Web Services

Library Services in Web Environment: A Study

Rohit Singh[1], Ashish Kumar Shukla[2]
Central Electronics Engineering Research Institute (CEERI)
Council of Scientific & Industrial Research (CSIR), Pilani (Rajasthan), India
Amity Business School, Pachgaon (Manesar)Gurgaon (Haryana)

ABSTRACT— This paper discusses the web-based library services and the impact of web technology on library and information services. Web technology is very useful and effective technology for the library and information center, and it support and promote to research and development activities. We provide the traditional services very fast and effectively through web technology. Library services play vital role in all type societies and in web environment library provide all traditional services and introduce new services according user needs. Web technology make library services more informative in order to attract more and more users.

KEYWORDS— Web based services; WWW (World Wide Web)

INTRODUCTION

Library is a place for storing knowledge under a system that facilitate identification and retrieval as needed, which is also a definition of a computer means library doing work like computer gathering of knowledge, organizing than retrieving. Libraries are mainly entrusted with a host of predetermined tasks like acquiring, organizing, preserving, retrieving and disseminating information to the users. Right from ancient times to the present Internet era, the primary objective of library has always been this. However, the way this purpose has been achieved has drastically changed. Web is a platform for transfer the knowledge.

Web-based information (or the internet) is a major force for changing the role of libraries and information centers. In fact, the Internet has broken new ground for finding and retrieving information and with it the end-user has found a way to become more information-independent. The rapid and high growth in the world's literature, curtailment in library budgets and advances in telecommunications and information technologies are progressing at a much faster pace. This now means that speedy (and increasingly electronic) access to materials held elsewhere is becoming more usual, with acquisition and relation policies based more on a "just in time" than "just in case" approach (Hopkins, 2000).

Web-based information retrieval (WIR) started out as information retrieval (IR). The phrase "information retrieval" largely replaced the older term "documentation" from about 1960 (Buckland and Liu, 1995). Perhaps the first example of the integrated approach to information retrieval was the TIP (Technical Information Project) run by at MIT under the sponsorship of the National Science Foundation. From 1962, it acted as a test bed for the evaluation of computerized libraries and methods of scientific communication. Exploitation of the retrieval speed of computers and the storage potentials of micro-files was still a subject of discussion and experiment. In 1972 Landau proposed an online retrieval system desk unit called COMICON, the design of which was based on an earlier idea by Vanevar Bush for a personal search and document retrieval terminal named MEMEX (Landau, 1972).

Since the arrival of the worldwide web in the early 1990s, web-based user interfaces have become a common method of accessing online databases. Since the mid-1990s, several IR systems have developed web-based operations, which have replaced or coexisted with earlier versions (Ahmed et al., 2005). Indeed, now almost all major IR systems offer web access to their databases. The fundamental characteristics of web-based IR (WIR) systems are that they are inherently interactive and provide low cost access to a variety of online databases. Web-based technology is based on the information technology. However, Third World countries are facing a poor information infrastructure, where the societies are influenced by the traditional information services. A few years ago, Gorman (2003), commenting on telecommunication facilities in 2003 and quoting Abdul Wahed Khan, noted that: ". . . 80 percent of the world's population lacks access to basic telecommunications facilities, which are the key infrastructure of the information societies, and the less than 10% has access to the Internet. Access to the information highways and to content, such as development data and information, is still a major problem in many countries".

OBJECTIVE OF LIBRARY

The primary objective of any library is to provide timely and quality services to its users. Advent of new technologies have made it imperative for libraries to absorb, adopt, adept, and to provide IT-based services to the users. Libraries are now providing web-based services at the users' desktops. Users, particularly, scientists, who are engaged in time-bound research and development projects, need to be alerted about the nascent literature in their respective fields. Librarians are providing current awareness services to meet these requirements. A few libraries and many e-publishers are providing alert services to keep the users informed about the latest additions to their collections and other information relating to the users' areas of interest.

WORLD WIDE WEB

The World Wide Web- know as WWW, W3 or simply, the Web - is one of several Internet resource discovery tools developed to help people publish, organize and provide access to information on the Internet. The Web was first developed by Berners Lec in 1989 while working at CERN, European Particle Physics Laboratory in Switzerland, and has since become the most powerful, and popular, resource discovery toll on the Internet. The WWW can be defined as a hypertext, multimedia, distributed information system that provides links to hypertext documents, as well as to many other Internet tools and databases. There are several features that are unique to the Web that make it the most advanced information system to appear on the Internet to date.

IMPACT OF 3W ON LIBRARY

The World Wide Web is important for libraries because it provides an extremely powerful method of organizing, and providing access to, information. It can provide one interface to a large variety of network information resources and systems. With the Web and its browsers libraries can:

- Electronically publish anything that they now publish on paper
- Provide access to in-house hypertext documents or to hypertext documents available out on the Internet
- Create electronic orientation services with floor maps and descriptions of services
- Provide access to Internet tools such TELNET, gopher, FTP, and WAIS through a single interface.
- Create interfaces to in-house databases or bibliographies
- Collect information from patrons through the forms feature.

With all its power, the Web is the most serious step yet towards creating electronic libraries. It provides a mechanism to present a wide variety of information resources to library users-and all Internet users-in a simple, efficient and effective manner.

WEB-BASED LIBRARY SERVICES

Through web we provides many types of information and services. These information and services are categories in different headings:

General Information

This category includes an introduction to the library, hours, general library policies, contact information, etc. Nearly all have an introduction, mention library hours, general library policies, contact information, and a staff directory. Under this category includes all information like:

- Library introduction (History, Vision, Objectives etc.)
- General library policies and Library hours
- Contact information and Staff Directory
- Library sections/divisions/departments
- Library and General suggestions / feedback form
- Liberary map and Virtual tour

Web OPAC

The web OPAC is a significant service offered by libraries through their websites. Under this service we provide library catalogue on web. An OPAC (Online Public Access Catalog) is an online bibliography of a library collection that is available to the public. OPACs developed as stand-alone online catalogs, often from VT100 terminals to a mainframe library catalog. With the arrival of the Internet, most libraries have made their OPAC accessible from a server to users all over the world.

Web-based Reference and Information Literacy Services

Provide general information about information literacy/ instruction services, general information about reference service, email reference, chat reference and covers information related to:

- General information about instructional-services/ user education/information literacy
- General information about Reference Service
- Reference Service using email/web-chat/web-form
- Text based tutorials for instructional-services/user education/information literacy
- Web-based/Interactive tutorials for instructional services/user education/information literacy

Library Collection and Access to Electronic Resources

This category had seven variables libraries provide access to online databases, digital collections, and descriptive information about their collections, and provide information about "New Arrivals." This category covers these points:

- Access to online databases journals and digital collection
- General information about collections
- Text based web-based recommendation form
- Collection development policies/New arrial list

Readers' Advisory

The next frontier is addressing readers' advisory. Catalog enhancements, such as book covers, reviews, and first chapters, are already standard attractions. Lists of suggested titles with links to the catalog and Web-based reserves are also part of the current fabric of library Web services. Web-based services, such as Book Browser, What to Read Next and others, help match readers with their next book. Online book discussion groups such as our own Yread? Are popular features that use bulletin boards or live chat rooms to share ideas? The next step must be entering the live environment to negotiate and guide the public as well as to personalize services.

Library Consortium

Through Internet we provide a special service to users that's call Library consortium. We define consortium in easy way that a consortium is basically a cooperative association of Libraries of different types. Set up with a purpose to share human and information resources so that the collective strengths of the institutions facilitate the research and learning of the member's constituents

Alternatively a Library Consortium consists of a number of libraries, preferably with some homogeneous characteristics by subject, institutional affiliation, or affiliation to funding authorities, that come together with an objective to do certain job collectively. Consortia-based subscription to e-resources is a way of maximising access to e-resources at minimum cost. A consortium supports resource sharing and provides services to users through programmes such as :–

- Cooperative acquisition,
- Access to electronic resources
- Enhanced interlibrary loan and document delivery
- Group licensing / purchasing of electronic resources is one of the key activities

Communication Service

Library provides to users many communication services through Internet. This is most important service, this service provide the facility to connect any one or group. These service categories in three broad sections:

1) Internet-based Communication Services
   a. Communication amongst Individuals: Electronic mail, Internet Telephony, Internet Chat
   b. Communication Services for groups: Internet Relay Chat, Electronic discussion groups, Listserv, News groups, Usenet and video conferencing.
2) Connectivity
   a. Telnet
   b. Remote Login
   c. File Transfer (FTP)
3) Information Resources
   a. Gopher
   b. WWW
   c. FTP Servers

Web-based Circulation

Web-based circulation services are attractive to patrons. In this study, many aspects of circulation services are analyzed. Shown that descriptive information about circulation policies is provided on their websites, also give the membership procedure. Manly four aspects are covered:

- Circulation policies
- Membership procedure
- Patron's access to their accounts
- Renewal and reserve facility

Interlibrary Loan (ILL) and Document Delivery

The libraries provide descriptive information about ILL and document delivery, or provide a text-based request form or web-based request form for this service. Under this three main points are:

- ILL and document delivery policies
- Text based request form
- Web based request form

CONCLUSION

The study shows that how the traditional library services change with application of web, and convert into web base library service. Web base library services help to library catch the objective that "the right knowledge to right user at right place in minimum time" There is an urgent need to develop dynamic websites with a direct hyperlink on the homepage. LIS professionals must recognize the significance of web-based library services and take the initiative to provide them. LIS programs, and professional associations should provide training opportunities to librarians. Further research is needed to identify the problems faced by libraries in introducing and maintaining web-based services. With the rapid development of the Internet and the World Wide Web, information exchange and information distribution can be transformed either by disparate formats or by dynamic channels. Knowledge, competence, and skills cannot follow up with the rapid advances of science and technologies. Many library users need assistance to access, locate, convert, synthesize, and evaluate information effectively and efficiently. It is very imperative for libraries and librarians to design, develop, enhance, implement, and deliver high quality user-centered information services, resources, and instruction at the fingertips of library users.

REFERENCES

[1] Ahmed, S.M.Z., McKnight, C. and Oppenheim, C. (2005), "A study of learning and retention with a web-based IR interface", Journal of Librarianship and Information Science, Vol. 37 No. 1, pp. 7-16.

[2] Buckland, M. and Liu, Z. (1995), "History of information science", The Annual Review of Information Science and Technology, Vol. 30, pp. 385-416.

[3] Gorman, G.E. (2003), "Sustainable development and information literacy: IFLA priorities in Asia and Oceania", IFLA Journal, Vol. 29 No. 4, pp. 288-92.

[4] Hopkins, M. (2000), "Managing the future growth of library collections", discussion document, available at: www.inf.aber.ac.uk/central/quest.asp

[5] Landau, H.P. (1972), "New economic factors in the system integration of computer of computer terminal online retrieval systems and large microform databanks", Journal of Micrographics, Vol. 5 No. 3, pp. 125-9.

[6] Wiederholt, G. (2003), "Increasing the information density in digital library results", available at www.sis.pitt.edu/dlwkshop/paper_wiederhold.html

[7] Ahmed, Syed Sajjad. "Managing change to enhance Web-based services in the Arabian Gulf libraries." *Online Information Review* 26, no. 4 (2002): 265-270

[8] Feldman, Sari and Tracy Strobel. "Advancing your library's web-based services." *ERICDigest* , (2002)

[9] Halub, Lisa Parker. "The value of Web-based library services at Cedars-Sinai Health System." *Bulletin of the Medical Library Association* 87, no. 3 (1999): 256-260

[10] Rhodes, Helen and Jacqueline Chelin. "Web-based user education in UK university libraries – results of a survey." *Program* 34, no. 1 (2000): 59-73

[11] Saeed, Hamid. "Internet use in university libraries of Pakistan." *Online Information Review* 24, no. 2 (2000): 154-160

[12] Church, J. (2005), "The evolving information commons", Library Hi Tech, Vol. 23 No. 1, pp. 75-81.

[13] Hayes, J., Stoll, C. and Tortorella, D. (2005), "XML: the DNA of the knowledge management evolution", Computers in Libraries, Vol. 25 No. 1, pp. 10-15.

[14] McDonald, J. and Van de Verde, F.F. (2004), "The lure of linking", Library Journal, Vol. 129 No. 6, pp. 32-4.

[15] Rogers, M. (2006), "European Digital Library in 2010?", Library Journal, Vol. 131 No. 6, pp. 28-9.

[16] Stephens, M. and Gordon, R.S. (2006), "IM ¼ FASTER virtual reference on the cheap!"Computers in Libraries, Vol. 26 No. 4, pp. 36-7.

Success of E-governance is with SOA

Aprna Tripathi[#], Bhawana Parihar[*]

[#]*Department of Computer Application, Kanpur Institute of Technology, Kanpur, India*
aprnatripathi@gmail.com

[*]*Department of Computer Science And Engineering, V.K.T. Kumaun Engineering College, Dwarahat, Almora, India*feeling2908@gmail.com

ABSTRACT— E Governance, especially in developing countries, is looked upon as a means to change the very concept of governance resulting in empowerment of the citizens and increased transparency in public dealings by the governments; increased efficiencies in delivery of public goods is an inherent underlying assumption. Last few years, several projects under the E-government initiative have emerged; they are aiming to bring public administration such as for i.e. taxation service, closer to citizens and industrials. Today most of governments are aware that the use of information communication tools represents a powerful means to increase collaboration and cooperation between different public institutions. The aims are to provide better and more efficient services to citizens. The ambitions of governments are to facilitate and encourage development of homogenous web-platforms providing to citizens just with one single authentication, access to several public services such as taxation service, social service, health care, etc.

