

## Harnessing user's knowledge in the construction of rating flows: The design of a collaborative system applied to academic repositories

Azevedo-Monteiro, Bruno Miguel<sup>a</sup>; Silva, Helder<sup>b</sup> & Tortosa, Rubén<sup>c</sup>

<sup>a</sup> Department of Drawing. Universitat Politècnica de València, Spain. [brunomiguelam@gmail.com](mailto:brunomiguelam@gmail.com).

<sup>b</sup> Department of Industrial Electronic. University of Minho, Portugal. [helderdavidms@gmail.com](mailto:helderdavidms@gmail.com).

<sup>c</sup> Department of Drawing. Universitat Politècnica de València, Spain. [rtortosa@dib.upv.es](mailto:rtortosa@dib.upv.es).

---

### Abstract

*Artifacts developed over several ages, such as libraries, encyclopedias, and databases, show the cultural evolution of information systems. Compiling, organizing and visualizing information is a task that has been carried out by mankind for thousands of years. The added difficulty in effectively communicating information in various sectors and services of our society reveals that an efficient communication of information is of the utmost importance in the current network society. The glut of information is directly related to the fact that the information we are exposed to is not subject to a filtering and organization process. This reveals an urgency to develop strategies that not only prioritize the organization and searching, but also increase the efficiency of the communication process, in order to promote an efficient framework to the user's cognitive and perceptual field. Therefore, the task of designing complex information systems in an accessible manner currently represents an important goal and an imperative task to the Designer.*

*The publication and the querying of papers, journals, books, is an integral part of the research process. However, the querying and information visualization process in a scientific academic repository often proves to be a complicated and inefficient task, as the wide range of results hardly fits in the user's specific subject. However, if we equate that the knowledgeable objects are accessed by a significant number of users with a specific interest in a topic and that, in the course of their research, each user handles a significant amount of results, it is, then, possible to consider the existence of an hierarchical and relational structure of evidences that emerges from the relationship established between the users, their specific interests and knowledge concerning a topic and the querying performed. Therefore, it is fundamental to consider the users' experience and the leading role that it plays concerning the information filtering process.*

*This paper aims to present key insights on the information glut problematic related to/associated with a massive amount of knowledge objects stored, and proposes a new approach/system applied to the academic scientific repositories. A collaboration system is designed, in order to filter and visualize the rating flows based in users' experience, instead of the usual citation "object" centered approach. The focus of this work is to describe one part of the system: the experimental implementation of an interactive hierarchical structure.*

**Keywords:** *collaboration, design, information, hierarchical structures, visualization.*

---

## 1. Introduction

Aspects such as the structuring and presentation of information, framework and content filtering are an urgent and ongoing challenge nowadays (Thackara, 2006), fact that is evident throughout the cultural evolution of information systems (Wright, 2008). However, due to Humanity's tremendous effort to collect and store information in earlier times (idem, 2008), the need to design and develop strategies to filter and reduce the information volume becomes evident (Card et al., 1999), (Wurman, 2001), (Thackara, 2006), (Wright, 2008), (Gleick, 2011) in the current digital information society (Castells, 2010). In fact, the current digital repositories of knowledge (DRK) only constitute an apparent solution to the problematic, as they allow the reduction of the distance and limits of the access to information and make a wide typology of knowledge objects available online (KO) (e.g. books, scientific journals, papers, thesis). Despite the referred advantages, the current problems addressed to DRK are directly related to filtering processes and the visualization of the retrieved information (Thackara, 2006). Although the KO is just one click away, the search for information on DKR proves to be a complex, inefficient and arduous task. This procedure is exacerbated by a slow query process of long lists of results (Marks et al., 2005, p. 57-59), which implies an individual analysis process of each KO (in the specific context of papers, books, thesis). In this sense, it becomes clear not only an intrinsic problem related to the wide range of results obtained that translates into a visualization difficulty due to the disturbing amount of data available in the users' cognitive and perceptual field, but also a problematic directly related with the filtering of the KO. Usually, the search engines of the DKR merely allow a search/filtering process based on topics like keywords, author, ISBN, subject, year of publication, title, among other similar examples. However, there are two other problems that are related to the lack of information concerning the characteristics of knowledgeable objects and users, because in both cases the available metadata is very limited, and, in the case of the users, they are virtually nonexistent.

