Dialogical elements in Harris, Dewey, Cutter, Otlet, Kaiser, and Ranganathan: Theoretical convergences in the history of Knowledge Organization

Elementos de dialogicidade entre Harris, Dewey, Cutter, Otlet, Kaiser e Ranganathan: um recorte teórico na história da Organização do Conhecimento

Rodrigo de SALES¹ 0000-0002-8695-9807
Daniel MARTÍNEZ-ÁVILA² 0000-0003-2236-553X
José Augusto GUIMARÃES² 0000-0002-0310-2331

Abstract

In this paper, we study the theoretical intersections and dialogues between some foundational authors on classification and indexing of the nineteenth and twentieth centuries that helped developing the theoretical-methodological framework of knowledge organization. More specifically, we highlight and analyze the theoretical convergences of Harris, Dewey, Cutter, Otlet, Kaiser, and Ranganathan as they can provide a clearer picture of the historical and theoretical contributions to the epistemological foundations of knowledge organization. Our methodology follows a critical-descriptive approach to the analysis of the main contributions of the authors and the critical reflections of some specialists and biographers. We continue with a discussion of the links between bibliographic classifications and knowledge organization drawing on the ideas of Bliss; then, we divide our historical narrative between the theoretical contributions during the nineteenth-century (Harris, Dewey, and Cutter) and the twentieth century (Otlet, Kaiser, and Ranganathan); and finally, we present a discussion of the history of knowledge organization from the point of view of the theoretical and methodological development of classification and indexing at the turn of the nineteenth century to the twentieth century. We conclude with some remarks on their main contributions to the development of the knowledge organization field.

Keywords: Classification. Indexing. Knowledge Organization.

Resumo

Neste artigo, são examinados diálogos e interseções teóricas entre alguns autores fundamentais para os estudos de classificação e indexação nos séculos XIX e XX, abordando estudos que ajudaram a desenvolver o arcabouço teórico e metodológico da organização do conhecimento. Mais especificamente, são destacadas e analisadas as convergências entre Harris, Dewey, Cutter, Otlet, Kaiser e Ranganathan, de modo a conceber um quadro esclarecedor de suas contribuições teóricas e históricas para a fundamentação epistemológica da organização do conhecimento. Do ponto de vista metodológico, adotou-se uma abordagem crítico-descritiva para analisar as principais

¹Universidade Federal de Santa Catarina, Centro de Ciências da Educação, Departamento de Ciência da Informação. Florianópolis, SC, Brasil.
²Universidade Estadual Paulista Júlio de Mesquita Filho, Faculdade de Filosofia e Ciências, Departamento de Ciência da Informação. Av. Hyginio Muzzi Filho, 737, Mirante, 17525-900, Marília, SP, Brasil. Correspondência para/Correspondence to: D. MARTÍNEZ-ÁVILA. E-mail: <dmartinezavila@marilia.unesp.br>.

Received on March 23, 2018, final version resubmitted on May 22, 2018, and approved on June 27, 2018.

Como citar este artigo/How to cite this article

http://dx.doi.org/10.1590/2318-08892018000300007
The discipline of knowledge organization is intrinsically linked to classification systems – here understood as systems that provide an organization of concepts and subjects – and how they contribute to the epistemological configuration of the field (Dahlberg, 1993). In addition to classification, we consider that indexing is another great contributor to the foundations of knowledge organization, as its intellectual efforts and philosophy seem to be very close to classification in relation to the organization of knowledge. However, the dialectics of the areas of classification and indexing as well as their main contributors do not seem to be explicit in the historical narrative of Knowledge Organization in the 19th and 20th centuries. In this sense, the research question of this paper is: What are the dialogical elements and theoretical convergences in Harris, Dewey, Cutter, Otlet, Kaiser, and Ranganathan in relation to the development of the Knowledge Organization field? Our historical approach aims to highlight the theoretical intersections between some of the most influential authors on classification and indexing of the 19th and 20th centuries, namely Harris, Dewey, Cutter, Otlet, Kaiser, and Ranganathan, that helped to develop the theoretical-methodological framework of knowledge organization. The aim of the present paper is to highlight the theoretical convergences of these authors in order to provide a clearer picture of the historical and theoretical contributions to the epistemological foundations of knowledge organization.

Our methodology follows a critical-descriptive approach to the analysis of the main contributions of the aforementioned authors and the critical reflections of some of their biographers and experts (Schneider, 1934; Shera; Egan, 1937; Leidecker, 1946; Mills, 1960; La Montagne, 1961; Foskett, 1969; Comaromi, 1976; Hunter; Bakewell, 1983; Wiegand, 1996, 1998; Dousa, 2010a, 2010b; Sales, 2014, among others). The analysis of the “dialogicity” between these authors is based on the structural bases and the theoretical-methodological principles that they proposed, such as the main classes in Harris and Dewey systems, the principles for the development of subject headings in Cutter, and the analytic-synthetic processes in Otlet, Kaiser, and Ranganathan.

