View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Repository of Nicolaus Copernicus University

Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 233–239
ISBN 978-83-60810-51-4

Information Retrieval across Information Visualization

Veslava Osinska
Institute of Information Science
and Book Studies
Nicolaus Copernicus University
ul. W. Bojarskiego 1
87-100 Toruń, Poland
Email: wieo@umk.pl

Piotr Bala
Faculty of Mathematics
and Computer Science
Nicolaus Copernicus University ul.
Chopina 12/18
87-100 Toruń, Poland
Email: bala@mat.umk.pl

Michal Gawarkiewicz
Faculty of Mathematics
and Computer Science
Nicolaus Copernicus University
ul. Chopina 12/18
87-100 Toruń, Poland
Email: garfi@mat.umk.pl

Abstract*—This article presents the analytical and retrieval potential of visualization maps. Obtained maps were tested as information retrieval (IR) interface. The collection of documents derived from the ACM Digital Library was mapped on the sphere surface. Proposed approach uses nonlinear similarity of documents by comparing ascribed thematic categories and thereby development of semantic connections between them. For domain analysis the newest IT trend - Cloud Computing was monitored across time period 2007-2009. Visualization reflects evolution, dynamics and relational fields of cloud technology as well as its paradigmatic property.

I. Introduction

USER'S information needs have made the Information Retrieval today the prominent and exciting field for scientists, physicians, enterprise and business analysts, information managers, librarians and any others who deals with a large-scale collections of data. Document retrieval systems are based on the theoretical models, where the most prevalent are Boolean, Vector Space, Probabilistic, and Language Modeling. The basic action in information retrieval is to compare an automatically produced index of the textual content of documents with the user's request [1,2]. This connection applies to text or content based (semantic) indexing.

From users perspective document search is still not a solved problem. Search engines find too many results that means too low precision of retrieval system or too few caused by their small knowledge how to formulate the best matching query. Nevertheless, the users show some common strategies, for example finding more documents similar to the one already found. This technique is known as pearl growing [3]. Current information retrieval systems provide this model by embedding it within the interface. Collection of suggested terms is derived from such units as synonyms, close indexing terms, thesaurus as well as the list of previously entered queries by users [4]. Examples of pearl growing models can be found in Google results "similar links", in Amazon's category "Customers Who Bought This Item Also Bought" and in any e-commerce service's with item: "related articles". Hence user-friendly information retrieval systems need to use option for associate context search.

Output results in the form of list ranking (Google, Yahoo) do not satisfy the searcher because of linearity. Some Web search engines besides page ranking allow advanced functions such as results grouping according to the topics or categories, visualization of results and social tagging. Linear ranking list is not sufficient for similar documents representation. In complex context the non-discrete property – similarity of documents must be described in more sophisticated way. Visual maps of retrieved results join such advantages as fuzzy representation, non-linear localization and topic diffusion. Maps provide a physical (geographical) structure for comparisons of measured objects as well as an understanding the organization of measured environment [5]. Further more maps also help us easy navigate the landscape of findings.

In this article we focus on retrieval versus topological characteristics of visualization maps. We have chosen a sphere surface as the mapping space of the collection of documents derived from the ACM Digital Library. Obtained visualization maps were tested as information retrieval (IR) interface. Studying map pattern across discrete years of documents publishing (longitudinal mapping [6]) it is possible to see the dynamics of changes within scientific domain.

For such analysis we have selected a newest IT trend which is Cloud Computing. Cloud, the most popular word/metaphor today presents both narrow, bigger and fuzzier meaning. This is at the same time a model of technology, model of computing providing web-based software as well as business model of providing resources to the user. Cloud is considered as a service giving access to the resources on demand. Some analysts define cloud computing as an updated version of utility computing: virtual servers available over the Internet. Others argue that anything we use outside the firewall is "in the cloud," including conventional outsourcing [22-29]. But this fashionable phrase has a long history and provokes controversy with regard to its source. Apparently longitudinal mapping of CS literature facilitates to study development of the concepts and new ideas in the interdisciplinary fields.