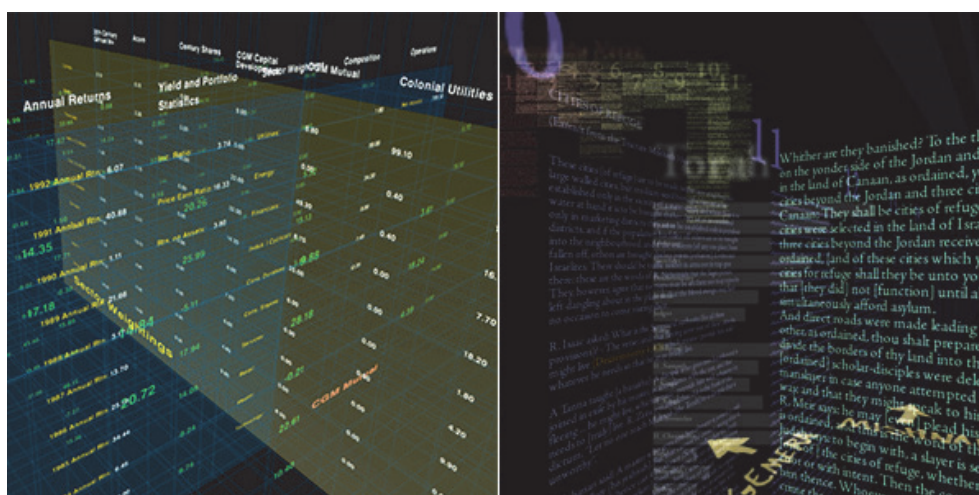
In this sense, the current knowledge retrieval systems are insufficient, due to the exponential amount of published scientific knowledge made available online. In fact, according to Börner (2010, p. 12), the current DKR do not allow a clear understanding of the various academic entities and their numerous and complex interdependencies. Therefore, the design and development of new visual languages and new communicative paradigms, whose purpose is the representation of knowledge structures at different scales, is an important and urgent issue in the complex field of academic research (Börner, 2010, p. 12). In this sense, this article aims to describe part of a system that intends to allow visualization of a wide range of KO, based on the rating flows, through a specific and interactive visual structure in order to provide relevant and efficient results for the user. A brief explanation of the rating process, which is based on a simplified evaluation factor/metric, is presented in point 2.1.

The problem with information flood reveals another issue concerning the exponential growth of stored KO, which is directly related to the user's inability to consult each KO individually, given the wide range of results obtained. This fact shows that the cooperation factor is a key component in the filtering process, through the rating of the KO by the academic community. Understanding and visualizing the collaborative structure of evidences reflecting the multiple perspectives and individual experiences of each user proves to be the key equation in allowing a more efficient insight of the KO with greater relevance, based in the wisdom of the academic community in a particular field of research.

## **2. Digital Knowledge Repositories And The Problematic Of The Information Glut: Brief Analysis Of The Context**

Taking into account the evolution of biological and cultural information systems (Wright, 2008), the current digital libraries are the main source for Humanity knowledge (Börner et al., idem, 2002). In this sense, it is fundamental to develop and redesign new interfaces focused on management, access, visualization and understanding of the various information types stored in the DKR. In fact, the DKR are key artifacts in the access to a wide typology of KO. The current research and development in the field of interfaces and Information Visualization (InfoVis) concerning DKR is directly related to the problematic of the information flood (Card et al., 1999), (Wurman, 2001), (Wright, 2008), (Castells, 2010), (Gleick, 2011), fact that is originated by increased storage and processing capacity, interconnection between different systems and development of new interfaces that facilitate both access and publication of contents.

According to Börner et al. (2002), the fast pace of scientific discoveries and technical development, and the appearing of new fields and themes in increasingly shorter periods, significantly contributed to a consequent increase of scientific publications. However, in spite of the published data being scientifically valid, the problematic of scientific information stored and available online also contributes to the current problem of the information abundance in academic contexts (Thackara, 2006, p. 163). That implies that this is a problem concerning scientific knowledge networks. In the field of DKR it is directly related to cataloguing, categorizing, structuring and allowing visualization of an exponential amount of produced and published scientific content and it clearly shows that users are experiencing difficulties to perceive and process such large volumes of available KO. Such fact stresses the need to conceptualize and develop new artifacts concerning the retrieval, visualization and communication of all stored data types, such as data resulting from users' interaction with DKR and KO. The current DKR interfaces are featured by advanced data analysis techniques, in order to display results, normally in the form of extensive lists of KO (Marks et al., 2005, p. 57), and organized according to a metric (e.g Association for Computing Machinery Digital Repositories) (Kim et al., 2011, p. 123). However, this is not an adequate solution, due to the growing volume of stored KO (Dushay, 2004). In this sense, InfoVis constitutes a viable response to difficulties concerning search and navigation tasks, as it allows a greater cognitive and perceptual efficiency (Marks et al., 2005, p. 57-69), (Kim et al., 2011, p. 123-136). This efficiency is ensured by the fact that InfoVis has the capacity to transform abstract data on visual attributes (e.g. shapes, colours, scales), therefore reducing the cognitive and perceptual effort required to process large volumes of information.