Classification and Knowledge Organization: Some remarks

Dahlberg (2006) states that when she founded the International Society for Knowledge Organization (ISKO) she adopted the term “knowledge organization” inspired by the term “organization of knowledge,” used by Henry Evelyn Bliss in his works “The Organization of Knowledge and the System of the Sciences” (1929) and “The Organization of Knowledge in Libraries” (1933). It should be noted that the term “knowledge organization” had already been used by Paul Otlet in 1903 in a paper entitled “Les sciences bibliographiques et la documentation” (translated by Boyd Rayward in 1990 as “The science of bibliography and documentation”). However, the knowledge organization referred by Otlet was about the external environment of science, i.e., the institutions, associations, journals, and events that, in a broader sense, organize knowledge. Knowledge organization, as it is understood today, for instance in the context of the aforementioned ISKO, would be closer to Bliss’s “knowledge organization” than to Otlet’s “knowledge organization.”
On the other hand, when we talk about knowledge organization in the context of bibliographic classifications we can also refer to Shera and his sociological approach to epistemology. When Shera said “order is heaven’s first law, not only in the world and in the universe but in our minds as well” (1976, p.69), it is clear that the organization of a library was not historically designed to follow the individualistic arbitrariness of the mind of the librarian. In practice, classification is also an essential part of Library Science. Traditionally, classification has been considered a means to access to the books arranged in the shelves. However, due to the rapid growth of the bibliographic production in the nineteenth and twentieth centuries, the design of classification schemes also adapted to this growth and demand, moving from an organization of subjects in books to an organization of knowledge. This concern, which Shera addressed in the 1970s, had its origins in a tradition of American scholars and librarians discussing the quick and efficient access to the books as a question of the order expressed in the classification system. The works of Ansteinsson (1934) and MacPherson (1939) in the 1930s are examples of this concern.

At the beginning of the twentieth century, the problem of library classifications was far from being solved. In this context, Bliss, in a very short paper of only three pages (first published in 1917), noted the lack of guiding principles and theory to establish principles about facts that are previously understood when faced with unsolved problems, drawing attention to the need to reformulate old methods and principles, namely “the correlation of class to concepts”, “relativity of classes”, and “economy of classification with expansibility and adaptation” (Bliss, 1917, p.201). Based on his views on the “natural order” of things, Bliss found his formative correlates in the systemic view. In a step further, it is possible to perceive in Bliss's formulation what the author calls the dominant tendency of the use of the term “organization”. The concept, hitherto referred to by the word “classification”, came to be denoted by the term “organization” in any instance in which disorganization could be considered disastrous. Therefore, it was argued that the universe of libraries should be better classified by putting a greater emphasis in the divisions and subdivisions of knowledge related to science and industry.

Bliss created his Bibliographic Classification based on a subject approach to knowledge. However, his classification system was neglected in favor of other more popular classification systems, such as the Dewey Decimal Classification and the Library of Congress Classification. John Dewey stated in the introduction of “The Organization of Knowledge and the System of Sciences” (Bliss, 1929) that Bliss’s work was permeated by a philosophy that went from the universal to the particular, from theory to practice, imposed by growth and change. The book was valuable for its academicism in the demonstration of the place that the modern library fills in the intersection of intellectual integration and practical applications, something that highlighted the strategic role of libraries. The correct organization of a library entails the combination of knowledge and experience, while providing support for the development of knowledge itself.

In the preface to the book, Bliss also questioned the meaning of organization of knowledge and proposed that, in a broad sense, it “comprehends not only the mental process, the development of concepts and the conceptual synthesis of knowledge, but also the intellectual correlation and systemization of valid knowledge, from the simple social synthesis to the more complex conceptual systems of Science and philosophy” (Bliss, 1929, p.xi). Bliss states that, one way or another, classification would be inherent to all these processes and methods. Classification is fundamental to the organization of knowledge, and that is why its principles, methods, and ways should be of interest to the community. Bliss also argues that classification is based on consensus, although it constantly changes. Thus, it is said that classification is based on two basic principles: the subordination of the more specific subjects to the more general ones, and the allocation of subjects together or nearby for reasons of convenience, allowing for maximum efficiency.

According to Bliss's second book (1933), a classification for libraries should be necessarily based on a systematic arrangement of subjects. He stated that alphabetical arrangements (such as those based on authors' names), although less complex in practice, were not appropriate for the requirements of a library collection and much less for the organization of catalogs. Bliss, as well as Cutter, thought that a classification for libraries should be first of
all a classification of subjects, i.e., it should be based on the logical relationships of the topics defined whether by philosophical classifications or scientific classifications. For Bliss, the useful life of a classification was directly linked to the stability of the relationships between subjects that were established by consensus. In this sense, and making clear the fact that the problem of classification in libraries should not be confined to indexes and classification notations (as these only serve as auxiliary devices), Bliss addressed the practical problems of catalogers, classifiers, librarians, and bibliographers in general. This book intended to provide a rationale for the study of these problems in relation to classification. In this sense, Bliss (1933) elucidated the problems of classification systems, ranging from the inherent aspects of the cataloging of subjects to the questions that affected bibliography, also including the criticisms to the history of library classifications.

Thus, it seems that the main concern of Bliss during his career was the necessity of providing guiding principles for the work of librarians and more specifically classifiers. Although he commonly talked about organization in a systemic perspective, Bliss mainly proposed the establishment of guidelines for a better organization of libraries in relation to their ultimate end – knowledge. In fact, Bliss's reflection on the organization of knowledge was manifested in his reflection on the organization of libraries, and more specifically on bibliographic classifications. In this sense, and since the knowledge organization field (at least in the context of ISKO) was originally inspired by the work of Bliss's, we highlight the inseparable relationship between the principles of knowledge organization and the principles of bibliographic classifications. For this reason, we aim to contribute to the discussion on the history of knowledge through the historical development of classificatory models.

19th Century: Subject description with emphasis in Classification

Although library classifications had already been developed and used in the United States and Europe in the 17th and 18th centuries, it was only in the 19th century that library classifications gained space and notoriety to the point of achieving more solid work dynamics and dialogues with other areas. In the United States, for example, William Torrey Harris, Melvil Dewey, and Charles Ammi Cutter made important efforts in the systematic organization of the subjects of books in libraries during this period.