However most of the governmental bodies are running their own information systems and connecting them together is not an obvious task. The main challenge relies on integrating efficiently all those heterogeneous public information systems by providing a unified environment. Governance with SOA should ultimately be about delivering our business and SOA objectives. It must link SOA investments to business goals and initiatives mitigate the risks associated with SOA, and fit into the context of an organization's overall IT Governance framework. Companies that have established governance to help individuals make good decisions within the context of the problem space have matured their SOAs successfully. These companies have also achieved an effective layering of SOA capabilities in areas such as architecture, technology infrastructure, operations, information, governance, people and organizational structure, portfolios, project execution, and finance. In this paper author defines an integration applications model based on service oriented architecture (SOA). And also outlines a framework and best practices for governance as it specifically relates to SOA.

KEYWORDS— WEB service, E-government, SOA, Services.

I. INTRODUCTION

E-governance is defined as the transformation of (governance) processes (resulting from) the continual and exponential introduction into society of more advanced digital technologies [5, 6]. E-governance is generally considered as a wider concept than e-government, since it can bring about a change in the way how citizens relate to governments and to each other E-governance is not just about government web site and e-mail. It is not just about service delivery over the Internet. It will change how citizens relate to governments as much as it changes how citizens relate to each other. It will bring forth new concepts of citizenship, both in terms of needs and responsibilities. E-governance will allow citizens to communicate with government, participate in the government's policy-making and citizens to communicate each other [3, 4]. The e-governance will truly allow citizens to participate in the government decision-making process, reflect their true needs and welfare by utilizing e-government as a tool. In the last few years almost all the Government agencies have their own e-Governance applications. In the initial stages organizations showed lot of enthusiasm and applications grew at a greater pace. But, with the passage of time when the need for interoperation was realized it was felt that such applications have been developed in total isolation, they were almost like islands.

In the present day the common problem faced by Government agencies is that they have many similar Citizen-centered applications. Every time a department wants something slightly different, the department builds its own version. As a result the organization ends up with different versions of more or less the same application. Service Oriented Architecture (SOA) is one of the best solutions for real time integration between systems, for reuse of services, where service implementation requires no advanced knowledge of the service. SOA allows different applications to exchange data with one another as they participate in e-Governance processes. The aim is to have loose coupling of services with different Operating Systems and programming languages .By implementing an e-governance initiative government departments will streamline their processes, connect all the stakeholders, cut costs, and improve the delivery of their services. Now with the wide

acceptance of SOA it is being adopted as a common framework around which e-Governance applications can be developed. Service-Oriented Architecture (SOA) facilitates provision of compound services covering whole customer processes, where a customer may be both a citizen and an enterprise. In this paper authors have described how the SOA gives a proper solution for e-governance applications.

II. SERVICE ORIENTED ARCHITECTURE

SOA consists of three words Service-Oriented-Architecture Here the Services are the -Web service, oriented is meant for association with services And Architectures is the Framework around which the services are building and frame work can be turned around services[19]. A service is an autonomous and reusable unit of business or administrative logic. Rules and methods of accessing services are specified in *service contracts*. At the one end a provider and at the other end a consumer connecting through messaging there might be a facility in between these named as Service broker -A program which invokes or interacts with a service is called a *service consumer*. The collection of web services that can be call through provider to bring over to the consumer this framework can be loosely coupled. There are some challenges over the SOA [1,2] in this Different application run over the different platforms, on different operating system by using different protocols. Enterprise Service Bus (ESB) gives s a common set of services that can be used a common platform for communication to the applications. Application can share information between one application to another by producing and consuming information on an ESB. Service-oriented architectures involve three different kinds of actors: *service providers*, *Service requesters* and service broker (As shown in fig.1).

Since service provider and service requester usually do not know each other in advance, the service broker is called by the service provider to bring over the consumer. In the Service-oriented architectures components and services are only loosely coupled and communicate according to standardized protocols; the process of communication involves either simple data-passing or two or more services coordinating some activity *service-oriented architecture* (SOA) as a group of services that communicate with each other[7,8]. Services can in turn be combined to produce composite services and business processes, in accordance with pre-defined policies, security. A business process is the composition of services, which are independent to each other. The service is the collection of some components[9].

III. E-GOVERNANCE CHALLENGES

E-government is defined as the use of ICT and its application by the government for the provision of information and public services to the people[12,14]. The primary aim of E-governance is to provide efficient government management of information and also giving better service delivery to the citizen. Secondly e-governance looks the empowerment of the people through access to information and participation

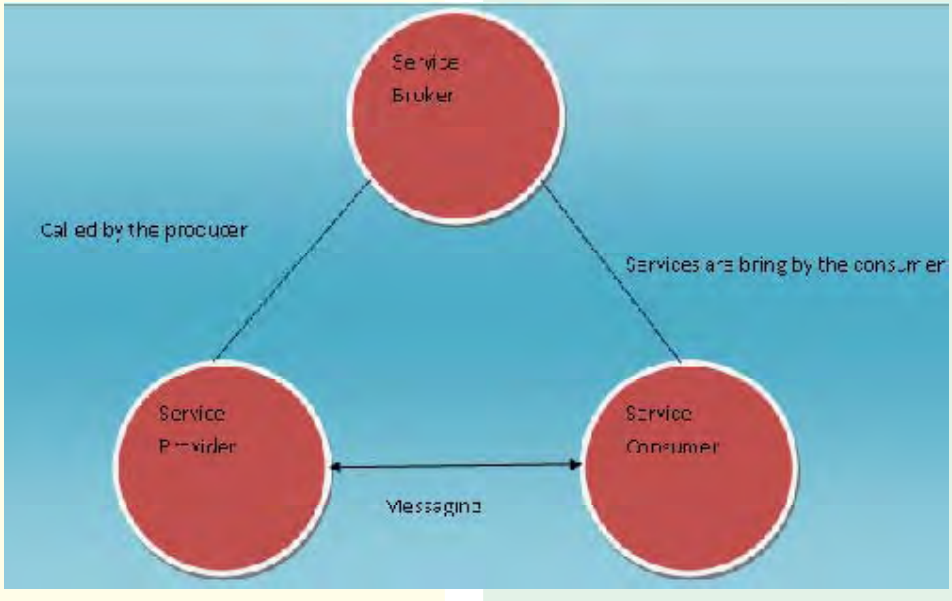


Fig. 1. Basic model of SOA

in public policy decision making. As defined in the MMPs the main mission of e-governance is to make all government services accessible to the common man in his locality, through common service delivery outlets and ensure efficiency, transparency and reliability of such services at affordable costs to realize the basic needs of the common man. And for achieving this authors require to increase locality of authority more openness in accessibility accountability and authorities E-governance is a big opportunity to bring services to all citizens, but there are some serious challenges for consideration.]The basic challenge is of the integration government, citizen and business community, which can be clearly seen in fig. 2. It may me categories into the three main categories [13]

A. Technical challenges

B. Economical challenges

C. Social challenges.

A. Technical Challenges

IT infrastructure is the backbone of E-governance. Interoperability with existing software and hardware platforms is a key success factor. Therefore interoperability is the key factor between existing software and hardware platforms. Some legal aspects like security and privacy must also be an important consideration. To cope with such requirements appropriate technical changes must be done. Multi-model application can make it more successful. The guarantee by the government could not suffice unless accompanied by technical solutions, transparency of procedures and possibly independent auditing. It is unlikely that available resources can support a full replacement of existing application. Hardware should be fully compatible with future technologies as well.

B. Economical challenge

Economical issues are mainly concerned with return of investment and safeguard of the previous ones. Implementation, operational and evolutionary maintenance costs must be low enough to guarantee a good cost and benefit ratio, E-governance must be seen as a nationwide plan, implemented applications must be reusable by other administrations. The independence from the hardware and software platforms is a primary concern for portable applications there is also a main issue of maintainability, this is the key success factor for long living systems in a rapidly changing technical environment.

C. Social challenge

It mainly concern with the usage made by the citizens.. Acceptance and usability by a large variety of people make e-governance successful. Since the social disparity is very high in India, so this issue needs a careful observation. Then there is a challenge of accessibility, usage and acceptance of the e- governance. Even if the internet population is growing exponentially there is a significant portion of the people who may not be able to access e-government for various reasons. This implies that interface must be usable by rich or poor, disabled or elderly people, understandable by low literacy or non-native language people, etc. Users are often not expert users they need the guidance to find the right way to perform their transactions. The successful implementation of e-governance requires a re-

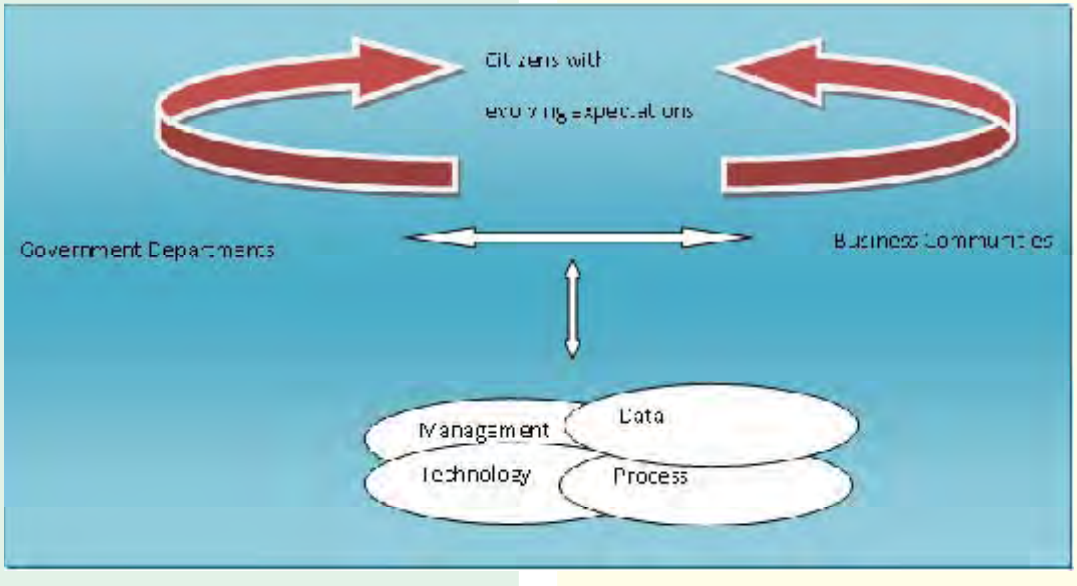


Fig. 2. E-Governance challenges view

conceptualization of the government. As e- government becomes a reality, the public sector organizational structure will change accordingly both internally and externally.

IV. SOA-GOVERNANCE

Implementing a Service Oriented Architecture helps an organization to align IT with business goals and to succeed in rapidly changing business environments. In order to achieve the benefits of SOA, authors need to implement a governance model.SOA Governance is all about finding the right balance for organization. Here authors need to create the right set of business services at the right level of granularity to support the business. If the services are too narrow and technically defined, then they may not have the right meaning for the business. In order to achieve business success with SOA, one needs to implement a SOA governance model that ensures business service reuse and business value. The emergence of SOA (service-oriented architecture) has changed the way enterprise IT needs to be viewed. An SOA approach for governance better aligns IT with service delivery goals and enables various government departments to re-use developed assets.

The first and the most essential condition is strong leadership of the national government. The national government must have a holistic view of electronic government. This view must be optimized from the global point of view to ensure interoperability at the level of the whole state. Interoperability is required to enable dynamically creating compound services covering the whole customers' processes.

The second condition is adjusting law to new architectures. It is necessary to separate responsibility of administrative units for data from responsibility for computer systems. Without necessary changes in law, cloud computing and service-oriented architecture in the public sector may appear illegal.

The third condition is to keep local innovators and enthusiasts in the public sector. This can be achieved through cloud computing and software as a service ability to adjust to specific needs and requirements of a particular user, i.e., a local government office or agency. Local developers can feel satisfied, because they will be able to optimize software from the local point of view, while simultaneously they will not disturb the global consistency and interoperability.

The fourth condition is to sponsor openness of the public sector Anyone who has an idea of how to better serve customers of the public sector is invited to create an enterprise, get basic e-government services as resources, integrate, mix or extend them and offer on the market. Such approach results in increase of competition, faster development, and price reduction. The author describes the SOA in public sector in fig. 3.

For achieving the maximum benefits of the SOA in E-governance there is a essential requirement that all the state level government, central government, notary public and banks must be come at some common platform for providing their services to the common man SOA governance, in such a way that common can take the benefits of the E-governance in a very simple manner. And for this authors require to mix the

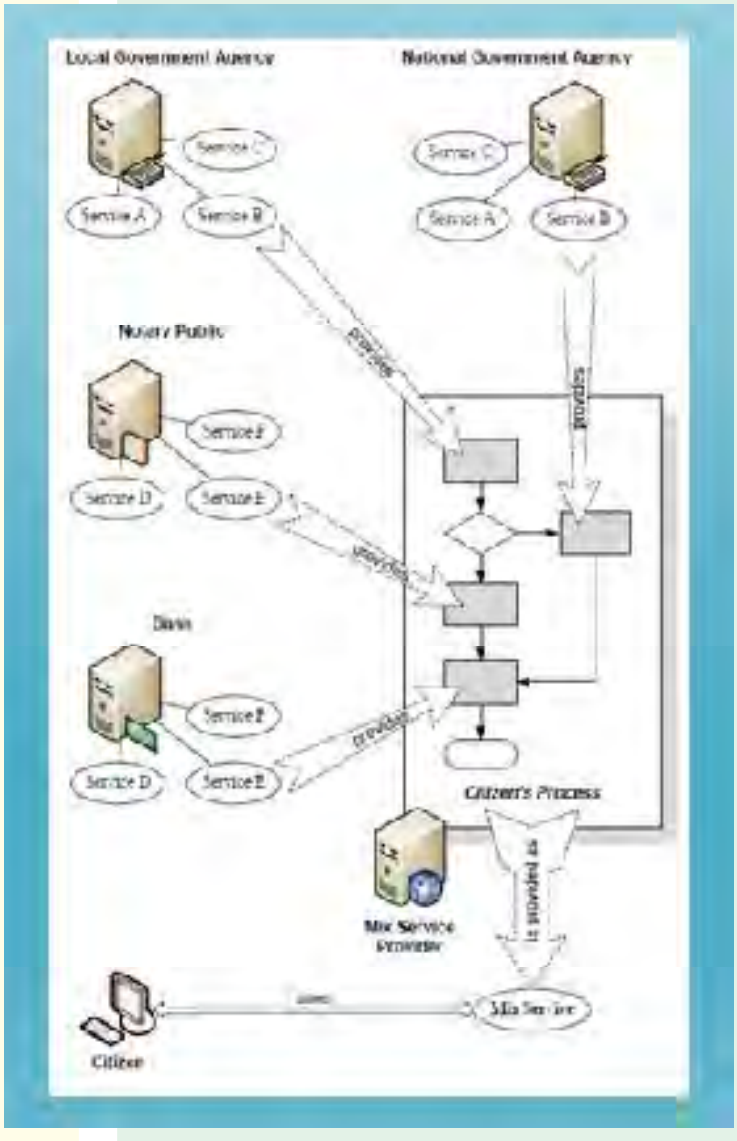


Fig. 3. SOA in public sector: a mix service composed of Services provided by both governmental and business units

services provided by the state level government and central government. For designing a effective SOA-Governance authors follow some specific step which are as follows

a. Approach executive management with a justification for SOA governance.

b. Create a comprehensive plan to create the right business services with executive support.

c. Establish process for organizational change since managing change is as important as creating SOA services.

d. Balance risk with oversight to find a proper balance for SOA governance.

e. Plan for the lifecycle of business services.