II. Mapping space review

Visualization 3D is current trend in graphic design, simulation and modeling. One can find a lot of arguments supporting the systems with dominance of spatial visualization. It is natural to say that we live in a four dimensional world

978-83-60810-51-4/$25.00 © 2012 IEEE

and our brain is designed to process information in three dimensions. Original image mapped to the spherical retina also has the characteristics of a spherical 3D structures [7].

It seems reasonable to use spherical data visualization methods adapted to the natural abilities of the human visual apparatus. But so far the two-dimensional visualization is still the easiest way of presenting the results.

By mapping multiscale data researchers face difficulties: large amounts of unstructured data must be displayed in a limited space with limited resolution. This main communication problem concerns spatial layout a graph drawing algorithms [8-10]. Last decade more often spatial graphs are used for spatializing, filtering, navigating, manipulating and clustering than plan visualization. To visualize the complex structure in three dimensions one needs to center the main node and place child nodes in all directions around it. There are a lot of open source software for graph and network analysis including wide interactive possibilities (Walrus , Gephi , Pajek etc.).

Maps of science are generated through a scientific analysis of large-scale scholarly datasets. They can be used to identify major research areas, experts, institutions, collections, grants, papers, journals, and ideas in a domain of interest [9, 10]. A lot of visualization layouts is limited to regular shapes, as circle or sphere. For example Science–Related Wikipedian Activity map [11-13] shows the structure and dynamics of the English Wikipedia based on 659,388 articles and their editing activity. The similarity of each article-article pair was calculated as the number of shared links to other articles. The final graph layout generated by VxOrd [14] routine is circular. According to the intention of authors, visualization has to highlight current trends and predict future editing activity and growth in Wikipedia articles related to science, technology, and mathematics. Next noticeable circular map layout produced by the same tool is visualization of the scientific paradigms. The authors used the VxOrd to recursively cluster the 820,000 most important papers referenced in 2003; in result they obtained 776 scientific paradigms [15].

Most frequently cited map of global science is Forecasting Large Trends in Science [11, 15, 16] within 3D space. The authors visualized 7.2 million papers and over 16,000 separate journals, proceedings, and series from a five-year period, 2001-2005. The metrics is based on a combination of the bibliographic coupling of references and keyword vectors. Using three dimensional graph they achieved spatial visualization of disciplines on a sphere, and to give two-dimensional version of map Mercator projection was used.

In the present work a sphere surface was selected as a target mapping space. Sphere surface has no edges and therefore it is possible to represent not only local similarity but also large-scale regarding to all space. The benefit of curved surface in comparison to a plane is more capacious exploration space. We have exploited digital library resources, so the aspect of ergonomic user interface for browsing, navigating and searching was also important. Such features as symmetry and continuity of sphere made it natural perceptional space. However for detail topological analysis of visualization maps we return to two dimensional representation through equidistant cylindrical projection. The reason is easy production and resemblance to geographic maps that everyone can read and interpret. Proposed approach uses nonlinear similarity of documents by comparing ascribed classes and thereby development of semantic connections between them.

III. The methodology of mapping

The articles [17-19] describe in detail the construction of graphical representation of original classification scheme adopted in ACM Computing Classification System. Thus we confine ourselves to a summary of steps.

Dataset consist of collected metadata of ACM Digital Library scientific articles regarding mainly computer science. Every article is ascribed to main thematic class/subclass and some additional ones. Overlapping classes and subclasses therefore appear simultaneously among documents collection. The author's idea consisted in estimation of co-occurrences of classes i.e. counting of common documents for every pair classes and subclasses. The larger number of common publications the larger thematic similarity of co-classes. The fact that authors of articles participate in classification is in favour of our procedure.

The final number of all possible classes and subclasses in collection was 353. This is the dimension of similarity matrix of co-classes. As similarity measure we used normalized IC-cosine [5]:

$$\cos_{i,j}=\cos_{j,i}=\frac{(RAW_{i,j})}{\sqrt{\sum_{k=1}^{n}C_{i,j}\sum_{k=1}^{n}C_{j,k}}}$$

To decrease matrix dimension we have used an MDS-based scatterplot selecting a sphere as output space. For this reason the nodes were considered as single particles under Morse potential [17].