In 1870, the North American philosopher and educator William Torrey Harris (1835-1909), responsible for the St. Louis Public School Library Network, developed a bibliographic classification that, based on Hegel's philosophy, reversed the logic defined by Francis Bacon in the seventeenth century. The Baconian influence in the classification developed for the St. Louis libraries was recognized by Harris himself (1870) and also highlighted by several authors (e.g., Leidecker, 1946; Eaton, 1959; Foskett, 1969; Olson, 2011). Wiegand (1996, 1998) also understood that this inversion of the Baconian classes had already been done by Hegel in his idealistic dialectics. On the other hand, other authors such as La Montagne (1961) claimed that the first person to adopt an inversion of Bacon's classes was Edward William Johnston for the College of South Carolina in 1836, while Graziano (1959) states that Harris's classification was fundamentally influenced by Hegel's idealism alone, barely reflecting the empiricist perspective of Bacon in the development of the subjects.

Regardless of the different points of view in the literature, it is undeniable that Harris's classification was a great influence for later bibliographic classifications. It also revived the studies on classification dialogues with the philosophical bases of classifications, notably, with the works of Bacon and Hegel. In other words, Harris combined his philosophical basis with the practical concerns of the classificatory universe of libraries. Understanding that every classification system should be based on a philosophical logic, Harris found his initial rationale in Bacon. For Harris (1870), although Bacon was not concerned with the classification of books, he eventually presented a principle of development of human knowledge that might be helpful for the fundamental distinction used for classifying knowledge in libraries. This principle in Bacon was based on the different mental faculties – Memory, Imagination,
and Reason – that correspond respectively to the areas of History, Poetry, and Philosophy. Influenced by Hegelian dialectics, Harris (1870) seems to have kept only the aspect of the Baconian narrative in which he was interested in – the formative idea of knowledge based on the mental faculties. Then, he followed Hegel, giving priority to Philosophy as the area that provides the archetypes of all other areas. In this sense, Harris reversed the Baconian logic and began his classification system with Reason (Science), ordering knowledge as follows: Science/Philosophy, Arts, and History. Regardless of the universal pretensions of any classification system, the influence of contemporary philosophy is also inevitable. As Harris (1870) pointed out, in Bacon’s time (during the 16th and 17th centuries), History was the most developed form of knowledge and had well-established developments; whereas in his own time, during the 19th century, Science reached a more defined structure and advanced developments.

Hegel’s Logic, that considers three categories of existence – Being, Essence, and Idea – was explained by Harris in “Hegel’s logic: A book on the genesis of the categories of the mind: A critical exposition” (1890). In this book, Harris understands Being as an illusory condition, Essence as a “condition of,” and Idea as a superior category. Here Harris correlates History with Being as illusory observations, Arts with Essence as disillusioned mediations, and Science, Philosophy, and Religion with Idea, as a category that transcends both illusion and mediation (Olson, 2011). We can say that Harris borrowed from Bacon the formative idea of a knowledge based on mental faculties, whereas he borrowed from Hegel the focus on the Idea and the modelling of Reason by Science, making his classification to be based on an idealistic dialectics. It must be emphasized that both Bacon and Hegel, although two philosophers with different stances (empiricism and idealism respectively), believed that it was possible to apprehend something (a form) that would be able to define both the structure and the dynamicity of knowledge. That is, they both pursued a method for the understanding of knowledge. Universality united both philosophers in the belief that classification has something natural and not artificial, something capable of reflecting or representing reality as it is in a rational way. This view, contrary to the recognition of artificiality and arbitrariness of classification, would influence Harris in the development of an essentialist, idealistic classification. This perspective, formalized in Harris’s classification, was expanded in the Dewey Decimal Classification, a system that could be considered the concretization of what we are calling here the period of subject description. In this vein, it is well accepted that Harris’s classification was fundamental for the design of the Dewey Decimal Classification (Barbosa, 1969; Foskett, 1969; Comaromi, 1976; Piedade, 1983; Wiegand, 1996; Olson, 2011). As pointed out by Comaromi (1976), Wiegand (1996, 1998), and Olson (2011), the development of the Dewey Decimal Classification main classes was based on the structure of Harris’s classification for the St. Louis Public School Library. Subclasses and sections of the Dewey Decimal Classification were based on the structures of the courses and readings used by teachers at Amherst College between 1870 and 1875. Wiegand even includes excerpts from a letter dated May 9, 1873, in which Dewey writes to Harris asking for more information regarding the arrangement of the classification at St. Louis Library.

In this period, the academic and professional environment in which Harris was immersed was heavily influenced by Hegelian idealism. This influence was also present among Amherst College scholars, including Dewey (Wiegand 1996; Olson 2011). According to Harris’s biographer, Leidecker (1946), Harris and Dewey were very close in 1873, united not only by their interest in library classifications but also by their interest in the English spelling reform in the United States. Based on the main classes of Harris’s system, Dewey (1851-1931) defined the ten main classes of his classification as follows: Generalities, Philosophy, Religion, Social Sciences, Languages, Pure Sciences, Applied Sciences, Arts, Literature, and History/Geography/Biography. The convergence between Dewey’s classes and Harris’s classes is evident: both systems begin with knowledge related to reason and scienticity, continue with knowledge related to the creative arts, and finish with records of memory. Thus, as Olson (2011) reports, Hegelian philosophy is also present in Dewey’s classification, whose original values can be found in several aspects, and yet not in the definition of its main classes.