By following the above five step authors can achieve the benefits of the SOA in the E-governance

V. CONCLUSION

Service Oriented Architecture for e-governance provides the transparent governance to citizens by IT enabling various government departments. Usage of SOA and web services optimizes the performance of government applications. The Service Oriented E-Governance (SOE) based solution transforms the existing applications, data, and content into web services without reengineering the applications[16,17]. As follows from this paper, due to the advances of Internet development and deployment, e-government solutions should be based on service-oriented architecture. This approach has significant technical, organizational, social and economical advantages. Using SOE, authors can automate the departmental process workflows with support for multistage approval, using standard web based workflow as well as building workflows using Email/ SMS/IVRS.

A solution based on SOE is highly scalable, which supports high availability, failover, and load balancing features of e-governance applications. It helps in providing the improvement in overall quality of service to citizens by increasing responsiveness, productivity, accountability, and transparency. However, a question arises: how to transform electronic government solutions based on traditional architectures that are already deployed into solutions based service-oriented architecture.

REFERENCES

[1] Endrei, M., et al.: Patterns: Service-Oriented Architecture and Web Services, IBM 2004,ISBN 073845317X. URL

[2] Microsoft (2006 Dec.) Learn about Software Oriented Architecture (SOA), Retrieved from http://www.microsoft.com/biztalk/solutions/soa/overview.mspx

[3] Anttiroiko, A. V. (Ed.) *Electronic Government: Concepts, Methodologies, Tools, and Applications*. Information Science Reference, Hershey, PA, 2008.

[4] Baird, S. A. *Government role in developing an interoperability ecosystem*. In Janowski, T., Pardo, T. A. (Eds.) *Proceedings of the 1st International Conference on Theory and Practice of Electronic Governance* (Macao, China, 2007). ACM Press, New York, NY, 2007, 65-68.

[5] Buyya, R., Yeo, Ch. S., Venugopal, S. *Market-Oriented Cloud Computing: Vision, Hype, and Reality for Delivering IT Services as Computing Utilities*. In Proceedings of 2008 10th IEEE International Conference on High Performance Computing and Communications, 2008, 5-13.

[6] Cellary, W. *Strategic impact of e-government on economy and society*. In Janowski, T., Pardo, T. A. (Eds.) *Proceedings of the 2nd International Conference on Theory and Practice of Electronic Governance* (Cairo, Egypt, 2008). ACM Press, New York, NY, 2008, 149-153.

[7] Davies, J., Harris, S., Crichton, Ch., Shukla, A., and Gibbons, J. *Metadata standards for semantic interoperability in e-government*. In Janowski, T., Pardo, T. A. (Eds.) *Proceedings of the 2nd International Conference on Theory and Practice of Electronic Governance* (Cairo, Egypt, 2008). ACM Press, New York, NY, 2008, 67-75.

[8] Erl, T. *Service-Oriented Architecture: Concepts, Technology, and Design*. Prentice Hall PTR, 2005.

[9] Leavitt, N. *Is Cloud Computing Really Ready for Prime Time?* IEEE Computer, vol. 42, no. 1 (Jan. 2009), 15-20.

[10] Leganza, G. *Why Is SOA Hot In Government?* Forrester Research, 2006.

[11] Rosen, M., Lublinsky, B., Smith, K. T., and Balcer, M. J., *Applied SOA: Service-Oriented Architecture and Design Strategies*. Wiley Publishing, Indianapolis, Indiana, 2008.

[12] Ruth, S., Doh, S. *Is E-Government Ready for Prime Time?* IEEE Internet Computing, vol. 11, no. 2 (Mar./ Apr. 2007), 80-82.

[13] "Cloud Computing for E-Governance" A white paper, IIIT Hyderabad, India.

[14] Thomas B. Riley and William Sheridan, "Electronic Government in Developing and Developed Countries: The Basics By Thomas B. Riley and William Sheridan," Commonwealth Centre for e- Governance , Ottawa, Canada, October 20,2006

[15] Sameer Sachdeva, "e-Governance Action Plan for India", December'2002.

[16] Bretschneider, S., Gant, J., & Ahn, M., "A general model of egovernment service adoption: Empirical exploration." Public Management Research Conference, Georgetown Public Policy Institute Washington, D.C., October 9-11, 2003.

[17] Michiel Backus, "E-Governance and Developing Countries Introduction and examples", RESEARCH REPORT No. 3, April 2001

[18] Md. Anwarul Kabir, "E-Governance Discourse: Bangladesh Reality"

[19]Aprna Tripathi,"SOA Designing with UML 2.0",March 2010

Information Retrieval using text detection from video by image processing and speech recognition for video search optimization.

Ajay Patel, Chirag Patel, Chintan Kanadia, Gaurav Patel
BCA Department, D. D. UniversityNadiad, Gujarat, India.
[1]*ajay_d_patel@yahoo.co.in*
[2]*chirag_mca_2005@yahoo.co.in*
[3]*kc_s008@yahoo.com*
[4]*gauravpatelmca@yahoo.com*

ABSTRACT— Nowadays, publishing a video and searching a video from internet has become more popular way for information retrieval. Too much information is available in the form of video over the internet. People are accustomed to search the videos and use it for getting the information of their interest. Youtube.com is the most popular example of it. But, these videos do not provide any identification of its content. Every video has its description attached as meta-information like title, author, and description. But this information is not enough to describe the whole video content. And sometimes this description can lead user to the wrong direction. Here we introduce a model which is applicable to retrieve the content from video such as title, subtitles and speech. This content can be converted into more meaningful information, further. This proposed model not only extracts the content from video but also converts it into search friendly format. So, the searching of information from videos becomes more relevant and credible.

KEYWORDS— HMM, OCR, Extraction Engine, Indexing, Speech Recognition

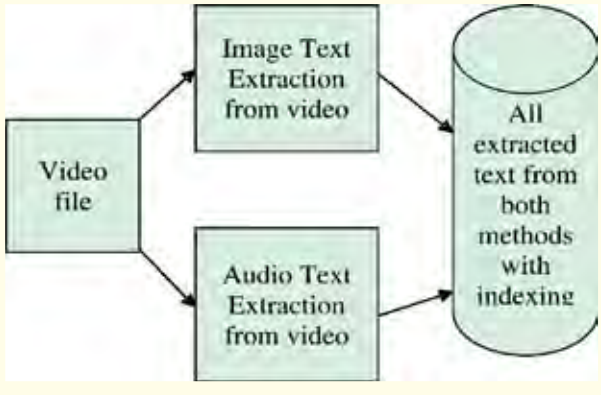
I. INTRODUCTION

Internet is the most popular way for information searching. This information is available in many formats like text, video, audio, images. Search Engines are the commonly used technology for searching the information. It crawls the world wide web, and maintains textual information. Although, it keeps information about all kinds of files like image, video, text, pdf, and office documents. But actual image and video contents (text covered in image and video) are never stored in any search engine index database. Here, we present an adaptive technique to extract the text information from video, images and video speech. In our approach, we have used OCR and Speech Recognition technique for extracting text from videos. The ongoing proliferation of videos over the internet has led to demand for system that can search the large video databases efficiently and accurately. Therefore, there is a strong need of automatic text extraction from videos and its speech. Text is abundant in videos with titles, subtitles and in speech. This text can be extracted and used for searching the videos of ones desire. Furthermore, this text can be used to classify the videos into different categories and catalogue. Currently, many text recognition algorithms are available for extracting text from video and also from the speech. In our research, we have used many efficient techniques for OCR and speech recognition and its extraction.

II. INFORMATION RETRIEVAL FROM VIDEOS

The content of any video is hidden until it has been viewed by any person. After watching the video, one can get idea about its actual content. By the technique which we are introducing here, we can get information about its content without watching it. Our main purpose behind this research is to make any video search friendly by its content. Text in video appears as either scene text or as superimposed [1] text.

A. Block *Diagram*



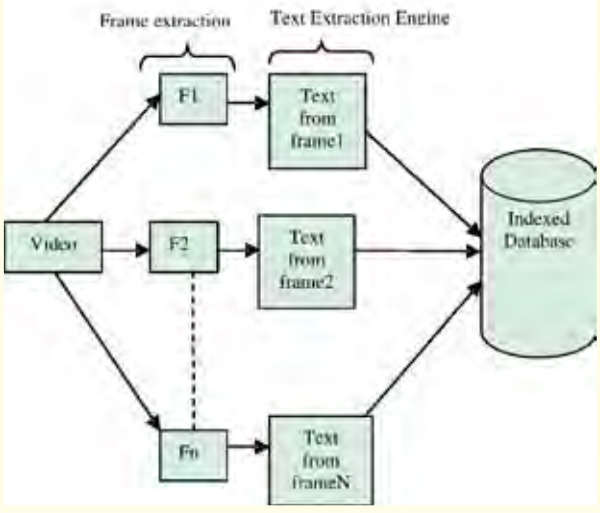
B. Our Approach

According to above model, Video file become the source for text extraction. We can extract it from images shown in video and speech spoken in video. Entire video file

is treated as stream of data. This stream will be used by two engine, simultaneously. One engine will extract the text from each video frame. And another engine will extract the text from audio speech. Finally both contents will be forwarded into indexed database.

III. IMAGE TEXT EXTRACTION

Videotext detection and recognition has been identified as one of the key components for video content retrieval [2]. To retrieve the text from visual content of video, first of all we need to divide each video into frames. These frames can be generated in the proportion of time. In our study, we have observed that one second delay for generating frame is far enough to get the new contents. Now, these frames will be transferred to image extraction engine for getting the contents.



C. Frame Extraction From Video

We are using a tool, VideoSubFinder, for extracting different frames from the running video. This programs uses text mining algorithm to separate the text from images. This algorithm produces the frame in jpg format. These pictures only cover the text information shown in the video. The other parts are removed from the generated pictures, automatically.

Video Frame

Pictures of subtitles captured from above video

D. Text Extraction Engine

Now, all the captured images are then converted to normal text using freeOCR tool. This tool converts any image text to normal text. The algorithm used in this tool does not provide good quality, if the picture is comprehensive. So, using the previous tool, we created the images which cover only text portion of a video frame. These images are fed into now freeOCR which will easily able to recover the text from the images with maximum accuracy.

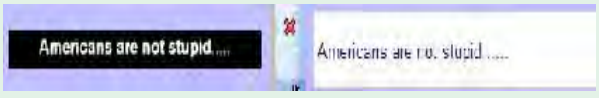
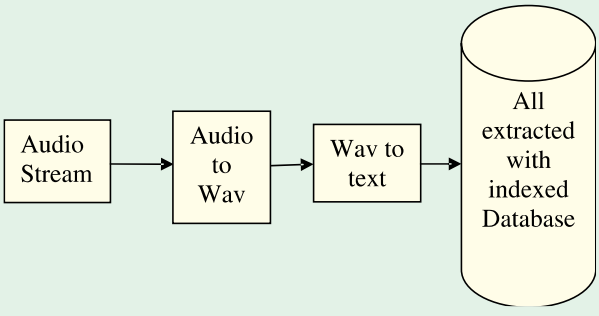


Image converted into plain text

IV. AUDIO TEXT EXTRACTION

We are converting the incoming audio stream to wav format. And finally this wav file will be converted into plain text.



Audio Text Extraction Engine

A. Audio to Wav Conversion

SwitchSound is a tool which we have used to convert incoming audio into wav file format. This tools converts various input audio format to major audio format. Our wave file frequency is 22050 Hz and 16bit is transfer rate. It is mono wave format. This is required for converting a wav file to plain text.

B. Wav to Text Conversion

WavetoText tool has been used to convert the Wav file into plain text. This tool only understands the wav file format with only 22050 Hz frequency and 16 bit transfer rate. But this limitation can be omitted in the future work. But as per our observation, using the wav file

provides good quality of output text. And many more converter supports wave to text conversion. We can also follow the Hidden Markov Model(HMM)[3].

V. INDEXING DATABASE

This is the last component of our model, which will store all extracted text in the form of keyword database. Here we can use inverted index to maintain this database. One more thing which we have introduced here that along with the keyword and it source of video, it also stores the exact location of keywords where it has been found in the video file during play.

Keywords	Video Files
Irish	[video1,5], [video2, 12],.....
Currency	[video1,5], [video3,25],.....
.....	.....