Among 353 (sub)classes nodes, positions of articles were calculated from topological relations between main and additional classifications with weights 0.6:0.4 accordingly. All documents nodes were marked by their main class color, thus the final visualization palette consists of 11 colors.

Thus the main classes are: A. General Literature; B. Hardware; C. Computer Systems Organization; D. Software; E. Data; F. Theory of Computation; G. Mathematics of Computing; H. Information Systems; I. Computing Methodologies; J. Computer Applications; K. Computing Milieux.

For convenient analysis cartographic projections of visualization layouts were used. Fig 1 represents visualization on a sphere and Fig 3 - its projections to plane according 2007 and 2009 years data. Application allowing for visual comparison of changes over years of publishing are accessible on-line[1].

[1]http://www-users.mat.uni.torun.pl/~garfi/vis2009/ - the best view is with Mozilla Firefox or Chrome browsers.

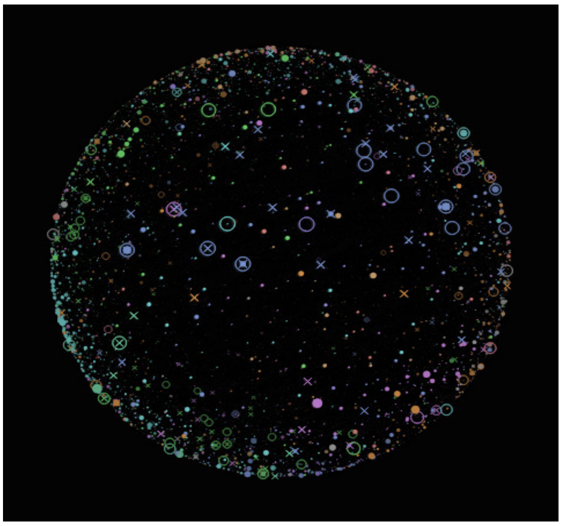

Fig 1. Visualization of classes (bigger circles) and documents (scattered points) nodes on a sphere surface. Year of publishing is 2009. Application is accessible on-line[1].

IV. Maps comparison[2]

A. CS analysis

Visualization process was repeated for different publishing years with ca. 10 years step: 1968, 1978, 1988, 2007 and 2009. Comparison of visualization maps (literature longitudinal mapping [6, 8]) allows to track and analyze dynamics of scientific domain. Another possibility of visual analysis refers to study how knowledge advances and knowledge organization changes [9,10].

ACM Digital Library's dataset changes in time is shown on Fig 2. Because of the performance not all articles were classified and for the preparation of the knowledge maps the representative part has been selected. CCS taxonomy falls behind the emergence of new thematic categories. The reason can be the crisis of classifications systems in the face of keyword searching. It is noticeable in last two decades the quantity of classified publications are similar and is about 30000.

In the previous papers [17-19] the distributions of documents nodes depending on time were characterized. The feature that the most ontologically different Hardware (B class) and Software (D class) are distributed in the opposite corners (poles in case of sphere) could be considered as verification of mapping. Maps revealed more or less uniform distribution of documents till 90th. The results from 1988 show how the category of Information Systems (class H) disseminates and CCS started to evolve. This is also the clustering time. Next the Computer Systems (class C) i.e. networks were quickly developed. However class C as networks category places between them because of both problems are represented. Comparing maps in time scale we concluded that in the last two decades classification evolves towards stronger adaptation of CCS structure in ACM digital library.

[2]For precise reading and interpretation we put the colored versions of generated maps with bigger resolution at the website: http://www.umk.pl/~wieo/infovis2009

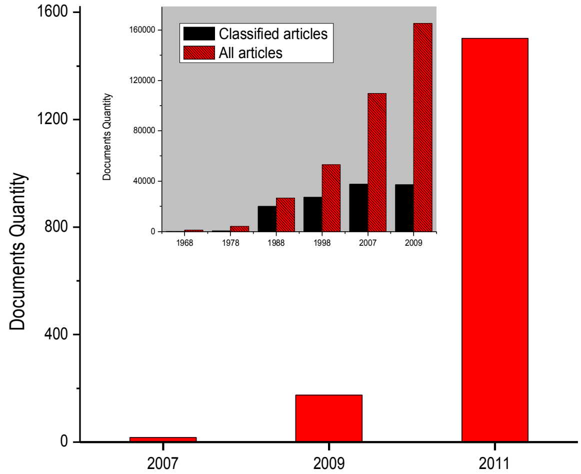

Fig 2. The quantity of documents with searching terms "cloud computing" since 2007 year of publishing. Miniature chart shows the quantity of publications according year of publishing in ACM Digital Library. Black columns relate to classified documents.

