access by Google and more people today are now renting books on Amazon kindle, and reading more books on other digital media like Play Books and Kobo. Perhaps these innovations may be a signal that will herald the library of the future as more libraries are now turning to digital tools for the management of their services with more and more reference work being done today over the internet especially among academic libraries.

Apart from the stories of Google and its mega friends, there are other projects stretching the mind regarding the future of knowledge organization outside of the library. Stephen Wolfram and his team at Wolfram Research Company have been engaged over the years in the extensive development of Wolfram Alpha – a computational knowledge engine that is built to provide answers to any query issued. Wolfram's knowledge engine is able to provide computational answers to queries from mathematics to people and history, physics, health and medicine, earth sciences, culture and media, music, astronomy, and several other bodies of knowledge. And about this knowledge engine called Wolfram Alpha, the goal is *to collect and curate all objective data; implement every known model, method, and algorithm; and make it possible to compute whatever can be computed about anything* (Wolfram Alpha LLC., n.d.). The world is sure in an exciting period of innovation that will see more and more inventions regarding knowledge organization outside of the libraries for easier access – indeed easier with just a click of a mouse.

Obviously, a few concerns that may arise are the fact that more and more knowledge that were previously accessible for free through the institution of the library will increasingly become monetized as the digital players assume the roles of knowledge organizers. While the library as it were provide the place and access to the resources including the reference services for free as a social organization, the companies now assuming this role are business concerns and will be expected to fashion out the business rationale for their investments. Another issue that comes with knowledge organization outside of the library is the fact that the more and more technology continue to make the process easier, the more information is organized for use and the more we continue to have access at a click of the mouse, and the more likely we are to continue to have information overload. At the moment, there are obviously more knowledge indexed on most of these platforms on certain subject that could be consulted by an individual per time. While availability is good, it sometimes also create a problem for choice – we would take more time to decide on what content to use and what not to. One other fundamental issue is the fact that majority of the work of organizing knowledge outside of the library is also a function concentrated in the hands of a few

global companies and the last concern just like it was briefly discussed earlier has to do with issues concerning the privacy of users' data.

In spite of some of the above concerns, knowledge organization outside of the library has tremendous advantages. Without having to step outside of one's apartment as it were, one can have access to millions of collections which is more like the *armchair's* philosophy of Paul Otlet according to Popova (n.d.). Once there is internet connectivity, location is no longer a barrier as access to any resource stored on any online platform can be made available regardless of where an individual is based per time as oppose to the fact that access to resources in the library has to be based on the patron being physically present. While most libraries have what is known as inter-library loans that allow access to resources from other libraries should any required material not be present in a local library, the global nature of the internet as a platform through which knowledge organization outside of the library is made possible allow users access to knowledge regardless of physical or location restraint and to several pool of providers ranging from multinationals to a server hosted by an individual in a remote village.

Conclusion

This paper set out to assess the concept of knowledge organization outside of the library with a look to the past as well as current initiatives. To build a background for the discussion, a brief historical review of knowledge organization in the library was carried out with a consideration of the works of some prominent librarians. Some of the librarians who have made invaluable contributions to the field of librarianship and, most importantly, to the core activities of knowledge organization were, for example, Melvil Dewey who left us with the Dewey Decimal Classification which is still being used in thousands of libraries and other bibliographic organizations around the world till date and S.R. Ranganathan whose Colon Classification is regarded as the first set of faceted classification that forms the basis for most classification efforts today including being used on several website and digital platforms. Paul Otlet, who has been regarded as the father of the modern day internet, also worked with Henri La Fontaine to build what is known as the Universal Decimal Classification that also has application in various libraries and bibliographic establishments around the world today.

Regarding the early efforts at knowledge organization outside of the library, this paper has briefly reviewed, based on Rayward (1994), the works of Samuel Hartlib and John Dury, who during the Mid-Seventeenth Century proposed an institution for information management with functionalities similar to what is done by institutions involved in information related activities today. Similar to Dury and Hartlib, Paul Otlet was also an information scientist who, more than developing a standard for knowledge organization, had also work to establish the Mundaneum with a mission to house all the world knowledge and every knowledge-related resources, the idea that have been widely acclaimed as a conceptual foundation for today's World Wide Web. Concerning the present, this paper has reviewed the current knowledge organization efforts by digital giants like Google, Amazon, Facebook, Netflix and a number of them who are involved in reference works much like the library but with the capability of information and communication technologies. The advantages provided by these platforms have been examined as well as some concerns relating to users' privacy, as well as the need for proactive measures on the side of libraries in other to stay relevant in the face of digital innovations. In all, it is projected that more and more platforms will emerge that will make knowledge organization outside of the library more prevalent and more appealing to the users as innovation in technology continues. For libraries to continue to reckon as the first stop for knowledge seekers, therefore, there will be a need for libraries to leverage the technologies that are making knowledge organization outside of the library more appealing to knowledge seekers.