Example Index of Keywords

VI. CONCLUSIONS

In our research, we have developed a model with the different components. Tasks carried out in each components are well-defined. In this paper, we have presented the task of each components using various open-sources tools and freeware. In future, we will develop our own framework which will accomplish the functionality of this model. Already we have started our effort in this direction.

REFERENCES

[1] Automatic Text Extraction from Video for Content-Based Annotation and Retrieval Jae-Chang Shimy, Chitra Doraiz, Ruud Bolle. IEEE

[2] Accurate Overlay Text Extraction for Digital Video Analysis. Dongqing Zhang, and Shih-Fu Chang, Electrical Engineering Department, Columbia University, New York, NY 10027.

[3] SOPC-Based Speech-to-Text Conversion, M.T. Bala Murugan, M. Balaji, National Institute of Technology, Trichy.

A Web Based Electronic Book (E-Books): An Overview

Gulshan Kr. Sachdeva
Library In-Charge
JK Padampat Singhania Institute of Management & Technology, Gurgaon
E-Mail: gksachdeva30@gmail.com

ABSTRACT— Electronic Books or E-books have significantly affected the publishing industry in recent years. The growing popularity of the web and public acceptance of new e-book technologies is facilitating their spread. However, initial optimism about the growth of the medium has been tempered by a measured uptake of the medium and the withdrawal of some e-books products from the markets. The Electronic books are one way to enhance the digital library with global 24-hours-a-day and 7-days-a-week access to reliable information and they enable users to quickly retrieve and access specific research information easily, quickly and effectively. This paper attempts to update the position with e-books and e-books technologies by types of e-books their advantages and disadvantages, hardware, software and their future. E-books into the digital library has created challenges and opportunities for librarians, publishers and e-books providers and library are being unclear.

E-BOOKS

The development of the e-books is even now still in its early stages. Formatting and other issues are clearly unresolved with a variety of competing commercial products in the market. The e-book can be defined variously as a text in digital form; as a book converted into digital form; as digital reading material; as a book in a computer file format or electronic file of words; or as images with unique identifiers-the metadata may be displayed on computer screen or read on a computer through a network or viewed on a desktop/notebook, or read on any type of computer or formatted for display on specific e-book readers (Sachdeva, op.cit.). The word e-books is often used simultaneously to describe content, format, reader software and reading devices. However, separate distinctions can be made. E-book content refers to the intellectual property component, e-book format refers to software that enable one to read varying file formats on a range of hardware types and e-book reading device refers to the portable hardware available for reading e-books

ADVANTAGES AND DISADVANTAGES OF E-BOOKS

Advantages

Many of the advantages of e-books are shared by any electronic format, and offer a level of features and flexibility that are impossible in printed books. The advantages for various target groups such as the end user or reader, libraries, publishers and authors are discussed in the following sections, and the potential effect on national literacy and educational standards suggested:

(1) For the end user/reader:

- Availability/access to titles: downloading a title from the Internet is quicker and more convenient than visiting a bookshop or library. The online shop offers 24/7 services and titles are less likely to become out of print (e.g. because of finite print runs) but can still be lent out like a library book (subject to copyright license methods).
- Searching of text: it becomes easy to find passages, keywords and definitions in texts.
- Customization: offers flexibility to change display brightness, font size and style (especially for the visually impaired) and to add markup, annotations and links.
- Portability: potentially a large number of electronic titles can be carried around at once either in memory or on a personal
- Virtual bookshelf on the network.
- Multimedia facilities: audio, video (voice, music, sound, graphics, images or video clips), and arbitrary extension of given features with added external applications that can enhance an e-book in an integrated way. The authorship paradigms might vary from the print model, however – for example implying a shift from individual author to team.
- Environment: despite production and operating resources, e-books may arguably require fewer environmental resources
- Such as wood and energy in the long run.

(2) For the library:

- Instant delivery of purchased books;
- Adjustable fonts instead of large print books;
- Lower prices due to lower production costs;
- Eco-friendliness;
- Shelf space savings;
- An end to lost or damaged titles and ability to create their own texts; and
- Integration of available e-book titles into library's online catalog offers a substitute for having to visit a library building.

(3) For the publisher:

- Publishing speed: publishing process becomes quicker with electronic media, if distributed through the Internet with easy to distribute updates.
- Publishing cost: cost of printing, binding, inventory or shipping disappear.
- Storage requirement: no physical storage room is required for copies.
- Usage studies: with the aid of e-books reading habits of users can be monitored.

(4) For an author:

- Publishing: easier for authors to publish directly to niche markets, without a publisher, on the Web.
- Feedback: enhanced direct contact with readers becomes a reality, through direct author publishing and integrated electronic feedback by readers.

(5) Raising national literacy and educational standards:

- E-books can be used to raise national literacy standards and to aid publishers and authors along the way. For example, a national library online can do this by offering a list of free downloadable readings and videos. Classical and other special titles can be bought from publishers at a fair price before distribution. Collectively, all of these costs will be far less than currently payable library services and yet the potential outreach will be far larger.

DISADVANTAGES

- The disadvantages are due to shortcomings of current e-book technology and its derivatives, including costs and mis-match with current user expectations about how books are handled. The typical issues that make e-books less user friendly than their paper counterparts are display, form factors, haptic feedback (relating to the sense of touch), e-book distributors and other factors:

(1) Display:

- Resolution: typical screen resolution is 100 dpi, which is far below the resolutions of 300+ dpi used in print. The Microsoft's ClearType solution offers acceptable readability without high screen resolutions.
- Contrast and brightness: these are still far better in printed media. The E-Ink projects may be able to improve these factors for electronic devices (www.eink.com/accessed 14 May 2004).
- Color: for portable devices such as e-books, color intensity and color ranges are not of the quality of printed documents

(2) Format factors:

- Weight: an e-book is easy to carry than a dozen printed books, but it is still heavier than a single paperback volume today.
- Dimensions: the physical dimensions of an e-book are fixed and cannot be changed individually for each title.
- Parallel use: to view several books next to each other may require several hardware devices.
- Power consumption: battery life determines access to material on an e-book and utilization of solar power as additional energy source can be useful.
- Fragility: e-books are far more liable to damage when dropped, bent or otherwise abused.
- Flexibility: e-book hardware is rigid and non-flexible.

(3) Haptic feedback:

- Thickness: visualizing the amount of pages that are behind or in front of current page is relatively easy, but conveying haptic feeling of this thickness is much more difficult to achieve in an e-book.
- Browsing: to quickly judge whether a book is suitable for buying, people tend to thumb through a book quickly. A lot of effort has gone into replicating this navigation metaphor on e-books, with variable results.

- Paper and print quality: e-book titles cannot use this haptic quality indicator for a large sales volume.

(4) E-book distributors:

- Need to find a business model that works and users find acceptable.
- Need to reassure publishers about their intellectual property rights by providing a very controlled user interface, without making it too cumbersome and unwieldy for users (electronic copyright management systems have found that if the environment is tortuous, users simply do not use the system, but this can be more acceptable if content is sufficiently massive relative to the effort of navigating copyright management system).
- Must provide reliable and speedy access to the right books.
- Need to persuade academics to recommend using such a service to their students.

(5) Other factors:

- Cost: dedicated e-book readers are expensive and there is as yet no break through in sales to reduce cost

E-Book Technologies

E-Book technologies comprise of e-book hardware, e-book software and formats

E- Book Hardware

At present, delivery of e-books through CD-ROM and the Internet is the most popular mode. In 1998 another type of e-book medium was introduced to the public: dedicated readers or hand-held devices and/or slates. The emergence of the e-book has resulted in many companies manufacturing electronic reading devices (dedicated readers) used for displaying, reading and storing electronic information. NuvoMedia Inc., SoftBook Press Inc., EveryBook Inc. and Sony(Sanders and Sanders 2000) are examples of manufacturers of such devices. Gemstar International Group, one of the leading companies involved in electronic technology, acquired NuvoMedia and SoftBook Press in early 2000.

The nature of the devices market has been evolving. Initially the companies only targeted professionals who require access to lots of reference material (Judge 1998). Now, they are targeting a more general mass market such as students, academics and individuals. These machines are produced purposely for reading downloaded electronic contents. They are lightweight devices aimed at duplicating the familiar experience of reading a paper book, yet contain electronic-age features to further enhance convenience and enjoyment (Thomson Multimedia 2000). Examples include:

- NuvoMedia Rocket eBook (now obsolete);
- RCA REB 1100 and 1200 (Figure 1);
- Microsoft IPM-NET Myfriend
- Cytale Cybook

Fig. 1. Reading devices: REB 1100 and REB 1200 by Gemstar

O'Donnell (1998) stated that all these devices have one common function in that they are dedicated to reading e-books only, and are not as sophisticated as handheld computers or personal digital assistants (PDAs). This statement no longer holds, however, as since 2000 hybrid devices (Wilson 2001) that contain address books, diary, calculator and PDA-associated functions have also been used to read e-books. Additionally, these devices can be used for emailing, Internet surfing, word processing and playing MP3s. An example is the Franklin eBookMan.

The latest addition to PDA technology is the all-in-one colour monitor handheld O2 xda which allows users to access the Internet, make phone calls, send emails, use Microsoft Word and Excel, read e-books, organise a diary, listen to music and play games. PDAs are rapidly becoming a common technology at school because they allow students and teachers to do essential tasks such as note taking, word processing, graph drawing, emailing and browsing. The ReBook™ from RealTimeTouch.com will allow print, video, audio, interactive touch and wireless communications.

Another approach, demonstrated by the EInk project, is a flexible paper-like electronic display on thin sheets of plastic. The basic idea behind electronic paper is that it is a reusable display device, allowing storage of visual content on a 'page' of plastic paper. This technology is in its very early stages, but EInk has expressed the desire to create the technology to allow the development of electronic devices that work like paper books.

Nuvomedia, SoftBook Press and EveryBook were the pioneers in terms of introducing reading devices (Manes

1999). They have a number of advantages and disadvantages. Some of the advantages over traditional paper books include:

- they can store tens to hundreds of titles, making them ideal for convenient, portable access to reading material from virtually any location (Thomson Multimedia 2000);
- they incorporate features such as interactive dictionaries, bookmarking, instant searching, note taking and cross referencing;
- the back-lit display and font size can be adjusted to improve readability (Selvidge and Phillips 2000).

Full descriptions of the advantages of these reading devices can be found in Williams (2000a). Although they have a number of advantages over paper books, reading from these devices is reported to cause more eye strain and glare(Williams 2000b). Display quality is poor (because of low resolution) compared with paper, and the devices are still too expensive for the average person to buy.

E-BOOK SOFTWARE

In addition to e-book reading devices, general-purpose software book readers can be used on personal computers or laptops. These function in similar way to dedicated readers but no special hardware is required. Microsoft Reader, Adobe Acrobat Reader and Adobe Acrobat eBook Reader are three examples of such software. One advantage of software-based readers (Lynch 2001) is that in addition to offering the functions of dedicated readers, they offer extra facilities through a keyboard and larger screen size.

Two versions of the Microsoft Reader are available as a free download from the Microsoft Web site:

- for desktop and laptop computers - can read encrypted e-books, such as those available at BarnesandNoble.com and Amazon.com, as well as unencrypted e-books usually found at electronic publisher Web sites;
- for Pocket PC devices - reads only unencrypted e-books.

The Microsoft Reader interface is designed specifically for reading online. It includes ClearType technology developed by Microsoft to enhance LCD screen clarity, and allows annotation, bookmarking, drawing and font size alteration. A text-to-speech feature can be activated when the associated software is downloaded from Microsoft. However, it supports only a one-page view (see Figure 2) compared with Adobe Acrobat eBook Reader.

Acrobat Reader™ is freely downloadable software that allows users to read Portable Document Format (PDF) files. Acrobat Reader is widely used and comes pre-installed on many computers, making PDF a popular e-book format.

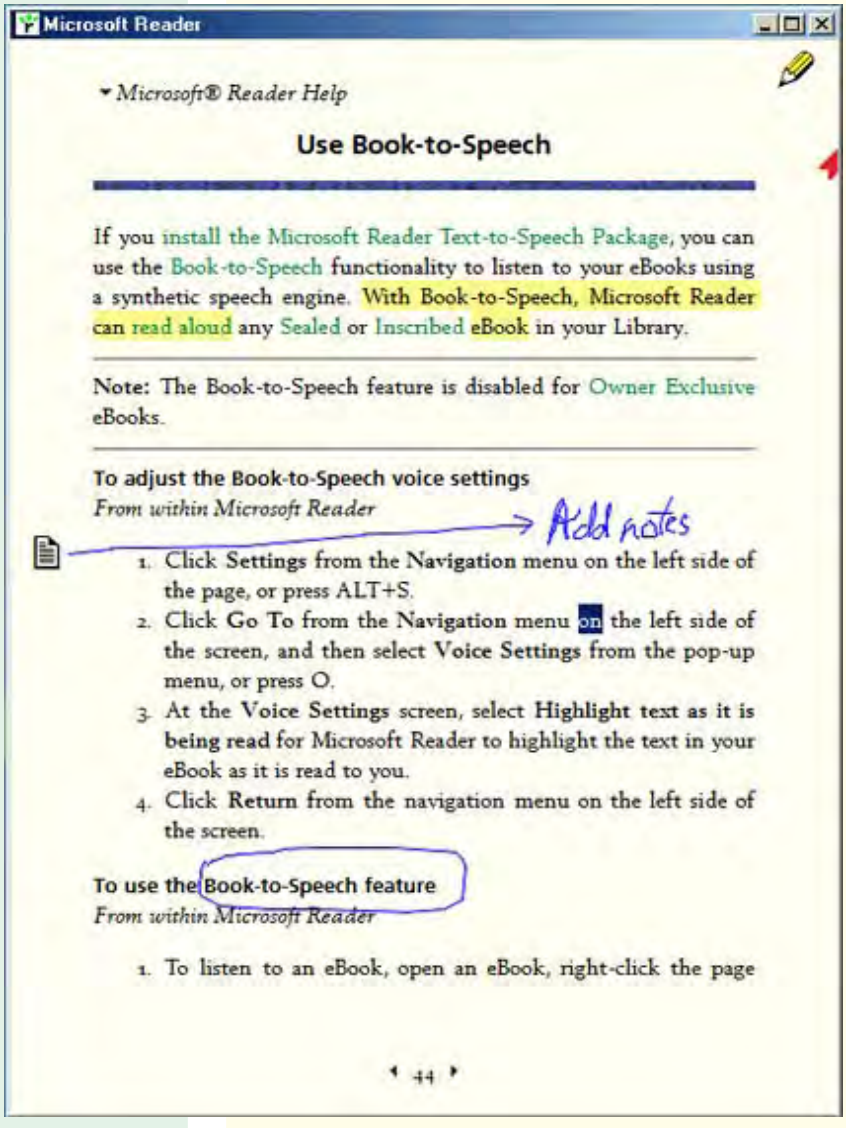


Fig. 2. Page display and personalising effects in Microsoft Reader

Glassbook Reader was recently renamed Adobe Acrobat eBook Reader and is a freely downloadable program that enables users to read PDF-based e-books on laptops, notebooks or desktop PCs. It allows easy buying and downloading of e-books at the Adobe eBook Mall and other online booksellers such as eBookTech.com, McGraw-Hill Primis Online, Taylor & Francis eBookstore, Harcourt College Publishers, CyberRead.com and BarnesandNoble.com. Features include (Figures 3 and 4):