Current paper concentrates on visualization and retrieval of documents published in 2009; the output layout is presented on Fig 3. Nevertheles with the similar quantity of documents (more than 37 000), the map reveals more uniform clustering than in 2007. Nodes of class I - Methodologies form continuous strip like a sinusoide. This category refers to problem solving and analysis using information technology. It covers: computer graphics, image processing and recognition, text processing, simulation and modeling, as well as artificial intelligence. The central arrangement indicates its present importance among other research fields. Information systems (H class) manifest the biggest changes in structure. Nodes "follow" I class nodes, that is computer scientists have comprehensive approach to describe methodology and need simultaneously to work out testing systems. Considering Hardware and Software nodes, the latter ones dominate according to its quantity and cohesion. Interesting observation is that these groups locate closer one another which points to integration software application and devices on every level. Reduced visibility in terms less significance of theme is characteristic of Applications (J) and Milieux (K).

B. Documents clustering and Cloud Computing

Spatial representation of publications nodes depicts their thematic closeness. Nonlinear approach through counting of co-classes determines similarity measure. Documents whose topics are similar must be located close each another regardless of the classes they are assigned. Articles nodes are arranged in the area around the proper node of main class. This distance depends on location of an additional class(es) and their quantity. For example, the node of document which belongs to the class C (Computer Systems Organiza-

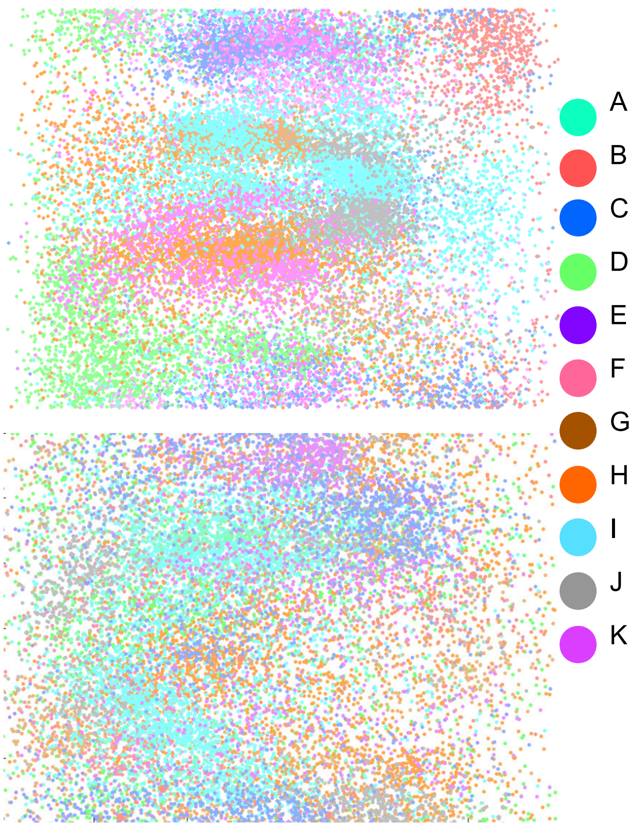

Fig 3. Visualization layout of documents published in 2007 (upper) and 2009 (bottom). From the right is the legend of main classes symbols with ascribed colors: A. General Literature; B. Hardware; C. Computer Systems Organization; D. Software; E. Data; F. Theory of Computation; G. Mathematics of Computing; H. Information Systems; I. Computing Methodologies; J. Computer Applications; K. Computing Milieux

tion) may be "expelled" near class J (Computer Applications) located in the different hemisphere.