*Fig.1 Information Landscapes. Visible Language Workshop. MIT Media Lab. Cooper (1994).*

In fact, the main purpose of InfoVis lies fundamentally in structuring complex information spaces (Cooper, 1994) [Fig.1], proving to be an asset to the interaction of users with DKR (Börner et al., 2002). This means not only a reduction on the degree of cognitive processing effort, by taking into account difficulties concerning the slow process of consultation and reading, but also translates into a more efficient perceptual process based on visual attributes, that allows efficient decoding and presentation of patterns and evidences imbued in the data, some of them unnoticeable until then (idem, 2002). In fact, the InfoVis makes use of the advantages and capacity of the human perceptive and cognitive system (Card et al., 1999), in order to assist the user with his mental organization and structuring of data while accessing and decoding complex information spaces (Shiri, 2008, p. 764-765). In this sense, it is stated by Börner et al. (2002) that the integration of InfoVis in interfaces such as digital libraries and repositories is fundamental, particularly in terms of search time; understanding and decoding a complex and broad set of data; visualization of relationships and evidences; simultaneous visualization of data and multiple perspectives; access to efficient sources of communication; quick access to contents and new forms of analysis.

According to Börner et al. (2002), the interfaces aimed to the DKR are defined by four implementation scenarios: identifying the composition of a particular result; providing an overview of the entire collection and facilitating of retrieval tasks; visualizing the interaction between user and data in relation to the documents available, in order to evaluate and upgrade the interaction properties; improving the sharing methodology of both information and collaboration (idem, 2002).

The current DKR are, in fact, the principal knowledge repositories, (Fox et al., 2002 p. 506-507), and the design of user-friendly interfaces for management, access and efficient understanding of the complex volume of stored data is nowadays both an imperative challenge and an essential task of the Designer (Wurman, 2001). To sum up, and according to Börner et al. (2001, p. 12-15), it is fundamental to study and develop interactive visualizations concerning the field of DKR, taking into account the following key points: intuitive interfaces; fast and efficient access to an increasing volume of KO; new ways to analyze the KO; addition of new data to existing information (e.g. metadata) and easier sharing of information.

## 2.1 Knowledge Networks: Brief Analysis Of The Context

Currently we live in a globally connected network society (Castells, 2001), (Wright, 2008), in which "billions of people produce trillions of connections" (Hansen et al., 2010, p. 3). Such fact is extensively expanded by the exponential development of Information and Communication Technologies, the appearance of the digital social networks (Web 2.0) being an example of it. According to Wright (2008, p. 9), if we consider only one focused perspective on the current cultural information systems, this proves to be shallow. In fact, the hierarchical and network systems and the tension between these two coexistent structures (idem, 2008) form the structural and organizational model that permeates in all layers of our biosphere, as well as in our infosphere (idem, 2008), (Castells, 2010), (Gleick, 2011). An example is the natural biological organization that follows an hierarchical logic (Simon, 1996, p. 172), (Wright, 2008). At the infosphere level, the World Wide Web is an hierarchical structure, as it obeys a physical system (hardware) organized hierarchically, and it simultaneously is a relational structure composed of an exponential number of hyperlinks (Yam, 1997), (Wright, 2008), (Gleick, 2011). In this sense, we can consider that hierarchies and networks are the basic structures of information, and that understanding them is fundamental to comprehend the biological and cultural evolution of information systems (Yam, 1997), (Wright, 2008).

In fact, the DKR is actually featured by hierarchical structures. One example of this is the organization per alphabetical order of the KO. Another example is a book or paper index. However, scientific papers or books are defined by a relational structure of hyperlinks (bibliographic references), namely a complex structure of citations between several KO. In fact, scientific publications are based on a set of references based in previous work, which holds a substantial impact on the actual research object, usually located at the final section of each paper or book, specifically the reference section. According to Börner et al. (2014, p. 170), the visualization of information networks is a branch of InfoVis, intending to perform the analysis of both natural/biological and digital/cultural networks, specifically, social networks, information science, bibliometrics, scientometrics, econometrics, infometrics, webometrics, communication theory, sociology of science and many other disciplines. In this context, the connections are featured for example by collaborations between authors, quotes from papers and patents (ibid., 2003, p. 4), (idem, 2014, p. 170). In the case of scientific knowledge networks, the main objective is to identify the authors, papers or knowledge domains with the largest number of connections, (e.g. citation and collaboration); network properties (e.g. size and density); structures (e.g. clusters) (idem, 2014, p. 170). It should be noted that the analysis around knowledge networks/domains is defined by three scales, specifically micro, meso and macro. In this sense, it is important to highlight that the presented experimental implementation is in the meso level. According to Börner et al. (2014, p. 3-7), the meso level is featured by values between 101 and 10,000 records, such as the number of researchers of a single university and/or a particular subject that is investigated.