Dewey’s main contribution to classification is the definition of his system in ten major subject classes, the subdivision of each major class into ten subclasses, and of each subclass into ten more sections, providing a systematic
arrangement of gradually more specific subjects. The pragmatic representation of these subjects using numerical notations based on decimal boxes gave the Dewey system several advantages, such as a relative location – before Dewey, books in libraries only had predetermined physical locations as the coding was assigned to the bookshelves, not the books – and a detailed specification of subjects – the subjects of the books began to be considered in a greater detail as they had a greater mobility within the collections and also notational mechanisms that allowed the specifications of subjects and locations (Foskett, 1969). Using a decimal numerical code, Dewey provided a clear and objective notational representation in which the hierarchy of the subjects was reflected in the hierarchy of the numbers that represent them while the decimal notation provided mnemonics. The first edition of the Dewey Decimal Classification was published anonymously in 1876 under the title, “A Classification and Subject Index for Cataloguing and Arranging of Books and Pamphlets of a Library”.

Charles Ammi Cutter (1837-1903) developed his Expansive Classification (published between 1891 and 1893) as a criticism to Dewey’s decimal notation. Cutter considered that this type of notation did not allow a detailed classification according to the diverse necessities of libraries. Cutter divided his own proposal of classification into seven levels of complexity that went from the most general one to the most detailed one. Aiming at a classification system that could be applied in libraries and collections of varying sizes, from city libraries to national libraries, Cutter called his system “Expansive Classification” due to the possibility of the system to expand according to the growth of the collections. His idea was to apply the first level of the system in the beginnings of the collection, and applying more elaborate classificatory levels as the collection grows. The seventh level of Cutter’s classification (the most detailed one) remained unfinished due to the author’s death, and also due the death of his nephew and biographer William Parker Cutter, who attempted to continue the Expansive Classification.

During his life, Cutter considered his classification as an evolutionary system that followed the ideas of Natural History, as each subject or part of a subject was allocated in an order that respected the manifestation of the respective subject (and its theory) in nature (Piedade, 1983). For example, in Zoology, Cutter’s classification goes from protozoa to primates. The Expansive Classification was a direct influence in the development of the Library of Congress Classification (LCC), especially in the definition of the main classes that partially follow the order of Cutter’s classification (Mills, 1960; Barbosa, 1969; Foskett, 1969; Maltby, 1975; Piedade, 1983; Coates, 1988). The Library of Congress Classification is the most important utilitarian classification reported in the literature, that is, it is a classification that is developed based on the needs of the collection of the Library of Congress and without any consistent scientific or philosophical basis in relation to the chain of subjects and logical organization. Cutter’s expansive nature in relation to the growth of the collection was kept in the LCC.

On the other hand, Cutter’s greatest contribution in relation to our research question was his “Rules for a Dictionary Catalog,” published in 1876. “Rules for a Dictionary Catalog” was the title Cutter chose because, as he put it, “headings (author, title, subject, and form) are arranged, like the words in a dictionary, in alphabetical order” (Cutter, 1904). The cataloging of a document consists of two activities, the descriptive cataloging - which deals with the formal representation of the document - and the subject cataloging - that deals with the representation of the content conveyed by the documents to be used as access points. Authors such as Shera and Egan (1937) and Taylor (1995) have stressed that the location of documents using subject access points was already the most common type of search among library users back in Cutter’s time. Regarding the standardization of the subject cataloging of documents, Cutter defined rules for the development of subject headings based on three basic principles that can be summarized as follows (Cesarino; Pinto, 1978, p.274): 3

3 Translated by the authors.
- Specific principle: the subject entry must be the most specific term and not the class to which it is subordinated. (Although this seems obvious, this principle made a great impact because at the time it was common to adopt general headings that did not represent the subject of the work that was being cataloged);

- Use principle: the headings must be those most likely to be searched by educated Americans, with cross-references to other forms of related headings. (This is also a principle of convenience, related to the user’s needs);

- Syndetic principle: the alphabetic order of subject headings makes absurd approximations of subjects and at the same time separates related subjects. Thus, Cutter proposed syndetic structures in the development of subject headings to overcome this problem using well-structured networks of cross-references.

In summary, the second half of the 19th century was characterized by a focus on the description of the subjects of the documents in libraries. While Harris devised a system for the St. Louis public school collections that described subjects based on Hegel’s idealism, Dewey adapted such a system into a classification that would suit any library. The predetermination and development of subject classes in Harris’s and Dewey’s schemes allowed for more than a preconceived description of the subjects of the books. The same idea applies to Cutter’s Expansive Classification. These predefined classificatory schemes were called descriptive schemes by Ranganathan. They were tools that formalized the hierarchies and the relations of the subjects in a predetermined way based on the knowledge within the literature, meaning that subjects were described according to what had been agreed at a certain historical moment. This was the case of Harris’s classification and Dewey’s Decimal Classification, that followed Hegel’s philosophy, and Cutter’s Expansive Classification, that followed an evolutionary model, all of them were used for the description of subjects. The intersection between Harris and Dewey was the Hegelian basis of their classifications, unlike Cutter who opted for natural history. However, the meeting point between these three authors lies in the descriptive stance of their classifications, as all of them, in the end, are mere ready and predefined tools. As in every case, the emphasis of knowledge organization for libraries during the second half of the 19th century, especially in the United States, was put on the descriptive classification of subjects.

However, if the main emphasis of this period was given to the description of subjects by bibliographic classifications, it could also be argued that Cutter was a transitional character as well who, in addition to elaborating a descriptive system of subjects, was also concerned with the standardization of another fundamental aspect of knowledge organization: the verbal representation of subjects in catalogs. When Cutter defined principles to guide the activities of the cataloguer, he also began a movement that was very alive during the 20th century: the standardizing of the construction of subject statements. This aspect makes Cutter “dialogue” with another foundational author of the field: Julius Kaiser.