IR possibilities of given graphical layout can be tested by tracking positions of thematically similar publications. To catch latter we retrieved collections by such metadata as keywords, title and abstract. We selected the term "cloud computing" because of young concept and quick expansion of this service in the world.

Cloud computing as a delivering computing resources through a global network has evolved through a number of services and concepts like grid and utility computing, application service provision, and Software as a Service [21]. Although the cloud computing is one of the hottest terms in the technology it has quite long history [25]. According to the sources [24, 25], the first scholarly use of phrase was in internal Compaq analysis titled "Internet Solutions Division Strategy for Cloud Computing" dated November 14, 1996. The power of term is its marketing value. It is generally considered to be born in 2006-2007 when cloud computing has to be provided to external customers and shows a steady increase in interest. Cloud computing is an evolving paradigm [19] that is rewarding object of study in information visualization because of distinct research frontiers are easy exposed due to scientific paradigms on domain maps [7-9].

Searching this expression across collections we found its occurrence since 2007. Fig 2 shows how the frequency of phrase increases rapidly from 2007 (13) through 2009 (175) to 2011 (1503). In the first case (on Figure 4) localization of a few objects is only interesting as source categories from which it starts to disseminate. Three clusters were noticed near the following themes (sub/classes nodes): Computer Systems Organizations (C), Software (D) and Information Systems (H). In order to learn more about the origin of "cloud" space we analyzed the closest neighboring of found nodes. Authors characterized these articles by following keywords shown in Table 1. The central, biggest cluster covers grid computing, Web services, data mining, peer-to-peer, semantic Web. In the area of neighborhood there is also a publication concerning condensed matter physics (the title: "Soliton trains and vortex streets as a form of Cerenkov radiation in trapped Bose-Einstein condensates" - (BECs). It is possible to find relationship with cloud computing if we recall the BECs experiments are used in quantum computations and quantum resonance imaging [22]. Next pair of keywords describes knowledge representation formalisms. Scientists consider a cloud concept as a way to solve problems of knowledge explosion and "information overload" over the past few decades.

Bottom cluster is specialized in information queries, algorithms, combinatory and imprecise computation. Upper cluster is computer systems security and stability oriented. Thus in 2007 grid computing publications discuss algorithmic, programming and technical aspects of information systems organization. The majority of documents reveals close relation with distributed computing issues. The most representative publication in filtered collection is "Computing in the clouds" [23].

Mapping of 2007 publications hits the origin of cloud. Till this year no search results for cloud computing in ACM Digital Library. The most important is possibility to see semantic source and paths of cloud idea by studying given clusters on 2007 map.

Map from 2009 represented on Fig 4 shows the new property for cloud topic: continuity through whole layout. Instead of clustering like in 2007 we see the coherent topics path of selected documents nodes. It is possible to discover cloud pattern follows the Methodologies (class I) nodes making a sinusoide like trajectory. Methodologies is the most diverse thematic category in ACM CCS covering contemporary problems of AI, knowledge representation, machine learning, image processing and computer vision, simulation and modeling as well as text processing.

Summarizing foregoing cloud articles are located nearby class I nodes and simultaneously are attracted to area with a big concentrations of H class nodes (Information systems). These two categories are crucial to cloud topic documents in early stage. One should be noted from right there is significant concentration of nodes in the same place of Applications (I) nodes cluster. It proves cloud concept rapidly evolved from methodologies to application. Innovation in the computing industry continues to enable new opportunities for information society. Business and government orga-