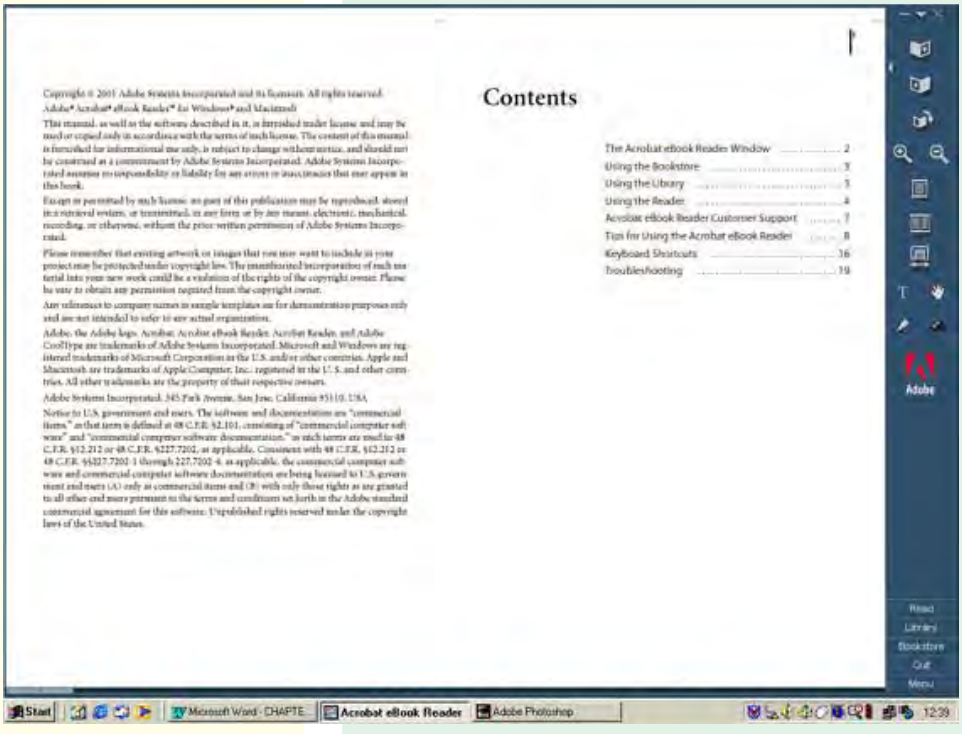


Fig. 3. Table of Contents in Acrobat eBook Reader

- single and two-page display to emulate the traditional reading experience;
- integrated bookstore access and downloading of e-books into a personalised library;

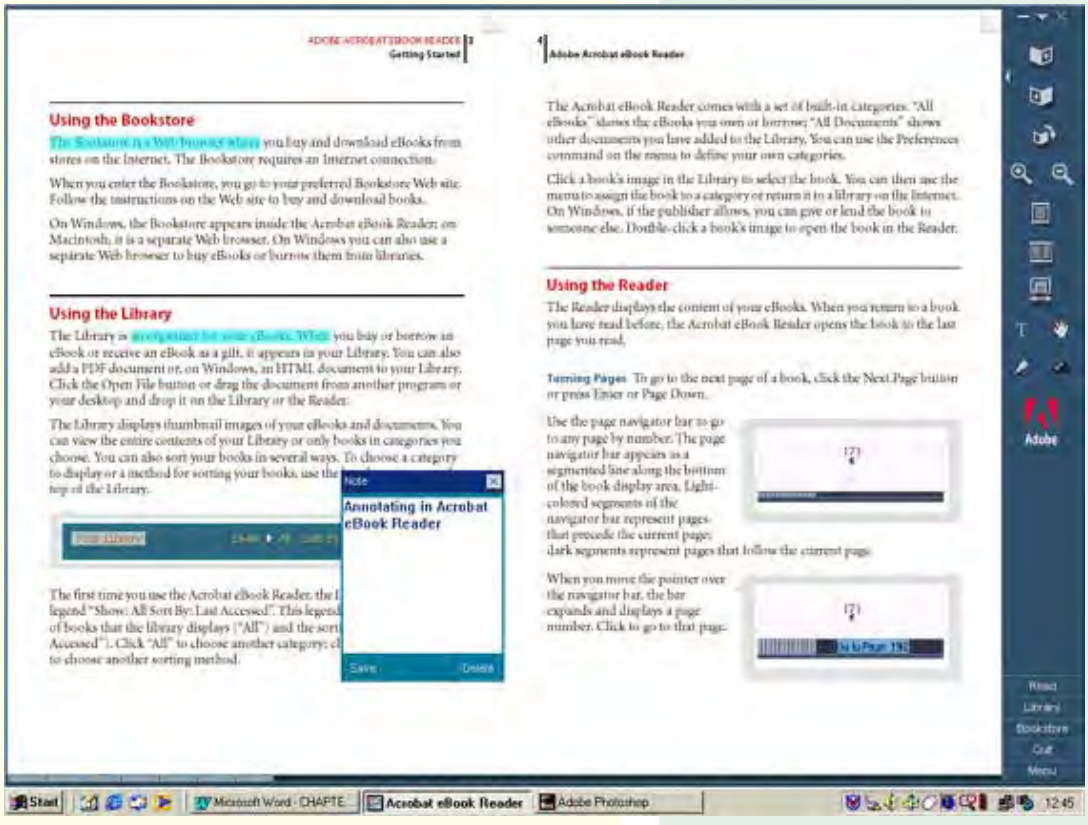


Fig. 4. Highlighting and annotating in Acrobat eBook Reader

- bookmark, highlight, annotation and search tools that make it easy to locate specific passages;
- rotating display and controls (i.e. able to change page orientation) for use on notebook computers;
- access to Web links contained in e-books;
- high-fidelity, full-colour display of e-book pages.

FORMATS

E-books are available in a wide range of formats, the simplest of which is plain ASCII-standard text. However, this format is extremely unappealing to read, cannot preserve formatting and cannot handle graphics. To solve these problems, the following formats can be used (Allen 2000, Armstrong and Lonsdale 1998, Hawkins 2000a):

- Adobe Acrobat's Portable Document Format (PDF);
- Microsoft Reader's Literature (LIT) ;
- Rich Text Format (RTF);
- Night Kitchen's Tool Kit 3 (TK3) ;
- Markup Language (e.g. HyperText Markup Language - HTML, Standard Generalised Markup Language - SGML, eXtensible Markup Language - XML);
- Software for PDAs such as AportisDoc for Palm Pilots and Pocketbooks, Palm Reader and Mobi-Reader for Palm Handheld, Handspring Visor, and Window CE devices.

The most popular journal formats according to Hitchcock et al. (1997) are either HTML or PDF. Although that study is now rather dated, the findings hold for existing e-books considered in this study. This may soon change as the LIT format is becoming widely used.

FUTURE OF E-BOOKS

Recent developments in information and computing technology have enabled higher bandwidth to be made available to support online multimedia applications. Wireless networking allows online access to information and content by mobile users on the move. Cheaper and higher capacity storage technology provides a means to contain large amounts of multimedia content. E-book is expected to create an environment where authors can circulate their works widely, producers can see their investments rewarded by high profits and distributors (information providers and librarians) can make cheap information widely available to all. This is evident by the number of e-Book initiatives. Today there are pre-made e-book covers are available in the Market (DataWorkZ.com), where an author can choose the design of the cover according to Marisa Peacock, 70% of publishing companies indicated that they felt ready for digital age, though many are doubtful that the digital age will adopt the e-reader and other related technologies over the printed word. Customers in US, followed by those in Japan and Europe are leading the sector in digitalization. The prediction is that in the next few years China will begin to edge the US out. Copyright, Digital Right Management, Standard for uniformity and retail price maintenance will be hot topic in this millennium. It is estimated that the amount of e-books will exceed 1 million by 2012 and the revenue will reach over 2 billion from library and other organization's purchase and more than 30% mobile phone users will read e-books with their mobile phone (Founder Apabi Technology Limited, Hong Kong). Technological Trends may be categories as – a) Related Technology including typeset technology, Resin Impregnated Paper (RIP) and output technology, Font e-technology and workflow management technology. b) New Technology including Full-text search, data mining, and semantic web, complex search over different resources and content related technology, Web 2.0 technology, POD technology based on Digital Right Management and Reading technology over different platform and devices.

CONCLUSION

The arrival of e-books presents many opportunities for users in new, convenient and cost-effective ways. At the moment viewing technology is a limiting factor but in future content will become paramount. The success or failure of e-books depends not only on their acceptance by users, but also on the publishing industry's ability to see beyond traditional business models and recognize that e-books are not just a replacement for paper-based books. They are a new medium with their own potential. E-book readers are still even now at an early stage of development and but the hope is that generally accepted industry standards will soon be adopted, creating a more homogenous product for a market whose potential is yet to be fully explored.

REFERENCES

[1] Allen, M. (2000) "E-Publishing FAQ". Writing-World.com http://www.writing-world.com/epublish/FAQ.html

[2] Almeida, E. (2001) eBooksN'Bytes Software Review. eBooksN'Bytes http://www.ebooksnbytes.com/compiler_review.html

[3] Armstrong, C. and Lonsdale, R. (1998) The Publishing of Electronic Scholarly Monographs and Textbookshttp://www.ukoln.ac.uk/dlis/models/studies/elec-pub/elec-pub.htm

[4] Borchers, J.O. (1999) "Electronic Books: Definition, Genres, Interaction Design Patterns". Conference on Human Factors in Computing Systems, CHI99 Workshop: Designing Electronic Books, Pittsburgh, May

[5] Bostock, S. (1998) Courseware Engineering - an overview of the courseware development process. University of Keele http://www.keele.ac.uk/depts/cs/Stephen_Bostock/docs/atceng.htm

[6] Carvajal, D. (1999) "Racing to Convert Books to Bytes". The New York Times, December 9th

[7] Carvajal, D. (2000) "Digital Publishing: From Arthur C. Clarke to Psoriasis Tales". The New York Times, February 7th

[8] Clister, J. (1999) "Electronic Books". Conference on Human Factors in Computing Systems, CHI99 Workshop: Designing Electronic Books (ACM)http://www.fxpal.com/ConferencesWorkshops/chi99deb/submissions/clister.htm

[9] Frauenfelder, M. (2000) "Digital Publishing: An Open E-Book?" The Industry Standard, August 7thhttp://www.thestandard.com/article/display/0,1151,17188,00.html

[10] Hawkins, D. (2000a) "Electronic Books: A Major Publishing Revolution, Part 1". Online, 15-28

[11] Hawkins, D. (2000b) "Electronic Books: A Major Publishing Revolution, Part 2". Online, 19-35

[12] Henke, H. (2001) Electronic Books and ePublishing (London: Springer-Verlag)

[13] Hitchcock, S., Carr, L. and Hall, W. (1997) "Web Journals Publishing: A UK Perspective". Serials, 10(3), 285-299 http://journals.ecs.soton.ac.uk/uksg.htm

[14] Judge, P. (1998) "E-books: A library on your lap". BusinessWeek, November 5thhttp://www.businessweek.com/1998/46/b3604010.htm

[15] Landoni, M. (1997) "The Visual Book System: A Study of the Use of Visual Rhetoric in the Design of Electronic Books". PhD thesis, University of Strathclyde

[16] Lynch, C. (2001) "The Battle to Define the Future of the Book in the Digital World". First Monday, 6(6), Junehttp://www.firstmonday.org/issues/issues/issue6_6/lynch/index.html

[17] Manes, S. (1999) "Gutenberg need not worry yet". Forbes, November 5th

[18] Modeland, P. R. (2000) "Leaping Onto the ePublishing Merry-Go-Round". eBookNet.com

[19] Moschella, D. (1999) "Electronic books are poised to become a key medium". CNN, April 23rdhttp://www.cnn.com/TECH/computing/9904/23/ebooks.idg/

[20] Noam, E. M. (1999) "The Dim Future of the book". Info: The Journal of Policy, Regulation and Strategy for Telecommunications Information and Media, 1(1), 5-10

[21] O'Donnell, B. (1998) "Is the world ready for digital books?" Info World Media Group Inc., July 6thhttp://www.everythingcomputers.com/1998%20Columns/July/july_6_1998_column.htm

[22] Prabha, C. (2000) "Understanding the elements of book publishing". OCLC Research http://www.oclc.org/oclc/new/n229/research.html (scroll down page)

[23] Sanders, Y. and Sanders, G. (2000) "eBooks in Japan, Part 1 of a special report". eBookNet.com