20th Century: Subject analysis with emphasis in Indexing

In the 20th century, there was a rise of studies on the subject approach to information (Foskett, 1969), especially regarding the tools for verbal representation that were developed in the English tradition of indexing (Guimarães, 2009). Indeed, it is in this context of verbal representation where a very important connection (apparently unstated) between Cutter’s subject cataloging and Kaiser’s systematic indexing exists.

Julius Otto Kaiser (1868–1927), a German librarian naturalized as a US citizen, had a professional career that also included specific information contexts such as museums, corporate archives, and factories in the United Kingdom and the United States. It could be said that his main concern was the knowledge organization of office documents,
that appeared in the most varied forms and conveyed the most varied subjects in the universe of specialization. During this period, approximately from 1896 to 1927, Kaiser developed and refined a method for systematic indexing that used the logic of categorization to guide the verbal representation of subjects (Rodriguez, 1984; Coates, 1988; Dousa, 2010a, 2010b; Sales, 2014; Sales; Guimarães, 2016). Kaiser’s work on systematic indexing was published in two volumes: “The Card System at the Office” (1908) and “Systematic Indexing” (1911).

In his work, Kaiser defined that subjects should be analyzed and synthesized on the basis of two fundamental categories: concretes and processes. Concretes referred to the most substantial aspects in the subject, that is, the commodities, what the document is about. Processes, on the other hand, refer to the actions that affect the concretes. For example, in the sentence “the Samsung Galaxy is sold in Brazil,” we could consider “Samsung Galaxy” as a concrete and “is sold” as a process. Kaiser complemented the category concretes with the category countries. Thus, in the above sentence, we could analyze the information as follows: Samsung Galaxy - concrete; is sold - process; Brazil - country. Kaiser’s idea was that the indexing process could be driven by the identification of these categories in the subjects.

A fundamental part of Kaiser’s systematic indexing was the definition of six rules for the creation of statements (Kaiser, 1911, p.348):

- Select that which is of real value for your purpose regardless of form or extent.
- Concentrate on the relatively specific information.
- Deal with each item absolutely independently.
- Do not tamper with names of concretes.
- Avoid inversions, prepositions, and plurals where possible.
- Test the accuracy of each statement both from the standpoint of the indexer and of the consulter of the index.

If we return to the principles defined by Cutter, we clearly find a correlation between those principles and the rules established by Kaiser. We observe that Kaiser not only resembles Cutter in the initiative of establishing rules for the creation of indexing terms, but especially in the foci approached. Cutter’s “specific principle,” in spite of being questioned by Kaiser (how specific is specific?) is addressed in the points on the statement. While Cutter stated that the subject entry should be the most specific term and not the class to which they are subordinated, Kaiser established that the focus of indexing should be relatively specific information. Although both authors were concerned with the specificity of the term used for the representation of subjects, Kaiser went a step further while discussing the need for providing references from specific subjects to broad subjects, a feature that was not available in Cutter’s dictionary catalog (Foskett, 1969).

As Mills (1960) reported, terms were related in Kaiser’s references using guiding cards that also specified, in addition to the relationships between broader and narrower terms, synonymous terms and any other term/concrete that was considered to be related at the moment of indexing. It should be noted that these relations among terms in the systematic indexing are a continuation of the syndetic principle that Cutter had proposed. This focus on the types of relationships between terms, originated in Cutter’s dictionary catalog with the cross-references and also continued in Kaiser’s system, was the beginning of a concern that characterizes the studies on the subject approach to information that were later developed by the English school of indexing.

In addition, Kaiser also addressed the problem presented by Cutter in his “use principle.” Both systems were not only oriented to the success of representation (in cataloging and indexing) but also to meeting the user’s needs. Here, we observe in Kaiser a continuation of the philosophy that Cutter had initiated. While Cutter presented a more deductive stance, stating that the heading should be chosen following a deduction of the terms that are likely to be
searched by the user, Kaiser proposed to extract the accuracy of the statements directly from the user. Although it is known that the representation of subjects is not a task that is free of subjectivities, it should be noted that both Cutter and Kaiser sought to minimize this factor while determining principles and rules that would provide a necessary standardization for the activities of cataloging and indexing.

If a dialogue between Cutter and Kaiser can be found in the definition of rules for the construction of subject statements, the main difference between them lies in the methods proposed for the subject analysis, an aspect that was much more developed in Kaiser. In fact, it is in his method of subject analysis that Kaiser gets close to another important author of the 20th century: Paul Otlet.

Dousa (2010b) also revealed another common point in Otlet and Kaiser in relation to the theoretical development of knowledge organization: the analysis of information. As introduced earlier, the turn of the 19th century to the 20th century was also characterized by an emphasis on the subject approach to information, especially in the fields of classification and indexing. Based on the works developed by classifiers and librarians during this period of the systematic organization of subjects, we would like to highlight the idea that the representation of the knowledge conveyed by the documents could be done not only by the decomposition in smaller bibliographic units (such as articles in journals or chapters in books), but mainly by the analysis of smaller information units, such as the concepts, facts, and evidences found in the texts. It is precisely here where we can theoretically and methodologically approximate the works and philosophies of Kaiser and Otlet.