References

- Ali, M. (2015, March 3). *Evolution of the Internet: from Decentralized to Centralized*. Retrieved December 12, 2015, from <http://blog.onename.com/evolution-of-the-internet/>
- Broughton, V. (2001). Faceted classification as a basis for knowledge organization in a digital environment; the Bliss Bibliographic Classification as a model for vocabulary management and the creation of multidimensional knowledge structures. *New Review of Hypermedia and Multimedia*, 7(1), 67-102.
- Cailliau, R. (1995, November 2). *A Short History of the Web*. Retrieved December 10, 2015, from Netvalley: http://www.netvalley.com/archives/mirrors/robert_cailliau_speech.htm
- Fox, Z. (2013, October 28). *The 10 Companies that Own the Internet*. Retrieved December 12, 2015, from mashable.com: <http://mashable.com/2013/10/28/google-monthly-traffic/#B59o1S0pukqx>
- Grunes, A. (2015, February 26). *Google's quiet dominance over the 'Ad Tech' Industry*. Retrieved December 12, 2015, from Forbes: <http://www.forbes.com/sites/realspin/2015/02/26/googles-quiet-dominance-over-the-ad-tech-industry/>
- Hjerpe, R. (1994). A Framework for the Description of Generalized Documents. *Advances in Knowledge Organization*, 4, 173-180.
- Kopstein, J. (2013, December 12). *The Mission to Decentralize the Internet*. Retrieved December 12, 2015, from <http://www.newyorker.com/tech/elements/the-mission-to-decentralize-the-internet>
- Laaff, M. (2011, July 22). *Internet Visionary Paul Otlet: Networked Knowledge, Decades Before Google*. Retrieved December 9, 2015, from SPIEGEL ONLINE INTERNATIONAL: <http://www.spiegel.de/international/world/internet-visionary-paul-otlet-networked-knowledge-decades-before-google-a-775951.html>
- Leiner, B. M., Cerf, V. G., Clark, D. D., Kahn, R. E., Kleinrock, L., Lynch, D. C., . . . Wolff, S. (n.d.). *Brief History of the Internet*. Retrieved December 10, 2015, from Internet Society: <http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet>
- Miksa, F. (1986). Melvil Dewey: the professional educator and his heirs. *Library Trends*, 34(Winter 1986), 359-382.
- Norman, J. (2015, May 27). *Thomas Bodley Founds The Bodleian, the First "Public" Library in England, & the First British National Library (November 8, 1602 – 1605)*. Retrieved December 7, 2015, from historyofinformation.com: <http://www.historyofinformation.com/expanded.php?id=44>
- Popova, M. (n.d.). *The Birth of the Information Age: How Paul Otlet's Vision for Cataloging and Connecting Humanity Shaped Our World*. Retrieved December 9, 2015, from brainpickings.org: <https://www.brainpickings.org/2014/06/09/paul-otlet-alex-wright/>
- Rayward, W. B. (1994). Some schemes for restructuring and mobilising information in documents: A historical perspective. *Information processing & Management*, 30(2), 163-175.
- Rosenberg, E. (2015, November 9). *The Business of Google*. Retrieved December 12, 2015, from Investopedia: <http://www.investopedia.com/articles/investing/020515/business-google.asp>
- Slavic, A. (2006). The level of exploitation of Universal Decimal Classification in library OPACs: A pilot Study. *Vjesnik bibliotekara Hrvatske*, 49(3-4): 155 – 182
- Teichmann, I. (2004). *Globalization*. North Mancato: Smart Apple Media.
- Tolzmann, D. H. (2001). *The Memory of Mankind: The Story of Libraries Since the Dawn of History*. New Castle: Oak Knoll Press.
- Uddin, M. N., & Janecek, P. (2007). Faceted classification in web information architecture: A framework for using semantic web tools. *The Electronic Library*, 25(2), 219 - 233.
- Wolfram Alpha LLC. (n.d.). *Wolfram Alpha: A computational knowledge engine*. Retrieved December 11, 2015, from <http://www.wolframalpha.com/>
- Wright, A. (2014). *Cataloging the World: Paul Otlet and the Birth of the Information Age*. Retrieved December 9, 2015, from [catalogingtheworld.com: http://www.catalogingtheworld.com/](http://www.catalogingtheworld.com/)