It should be noted that this subfield of InfoVis is intrinsically linked to the metric of the impact factor of a journal or paper, meaning that the importance of a scientific publication is directly related to the number of citations: the greater is the number of citations, the bigger is the impact/relevance of the publication. It is therefore important to highlight Garfield and his fundamental contribution to the study of communication/scientific dissemination (field of bibliometrics, scientometrics). In fact, Garfield (1963, p. 5) revealed concerns with the aesthetics of scientific communication, taking into account its chaotic state in the 1963 period (Garfield, 1964, p. 88). According to Wright (2008, p. 203), the article published by Vannevar Bush, *As We May Think* (1945) (1996, p. 35-46), inspired Garfield to explore and develop new forms of access to scientific journals (Wright, 2008, p. 203). Therefore Garfield develops a methodology called citation ranking, that is a tool to assess the impact factor of academic publications based on the number of citations (Garfield, 2003, p. 363-339), (Wright, 2008, p. 203). The Science Citation Index then



allowed measuring of the impact factor of the KO determined by the cumulative value of citations. In fact, the bibliographic citation is a common practice in various types of academic publications and an important measure of credibility and popularity for research projects, journals, papers, researchers and institutions (idem, Garfield, p. 363-339), allowing equating the existence of a vast relational structure and/or similarities between subjects (Lima, 2011, p. 102).

Taking into account the vast scenario of academic publications, it is possible to infer the existence of a vast hierarchical and relational structure in which it is, for example, possible to gain insight about the proximity between distinct areas (Lima, 2011, p. 102) and /or citation patterns between different areas: which papers are most cited in a given area; which area has the highest number of citations; and if an author of a paper is cited by other authors. In fact, a large part of quantitative studies about science is based in the analysis of hierarchical and relational structures that are based on the reference or citation of publications or coauthoring, namely cooperation structures between researchers (Staudt, 2011, p. 1). This means that two researchers are interconnected when they are co-authors in one or more KO (Newman, 2001b). According to Newman (2001b), quantitative analyzes of relational structures are defined by the number of papers written, the number of authors of a paper, the number of contributors, the existence and the extent of a researchers network and the degree of network clusters.

According to Meirelles (2013, p. 49), individuals are actors (vertexes or nodes) and links (edges) between individuals are ties. This designation might refer to trust and cooperation ties between two or more individuals, or from an ordinary member between groups, among other examples (idem, 2001), (Hansen, 2010, p. 34-35).

In short, in the academic collaborative social networking actors are researchers, and the bonds that emerge from their collaborative relationships represent the co-authoring linkages between one or more papers. According to Börner (2015, p. 60-61), network analysis and techniques that enable the visualization of relational structures allow to answer the question "with whom?". However, it should be noted that the hypothesis here presented highlights one fundamental question: "Which?". It stresses the need to promote an approach around the retrieval problematic and the obtained results, through the visualization of the hierarchical structures (the main object of study of this article) and the relational structures that emerge from each user's interaction with researched objects.

### **3. Material and Methods**

The equated hypothesis is a new paradigm that determines a change on the approach focus, usually centered in the citation of KO or authors. That means that this article presents an experimental implementation of one part of the system, specifically the design and computation of a contention hierarchical structure using the programming language Processing, intended to assure the visualization of KO with greater relevance within a particular branch of knowledge. As mentioned before, it emerged from the relationship of a problem concerning the retrieval and filtering of KO and the visualization of the structure of evidences that results from the relationship established between the users' queries and the enrichment process (rating) of the KO. Therefore, instead of the usual "object" centred approach, it establishes an approach based on user's experience.

The main goal of this point is to present and describe one part of the system, that is the experimental implementation of an interactive hierarchical structure. It should be noted that in the absence of data concerning the rating of KO, it was decided to simulate hypothetical contexts of interaction by using a set of structured data between users, KO, ratings, knowledge domains and knowledge subdomains. In this context, a set of fictitious data (meso level) was generated in order to simulate the interaction of users

while performing their consultation and rating of the stored KO in the DKR. The use of fictitious data concerning the metadata of the KO (e.g. ISBN, Title, Year) should also be highlighted. MySQL was used to implement the database system, as it is an open-source relational database management system. A more detailed description of the modeling of relations in the database is out of the scope of this article.

The decision to maintain the evaluation system centered in and closed to the academic community is related to the advantage of being able to identify the type of user (Student, Professor, Researcher). As the system can only evaluate the KO one time, and taking into account that in open systems such as Amazon or Ebay the user remains anonymous and normally uses a pseudonym making it impossible to know what type of user it is (Rheingold, 2002), in the particular case of the formulated hypothesis the access to the institutional repository and rating of the KO is conducted in accordance to each user's access credentials.

### **3.1 Simplified Weighting Factor: Brief Explanation**

Regarding the KO rating process, the evaluation was made based in a range of integers from one to five, which is directly related to the user's knowledge subdomain and with the KO subdomain. Therefore, at the relational level, the sub-levels of the knowledge branch of the users and of the KO were considered. This means that in the hierarchy of relations, particularly between the users and the KO, a linkage between the user's knowledge subdomain and the KO subdomain is considered. Thus, in knowledge domain, Design is taken into account on various subdomains, such as, for example, Communication Design and/or Information Design. Considering the weighting factor, it should be noted that a greater weight is assigned to users whose subdomains are directly related to the subdomain of the consulted article. Therefore, the rating of a user who belongs to the subdomain of the consulted KO has more weight than a user who does not belong to the general domain or subdomain of the KO.