It is said that information units (concepts and facts), once identified, can be reconfigured in a new arrangement for the purpose of facilitating document retrieval (Dousa, 2010b). This is something that is evident in the works of both Otlet and Kaiser. This idea of representing the knowledge of the subjects in smaller units of information, aiming at a more precise document retrieval, would have a significant influence on the works and theories that came later. Some of these theories include: the facet-analytic approach to classification, in which the development of classification systems is based on facets; the work by documentalists and specialized librarians (during the first half of the 20th century) that, unlike general librarians that classified books based on the whole content of the document, began to classify the specific information of the content using information units; the distinction that emerged in the 1950s and 1960s between document retrieval systems and fact/object-based retrieval systems; and, more recently, the works on extraction of information units in digital documents for information retrieval systems that emerged in the 1990s, whose objectives were based on the possibility of identifying, retrieving, and structuring smaller units of information within digital texts.

On the other hand, it seems necessary to ponder the differences between Otlet’s philosophy and Kaiser’s before discussing their convergences. Based on Dousa (2010b), we summarize the differences between both authors in two aspects: the structure of the knowledge organization - classificatory order versus alphabetical order - and the scope of the structure - universalism versus localism. While Otlet held a preference for the classificatory order of subjects, as shown in the Universal Decimal Classification (UDC), assuming that the alphabetical order gave room to the indexer’s subjectivity in the choice of words (terms) to be used in the indexes, Kaiser developed his method in such a way that subjects were arranged alphabetically.

The second difference concerns the scope of the index. As a vocational bibliographer and a visionary of bibliographic internationalization, Otlet dealt with subjects in a universal and encyclopedic way (as did the classificationists of his time). On the other hand, as a librarian specialized in institutions such as companies, industries, museums, and specialized libraries, Kaiser aimed to account for the systematic organization of the specific (specialized) knowledge conveyed in the technical and professional literature. This focus on the development of indexes for particular organizations (localism), based on expertise and literature, distinguished Kaiser not only from Otlet but also from other contemporary fellows such as Dewey, Cutter, and Ranganathan. Although Otlet also developed specialized information systems for specific areas during his time at the International Institute of Bibliography, it was with the
UDC, a tool that followed a universal approach to the representation of subjects, that he became well-known in the field of bibliographic classification.

On the other hand, this difference in scope between Kaiser’s and Otlet’s systems does not necessarily seem to be the result of a divergence of philosophies, but rather of historical and conjunctural circumstances, as it was Kaiser’s professional profile what made him concerned regarding the construction of localized indexes for specific organizations, in opposition to Otlet’s academic vision and universal scope (Sales, 2014).

Indeed, the theoretical connection between Otlet and Kaiser lies in the fact that both authors distanced themselves from an organization of knowledge based on the documents in favor of an organization of the knowledge that is contained in them. For Otlet, knowledge would encompass all that is known about the objects of the external world or of the mind, whether they are physical objects (natural or artificial) or non-physical (laws, thoughts, feelings, etc.). The primary (main) elements of these objects were the “facts” or “ideas” that, conceptually structured by the authors, are recorded in documents in a way that reflects the view of each author regarding the phenomena that are presented and discussed (Dousa 2010b).

In Kaiser’s view (1911), knowledge, that would be the result of observation and reflection on something, would be materialized in the specialized literature and could be better understood by the analysis of the concrete (entity), country (place), and process (action) categories. It is important to note that as early as the 20th century, both Otlet and Kaiser were aware that users/readers/researchers were effectively seeking information, facts, and ideas that provided some kind of knowledge. It is also noticeable that both authors argued that it was necessary to dissociate the information from the document and to analyze the elements that compose this information in order to better organize the knowledge and consequently facilitate the retrieval of more specific information. It would be possible to relate both views and state that Kaiser’s concrete, place, and process categories could have perfectly well served as the basis for the identification of the facts, concepts, and ideas pursued by Otlet.

In addition to agreeing on the necessity of organizing subjects based on the analysis of the information contained in the documents, Otlet and Kaiser also converged on the ways to achieve such an organization. For Otlet, the function of the documentalist was to extract the relevant informative components (information units) from the documents, separating them from the original bibliographic context and recording them individually. This enabled records to be established analytically in “encyclopedic repertoires”, which in turn, provided an interconnection between the information units and the various publications (Sales, 2014). For Kaiser, the function of the indexer was virtually the same. The indexer should analyze the indexable literature in order to also extract the informative components (indexable information), separating them from the literature, and recording them in subject statements that would compose a systematic index (that in turn would enable the interconnection between the subject statements and the various materials – documents).

In this sense, it is possible to say that the knowledge organization field gained a new approach based on the works of Otlet and Kaiser, a perspective that established the organization of subjects based on the analysis of its informative components, or rather, an organization that was no longer of documents but of information about knowledge. As Dousa (2010b) puts it, both Kaiser’s systematic index and Otlet’s encyclopedic repertoires were developed to provide an immediate access to information.

Just as it is possible to state that there is a theoretical intersection in the analysis of the information (analysis of smaller units of information) developed by Otlet and Kaiser, we also observe a similarity regarding the technological viability of their systems. Both Otlet and Kaiser embraced the standardized card system. In the late 19th and early 20th centuries, such technology was seen as a “modernist” approach that raised several practical questions regarding the relationship between records and documents (compared to the systems of record books – catalogs in the form of books – that were common until then). A system based on cards of standard size, containing standardized transcripts of information, physically allocated in drawers and binders, related to specific furniture, and conceptually organized
using a classification scheme, was something really advanced and efficient in the time of Otlet and Kaiser (with the latter being responsible for its refinement for the field of indexing).

Both in Kaiser’s and Otlet’s systems, each card meant the record of a unit that represented an information item (Dousa, 2010b), cards were considered ideal tools to record the results of this analysis of information. Each card should serve as the bearer of a single information unit extracted from a given document, establishing a one-to-one relationship that would correspond to the relationship between the card and the information unit (the informative component extracted from the content of the document). Otlet called this principle the “monographic principle”, while Kaiser called it “individual reference”.

