Globalization Of Knowledge Discovery And Information Retrieval In Teaching And Learning

Mark Zaidel, (Email: zaidel@northern.edu), Northern State University
Osiris Guerrero, Graduate Student, Northern State University

ABSTRACT

Developments in communication and information technologies in the last decade have had a significant impact on instructional and learning activities. For many students and educators, the Internet became the significant medium for sharing instruction, learning and communication. Access to knowledge beyond boundaries and cultures has an impact on methods of teaching and learning, affecting professional education in corporations worldwide. The blending of science with commercial applications, new methods of conducting business, and access to published and preliminary research is no longer limited to leading educational institutions and the northern hemisphere. This paper discusses educational opportunities, reviewing the latest trends in information retrieval, along with new teaching methodologies related to availability and globalization of information. The good news is that the multicultural melding of various intelligences and the global linking of intelligent knowledge systems promises to lend a multicultural wisdom to the selection and provision of information. We hope that this may ultimately contribute to a lessening of the gap between educationally rich and poor nations. In this paper the “digital divide” is seen to be a less vexing problem for the future of globalized intelligent information systems. This system, called the “Global Brain,” ameliorates existing differences in national access to educational content repositories used for teaching and learning.

BACKGROUND / INTRODUCTION

The last decade opened new opportunities for college teaching and learning related to the rapid development of communication and information technologies. Our students are no longer limited to accepting knowledge in classrooms and laboratories; they are instead encouraged to expand their intellectual horizons by using multiple means of computer mediated communication. Using free Internet sources for learning and research has become habitual for students and faculty not only in traditional technology areas but also in every area of science, medicine and art. Data from the Association of Research Libraries shows that reference queries at university libraries have decreased since the late 1990s. As of 2006 about three-quarters (76%) of college students said they use the Internet more than the library, while only 8% said they use the library more than the Internet for information searches. The convenience of Internet searches causes students to rely on it when looking for informational and academic resources. In citing our research and that of our students, we are acknowledging Internet links, rather than the library, as an academic source.

EMERGENCE OF THE GLOBAL BRAIN

The growth and evolution of informational and academic networks, according to Russell and Gotfried, will eventually lead to an emerging artificial superstructure commonly called the Global Brain. This structure could be accessed using communication and information procedures through intelligent agents (Heylighen) providing interactions with humans and communication between computers. Philosophically, the global computer network functioning as a Global Brain will eventually execute information processing tasks on the Internet. An analogy between learning in networked society and the neurological processes of the human brain leads to the conclusion
that Internet-based learning is most effective when students explore the most commonly used links. These links will gain strength thus generating new neural connections based not only on similarity between documents but also on the topology of hyperlinks. As described this mechanism incorporates the collective knowledge of all users using the Web, and this leads to the “thinking Web” concept where knowledge discovery is implemented by intelligent Web agents.

**FEEDBACK LOOP BETWEEN THE GLOBAL KNOWLEDGE DEPOSITORIES AND EDUCATIONAL ACTIVITIES**

We propose to extend the applicability of the Global Brain concept to the educational function of dissemination of knowledge. In our concept, searching for specific content or asking a question to the global network of search engines will trigger the review of a large numbers of documents. The uniquely generated response will be merged into the global knowledge data set by intelligent knowledge agents operating within the Global Brain. The dissemination of this enhanced content becomes practically free and competitive with “for profit” knowledge providers using private databases or knowledge management systems. It is especially beneficial if this whole process happens in the multicultural environment because the various multicultural perspectives are incorporated as significant factors and thus relates to complex and sophisticated verification of the published content. The oft-cited Semantic Web project demonstrates some preliminary results of this process. The importance of this Global Brain philosophy rests not only in the ability to collect and process content but also on its multicultural verification by users exploring and building a topology of “neural connections” through the process of using and strengthening the most important links. The evolution from the mere retrieval of collected content into its dissemination as an interactive process, i.e. the globalization of culturally enhanced knowledge access, will not only benefit students but also strengthen the topology of the Global Brain concept.

As we demonstrate in this paper, there are many global initiatives related to the sharing of academic and research content in the present structure of the Internet. However, instead of fully utilizing such global sharing, the majority of the world’s 59 million teachers are “reinventing the wheel” by producing their own curricula, to be used only by a handful of students through proprietary or private knowledge management systems. In the northern hemisphere (West) one of the underlying reasons for this is the commercialization of on-line academic instruction, and another reason relates to the protective policies of intellectual property regimes and regulations implemented by both governments and educational institutions. In the developing world, students have largely been shut out of the communications revolution by the cost of high-priced computers, software and lack of access to scientific publications.

