

Trends Discovery in the Field of E-Learning with Visualization

Muhammad Salman Khan, Martin Ebner, Herman Maurer
Technical University of Graz
Graz, Austria
skhan@iicm.edu, martin.ebner@tugraz.at, hmaurer@iicm.edu

Abstract: In this paper we have performed a content analysis using a collection of papers from Ed-Media conferences archive. Various current trends in research topics, change of interest in contributions from across the world in the field of e-learning have been described using our internally built visualization tool. This visualization tool reveals the evolution or decline of research themes, hot research topics and key researchers across the world. In addition to this, evolution of countries, continents and institutions across research topics has been described as well.

Introduction

In any academic discipline the research publications represent the knowledge structure of that discipline. This knowledge structure represents the research trends, landmark papers, key researchers, network of scholarly papers and researchers, collaborations and contributions from different institutions and regions. Much can be learned by analyzing the research contributions in a journal or conferences of any discipline about a given field of study (Taylor, 2001). The benefits of such analysis is twofold; first it can be used to evaluate individuals, organizations, groups and nations which in turn can be used to know the impact of decisions and policies made for allocating resources and funds. Secondly it reduces the researchers' menial efforts to conduct surveys themselves and shows them a broader picture to understand the field of their interest (Boerner et al., 2003).

Traditionally, researchers have tried to analyze different patterns in the scholarly publications using normal tables and statistical charts (Taylor, 2001), (Marcouiller & Deller, 2001). Interactive visualizations have been used by (Ahmed et al., 2004), (Ke et al., 2004), (In-SPIRE, 2004) to realize different patterns such as citations networks of papers, number of papers over time, trends of research areas and the correlated research areas in the eight years of InfoVis conferences. In (Erten et al., 2004), the authors have used 10,000 unique ACM computer science papers to analyze research trends and collaborative networks of researchers in the field of computer science. All the systems mentioned above are good tools to understand the research trends, networks of papers, authors and research topics. But they do not demonstrate and compare the change of interest in publications contributions and research areas across different regions and institutions over the period of time.

Over the years many studies have been conducted regarding the role of hypermedia, multimedia and telecommunications for providing education. The World Conference on Educational Multimedia, Hypermedia & Telecommunications (Ed-Media) is an international conference, organized by Association for the Advancement of Computing in Education (AACE) annually since 1998. The conference aims at providing a multidisciplinary forum for the discussion, exchange of information regarding the new research, developments and applications on all topics related to hypermedia, multimedia and telecommunications/distance education (Ed-Media, 2008).

The importance of analyzing trends in distance education or e-learning has also been realized by many researchers. In [Lee et al., 2004], the authors have exposed the hidden trends patterns by examining four well known distance education journals from 1997 to 2002. In [Shih et al. 2007] the authors considered five e-learning journals for their analysis. The studies provided insightful information about overall research themes, research methods, research trends and important papers. Our work is also among these lines where we seek to uncover different research trends in the field of e-learning. But instead of finding overall trends in research topics or themes we are also considering locations (continents, countries and institutions) to determine in detail that how the world is progressing in the field of e-learning.

The aim of this study is to allow the researchers in our field to understand what kind of different research areas exist and to identify different patterns over the last six years of Ed-Media conference using our interactive visualization tool. The visualization tool which is primarily based on animated bubble chart and pie chart in particular will help to identify the trend of contributions in the field over the years, different research areas that have evolved or diminished over the period across the globe, hot research topics, key researchers across the globe, leading institutions and nations in any particular research area, evolution of countries, continents, and institutions in comparison to each other. This will provide an overview of how the world is progressing in the field, international standing and impact of the field.

In the next section we describe the techniques we adopted to extract and normalize the papers metadata of the Ed-Media conferences.

Ed-Media Data Extraction

Each paper published in Ed-Media has a well formatted html page in the AACE digital library. We parsed these html pages from year 2003 to 2008 using regular expressions to extract metadata about each paper. This includes paper title, publication year, abstract, keywords, paper length (page start, page end), authors' names, institutions and countries and stored them in a relational database.

The country information of the authors needed to be cleaned and standardized as the data contains various representation of the same location (e.g. US, United States). In order to rectify these, we compared the countries data with GeoBytes database (GeoBytes, 2008) containing the names of all countries and cities across the globe. Countries that had no match with the GeoBytes database were then identified and corrected accordingly. There were 33 authors whose country information was found missing. In order to cope with this problem, we first parsed the names of the institutes to locate the name of the country. In cases where this method failed we searched for the same institute name for other authors to look for the existence of country information. By using the above mentioned approaches, we succeeded to find the countries information of 16 authors.