Given this, we can stress the link between Kaiser and Otlet in relation to the analysis of the information in constituent elements, moving away from an organization of documents to an organization of knowledge using a specific technology. Although Otlet developed his analysis of information as a means of perfecting a classification system, his subject analysis, together with Kaiser’s category analysis, also contributed to the foundation of what we are calling a period of subject analysis with emphasis in indexing.

20th Century (ii): Analysis and synthesis of subjects

If Otlet and Kaiser took the first steps in the subject analysis approach to knowledge organization, the combination of analysis and synthesis of subjects reveals another dialogue between authors, in this case between Kaiser and Ranganathan.

Both Kaiser and Ranganathan sought to develop their subject analyses without being restricted to the subjects of the documents, but rather focusing on the fundamental aspects (categories) of the universe of knowledge that generated the subjects. Regarding this, Ranganathan went a step further. While Kaiser stated that all specialized subjects could be analyzed based on concretes and processes, Ranganathan enriched his classification of subjects adding a level of facets to his fundamental categories. Ranganathan, like Otlet and Kaiser, understood that the universe of subjects should be analyzed based on its most particular aspects, i.e., its constituent elements. Ranganathan argued for an analysis that went from the most specific elements to the most general ones, until reaching five generic, final, indivisible, and fundamental ideas. These ideas had no value of true or false, being only utilized for classification purposes, therefore, they are not definable, only assumed (Ranganathan, 1967).

Although the facet-analytic approach would be the main innovation of Ranganathan in relation to Kaiser, we should also highlight the characteristics of the categories that both authors used for the subject analysis. Both Kaiser and Ranganathan used categories as general theoretical principles and this is the first point of convergence between them. While Kaiser had defined the concrete, process, and country categories, Ranganathan defined the Personality (P), Matter (M), Energy (E), Space (S), and Time (T) categories, known by the acronym PMEST. For instance, in the case of a document about “export rates of agricultural products in Brazil,” we could analyze it using Ranganathan categories as follows: Personality (P) – Agricultural Products, Energy (E) – Export, Space (S) – Brazil.

The terms “agricultural products,” “export,” and “Brazil” are components of a complex subject that grouped by facets are manifestations of the categories P, E, and S. Similarly, these categories would be manifestations of the concrete, process, and country categories in Kaiser’s subject analysis. From the perspective of the PMEST category analysis, of course, some extra information can be described. For example, if the document was about “agricultural export rates in Brazil during the 20th century,” the category referring to the temporal aspect can also be included in the analysis: (P) agricultural product, (E) export, (S) Brazil, and (T) 20th century. Likewise, a document on the oscillation of the US dollar in the year 2018 could be analyzed as follows: (P) Dollar, (E) Oscillation, (S) US, and (T) 2018. For a better visualization, Chart 1 draws a parallel between the possible subject analyses of these examples using Kaiser and Ranganathan categories:
Chart 1 shows a correspondence between the concrete and personality, country and space, and process and energy categories. However, it should also be noted that Ranganathan, while trying to explain the personality category, took a different approach than Kaiser, which stated that every process is concerned with something (either a concrete or a country). In other words, for Kaiser, the main category of any subject was concrete, while in some cases it was possible that there were no concretes, but only countries. There was no point in having a process without “something” being processed (or processing). For Kaiser, every process presupposes something concrete and/or spatial. On the other hand, Ranganathan (1967), from a perspective of eliminative analysis, stated that every facet that was not related to matter, energy, space, or time, would necessarily be related to personality. However, Ranganathan himself (1967) stated that personality had to do with entities, things, and kinds of things, thus it was also related to Kaiser’s concretes. Indeed, in both analyses, concretes and personalities play similar roles for the understanding and definition of complex subjects. In the correspondences between place and space, and process and energy, there seems to be a clear reciprocity, as the former refers to spatial aspects and the latter to actions and operations.

Obviously, because they are general principles and, as Ranganathan (1967) pointed out, indefinable fundamental categories, the categorization process does not guarantee a uniform analysis without variations. However, it does guarantee a guideline, that is, a definite way of analysis. Regardless of the different natures of the categories themselves – semantic categories in the case of Kaiser (Svenonius, 2000) and ontological categories in the case of Ranganathan (Aranalde, 2009) – the truth is that both authors followed a similar pragmatism (practicalism) in the development of their subject analyses: an analysis based on their fundamental constituent elements or, in other words, based on predetermined categories. This way of analysis, based on the identification of elementary aspects of the universe of knowledge and the universe of subjects, “crystallized by categories”, reinforces the analytical dimension of the analytic-synthetic method.

By understanding the analytical dimensions of Kaiser and Ranganathan, we can understand the synthetic dimensions established by both authors. Kaiser (1908, 1911) used the term “statement” to designate the subject statements or headings of his system. For the German librarian, statements essentially were standardized verbal expressions, composed of terms that expressed the categories (concretes, countries, and processes) of the analyzed subjects. In other words, statements represented the most relevant information extracted from the subjects of the documents, based on the concretes and the processes in them. If, for the understanding of the analytical dimension of Kaiser’s system, the focus was directed to the understanding of the categories, for the understanding of the synthetic dimension, the attention should be turned to the modus operandi of the development of indexes. More specifically, the attention should be directed to the development of statements, as it was with them that Kaiser sought to establish how the analyzed subjects could be synthetically stated.