Despite being outside of the scope of this article, a more rigorous approach will be considered in the rating of the KO consulted. In this context, a distinct weighting factor between users belonging to different knowledge subdomains of the same domain must be considered. This means that, if the KO consulted belongs to the subdomain Multimedia, the weight of an evaluation of a user belonging to the subdomain Communication Design should be inferior to the weight of an evaluation from a user belonging to the subdomain Multimedia. It should also be noted that both users belong to the same field of knowledge, namely Design. It would also be important to consider each different type of user, making it essential to assign different weighting factors to Students, Professors and Researchers. The modeling of relations between users of the same subdomain, although described, remains an open question that will be the subject of future studies and work.

Concerning the simplified weighting factor, the rating of an user that belongs to the knowledge subdomain of the KO consulted has a greater weight than the rating of a user who does not belong to the KO subdomain. It should be highlighted that the evaluation weight of users who belong to the same subdomain of a knowledge domain of the KO is equal. However, it presents a higher weight in comparison to a user that does not belong to the knowledge domain or subdomain of the consulted KO, as a value of one (+1) is assigned to the evaluation performed. This implies that the simplified weighting factor is associated with a correlation of the subdomain of the user and the subdomain of the KO consulted.

Thus, the weight of an evaluation would function in two directions, to improve or decrease the rating impact. A study on a parameter that balances the final evaluation is outside the scope of this article. The example included on the previous paragraph reveals the kind of complexity that could be implemented at the level of relations between the knowledge subdomains of each knowledge domain and the knowledge subdomains and the typology of users. It should be noted that the ratings assigned by users of the

knowledge subdomain of the KO have greater resistance to change when compared to the ratings assigned by users that do not belong to the KO subdomain. However, if the number of evaluations performed by users who do not belong to the KO subdomain increases, the value can tend to the evaluation allocated by these users due to the increased number of evaluations.

The approach stated in this section considers a relationship based on the KO subdomain and in the user's knowledge subdomain. It should, however, be noted that in the present approach a different weighting factor was not implemented, neither in terms of subdomains nor in terms of user's type. This means that only the user knowledge subdomain was considered, despite belonging to a knowledge domain. It should also be highlighted that, in case of equated hypothesis, the views of the structures will be available only if there is a participation of the user.

### **3.2 Contention Hierarchical Structure: Experimental Implementation with Ordered And Squarified Treemap**

The main objective of the experimental implementation is the visualization of the KO with greater relevance within a particular subdomain of knowledge. It should be noted that this experimental implementation is based on the treemap algorithm developed by Fry (2007, p. 182-219) in the programming Processing language, specifically the Ordered Treemap (SHNEIDERMAN et al., 2001), (BEDERSON et al. 2002) and the *Squarified Treemap* (Bruls et al., 2000, p. 33-42) algorithms. In this sense, the main objective of this point consists in the experimental implementation of both structures. In both cases the size of the squares varies according to the relevance of the KO in a particular knowledge domain. As previously mentioned, the proportion of the areas is based on an evaluation factor, as described in section 2.1 (Simplified Weighting Factor). The two knowledge domains are defined by two colours (grey and blue): a range of grey and blue shades define the various subdomains of the two principal knowledge domains.

Regarding the treemap context, this is a visualization technique that is originally based on the algorithm developed by *Johnson et al.*, (1991) and *Shneiderman* (1992). It consists on a rectangular hierarchical structure of containment, aimed to the visualization of a large hierarchical structure of quantitative data (Card et al., 1999), (CHEN, 2006, p. 190-194). The Treemap technique is fundamentally characterized by a rectangular layout, divided into a sequence of rectangles, in which the area of each rectangle corresponds specifically to a given attribute of the data (BEDERSON et al., 2002). It is also characterized by an efficient use of the layout space when compared to horizontal or vertical relation hierarchical structures, which are very extensive structures. Thinking of disadvantages of the Treemap algorithm it is important to highlight that the rectangular shapes do not allow an efficient comparison between areas of identical proportions when randomly positioned in space, and that in the case of rectangles with very small proportions it is not possible to perform an efficient selection of the areas (BRULS et al., 2000).