In many cases different institutions can have the same name and same institution can have different names. For accuracy of results there is a need to disambiguate the institutions names. In our case we have applied a simple string matching algorithm called n-grams found in (Alvisio et al., 2007) to disambiguate the institutions names. To get better results advanced techniques also exist that include dictionary and matching rules (Yang et al., 2007). However in many cases human intervention is also required to resolve this problem.

Moreover in our study we have considered 99 unique keywords assigned to the papers in Ed-Media as research topics. After applying all above approaches our database comprise a total of 4607 papers containing 8186 authors and 2911 institutions. This data is given as input to our internally built visualization tool. In the next section we describe the working of the visualization tool and some interesting results drawn from it have been discussed.

Experimental Results

In this section, some interesting results are presented that can be obtained by the visualization tool. The main interface of the visualization tool is shown in Fig. 1. The user has the choice to view the hidden patterns in the publications across the world either as a single entity or across the continents, countries and institutions. The results can also be filtered by selecting any particular location or topic from the given lists. A temporal slider has been provided to scroll across different years. Each bubble on the animated bubble chart represents a location which can be a country, continent, institution or the entire world as a single entity. The color and size of the bubble represents the location and the number of papers respectively. The axis of the animated bubble chart contains various options in which the user might be interested, such as number of institutions, number of authors, number of papers, average length of papers and average number of authors per paper. The pie chart represents the distribution of publications across topics for any particular country, continent, institution or the whole world selected by the user from the bubble chart at a particular year. By clicking on any bubble, a list of authors appears which provides the affiliation information and respective publications of authors corresponding to the selected bubble. In this way one can also find an

expert based on total number of publications over the years for any particular field and location. In order to better understand and facilitate an easy analysis of the results, we have divided the publications in three groups each spanning to 2 years i.e. 2003-2004, 2005-2006 and 2007-2008 inclusive.

The following sub-sections represent some interesting results based on four different kinds of views i.e. world, continents, countries and institutions.

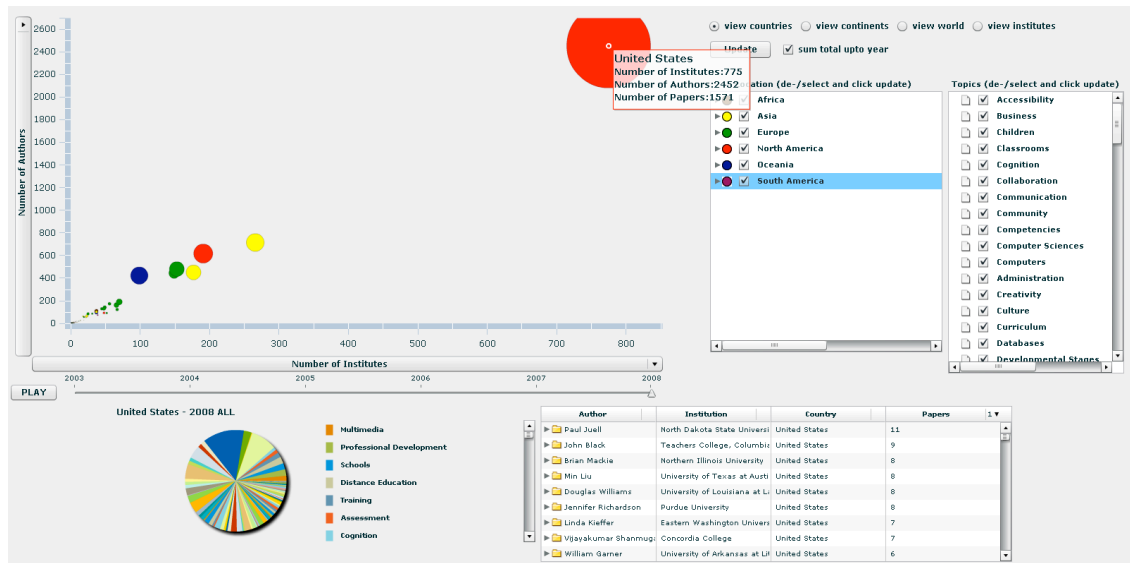


Figure 1: Main Interface.

World View

This view reflects all publications as a single entity. The results revealed that up to year 2004 the total number of publications, authors and institutions were 1700, 3210 and 1358 respectively. It was also observed that there is a sudden decline in them for the time period 2005-2006 (1311 paper, 2233 authors, and 693 institutions) and again a rise for the time period 2007-2008 (1596 papers, 2743 authors and 860 institutions). These statistics also reflect the inclusion of new authors and institutions in the field instead of being occupied by some groups of authors.