For an efficient development of subject indexes, it was necessary to determine the possible combinations between categories, i.e., the order of importance of the categories to establish a proper citation order according to the statements. For Kaiser (1911), there were three possible combinations: (a) concrete – process; (b) country – process;
and (c) concrete – country – process. It should be noted here that the order of importance defined by Kaiser privileged the concreteness of the aspects of the subjects, in other words, it privileged the element that was more static. As a consequence, the concrete would always be in a privileged position compared to the country and process categories. It should also be noted that the categories that designate things (concrete and country) would always appear as entry terms in the citation order established by Kaiser. Here, it is also worth noting that the citation order based on “decreasing concreteness” was also adopted by Ranganathan in the arrangement of the PMEST categories (Ranganathan, 1967).

However, for the sake of our analysis, it also seems necessary to explain the different perspectives and notions of concreteness in Kaiser and Ranganathan. Concreteness in Kaiser’s system is related to how static a subject element can be. Once the most static element of the subject is identified, all other information would relate to this (concrete) element. For Ranganathan, the most concrete element may stem from a relation of cause and effect. As pointed out by Aranalde (2009), for Ranganathan, the cause (the reason why) can often be more concrete than the effect itself (the what). For instance, in an article about indexing, indexing could be understood as an effect of Library Science, while Library Science could be considered its cause. Thus, the domain of Library Science would be more concrete than indexing itself. Although Kaiser’s notion of concreteness is based on the relation between statics and dynamics and Ranganathan’s notion is based on the relation between cause and effect, the fact is that both authors sought to privilege the most concrete element in the subjects in order to dictate the order of importance of their syntheses.

The main aspects that made Kaiser’s synthetic dimension a starting point for the analytic-synthetic method were the reorganization of the information that was collected during the subject analysis and the development of rules for a controlled synthesis (Sales, 2014). Even considering that Ranganathan’s synthesis (classification notation) resulted in something different than Kaiser’s synthesis (alphabetic statement), it is possible to relate both syntheses regarding the pragmatism (the path) followed by both authors. In fact, it is possible to verify a considerable overlap between the analyses and syntheses developed by Kaiser and Ranganathan. Saving the differences between the results of the syntheses, it can be said that both authors conducted their analyses and syntheses in a similar way. Both librarians conducted their analyses through the decomposition of the subjects into their constituent parts, moving between the idea plane, that defines and supposes the categories of analysis, and the verbal plane, that provides the terminological outlines to the ideas in the subjects. In order to do the synthesis, both authors represented the subjects using their reconstruction in the development of subject statements (that are verbal in the case of Kaiser and notational in the case of Ranganathan). In both cases, the orders of importance are guided by the degree of concreteness of the categories of analysis. Thus, both Kaiser and Ranganathan developed analytic-synthetic methods guided by category definitions. They both envisioned analyses and syntheses of subjects using categories that functioned as theoretical principles. In this way, by putting the perspectives of indexing and classification next to each other, we understand that Kaiser and Ranganathan confirmed or, at least, established together the period of subject analysis and synthesis.

**Conclusion**

We have analyzed the contributions of some of the most important authors in the history of knowledge organization, namely Harris, Dewey, Cutter, Otlet, Kaiser, and Ranganathan, in order to reveal some theoretical dialogues and aspects that are not always explicit in the literature. Assuming these authors and their works are foundations of the theoretical and methodological frameworks in knowledge organization, we studied their main convergences in order to historicize their contributions in the second half of the 19th century and the first half of the 20th century, drawing attention to potential dialogues in the epistemological development of the field. For each author, we focused on their approaches to bibliographic classification and indexing in order to better understand the role they played in the history of knowledge organization.
In the second half of the 19th century, we revealed a theoretical dialogue between Harris and Dewey that, although with different purposes, marked the influence of Bacon and Hegel's philosophy in the development of library classifications. This dialogue conceived the foundation of what we have called the period of subject description formalized in classification schemes. During the same period, and also for the purpose of describing the subjects of books, Cutter created his expansive classification following an evolutionary approach. Thus, notably in the North American scenario, there was an emergence of a knowledge organization system based on the construction of descriptive classification schemes following idealist and evolutionary approaches. The dialogue between these three authors allows us to conclude that the spirit of the library discourse of the time was the subject description of library collections.

We also studied Cutter as an author who oscillated between his work on bibliographic classifications and subject indexing, dealing with both notational representations and verbal representations of subjects. Cutter’s pioneering work on establishing principles for the development of subject headings can be seen as an important link between the approach to classification during the 19th century and the approach to indexing during the 20th century, something that, in the context of Cutter’s definitions, was called subject cataloging.

The link between Cutter and the indexing of the 20th century becomes more evident with the analysis of the contributions to systematic indexing developed by Kaiser. In some way, Kaiser continued the definitions of rules and principles for the standardized development of subject statements. However, Kaiser’s verbal representation went further when he proposed a subject analysis of documents based on semantic categories. This analysis of the information contained in the documents reveals a link between Kaiser and Otlet, as they both had in common an analysis of information based on the constituent elements – facts, names, and ideas, in the case of Otlet, and concretes, processes, and countries in the case of Kaiser. With Kaiser, we identified the emergence of another fundamental dialogue in knowledge organization, in this case with Ranganathan’s facet-analytical approach. Kaiser and Ranganathan, while focusing both on the analysis and synthesis of complex subjects, defined the so-called analytic-synthetic method.

In conclusion, we summarize the dialogues as follows: Harris, Dewey, and Cutter – idealist and evolutionary approaches to descriptive classifications; Cutter and Kaiser – definitions of rules for the standardization of subject statements; Kaiser and Otlet – analysis of information based on the constituent elements; and Kaiser and Ranganathan – development of the analytical-synthetic method.

Contributors

All authors contributed to the conception, design of the study, data analysis, and find editing.

References