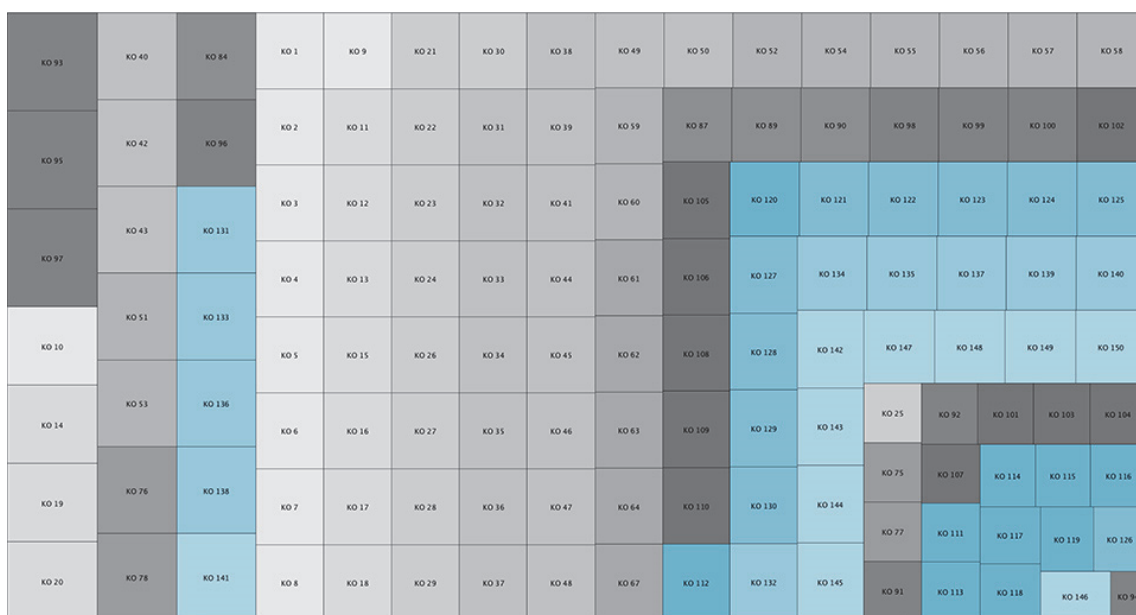
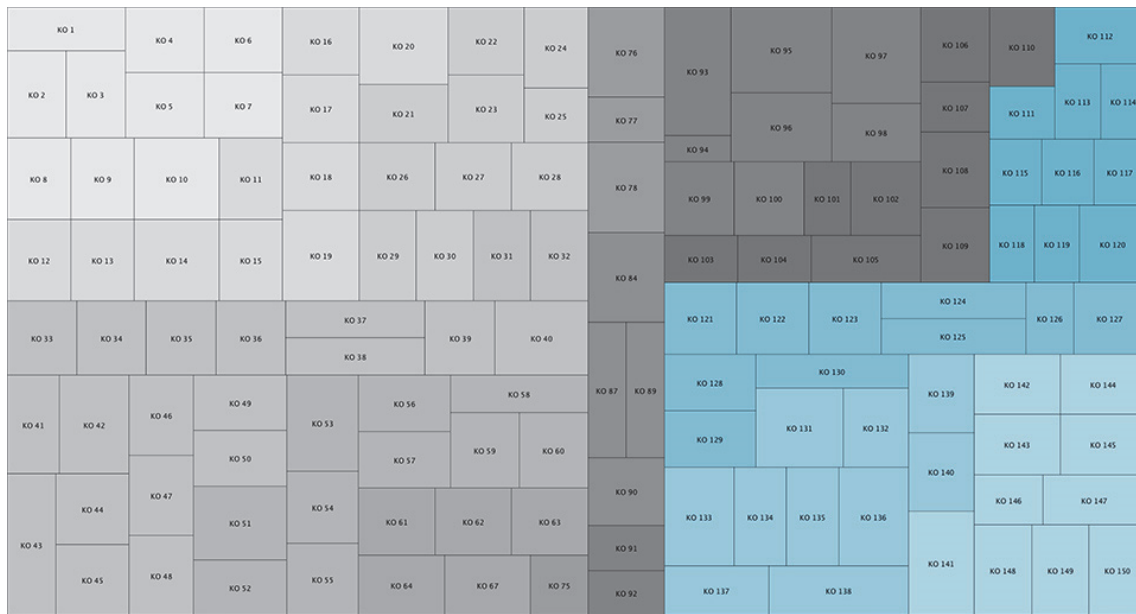


Fig. 2. Experimental implementation of the Squarified Treemap Algorithm. Bruls et al. (2000)

According to Bruls et al. (2000), the Squarified Treemap algorithm [Fig. 2.1] allows the transformation of rectangles with similar proportions into square shape proportions. That said, the main advantages of this algorithm are the efficient use of available layout space; the easier distinction and selection between square shapes when in comparison to rectangular shapes, as, even if the shape proportions are similar, it is easier to establish comparisons; and the improved presentation accuracy. As it is pointed out by Bruls et al. (2000), the main disadvantage of this technique lies in the ordering of subadjacent data (sibling data), because it is not possible to establish comparisons between the subadjacent data, as they are not organized according to a relation of proximity, such as by area or by colour. This means that it is not possible to establish comparisons with KO with most relevance in a specific knowledge subdomain.