The user can also view the performance of each author across the world based on the number of publications. The top authors across the world can be identified by sorting the list using number of papers. It can be seen in Fig. 2 that the author with the highest number of papers is “Kanji Akahori” from “Tokyo University of Japan”. The user can also view the per year performance of any author. The Fig. 2 also demonstrates the performance of “Kanji Akahori” over the years. It is clear from the figure that he has been publishing consistently in the field. In this way a user can find the expert of a field at any particular location.

The user also has the choice to view and compare the evolution or decline of the topics over the period of time. The Fig. 3 illustrates the distribution of publications across top ten topics, it is clear from the figure that the top two topics are “Students” (2003-2004: 380, 2005-2006: 343, 2007-2008: 431) and “Educational Technology” (2003-2004: 298, 2005-2006: 212, 2007-2008: 325) throughout the entire period. Moreover papers are consistently being contributed in all other topics. Such information is necessary for the students and practitioners of the field as it gives an overview about the emergent or hot research areas of their field.

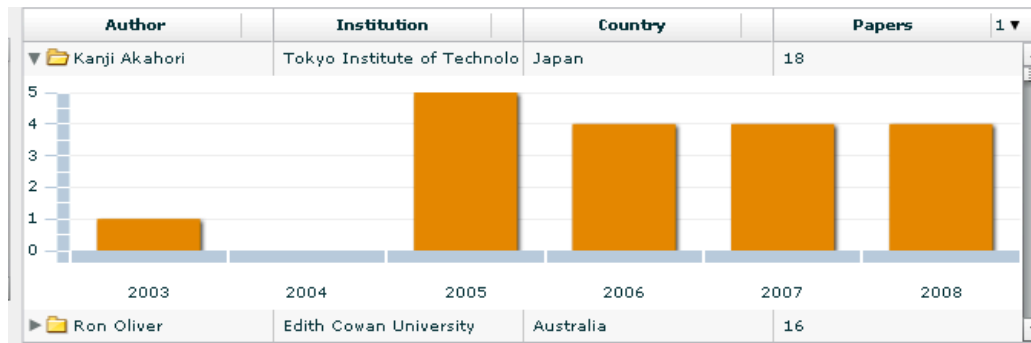


Figure 2. Performance of an author up to 2008.

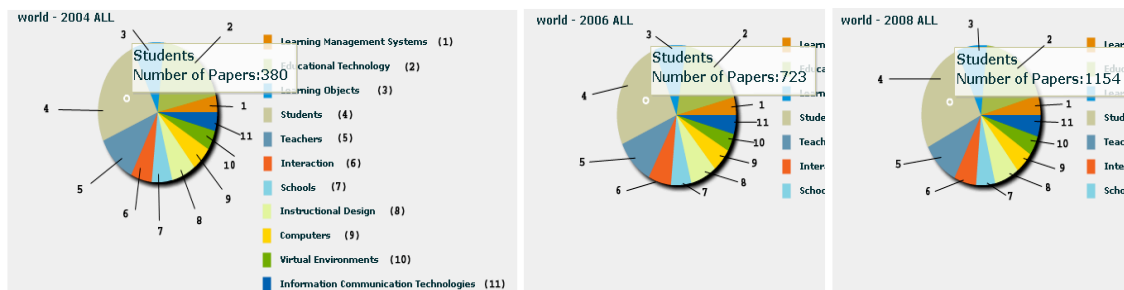


Figure 3. Distribution of publications across top 10 topics.

Continent View

This view represents the distribution of publications across different continents. It can help the users to understand how different regions have evolved with the passage of time, which region occupies the field and which region is active or passive as a whole or in any particular topic, how a particular region is performing as compared to others over the period of time.

The results demonstrated that North American countries which include Barbados, Canada, Mexico, Trinidad and Tobago, and United States remain as the main source of publications in the field for all the time periods followed by Europe (2003-2004: 493, 2005-2006: 338, 2007-2008: 534), Asia (2003-2004: 265, 2005-2006: 300, 2007-2008: 371), Oceania (2003-2004: 169, 2005-2006: 107, 2007-2008: 130), Africa (2003-2004: 39, 2005-2006: 37, 2007-2008: 25) and South America (2003-2004: 23, 2005-2006: 16, 2007-2008: 33). Interestingly, the Asian countries contributions are increasing over all the periods. The emergence of Asia as a big player in distance education has also been predicted by McIssac (McIssac et al., 1996). This analysis also confirms that distance learning or e-learning is rising up in Asian countries.