**KNOWLEDGE AVAILABILITY AND INFORMATION EXTRACTION IN A GLOBALLY INTERCONNECTED SOCIETY**

Connectivity between computers and the status of a communications infrastructure in the world is one the most important factors in the global brain concept. The issues of information blocking, censorship and selective retrieval are of concern in many countries (Rugaber), but in this chapter we will focus rather on methods used for knowledge categorization, data mining and selective information retrieval. The educational benefits originate from direct use of these technologies by teachers, especially in preparing on-line instruction, and by students seeking answers in complex or multidisciplinary problems. The large amount of data becomes a barrier in knowledge availability; we need tools that could process data and present results in a suitable format. These tools include cleaning a data, patterns detection, semantic web ontology, rule inductions techniques and many others, leading towards automatic text summarization and cross-reference of discovered entities, processes and theories. The ultimate goal is more effective use of human’s time in the process of knowledge access and assimilation. In order to improve the access to knowledge computers have to “understand” the digitally stored content of websites, papers, journals and books. Understanding the content by computer means the ability to relate newly processed information to information already accessible from the knowledge depository identified here as Global Brain. As an example of solutions provided by the knowledge discovery system we could use a new paper analyzing current trade relations between two countries, and with the help of system we would draw further conclusions based on other earlier processed documents associated with the subject of international trade. When the amount of stored information in
the networked structure increases toward infinity, the single knowledge discovery system that uses grammar rules, conceptual-semantic and lexical relations, clustering, and pattern recognition would be able to relate facts, processes and ideas and eventually distribute them as textual summarizations.

To verify the above outlined concept we are developing computer software that analyzes and processes text. This effort is facilitated as a part of institutional research involving graduate faculty and eight graduate assistants. The software builds a library of structural templates used to recognize meaningful relations between English words that are embedded in the lexical database known as WordNet®. Based on research experience, we believe that natural language processing and knowledge discovery systems can be a sharable resource in the e-learning environments providing better access to knowledge repositories in the networked world.

THE DIGITAL DIVIDE AND THE GLOBAL BRAIN

In recent years, as information and communication technologies have become the backbone of the global information economy and, according to Castell, are referenced to the rise of the information society, more attention has been focused on the gap between developed and developing countries in the specific context of information and communication technologies. This gap has come to be known as the “digital divide.” We regretfully read of expectations that the “information world” will stay fragmented into developed and developing regions. North America and Europe, for example, will have advanced Internet usage, as the Internet is used for everything from business to leisure and entertainment. We worry that the African continent and parts of central Asia will be dead zones where comparatively little or no Internet usage will be found. The technological differences between poor and rich regions is predicted to multiply with advancements in connection speed and new technologies, such as wireless or two-way satellite connectivity.

The UN Economic and Social Council (cited in the references) propose measuring this gap as a regional difference between the level of penetration of different ICT services (telephone, mobile phone, Internet) and personal computers over the course of the past decade. Data analysis shows, however, that the gap has narrowed markedly, with particularly rapid progress in the field of mobile phone and Internet use; in the last 10 years, the relative number of users has increased about 30 times in developed countries, 45 times in the world and about 100 times in developing countries. The average level of Internet and mobile phone usage in the developing world in 2002 (4.1 Internet users and 10.7 mobile phones per 100 inhabitants) was equivalent to the level reached in the developed countries around five years hence. By contrast, the average level of fixed-line usage in developing countries at the end of 2002 (just under one in every 100 inhabitants) was reached in the developed world as long ago as the 1960s.

The Global eSchools and Communities Initiative (GeSCI), launched in 2003 by the United Nations, supports the development of information and communication technologies used in education in developing countries. The main task of this effort is to create effective end-to-end local systems needed to deliver, operate and support the education solutions identified by the partners based upon the needs of their students. Since its creation, this initiative has established several collaborations between regional governments, corporations and supporting educational institutions in the developed world.

The catalytic role of higher education in these information and communication initiatives was also recognized by UNESCO, which started in 1998 the Open and Distance Learning (ODL) Knowledge Base project. The objective is to support decision makers and practitioners with ready access to information and tools that will assist them in more effective policy planning, development and management of ODL in undergraduate and graduate curricula. The Regional Knowledge Base Resources were created to provide storage of information related to various aspects of ODL, tailored to specific regional needs. Currently there are three recognized regions: Africa, Asia and Pacific, and CIS/Baltic countries. For these regions UNESCO also provides the databases offering both primary and secondary information on topics related to political and social implementation of ODL. The technical issues are supported by the so called Knowledge Finder providing access to virtual academic and research libraries from selected educational sources. This service is accessible directly from the participating Knowledge Base sites as an advanced search query builder and knowledge management tools which may be used to prepare, present and save
results. A parallel UNESCO initiative coordinates efforts related to the renewal and internationalization of higher education, its systems, institutions, programs and courses by the UNITWIN projects. These efforts are helping to establish sub-regional, regional and international interuniversity networks involving academics, scholars and graduate students in all major areas of knowledge in nearly 120 countries.