Schneiderman et al. (2001) and Bederson et al. (2002), taking into account the several algorithms developed (*Strip/Clusters Treemaps, Squarified Treemaps,*) (MEIRELLES, 2013, p. 32), emphasize several disadvantages, despite the improvements in areas such as visualizations modes and integration of smaller proportions in a single layout. In fact, they prove to be unstable when the data is changed (Updates) and disadvantageous in comparison to the ordering and agglomeration of adjacent data, originating layouts that do not allow the establishing of comparisons, as well as efficient visualization of patterns.



*Fig. 2.1 Experimental implementation of the Ordered Treemap Algorithm. Bruls et al. (2000).*

In the *Ordered Treemap* algorithm [Fig. 2.1] (Shneiderman et al., 2001), (Bederson et al. 2002) previously ordered data preserves the proximity/adjacency in the layout. In this sense, we highlight the algorithm adaptability, taking into account the problem of dynamic data representation, allowing the previously ordered data to maintain a position of proximity on the layout, as well as a balanced proportion ratio of the rectangles.

To sum up, the *Ordered Treemap* algorithm is distinguished by the following characteristics: in the dynamic updates the changing of the forms occurs relatively smoothly; it preserves the subjacent data order in the layout; and it generates rectangles with reduced proportion ratio.

#### **4. Discussion And Future Work**

In terms of results, and although being preliminary, it can be deduced that treemaps constitute the most efficient search solution, by translating synthetically a large structure of evidences. In this sense, the two hierarchical containment structures tested (Squarified and Ordered Treemap) allowed to provide a birds-eye perspective of the KO with greater weight within a knowledge domain. This means that the greater the rating of the KO is, the bigger is the area it occupies. In fact, each area represents a KO, and each area's size varies according to the weight of the global evaluation assigned. Taking into account the advantages and disadvantages of the studied algorithms, one must conclude that it is fundamental to establish a fusion between this two types of Treemaps, an issue that it will be studied and implemented in future work.

KO 12	KO 23	KO 32	KO 41	KO 60	KO 105	KO 120	KO 121
KO 13	KO 24	KO 33	KO 44	KO 61	KO 106	KO 127	KO 134
		KO 24. title description, domain description 001, subdomain of domain 001					
KO 15	KO 26	KO 34	KO 45	KO 62	KO 108	KO 128	KO 142
KO 16	KO 27	KO 35	KO 46	KO 63	KO 109	KO 129	KO 143

Fig. 3 Experimental implementation of the interactive tooltip (Ordered Treemap Algorithm)

Given the limited space of the areas to show relevant information about the characteristics of the KO (metadata), an interactive tooltip was implemented [Fig. 3] in order to provide specific details on the KO, specifically metadata such as ISBN, Title, Year, Author/s, title, type of KO and subdomain and the average of the ratings assigned by the community. In this sense, the tooltip is shown when a specific area of a particular KO is clicked. Tufte (2009, pp. 178-182) points out that the integration of tables, graphs and words (legends) is fundamental, because even though they belong to different systems, they have a single purpose: the presentation of information. It is nevertheless necessary to pursue improvements in areas such as presentation and communication of metadata, a theme that will be addressed in future work.

To summarize, the advantage of using this structure lies in the fact that it gives, at first glance, a panoramic view of the most relevant KO within a particular area of knowledge. It is important to emphasize that the collaborative filtering mechanism conceptualized and implemented plays a key role in allowing users to perform a more sustained and directed research on their specific theme, based on the relevance of the KO that is determined by the "wisdom of crowds."

Intending to optimize the preliminary results of this study, it is fundamental to maintain open lines for future work, the implementation of the interactive zooming techniques being one of the key components.

## 5. References

- BAR-YAM, Y (1997). "Complexity Rising: From Human Beings to Human Civilization, a Complexity Profile" In Encyclopedia of Life Support Systems (EOLSS UNESCO Publishers, Oxford, UK, 2002); NECSI Report, Vol. 01, n°. December, p. 1-33.
- BEDERSON, BENJAMIN; SHNEIDERMAN, BEN; WATTENBERG, MARTIN. "Ordered and quantum treemaps: Making effective use of 2D space to display hierarchies." In ACM Transactions on Graphics, Vol. 21, n°. 4, p. 833-854.
- BÖRNER, K (2010). *Atlas of Science: Visualizing What We Know*. Massachusetts: MIT Press.
- BÖRNER, K (2015). *Atlas of knowledge: Anyone can Map*. Massachusetts: MIT Press.
- BÖRNER, K., CHEN, C., (2001). "Visual interfaces to Digital libraries". In ACM

SIGIR Forum, Vol. 35, no. 1, p. 12–15.

BÖRNER, K; CHEN, C (2002). *Visual Interfaces to Digital Libraries*. Springer-Verlag.

BÖRNER, K; POLLEY, D (2014). *Visual insights A Practical Guide to Making Sense of Data*. Massachusetts: MIT Press, 2014. ISBN 978-0-262-52619-7.