TABLE 1
THEMATIC CHARACTERISTIC OF THREE CLUSTERS (SEE FIG. 4) WITH "CLOUD" DOCUMENTS PUBLISHED IN 2007

Cluster	**Categories**	**Keywords**
central	Systems and so-ftware, online information services, informa-tion retrieval, project and people management	retrieval, latent semantic analysis, lexical disagreement problem, query log clusters, clustering, data mining, peer-to-peer, Grid services, collaboration, portlets, environmental data distribution, grid computing, resource broking, Grid portal, GridPortlets, GridSphere, OGCE, comparison, business process management, Semantic web, specification integration, UML meta-models, Web services, counting disjoint, optimal partition, sortability, Bose-Einstein condensation, Matter waves, nucleation, solitons, vortices, generalized implicatures, literal meaning, nonmonotonic logics.
bottom	distributed systems, process management, public policy issues	Information query, mobile ad hoc networks, time indexed information, distributed system, insider threat, autonomic computing, cellular automata, grid computing, algorithms, analysis of algorithms, combinatorial problems, imprecise computation task, polynomial time algorithms, preemptive scheduling, uniform processors .
upper	Network architecture, decision problems	audit, security, service learning, asymptotic stability, congestion control, heterogeneous delay, overlap-free words, formal languages, Thue--Morse word, rewriting logic, semantics and analysis of programming languages.

nizations as well as engineering and medical centers looking to drive down costs implement cloud solutions in IT infrastructure. Large data centres offer economies of scale, providing cheaper computing power, with the flexibility to pay. The critical issues for interoperable cloud applications concentrate around security, privacy, portability and availability of data.

The same state of cloud articles connection with K class (Milieux) nodes is observed on the top of layout. This minor category covering imprecise fields concerning computing and society, social issues of computation, organizational impacts, legal aspects and so on.

It is possible to observe from Fig 4 cloud computing relies on grid computing as its backbone and infrastructure and generally grows out distributing systems (C class). Cloud clusters overlap with distributed computer-communication networks (C class). This evolution may be considered as a result of a shift in focus from an infrastructure that delivers storage and compute resources to an economy based aiming to deliver more abstract resources and services [28, 29].

The pictures show that within two years cloud computing has been applied in the more global scale of CS. Big concentration of both H, C and J nodes in the cloud range indicates researchers successfully implement cloud idea in an information systems, networks and computer applications. Definitely cloud can not be limited to primary grid category, i.e. Computer Systems class (C). We see a weakness of ACM CCS current taxonomy. Cloud as a multifaceted concept became adapted in a wide spectrum of technological, business, social and education problems.

V. SUMMARY AND CONCLUSION

Information space must illustrate semantic correlations between scholar publications regarding computer science. We have used nonlinear similarity measure of co-classes from ACM classification system that allow to arrange their in output space and then on visualization layout. Documents are represented by many subclasses and have no one occurrence on the classification tree. Their nodes were located between classes/subclasses nodes by using geometrical principle. If place of classes/subclasses determine wide or narrow research and technology, articles with similar topics positions close each other on visualization layout.

Thus information space depicts semantic similarity between documents belonging to ACM digital library. Topology of neighboring documents nodes on a visual map decides about their semantic relations. This property may be useful in associate context retrieval. For example when we search an article on-line any search engine now offers option "related" or "similar". If output visualization was enriched by feedback regarding the user it could be served as documents retrieval interface.

We constructed visual IR system with inverse data processing: first stage includes visualization, second – semantic retrieval by using maps. From the perspective of IR systems evaluation, effectiveness of thematically similar documents retrieval can qualify visualization results. To test visualization-searching system we investigated the neighborhood of filtered documents with the query "cloud computing" by following fields: keywords, title and abstract. Increased number of data from 16 in 2007 year to 175 in 2009 and 1500 in 2011 points to cloud methodology spreads very rapidly that time. This is agreed that the term is starting to be known in global technical and economical scale since 2007. Documents distribution varied from three separate clusters (2007) to "satellite path" along computing methodologies which cover diverse spectrum of current CS issues (2009). In 2007 cloud clusters show significant coherence with network architecture, distributed computing, systems and software issues, process management. Two years later cloud computing nodes quickly disseminate over whole map and are strongly attracted by methodologies class.