Countries View

This view further provides insights into the publications patters. It enables the users to understand the participation of each country in the field and provides some meaningful statistics such as when a country started to contribute, when it stopped contributing, which is contributing more or less, which country is strong or passive in any particular research area and how different research topics have evolved in each country, how a particular country is performing as compared to others with the passage of time. The Figs from 4a to 4c demonstrate that Unites States (2003-2004: 684, 2005-2006: 437, 2007-2008: 450) with the highest number of internet users contributed most of the publications and is followed by Canada (2003-2004: 87, 2005-2006: 140, 2007-2008: 139), Japan (2003-2004: 104, 2005-2006: 96, 2007-2008: 144), Australia (2003-2004: 136, 2005-2006: 84, 2007-2008: 113), United Kingdom (2003-2004: 94, 2005-2006: 64, 2007-2008: 130) and Taiwan (2003-2004: 69, 2005-2006: 119, 2007-2008: 99). Interestingly, Canada was behind in publication up till year 2004 but for the rest of the periods it started to evolve and succeeds

other countries. Surprisingly two Asian countries Japan and Taiwan are moving side by side with Australia, Canada and other top countries of Europe.

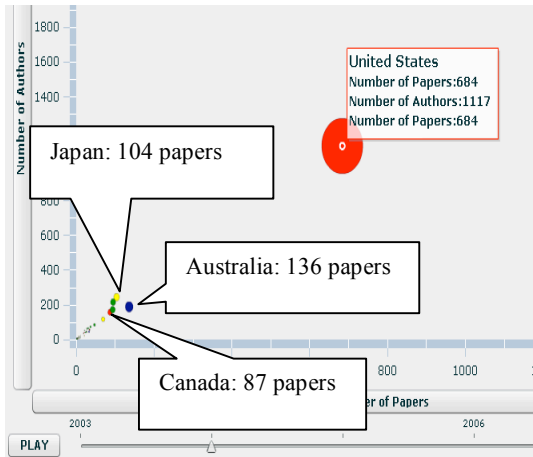


Figure 4.a. Publications across countries up to 2004.

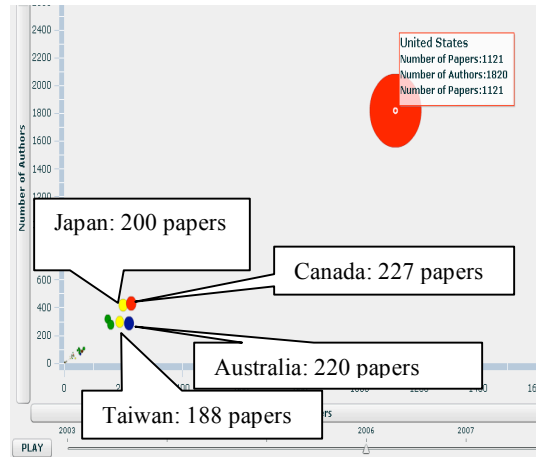


Figure 4.b. Publications across countries up to 2006.

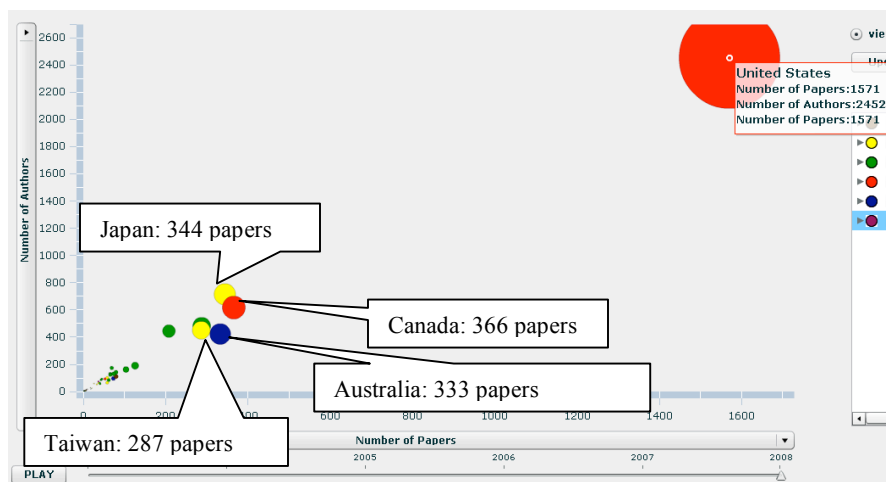


Figure 4.c. Publications across countries up to 2008.