BRULS, M; HUIZING, K; WIJK, J. V.(2000). “Squarified Treemaps”. In *de leeuw, Willem Cornelis and Van Liere, Robert (eds.), Proceedings of the Joint EUROGRAPHICS and IEEE TCVG Symposium on Visualization in Amsterdam*. Amsterdam, Netherlands: Springer Vienna, 2000. 33–42.

BUSH, V (1996). “As we may think.” In *Interactions*, Vol. 3, n°. 2, p. 35–46.

CARD, S; MACKINLAY, J; SHNEIDERMAN, B. (1999). *Readings In Information Visualization: Using Vision To Think*. San Francisco: Morgan Kaufmann.

CASTELLS, M. (2010). *The Rise of the Network Society*. United Kingdom: Wiley Blackwell.

CHEN, C. (2006). *Information Visualization Beyond the Horizon*. London: Springer-Verlag.

COOPER, M. (1994). *Information Landscapes*. Visible Language Workshop. MIT Media Lab”. *You Tube* <<https://www.youtube.com/watch?v=Qn9zCrIjzLs>> [Consulted: 13 May 2016]

DUSHAY, N. (2004). “Visualizing bibliographic metadata - A virtual (book) spine viewer.” In *D-Lib Magazine*, Vol. 10, n°. 10.

FOX, E; URS, S. (2005). “Digital libraries.” In *Annual Review of Information Science and Technology*, Vol. 36, n°. 1, p. 502–589.

FRY, B. (2007). *Visualizing Data*. Sebastopol: O’Reilly Media, Inc.

GARFIELD, E. (2003). “The meaning of the Impact Factor.” In *International Journal of Clinical and Health Psychology*, Vol. 3, p. 363–369.

GARFIELD, E. (1963). “Aesthetics In Scientific Communication.” In *Essays of an Information Scientist: 1962-1973*, Vol. 1, p. 5

GARFIELD, E. (1964). “Towards the World Brain.” In *Essays of an Information Scientist: 1962-1973*, Vol. 1, p. 8

GLEICK, J. (2011). *The Information, A History, A Theory, A Flood*. New York: Vintage Books.

HANSEN, D; SHNEIDERMAN, B. & SMITH, M. (2010). *Analyzing Social Media Networks With NodeXL: Insights From a Connected World*. USA: Morgan Kaufmann. <[https://i11www.iti.uni-karlsruhe.de/\\_media/teaching/theses/dastaudt-11.pdf](https://i11www.iti.uni-karlsruhe.de/_media/teaching/theses/dastaudt-11.pdf)>. [Consulted: 14 June 2015]

KIM, B; SCOTT, J; KIM, S. (2011). *Exploring Digital Libraries through Visual Interfaces*. Digital Libraries - Methods and Applications. InTech, p. 123–137. Available from: <<http://opus.ipfw.edu/cgi/viewcontent.cgi?article=1034&context=compscifacpubs>> [Consulted: 4 July 2015]

LIMA, M. (2011). *Visual Complexity. Mapping Patterns of Information*. New York: Princeton Architectural Press.

MARKS, L; HUSSELL, J; MCMAHON, T; LUCE, R.(2005). “ActiveGraph: A digital library visualization tool.” In *International Journal on Digital Libraries*, Vol. 5, no. 1, p. 57–69.

MEIRELLES, I. (2013). *Design for Information: An introduction to the histories, theories, and best practices behind effective information visualizations*. USA: Rockport Publishers.

NEWMAN, Mark (2001). “Scientific collaboration networks. Network construction and fundamental results ” In *Physical Review*, Vol. 64, n°. 1, p. 1–8.

RHEINGOLD, H. (2002). *Smart Mobs The Next Social Revolution*. USA: Perseus Basic Books.

SHIRI, A. (2008). “Metadata-enhanced visual interfaces to digital libraries.” In *Journal of Information Science*, Vol. 34, no. 6, p. 763–775.

SHNEIDERMAN, B. (1992). “Tree visualization with tree-maps: 2-d space-filling approach.” *ACM Transactions on Graphics*, Vol. 11, n°. 1. 92–99.

SHNEIDERMAN, B. & WATTENBERG, M. (2001). “Ordered treemap layouts.” In *IEEE Symposium on Information Visualization, INFOVIS 2001*. IEEE. 73–78.

SIMON, H. (1996). *The Sciences of the Artificial*. Massachusetts: The MIT Press.

- STAUDT, C. (2011). "Analysis of scientific collaboration networks: social factors, evolution, and topical clustering." Germany: Karlsruhe Institute of Technology. Available from:
- THACKARA, J. (2006). *In the Bubble: Designing in a Complex World*. London: MIT Press.
- TUFTE, E. (2009). *The Visual Display of Quantitative Information*. USA: Graphics Press LLC.
- WRIGHT, A. (2008). *Glut Mastering Information Through the Ages*. Ithaca and London: Cornell University Press.
- WURMAN, R. (2001). *Information Anxiety 2*. Indiana: QUE.