The authors of ACM publications have assured them the best topic characteristic by introducing keywords. Investigated neighborhood of cloud articles by reviewing keywords depicts common roots with grid computing: distributed sys-

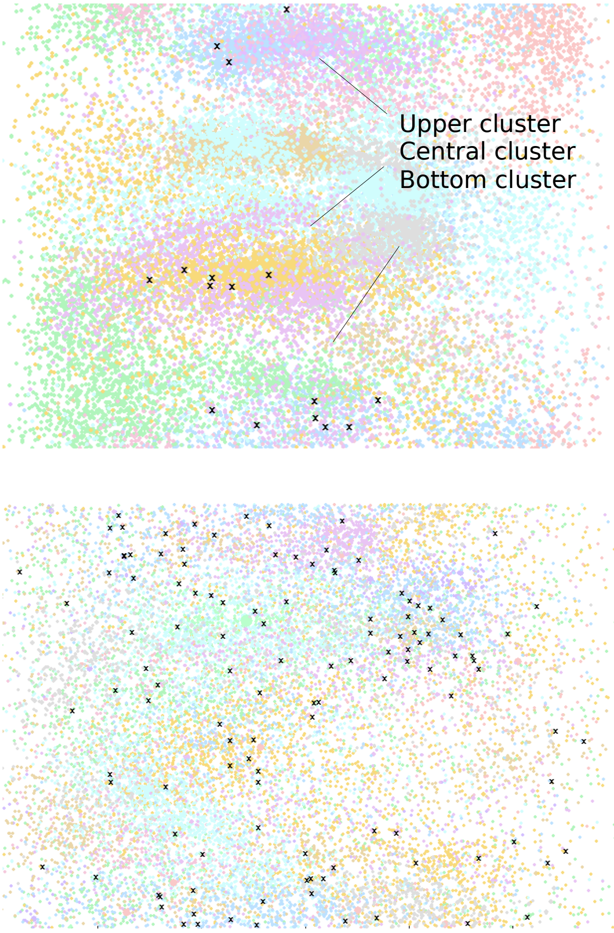

Fig 4. Distribution of articles with terms "cloud computing" on visualization maps generated for 2007 (upper) and 2009 (bottom) year of publishing.

tems category. Cloud is the technology which exploited grid concept especially in early stage.

Thus by using clustering patterns on visualization maps we show how emerged and evolves the current multifaceted concept - cloud computing, including a wide spectrum of their technological, business, social and education issues. Therefore described tests allow to discover historical basis, etymology and relative concepts of initial subject and thus research fields. The question is: how the visualization pattern may facilitate current domain/field development prediction?

VI. Discussion

Beside the visual analysis of all documents distribution we studied thematic clusters organization by counting keywords frequency [18-20]. This way keywords map of dataset regarding computer science was obtained and published. Visualization maps including all layers of mapping like keywords map, semantic map, co-authors map have a big potential in domain analysis. Future research plans concern comparison tests between proposed approach and traditional visualization methods.

Due to visual patterns analysis it is possible to study semantic similarity of documents as well as track where scientific paradigms or technological jumps were appeared.

Exploring visual map which is a mine of semantic knowledge could be considered a new research field "map-mining" equally to data-, text- or webmining.