[24] Schilit, B. (1999) "Why e-Read? Finding opportunities in the merger of paper and computer". The Future of Print Media,Spring http://www.futureprint.kent.edu/articles/schilit01.htm

[25] Selvidge, P. and Phillips, C. (2000) "E-books: Are We Going Paperless?" Usability News, Usability Research Laboratory, Wichita State Universityhttp://psychology.wichita.edu/surl/usabilitynews/2W/ebook.htm

[26] Stork, P. P. (2000) "The Promise of eBook Publishing". The Internet Writing Journal (IWJ), 4(9)http://www.writerswrite.com/journal/oct00/stork.htm

[27] Thomson Multimedia (2000) "Thomson Will Mass-Produce Next Generation Rocket eBook, Soft-Book". eBookNet.com

[28] van Buren, C. and Cogswell, J. (2001) Poor Richard's Creating E-Books: How Authors, Publishers, and Corporations Get into Digital Print (Colorado: Top Floor Publishing)

[29] Williams, G. (2000a) "Where Reading Devices Make Sense". eBookNet.com

[30] Williams, G. (2000b) "A Few Things that Absolutely Must Change in the eBook Industry". eBook-Net.com

[31] Wilson, R. (2001) "Evolution of Portable Electronic Books". Ariadne, issue 29 http://www.ariadne.ac.uk/issue29/wilson/

E-governance challenges and cloud benefits

Aprna Tripathi[#], Bhawana Parihar[*]

[#]Department of Computer Application, Kanpur Institute of Technology, Kanpur, India [*]Department of Computer Science Kumau Engineering College, Dwarahat, India

[1]*aprnatripathi@gmail.com*[2]*feeling2908@gmail.com*

ABSTRACT— The worldwide revolution in Internet is changing our lives in terms of the way we work, learn and interact. These changes naturally should reflect the way government functions in terms of the organization of the government, its relationship with its citizens, institutions and businesses and cooperation with other governments. E-Governance, especially in developing countries, is looked upon as a means to change the very concept of governance resulting in empowerment of the citizens and increased transparency in public dealings by the governments; increased efficiencies in delivery of public goods is an inherent underlying assumption. This increasing generalization of technology access by citizen and organizations brings expectations and demands on government. Government aims to deliver more interactive services to citizens and businesses through E-Governance. For this, cloud computing may lead to significant cost savings. It entails use over the Internet of computing hardware and software infrastructure and applications that are remotely hosted. Cloud computing is the future generation of computing [1, 2]. Cloud computing permits to uniformly cover the whole country with e-government solutions, independently of divergence of local administrative units that may be better or worse prepared to provide e-services. Service-oriented architecture facilitates provision of compound services covering whole customer processes, where a customer may be a citizen or an enterprise. And also a analysis of cloud computing and its application in the context of e-government.

KEYWORDS—Cloud, Service, Architecture, System, Cloud Computing. E-Governance.

I. INTRODUCTION

An e-government system should be able to selecting the services and provide the services in an efficient and effective way. An effective e-government system should be reliable, cost effective, ease to maintenance, satisfying other nonfunctional. Presently two main trends in the area of information technology influence e-government. The first trend is constant development of computer infrastructure which becomes more powerful with the less expense. The second trend is constant increase of users' skills and knowledge of operating computers and the Internet. These two trends enhance possibilities of providing electronic services both in the public and the private sector. The private sector already noticed that chance – development of e-economy and e-business, both B2B and B2C, accelerated. Cloud computing is a very new concept of computing. It has been thought of as the future of computing. In cloud computing, software, hardware and network play the main role [3, 4]. The collective efforts of these entities make cloud computing possible. We can visualize the cloud as a cluster of computers which are based upon distributed systems that provide services in real time over a network. E-Government aims to provide required services to required users on the required time. An e-government system should be able to selecting the services and provide the services in an efficient and effective way. An effective e-government system should be reliable, cost effective, ease to maintenance, satisfying other nonfunctional properties as well. Unfortunately, in spite of many available resources and technologies, many challenges have been encountered in developing and implementing e-government systems [5, 6].

Cloud computing provides a new service consumption and delivery model inspired by Consumer Internet Services. Cloud computing drives down costs and accelerates cost reduction benefit. Cloud is making rapid inroads. E-Governance with cloud computing offers integration management with automated problem resolution, manages security end to end, and helps budget based on actual usage of data. At a global level, Cloud architectures can benefit government to reduce duplicate efforts and increase effective utilization of resources. According to Richard Heeks [7], more than 80% projects are in the failure category. Existing technologies is not enough to address all challenges. With the emergence of cloud computing, there are good basis to hope that some of the traditional challenges can be addressed. This paper describes the role of cloud computing standards in framing a good E-Governance strategy to realize e-Government. Governments have been slower in realizing the potential benefits of the Information Technology to provide e-services. E-services are delivering cost-effective services, which can drive the growth of the economy and government productivity. **Governance, Cloud Computing.**

II. CLOUD COMPUTING

What is cloud computing? Cloud computing encompasses a whole range of services and can be hosted in a variety of manners, depending on the nature of the service involved and the data/security needs of the contracting organization .Cloud computing is fast creating a revolution in the way information technology is used and procured by organizations and by individuals. According to the IEEE Computer Society Cloud Computing is:"*A paradigm in which information is permanently stored in servers on the Internet and cached temporarily on clients that include desktops, entertainment centers, table computers, notebooks, wall computers, handhelds, etc.*"Cloud computing is the collection of scalable, virtualized resources, which is capable of hosting application and providing required services to the users and can charge as per the uses like utility. He basic model of cloud computing is shown in fig 1.

The main goal of cloud computing is to provide ICT services with shared infrastructure and the collection of many systems. In cloud computing every facility is provided in terms of service. It provides infrastructure as a service, software as a service, platform as a service, network as a service, and data storage as a service. The main philosophy of cloud computing is to provide every required things as a service. In order to be clearer, the services in the cloud can be thought in layer architecture where various resources are available in different layers.

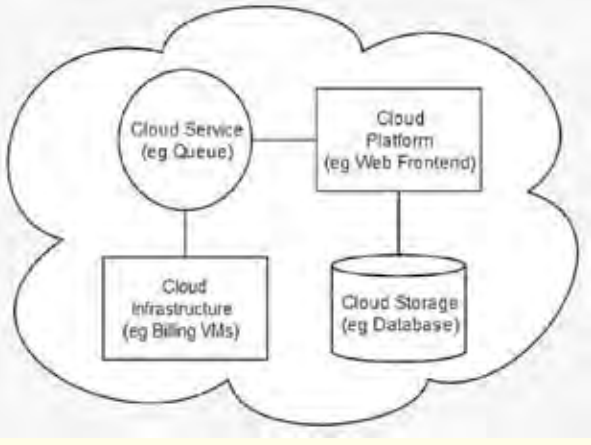


Fig. 1. Basic model of cloud computing

For individuals, cloud computing means accessing web-based email, photo sharing and productivity software, much of it for free [8]. For organizations, shifting to the cloud means having the ability to contract for computing services on-demand, rather than having to invest to host all the necessary hardware, software and support personnel necessary to provide a given level of services [9]. And for governments, the value proposition of the cloud is especially appealing, given both changing demands for IT and challenging economic conditions [10]. According to the concept of cloud computing, instead of purchasing hardware or software, a user purchases remote access to them via the Internet. There are three levels of cloud computing as shown in fig 2:

1 Infrastructure as a Service – IaaS
2 Platform as a Service – PaaS
3 Software as a Service – SaaS

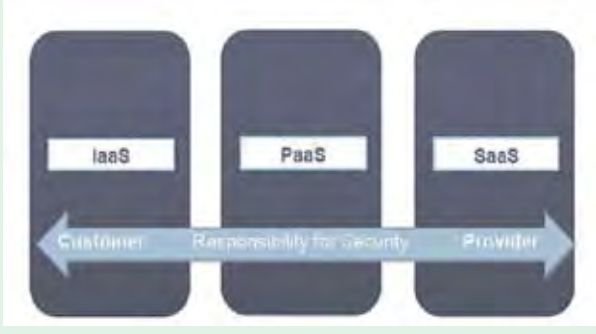


Fig. 1. Cloud Architecture

A. Infrastructure as a Service

It consists in delivering computer infrastructure as a service. The infrastructure can include servers, storage space, network equipment and system software like operating systems and database systems. The infrastructure is provided in the form of virtual environment. The applications are accessible from various client devices through a thin client interface such as a web browser From the client's point of view it looks and operates exactly like standard infrastructure, while in fact it is one of many virtual environments hosted simultaneously on the same physical infrastructure resources.

B. Platform as a Service

It consists in delivering application development environment. It supports the full life cycle of designing, implementing, testing, and deploying web applications and services. Developers, project managers, and testers are not required to download or install any development software on their local computers. The capability provided to the consumer is to deploy onto the cloud infrastructure consumer-created or acquired applications created using programming languages and tools supported by the provider.

C. *Software as a Service*

It consists in delivering complete applications such as customer relationship management or enterprise resource planning over the Internet. A client purchases an access to these applications instead of purchasing licenses and exploiting them locally. The capability provided to the consumer is to provision processing, storage, networks, and other fundamental computing resources where the consumer is able to deploy and run arbitrary software, which can include operating systems and applications.

However, Cloud computing enables providers to use distant data centers for cloud computing. Still, while some have predicted the end of the PC era with the rise of the cloud computing model, many believe that most organizations and even individuals will continue to make use of traditional PCs and laptops, even if more and more of their use will be to access the cloud [11].

III. E-GOVERNANCE CHALLENGES

In many senses the primary advantages the cloud brings are to do with cost and efficiency, which are closely intertwined. Essentially the capital costs of computing can be done away with if an organization relies on the public cloud, buying virtual server time and storage space on demand. Expenditure on IT becomes operational, rather than capital. Moreover, the physical space required for racks of servers is no longer necessary and the organization no longer incurs energy costs for running and cooling its servers. For many start-up businesses, cloud computing offers access to computing power that would otherwise be beyond their reach. Cloud computing holds a number of advantages for the government. These include "reduced cost, increased storage, higher levels of automation, increased flexibility, and higher levels of employee mobility." E-governance is a big opportunity to bring services to all citizens, but there are some serious challenges for consideration. They may me categories into the three main categories [13].

A. Technical challenges

B. Economical challenges

C. Social challenges.

A. *Technical Challenges*

Any project could not be start from the scratch. But there in a requirement of proper investments resulting in legacy systems. Some of them can be rewritten in new environment, while in some other case this could be too expensive. Therefore interoperability is the key factor between existing software and hardware platforms. Some legal aspects like security and privacy must also be an important consideration. The guarantee by the government could not suffice unless accompanied by technical solutions, transparency of procedures and possibly independent auditing.

B. *Economical challenge*

Economical issues are mainly concerned with return of investment and safeguard of the previous ones. Implementation, operational and evolutionary maintenance costs must be low enough to guarantee a good cost and benefit ratio, E-governance must be seen as a nationwide plan, implemented applications must be reusable by other administrations. The independence from the hardware and software platforms is a primary concern for portable applications there is also a main issue of maintainability, this is the key success factor for long living systems in a rapidly changing technical environment.

C. *Social challenge*

It mainly concern with the usage made by the citizens. India has a wide range of languages, and the technical literacy rate is also not so good. Then there is a challenge of accessibility, usage and acceptance of the e-governance. Even if the internet population is growing exponentially there is a significant portion of the people who may not be able to access e-government for various reasons.

Users are often not expert users they need the guidance to find the right way to perform their transactions. The successful implementation of e-governance requires a reconceptualization of the government. As e-government becomes a reality, the public sector organizational structure will change accordingly both internally and externally.

IV. CLOUD BENEFITS OVER THE E-GOVERNANCE

As in the previous section author discussed the main issues of e-governance. The cloud computing in capable of resolving such issues.

A. *Data Scaling*

The databases should be scalable, to deal with large data over the years for E-Governance applications. Where relational databases ensure the integrity of data at the lowest level, cloud databases could be scaled and can be used for such type of applications. Cloud databases available for deployment offer unprecedented level of scaling without compromising on the performance.

Cloud databases must be considered if the foremost concern is on-demand, high-end scalability, that is, large scale, distributed scalability, the kind that can't be achieved simply by scaling up[12].

B. Auditing and logging

Traceability to any changes to information content in E-Governance services is required. Corruption in government organizations can be controlled by using Information Technology services, by keeping the providers of the services accountable. Process audits, security audits must be done periodically to ensure the security of the system. Cloud can help in analyzing huge volumes of data and detecting any fraud. It can help in building and placing defense mechanisms to enhance the security, thereby making the applications reliable and available.

C. Performance and Scalability

The architecture and technology adopted for the E-Governance initiatives should be scalable and common across delivery channels .It is required to meet growing numbers and demands of citizens. If implemented, the E-Governance portals could become the biggest users and beneficiaries of Information Technology. With cloud architectures, scalability is inbuilt. Typically, E-Governance applications can be scaled vertically by moving to a more powerful machine that can offer more memory, CPU, storage.

D. Reporting and Intelligence

Data center usage (CPU, storage, network etc), peak loads, consumption levels, power usage along with time are some of the factors that needs to be monitored and reported for better utilization of resources. It minimizes costs and plan well. Profiling data enables better visibility into various services provided by the government. Cloud offers better Business Intelligence infrastructure compared to traditional ones because of its sheer size and capabilities. Applications can mine huge volumes of real time and historic data to make better decisions to offer better services.

E. Policy management

E-Governance applications have to adhere and implement policies of the governments in terms of dealing with citizens. Along with the infrastructure and data center policies has to be enforced for day to day operations. Cloud architectures help a great deal in implementing policies in data center. Policies with respect to security, application deployment etc can be formalized and enforced in the data center.