Institutions View

This view demonstrates the contributions of each institution in the field. It enables the users to understand which institution is participating more or less, which institution has more authors or experts of a particular research area, how any particular research area has developed or diminished in an institution and how the institutions are performing in comparisons to each other over the period time. We performed the experiment about the participation of all the institutions from United States. It was observed that the “Teachers College, Columbia University” (2003-2004: 11, 2005-2006: 11, 2007-2008: 13) was not a leading institution in terms of number of publications and authors. But it started to grow both in number of papers and authors gradually and is currently one of the leading institutions in the field of e-learning from the United States. Other leading institutions from United States are “University of Georgia” (2003-2004: 19, 2005-2006: 4, 2007-2008: 4), “University of Texas at Austin” (2003-2004: 10, 2005-2006: 6, 2007-2008: 10), “University of Hawaii at Manoa” (2003-2004: 15, 2005-2006: 3, 2007-2008: 3) and “DePaul University” (2003-2004: 8, 2005-2006: 3, 2007-2008: 13).

Conclusion

We performed a content analysis to find the hidden research patterns in the field of e-learning from 2003 to 2008 using Ed-Media conference papers and our internally built visualization tool. From this we determined trends of contributions in the field, hot research topics, key researchers, evolving regions and institutions in a country. During the extraction of papers' metadata we encountered the problems of institutions names ambiguity. More exciting and precise results can be gained by disambiguating the institutions names using advanced techniques and classifying the papers in appropriate and meaningful categories.

References

- Ahmed, A. Dwyer, T., Murray, C., Song, L. & Y. X. Wu. (2004). InfoVis 2004: WilmaScope Graph Visualization. In InfoVis'04, IEEE Symposium on Information Visualization.
- Alvisio, M. , Nathe, C., Kim C. & Kungurtsev, V. (2007). Disambiguation of the DBLP Database, [Available at http://www.micsymposium.org/apache2-default/mics_2007/papers/Alvisio.pdf]
- Boerner, K., Chen, C. & Bayak, K. W. (2003). Visualizing Knowledge Domains, Blaise Cronin (Ed.) Annual Review of Information Science and Technology, Inc/American Society for Information Science and Technology, chapter 5, pp. 179-255.
- Ed-Media. (2008). <http://www.aace.org/conf/edmedia/>
- Erten, C., Harding, P. J., Kobourov, S. G. & Wampler, K. (2004). Exploring the computing literature using temporal graph visualization. (Tech. Rep. TR0304), University of Arizona.
- GeoBytes (2007). <http://www.geobytes.com/FreeServices.htm>
- Hawkins, D. T. (2001). Bibliometrics of electronic journals in information science. IR, 7, 1, [Available at <http://InformationR.net/ir/7-1/paper120.html>]
- In-SPIRE. (2004). <http://in-spire.pnl.gov/>
- Ke, W., Boerner, K. & Vishwanath L. (2004). Major Information Visualization Authors, Papers and Topics in the ACM Library. In InfoVis'04, IEEE Symposium on Information Visualization.
- Lee, Y., Marcy P. D. & David W. N. (2004). The Past, Present, and Future of Research in Distance Education: Results of a Content Analysis. The American Journal of Distance Education, 18, 4, pp. 225-241
- Marcouiller, D. W. & Deller, S. C. (2001). Thirty years of academic publishing in regional studies: A content analysis of MCRSA's scholarly output, JRAP, 31, 2, pp. 33-43.
- McIssac, Stock, M. & Gunawardene, C. N. (1996). "Distance Education." In Handbook of Research for Educational Communications and Technology: A Project of The Association for Educational Communications and Technology. pp. 403-437.
- Shih, M. et al., 2007. Research and trends in the field of e-learning from 2001 to 2005: A content analysis of cognitive studies in selected journals. *Computers & Education*. Vol. 15, No. 2, pp. 955-967.
- Taylor, E. W. (2001). Adult Education Quarterly From 1989 To 1999: A Content Analysis of All Submissions. AEQ, 51, 4, pp. 322-340.
- Tutarel, O. (2002). Geographical distribution of publications in the field of medical education. BMC ED, 2, 3.
- Yang, L., Morris, S. A., Barden, E. M. (2008). Mapping institutions and their weak ties in a specialty: A case study of cystic fibrosis body composition research. *Scientometrics*. [Preprint available at http://www.conceptsymbols.com/web/publications/yang_y1_2007_ISSI.pdf]