References

[1] E. D. Liddy, "Automatic Document Retrieval" in *Encyclopedia of Language and Linguistics,*E. D. Liddy, 2nd ed, Elsevier Press, . 2005.
[2] A. Singhal, "Modern Information Retrieval: A Brief Overview". Bulletin of the IEEE Computer Society Technical Committee on Data Engineering 24 (4): 35–43, 2001, http://singhal.info/ieee2001.pdf.
[3] P. Morville, S*each Patterns*. Sebastopol, CA: O'Reily, 2010.
[4] J. Kalbach, *Designing Web Navigation*. Sebastopol, CA: O'Reilly, 2007.
[5] K. W. Boyack et al., "Mapping the backbone of science". *Scientometrics*. Vol. 64, no. 3, pp. 351-374, 2005.
[6] E. Garfield, "Scientography: Mapping the tracks of science". *Current Contents: Social & Behavioural Sciences,* no. 7(45), pp. 5-10., 1994
[7] J. C. A. Read and B. G. Cumming, "Does depth perception require vertical-disparity detectors?", *Journal of Vision,* Vol. 6 (12) , A. 1, pp. 1327, 2006.
[8] K. Börner, Ch. Chen and K.W. Boyack, "Visualizing Knowledge Domains", In: B. Cronin, Ed. *Annual Review of Information Science & Technology,* Medford, NJ: Information Today, Inc./American Society for Information Science and Technology, Vol. 37, pp. 179-255, 2003.
[9] K. Börner, "Extracting and Visualizing Semantic Structure in Retrieval Results for Browsing", In Proceedings of the fifth *ACM conference on Digital libraries,* NY, USA:ACM, 2010.
[10] Ch. Chen, *Information Visualization*. *Beyond the Horizon*. 2nd ed. London: Springer, 2006.
[11] "Exhibit Purpose and Goals" [on-line]. *Places@Spaces: Mapping Science,* Accessible at World Wide Web: http://scimaps.org/.
[12] K. Borner, *Atlas of Science*, MIT Press, 2010
[13] T. Holloway, M. Božičević, and K. Börner, "Analyzing and Visualizing the Semantic Converage of Wikipedia and Its Authors." *Complexity* 12 (3), pp. 30-40, 2007.
[14] "DrL: Distributed Recursive (Graph) Layout (was VxOrd)". https://nwb.slis.indiana.edu/community/?n=VisualizeData.DrL.
[15] R. Klavans, and K. W. Boyack. "Is There a Convergent Structure to Science? A Comparison of Maps using the ISI and Scopus Databases." In Proceedings of the 11th International Conference of the International Society for Scientometrics and Informetrics, D. Torres-Salinas and H. F. Moed, Ed. 437-448. Madrid, Spain: *Society for Scientific Information and Documentation*, 2007.
[16] R. Klavans and K. W. Boyack, "Quantitative evaluation of Large Maps of Science" *Scientometrics* 68 (3), pp. 475-499, 2006.
[17] V. Osinska and P. Bala, Classification Visualization across Mapping on a Sphere. In: *New trends of multimedia and Network Information Systems*. Amsterdam: IOS Press, pp. 95-107, 2008.
[18] V. Osinska and P. Bala, Nonlinear approach in classification visualization and evaluation. In: *New perspectives for the dissemination and organization of knowledge:* Proceedings of the IX Spain Group ISKO Congress 11-13 March, 2009, Valencia, Spain. pp. 222-231, 2009, Accessible at World Wide Web: http://dialnet.unirioja.es/servlet/fichero_articulo?codigo=2923178
[19] V. Osinska and P. Bala, New Methods for Visualization and Improvement of Classification Schemes – the case of computer science. *Knowledge Organization*. 37 (3), 2010.
[20] V Osinska, Visual Analysis of Classification Scheme. *Knowledge Organization*, 37(4), 2010.
[21] The NIST Definition of Cloud Computing". National Institute of Science and Technology, Accessible at World Wide Web: http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf. Retrieved 24 July 2011.
[22] T. Hacht. "Quantum Computation with Bose-Einstein Condensates", *Ph.D. dissertation, Technische Universitat Munchen, Max-Planck-Institut fur Quantenoptik.,Munchen,2004.*
[23] A. Weiss. Computing in the clouds. In *netWorker*, NY, USA: ACM, Vol. 11 Issue 4, Dec. 2007.

[24] M. Armbrust, “A View of Cloud Computing”, C*ommunications of the ACM,* Vol. 53 No. 4, pp. 50-58, 2010.

[25] U. Banerjee, “Cloud Computing – Important Events till 2010”. *Technology Trend Analysis*. Accessible at World Wide Web: http://setandbma.wordpress.com/ March 8, 2011.

[26] A. Regalado. “Who coined the term “Cloud Computing”? *Technology Review,* Oct. ,2011, Accessible at World Wide Web: http://www.technologyreview.com/business/38987/?mod=chfeatured

[27] C.A. Julien, J.E. Leide and F. Bouthillier, Controlled User evaluations of Information interfaces for Text Retrieval: Literature Review and Meta Analysis, *Journal of American Society for Information Science and Technology*, 59(6): pp. 1012-1024, 2008.

[28] E. Knorr and G. Gruman, “What cloud computing really means”, Accessible at World Wide Web: http://www.infoworld.com/d/cloud-computing/what-cloud-computing-really-means-031.

[29] I. Foster et al., “Cloud Computing and Grid Computing 360-Degree Compared”, In 2008 *Grid Computing Environments Workshop*, pp. 1-10, 2008.