F. Systems Integration and Legacy Software

Not only the applications that are already deployed and providing services are to be moved to the cloud, but also integrate with applications deployed in the cloud. The power of Information Technology comes in co-relating the data across applications and pass messages across different systems to provide faster services to the end users. Cloud is built on SOA principles and can offer excellent solutions for integration of various applications. Also, applications can be seamlessly easily moved into cloud.

V. CONCLUSION

Cloud computing is an emerging technology in which every services are available in the cloud. Cloud is the collection of distributed computing devices. Cloud provides a solid foundation for the introduction of widespread provision of services to various stakeholders Cloud provides service through public and private clouds with the help of required technology like, system approach, distributed system, service oriented architecture, grid computing and virtualization. The domain of cloud application is very big. E-Government system requires entities like, software, hardware, service, management, network, business, policy, security etc to survive and function properly. Unfortunately current approaches or technology is insufficient to manage all these entities. Cloud computing which treats all these entities as a service can be used in e-government system. Cloud computing can handle the above mention challenges and finally address global challenges of e-government system. In order to make e-government system sustain and survive for a long time in entire world, the cloud computing is only solution for today and tomorrow. Cloud helps enabling E-Governing services faster and cheaper thereby accelerating the adoption and use of Information Technology for e-services. Cloud architectures allow rapid deployment of turnkey test environments with little or no customization. For government, there simply may be no choice but to turn to cloud-based models for computing.

REFERENCES

[1] Grossman, R. L. The case of cloud computing, proc. of IEEE Educational Activities Department, Piscataway, NJ, USA vol. 11, Issue 2, pp. 23-37, March, 2009

[2] Maria, A. F., Fenu, G., and Surcis, S., "An Approach to Cloud Computing Network," Proceedings of the 3rd International Conference on Theory and Practice of Electronic Governance, Bogota, Colombia, vol. 322, pp. 409-410, 2009.

[3] Youseff, L., Butrico, M., and DaSilva, D., "Towards a Unified Ontology of Cloud Computing," Proc. of Grid Computing Environments Workshop 2008,

[4] Leavitt, N., "Is Cloud Computing Really Ready for Prime Time," J. of ACM, vol. 42, issue 1, pp. 15–20, 2009.

[5] Janssen, M., "Cressworld Anthony Enterprise Architecture Integration in E-Government," Proc. of 38th Hawaii International Conference on System Sciences, 2005

[6] Armbrust, M et al., "Above the Clouds: A Berkeley View of Cloud Computing, Technical Report" No.UCB/EECS-2009

[7] Heek, R., "Implementing and Managing E-Government," Vistaar Publication, 2006

[8] C. Anderson, *Free: The future of a radical price*. New York: Hyperion, April 2009.

[9] Knowledge@Wharton, "How about free?: The price point that is turning industries on their heads," *Knowledge@ Wharton*, University of Pennsylvania, March 4, 2009.

[10] D.C. Wyld, *Moving to the cloud: An introduction to cloud computing in government*. Washington, DC: IBM Center for the Business of Government, November 2009.

[11] L. Ulanoff, "Google's cloud: 8 key questions," *PC Magazine*, February, 2009.

[12] David C. Wyld1, "The cloudy future of government IT: cloud computing and public sector around the world," International Journal of Web & Semantic Technology, vol 1, issue 1, January 2010

[13] "Cloud Computing for E-Governance" A white paper, IIIT Hyderabad, India.

Internet and Web as the source of High performance/Low cost Software Solutions

Major Singh[1], Lakhwinder Kaur[2]
[1]*Sant Longowal Institute of Engineering & Technology, Longowal, India*
[2]*University College of Engineering, Punjabi University, Patiala, India*
Corresponding author E-mail: mjrsingh@yahoo.com

ABSTRACT— Ubiquity of Internet and Web has provided the opportunity to devise largely distributed software solutions with enhanced capabilities and capacities. The earlier monolithic applications which were to be executed on a single machine are replaced with distributed applications where functionality is partitioned to achieve the parallel execution capabilities for high performance. For example Client-Server computing, cluster systems, peer to peer systems, grid computing and utility computing etc. Availability of the standardized protocols and services over internet has facilitated the software designers by providing a readymade functionality and infrastructure. This paper encompasses a detailed overview of the technological development of Internet and Web. Integration of Internet and Web with the Grid computing to implement Service Oriented Architecture (SOA) is also discussed.

KERWORDS— Internet, WWW, High performance computing, Internet protocols, Standardizatiobn of Internet protocols

I. INTRODUCTION

The earlier monolithic computing applications suffered with the redundancy of data/hardware/software resources across the organizations. Resource sharing across organization was possible only through the physical or manual transportation of data/software/hardware across the organizations. Thanks to the emergence and exponential growth in networking protocols and tools that the organizations started sharing resources locally as well as at the global level. LAN provided the means to setup and install cluster and multiuser high performance systems at the organization level whereas WAN provided the opportunity to setup Client-Server and large scale distributed computing system. In today's scenario an organization is no more restricted with the computing resources it owns instead it can use the computing resources of its peers remotely.

Along with the facility of resource sharing, networking infrastructure also revolutionized the software development process. The earlier batch mode applications which were progressed through the structured and object oriented programming paradigm are partitioned into component and module based programming where individual component and modules can resides on independent systems connected through high speed networks. Theses components and modules are assembled dynamically to form the software applications as and when needed. hence the reusability of code is provided to minimize efforts in the creation of future applications. The individual modules are mutually independent, extensible, platform independent and have standard interfaces to form connectivity. Before the advent of SOA and web services the technologies such as Object Request Broker and DCOM, COM, OLE were used which were based on Common Object Request Broker Architecture (CORBA) specifications[1]. The CORBA objects were available on the network from where clients can access them transparently by message passing. Clients need not to know the implementation details and whereabouts of the objects. Instead the clients know only the interfaces of theses objects as published by the server object. Through interfacing a level of abstraction is provided to hide the implementation details at the component level.

Internet and web has proved to be a readymade infrastructure to form the basis of highly distributed service oriented architecture (SOA) and high performance computing. Due to the integration of a variety of platforms of heterogeneous configurations and architectures to access internet and web, efforts are already put in to standardize the protocols and routines for interoperability.

Rest of the paper organized as. Next section covers the emergence of Internet and Web. Section 3 discusses realization of service oriented architecture through Internet and Web. Section 4 gives an overview about the integration of internet, web and grid. Last section concludes the work.

II. EMERGENCE OF INTERNET AND WEB

Information dissemination and computing has been revolutionized with the fast pace development in internet and web technologies. Internet, originally

called ARPANET was launched in 1969 by the US Department of Defense's Advanced Research Project Agency (DARPA) and was restricted to a few nodes until 1989 [2]. Ethernet idea was flown in 1973 and in 1976, it came into practice. ARPANET in the form of Internet could be expanded exponentially only in 1989 with the academic oriented research on the infrastructure. After the invention of TCP/IP models and protocols in 1974, many regional, heterogeneous local area networks were combined with the ARPANET and internetwork communication became feasible.

A breakthrough in information sharing and presaentation was the invention of Web in 1989. The World-Wide Web began in March 1989 at CERN [3]. The initial objective was to sahre information among internationally dispersed teams of researchers at CERN. But it got a rapid growth and the number of computers connected grew exponentially and it provided a plateform for business marketplace. Today, The World Wide Web Consortium (W3C) [4] is responsible to develops standards ensure the long-term growth of the Web

Web has provided an unprecedented means of information sharing across wide areas through the Internet. In early 1970s, with the advent of networking the entire computing scenario got changed. Distributed computing was experimented to harness the computing cycle on the idle machines through the networks. Evolution of high speed, high bandwidth networks, powerful computers and workstations lead to the emergence of clusters to fulfill the needs of high performance computing users [2] in low costs.

III. REALIZATION OF SERVICE ORIENTED ARCHITECTURE THROUGH INTERNET AND WEB

With the availability of global internetworking medium in the form of Internet and protocols/routines for communication, resource sharing and remote access, service oriented architecture is perceived. A web service is an abstract entity which is self contained and stateless and can be discovered and accessed on the network as and when needed. In simple terms it is a functionality available on the net along the provision of protocols to discover and access the service. Directory service is provided for a dynamic discovery of service. Service providers and consumers communicate through well defined set of messages.

Through the standard protocols UDDI [5], WSDL[6], SOAP[7] and XML, web services can interact automatically. The SOA includes the entities of Service Provider, Service Consumer and Service Registry [1], Service consumer discovers registry for required service and after getting the source of service can directly interact with service provider. Web service is described using WSDL. The syntax of service invocation along with parameters of invocation is completely described using WSDL. In response to the SOAP message of the client, web server respond back in SOAP message to provide the required functionality. The transport protocol used in web services is basically the HTTP.

The increasingly complex scientific, engineering, and commercial applications are facing shortage of resources. A single organization is unable to fulfill the resource demands of the resource intensive computing applications. Ways are being explored to integrate the resources across organizations through internet which otherwise remains idle to form a virtual high performance computing infrastructure which can take on the highly complex applications. A recent example is LHC grid [8].

In this regard Open Grid Service Architecture (OGSA) [9] is developed by Global Grid Forum (GGF). OGSA has provided a common, standard and open architecture of grid based applications. It is a web service based architecture that defines interfaces to provide the features for resource sharing and accessing in heterogeneous and highly distributed grid environment. GGF defined the framework Web Service Resource Framework to integrate the notion of web service with grid service.

IV. INTERNET, WEB AND COMPUTING GRID

Compting grid[10,11,12] aims at large scale resource sharing among geographically dispersed organizations. Service oriented grid environment is created through Virtual organizations [10] by integrating the geographically dispersed resources. For example Fig.1. shows the formation of VO to provide computing resources for weather forecasting application. Resource belongs to multiple organizations and are integrated together dynamically for the the grid service. The resource intensive scintific, engineering and commercial applications are benefitted from grid as it makes available vast resources to them at low costs.

Where the grid is a network of computing resources for resource sharing purpose world wide web is network of computers for information sharing. At the backbone of both implemenations is the large availability of internet as a medium. In grid as well as web due to

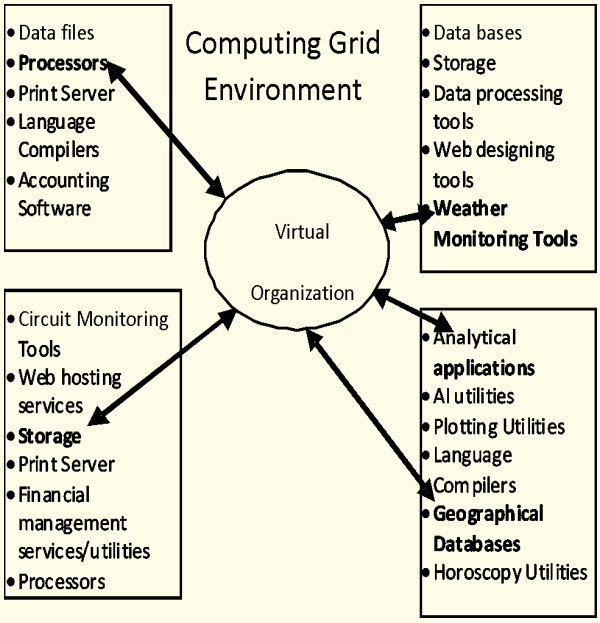


Fig. 1. VO for weather forcasting application

heterogeneity of the underlying platforms and networks standardization of protocols and interoperability needed to be addressed.

To share and access resources in the distributed environment of grid computing the services on the internet are needed which can discover the required resources, can obtain their configurations, submit requets for resource usage to execute applications, assemble results from dispersed resources, move data and results on the network etc. standardised protocols and definition language exists to implement web services hance through the integration of web service and grid a high performance computing capability at the back end can be implemented.

Internet and web technologies have many features and standardised protocols which serve as a basis of grid middleware implementaion [13]. XML provides a strong interioperability feature in the heterogenous distributed environment which is the primary requirement of grid. Various grid services can be implemented and composed on top of web services and XML.

Web can be considered as information Grid and the Grid an extended Web [14] which allows the sahring and usage of remote resopurces. Where the web has proved to be an effective and largely acceptable communication medium the grid technology aims at the large scale distributed computing in scientific, engineering and commercial applications.

V. CONCLUSION

This is a well established fact that in the past century Internet and Web has revolutionized the information technology. Along with opening the new vistas of information gathering, information access, information maintenance, presentation etc. Internet and Web has also provided the mechanisms of software engineering. Software development has underwent huge advancement in the form of code generation, execution and house keeping. A brief sketch about the development in Internet and web is drawn through this paper. How the Internet and web and related technologies are exploited to form the high performance and service oriented computing infrastructure and their integration with grid is also covered in this paper

REFERENCES

[1] Prabhu C.S.R., Grid and Cluster Computing, PHI Pvt. Ltd., New Delhi, 2008.

[2] Chetty M., Buyya R., "Weaving Computational Grids: How analogous are they with Electrical Grids?" Computing in Science and Engineering, IEEE, July/ August 2002.

[3] The European Organization for Nuclear Research http://www.cern.ch

[4] The World Wide Web Consortium (W3C) http://www.w3.org

[5] Universal Description Discovery and Integration(UDDI) http://www.uddi.org

[6] Web Service Definition Language (WSDL) http://www.w3c.org/TR/wsdl

[7] Simple Object Access Protocol http://www.w3c.org/TR/soap

[8] The Large Hadron Collider (LHC) Computing Grid http://www.cern.ch/lcg/

[9] Open Grid Services Architecture (OGSA) www.globus.org/ogsa/

[10] Foster I., Kesselman C., The GRID 2: Blueprint for a New Computing Infrastructure, Morgan-Kaufmann, 2003.

[11] Leinberger W., Kumar V., "Information power grid: The new frontier in parallel computing?" IEEE Concurrency 7 (4) (1999).

[12] Foster I., "What is Grid? A Three Point Checklist", Grid Today, July 22, V1 (6), 2002.