CURRENT ISSUES AND TRENDS IN GLOBALIZATION OF TEACHING AND LEARNING

Individual students seeking undergraduate and graduate degrees in the USA from accredited on-line educational institutions often confirm the need for the broader dissemination of knowledge through enhanced educational opportunities via the Internet. There is a trend of growing on-line enrollment of undergraduate and graduate students in major educational institutions in the USA, and, also the ratio of foreign students seeking on-line degrees has increased in major on-line educational institutions in the last decade.

The National Science Digital Library (NSDL) is one of the positive examples of federal government involvement in implementing and in expanding access to research and academic content, thus opening up competition with the various proprietary publishing systems. The Public Library of Science is a similar collaborative effort of colleges and universities. In comparison, the scientific, technology and medical publishing industry in 2006 in the USA currently constitutes yearly revenues of more than $14 billion dollars, when for the same year the total revenue of the U.S. industries associated with e-learning is assessed on the level of 60-70 billion dollars. Scholarly publishers provide on-line access to electronic journal subscriptions and paid search services to further increase their revenues. Their customers – mostly academic institutions and research organizations – are accessing protected data resources by using private permission-based search gateways to the contents of their scientific journals. This proprietary policy tends to significantly limit the concept of sharing educational resources through the Global Brain superstructure. Legislation is currently pending in Congress that would at least ensure that all research funded by federal taxpayers will be made available free of charge to the public over the Internet.

Education is one of the strongest forces driving development of the Internet technologies. On the other hand, the Internet will have significant impact on teaching and learning as students will increasingly choose virtual classrooms and virtual learning communities that match their interest, knowledge and skills. According to [Fox], learning will be a second factor (after commercial news and publishing) that will drive development of Internet in the next decade.

The ability to utilize global Internet resources for teaching and learning in the next 10 to 15 years will still depend upon the geographical location of the student, but it will not be constrained as much as might be expected due to the existing digital divide. A good comparison of regional differences that will have an impact on education is the ratio of college students who own personal computers. In the least fortunate societies (dead zone of Internet usage) less than 0.5% of college students have computers in their home. In eastern European countries this ratio is between 15% and 40%, while in the USA 84% of college students own a personal computer. However, in the USA about 33% of these students use public computers on campus almost exclusively for their learning activities. In addition, about 80% of students use public computers for social and entertainment purposes even if they have their own computer in their home or dormitory [Featured Reports]. These statistics underline the importance of the world communication initiatives as described above in this paper because these initiatives are the most effective processes providing unrestricted access to global educational resources through higher education facilities.

It is expected that the quantity of educational content freely available on the Internet will grow much faster than the number of users. Growing numbers of non-government local initiatives generating or managing instructional and research content on the Internet support this statement. The MIT Open CourseWare is an example of a large-scale, Web-based electronic publishing initiative funded jointly by the Hewlett Foundation, the Mellon Foundation and MIT. The OCW’s goals are to provide free, searchable, access to MIT’s course materials for educators, students and self-learners around the world and to extend the reach and impact of quality instruction using the "open courseware” concept. With 1700 courses published as of September 2007 and another thousand expected to be published by 2010, this MIT initiative shares intellectual property such as research, pedagogy and knowledge for the benefit of others. The exemplary list of other world-wide initiatives providing shared educational
content include the following (in no-particular order): Sloan Foundation, SCATE National Resource Center, Multimedia Educational Resource for Learning and On-line Teaching, European Experiences on eLearning for Economic Development (contains database of projects), Association Easy Spanish: British Members (free Spanish course in five languages delivered by now to 25000 students).

Access to educational and informational resources on the Internet has already changed the traditional approach to teaching and learning. In the USA 5% of adults who already received a college degree have taken at least one on-line course for college credit. On-line research is conducted on a daily basis by more than 50% of Internet users in response to job or college assignments; every day about 8% of Internet users are on-line for topical research or job training. About 98% of public schools in the USA provide Internet connectivity for their students, and, in college education on-line activities are commonplace.

CONCLUSION

The Data Mining and Knowledge Discovery implementation models represent the modern approach to the process of identifying complex and understandable information from global computer networks. For everyone who is willing to consider that international and national efforts will lead to improvements in access and quality of educational content on the Internet, the obvious conclusion is that the concept of the Global Brain might quickly become a reality.

REFERENCES

2. Featured Reports: Pew Internet & American Life Project; http://www.pewinternet.org/
4. Heylighen Francis, Bollen Johan: The World-Wide Web as a Super-Brain: from metaphor to model; Center Leo Apostel, Free University of Brussels, Belgium
11. Semantic Web: http://www.semanticweb.org/