[13] Xue G., Fairman M.J., Cox S.J., "Exploiting Web Technologies for Grid Architecture", Proceedings of the 2nd IEEE/ACM International Sysmposium on Cluster Computing and the Grid (CCGRID 2002)

[14] Talia D., "The Open Grid Services Architecture: Where the Grid Meets the Web", IEEE INTERNET COMPUTING, Nov-Dec 2002. http://computer.org/internet/

Web Site Evaluation Through Web Mining

Sandeep Kumar Bothra,
MCA, M.Tech., Assistant Professor & Web Incharge
Jain Vishva Bharati University, Ladnun, Dist : Nagour, Rajasthan
Mob. 09352856105, Mail-id : bothrajain@gmail.com

ABSTRACT— The rapid e-commerce growth has made both business community and customers face a new situation. Due to intense competition on one hand and the customer's option to choose from several alternatives business community has realized the necessity of web Site Evaluation. Web usage mining attempts to discover useful knowledge from the secondary data obtained from the interactions of the users with the Web. Web usage mining has become very critical for effective Web site management, creating adaptive Web sites, business and support services, personalization, network traffic flow analysis and so on. In this paper, I propose the study of evaluation of a site, based on the data automatically recorded on it. Each site is electronically administered by a Web server, which logs all activities that take place in it in a file, the Web server log. All traces left by the Web users are stored in this log. From this log, we can extract the information that indirectly reflects the site's quality by applying data mining techniques.

INDEX TERMS – data mining, web usage mining, log file analysis, usage pattern discovery.

The continuous growth in the size and use of the World Wide Web imposes new methods of design and development of online information services. Most Web structures are large and complicated and users often miss the goal of their inquiry, or receive ambiguous results when they try to navigate through them. On the other hand, the e-business sector is rapidly evolving and the need for Web marketplaces that anticipate the needs of the customers is more evident than ever.

Therefore, the requirement for predicting user needs in order to improve the usability and user retention of a Web site can be addressed by personalizing it. Web personalization is the process of customizing a Web site to the needs of specific users, taking advantage of the knowledge acquired from the analysis of the user's navigational behavior (usage data) in correlation with other information collected in the Web context, namely, structure, content, and user profile data.

The objective of a Web personalization system is to "provide users with the information they want or need, without expecting from them to ask for it explicitly".

There are three factors affecting the way a user perceives and valuates a site: content, Web page design, and overall site design. The first factor concerns the goods, services, or data offered by the site. The other factors concern the way in which the site makes content accessible and understandable to its users. We distinguish between the design of individual pages and the overall site design, because a site is not simply a collection of pages—it is a network of related pages. The users will not engage in exploring it unless they find its structure intuitive.

EVALUATING A WEB SITE

Before the web site can be improved, its current usage needs to be evaluated. According to Preece et al., "Usability evaluation is the process of collecting data about the usability of a design by a specified group of users for a specific activity within a specified environment"[10]. .For this, Using traditional methods, in order to evaluate the usage of the web site, a group of users would have to be selected. Certain activities that they are expected to perform would have to be identified and their actions recorded while performing these tasks. It is impractical and costly to carry this out each time the quality of the web site is evaluated. It would be ideal to evaluate the web site based on the data that is automatically recorded by the web server. All activities that take place on a web site are recorded in a file called the web server log. By using data mining techniques, it is possible to extract information from these files which reflect the web site's quality.

WEB MINING FOR SITE EVALUATION: FORMULATING THE PROBLEM

Web mining is the application of data mining techniques to find interesting and potentially useful knowledge from Web data.

Data mining is a methodology for the extraction of knowledge from data. This knowledge is not arbitrary; it relates to a problem, the problem we want to solve. We can perform data mining to optimize the performance of a Web server, to discover which products are being purchased together, or to identify whether the site is being used as expected. The concrete specification of the problem guides us through different preparation and analysis steps of the same Web server

log. The problem of site evaluation leads us to the next questions: What data is appropriate for the analysis? How do we express a concept like "expected usage" to the miner, so that it discovers expected and unexpected navigation patterns? The first question is dealt with in the data preparation phase; the second one is resolved in the mining phase.

PREPROCESSING

The first issue in the preprocessing phase is data preparation. Depending on the application, Web log data may need to be cleaned from entries involving pages that returned an error or graphics file accesses. In some cases such information might be useful, but in others such data should be eliminated from a log file. Furthermore, crawler activity can be filtered out, because such entries do not provide useful information about the site's usability.

Most important of all is the user identification issue. There are several ways to identify individual visitors. The most obvious solution is to assume that each IP address (or each IP address/client agent pair) identifies a single visitor. Nonetheless, this is not very accurate because, for example, a visitor may access the Web from different computers, or many users may use the same IP address (if a proxy is used). A further assumption can then be made, that consecutive accesses from the same host during a certain time interval come from the same user. More accurate approaches for a priori identification of unique visitors are the use of cookies or similar mechanisms or the requirement for user registration. However, a potential problem in using such methods might be the reluctance of users to share personal information.

Assuming a user is identified, the next step is to perform session identification, by dividing the clickstream of each user into sessions. The usual solution in this case is to set a minimum timeout and assume that consecutive accesses within it belong to the same session, or set a maximum timeout, where two consecutive accesses that exceed it belong to different sessions.

The next source of difficulty is caching. The Web cache is a mechanism for reducing latency and traffic on the Web. A web cache keeps track of web pages that are requested and saves a copy of these pages for a certain period of time. Thus, if there is a request for the same web page, the cached copy is used instead of making a new request to the web server. Web caches can be configured either at the users local browsers, or at intermediate proxy servers. The obstacle that is introduced is the same for both types of cache. If the requested web page is cached, the client's request does not reach the corresponding web server holding the page. As a result, the server is not aware of the action and the page access is not recorded into the log files. One solution that has been proposed is cache-busting, i.e., the use of special HTTP headers defined either in web servers or web pages, in order to control the way that those pages are handled by caches. These headers are known as Cache-Control response headers and include directives to define which objects should be cached, how long they should be cached for, etc. however this approach works against the main motivation for using caches, i.e., the reduction of web latency.

Cooley et al. exploit the knowledge on the site's organization to resolve the caching problem [6]: If two pages not directly connected to each other are visited in sequence, then a previously visited page connecting them has been accessed again. If no such page exists, then the accesses come from two different users.

Privacy concerns and, occasionally, security considerations, make some users avoid sites that install cookies or execute scripts on their behalf. Hence, the distinction among users is mostly based on heuristics, a collection of which can be found in [6].

Preprocessing consists of converting the usage, content, and structure information contained in the various available data sources into the data abstractions necessary for pattern discovery.

PATTERN DISCOVERY

In this stage, machine earning and statistical methods are used to extract patterns of usage from the preprocessed web data. A variety of machine earning methods have been used for pattern discovery in web usage mining. These methods represent the four approaches that most often appear in the data mining literature: clustering, classification, association discovery and sequential pattern discovery. Similar to most of the work in data mining, classification methods were the first to be applied to web usage mining tasks.

Unlike the data preprocessing tools, the methods used for pattern discovery are domain-independent, meaning that they can be applied to many different domains, e.g. any web site, without concern about the content of the web site. Furthermore, most of the pattern discovery methods that have been used are general purpose, i. e. they are the same that have been applied to other data mining tasks. However, the particularities of web data have also led to the development of some new methods.

Dedicated Web log miners 123LogAnalyzer [11] and of the Web Utilization Miner WUM [2] are used. Both systems have been designed in accordance with the increased demand for intensive interaction with human users. This interaction is based on a powerful mining language in which expert users can express their background knowledge, guide the miner and gradually refine or refocus the discovery process, according to the mining results obtained after each query.

EVALUATION AND PRACTICAL APPROACH

Each query to a Web usage miner returns a set of navigation patterns. Then, the analyst faces the nontrivial problem of evaluating these patterns and deriving reliable conclusions from them. A navigation pattern describes one or more routes among given Web pages, along with statistics on how often each page of each route has been accessed. The science of statistics provides rules with which the analyst can determine whether a pattern or a component of it is significant or the output of coincidence. For example, assume a site contains some very popular page P, which users can reach from many other pages. Then, all pages containing links to P will be accessed quite frequently. However, it is not safe to conclude that these pages are also of interest to the users, since their popularity is a reflection of the popularity of P. On the other hand, if only one page leading to P is accessed frequently, while other pages are never used to reach it, this page may itself be of importance to the users. Statistical testing of the mining results is indispensable. However, site evaluation goes well beyond this test. The reader may recall that the site's designer needs insights regarding which pages should be improved and how. The combination of criteria such as frequency and expectedness of a mining result can help to this end. If the designer sees a frequently followed route and characterizes it as expected, this implies that many users perceives this part of the site as modeled by the designer. If a frequent route is surprising to the designer, this signals that many users navigate differently than originally anticipated when the site was designed. By studying this route closer and comparing it to other routes crossing it, the designer can detect pages that are not properly designed or linked and redesign them.

If the site is an online shop, it is successful if the users purchase products; if it is built for product promotion, it is successful if the users click at the advertisements. Then, borrowing from marketing theory, we can measure the conversion efficiency of a Web page as its contribution to the success of the site: For an online shop, the conversion efficiency of a Web page is the ratio of visitors that purchased a product after visiting this page to the total number of visitors that accessed the page. For a promotional site, the conversion efficiency of the page could be measured as the ratio of visitors that clicked on an advertisement after visiting the page. With this measure of "success" as a basis, the analyst can concentrate on discovered patterns containing pages with low conversion efficiency. These pages should be redesigned to better serve the purposes of the site. Ultimately, navigation pattern discovery should help the designer in improving the site.

I applied the 123loganalyzer and miner WUM on www.jvbu.ac.in Web site. The JVBU Web site accommodates the largest and most comprehensive database of Jain Vishva Bharati University, Various Courses, Study materials, Exam centers, Contact Class Centers, collection of online resources, and communication services. Access to the site's resources occurs via form-based queries. The query results are records retrieved from the database and placed in dynamically generated pages. For the analysis of this site, I have concentrated on the group of users looking for a particular Contact Class Center. The corresponding query interface allows the specification of (a) the region, (b) the type of contact class center, and (c) a text string that should be contained into an attribute of the contact class center, for example, the city, the person name and so forth. The concept hierarchies built for the analysis reflected the possible combinations of search parameters. In particular, the search options were abstracted into "RegionalSearch," Contact Class TypeSearch" and "TextSearch," and concepts were devised for the permissible parameter combinations. According to the site's design, three steps suffice to get material.

The discovered navigation patterns showed that users prefer search parameters that allow them to select a value by clicking, instead of typing text themselves. In particular, WUM was used to discover the various types of frequent navigation patterns that reflect the conversion efficiency of the search strategies offered by the JVBU.

CONCLUSION

Personalized Web access services are a demand of many users that feel overwhelmed with the information available on the Web. Building a site that satisfies this demand presupposes knowing the site as it is perceived by its users. This knowledge is not trivial to acquire, because the site designer has a different perception of the site's content and intended use than the occasional, the regular, the novice, or the expert user.

To personalize a site according to the requirements of each user, user navigation patterns must be discovered and analyzed. The navigation patterns reflect how the site is being perceived by different groups of users. It is not possible to establish a static site that satisfies all groups. Instead, a service should assist the user by finding expected and unexpected patterns that contain the user's trace thus far and adjust the content of each page in such a way that the expected route becomes apparent and the unexpected route is still possible. Services for adaptive page generation do exist. They must be adapted to this new use. Web usage analysis extracts knowledge from a Web server log. The research on data preparation and data mining in this domain already contains many remarkable contributions. It is anticipated, though, that Web usage mining is particularly difficult due to a gap between the advances of human-computer interaction, the practice of Web site organization, and the support offered by Web usage miners in evaluating the quality of a site. To close this gap, mining should better reflect the knowledge of the designer on Web site usage. This knowledge must be injected in the preparation of the data, in the instructions of the analyst to the miner, and in the interpretation of the results. To the latter task, the analyst and the designer are currently assisted by visualization tools and statistic theory. This must be enhanced by a better understanding of the results by the mining software itself.

REFERENCES

[1] BECKER, K. and VANZIN, M. (2003): Discovering interesting Usage Patterns in Web-based Learning Environments. Utility,Usability and Complexity of e-Information Systems:57-73.8-9 December 2003.

[2] Berendt, B. and Spiliopoulou, M. Analyzing navigation behavior in Web sites integrating multiple information systems. VLDB Journal, Special Issue on Databases and the Web 9, 1 (2000), 56–75.

[3] Berry, M. and Linoff, G. Data Mining Techniques for Marketing, Sales, and Customer Support. Wiley, NY, 1997.

[4] Büchner, M., Baumgarten, M. Anand, S.S., Mulvenna, M.D., and Hughes, J.G. Navigation pattern discovery from Internet data. In WEBKDD'99, San Diego, CA, (August 1999).

[5] EIRINAKI, M. and VAZIARGIANNIS, M. (2003): Web mining for web personalization. Acm Transactions on Internet Technology(TOIT) 3(1):1-27.February 2003.

[6] Cooley, R., Mobasher, B., and Srivastava, J. Data preparation for mining World Wide Web browsing patterns. Journal of Knowledge and Information Systems 1, 1 (1999).

[7] Eighmey, J. Profiling user responses to commercial Web sites. Journal of Advertising Research 37, 2 (May–June 1997), 59–66.

[8] Mannila, H. and Toivonen, H. Discovering generalized episodes using minimal occurrences. In Proceedings of the 2nd International Conference KDD'96, (1996), 146–151.

[9] SRIVASTAVA, J., COOLEY, R., DESHPANDE, M. and TAN, P.-N. (2000): Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. ACM SIGKDD Explorations Newsletter Vol I(2):12-23. January.

[10] Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S. and Carey, T. Human-Computer Interaction. Addison Wesley, 1994.

[11] ZY COMPUTING, I. (2003): 123 Log Analyzer. San Jose, USA. http://www.123loganalyzer.com