The World Wide Web, Hyper-G and Microcosm: Historical Analysis and Future Integration Directions

Noura Nassir Alomar
Software Engineering Dept., CCIS
King Saud University
Riyadh, Saudi Arabia

ABSTRACT
Hypertext systems went through several stages of development and dedicated research efforts. Currently, the World Wide Web is the backbone which links massive amount of hypertext and hypermedia documents published by communities and individuals working from all parts of the universe. However, the history of hypertext systems reveals that there were a competition between the Web and two open hypermedia systems which are Hyper-G and Microcosm. The aim of this paper is to investigate the success factors of the Web and present the ways of including open hypermedia features in the Web.

General Terms
Web Technology, Information Systems.

Keywords
Microcosm, Hyper-G, Open Hypermedia, the World Wide Web.

1. INTRODUCTION
Hyper-G, Microcosm and the World Wide Web were developed in parallel since 1989 [1, 2 and 3]. The development of the Web is accredited to Tim Berners Lee and other researchers who were working at CERN laboratory, Switzerland [1 and 4]. Moreover, Hyper-G was implemented at the Graz University of Technology in Austria while Microcosm's development was at the University of Southampton in the United Kingdom [3 and 5]. Hyper-G and Microcosm were retrieving links from link bases rather than embedding them inside documents. Moreover, Microcosm differs from Hyper-G and the Web in that it was originally designed as a peer-to-peer system while the latter two were following the client/server architecture [4]. Therefore, knowing the motives behind the Web's tremendous victory is essential. In this paper, the reasons are discussed from many perspectives in Section 1 and Section 2. The following section investigates the ways of exploiting open hypermedia features in improving the Web.

2. TECHNICAL PERSPECTIVE
2.1 Viewers
Open hypermedia allows the use of data processed by the system in any other system without restrictions and implies the possibility to access them using any viewer [6]. As the entry point to any hypermedia system is its viewer, the release of the Web was followed by the development of several browsers that were free of charge, easy to use and have simple user interfaces that hide the implementation details [11]. In 1993, Mosaic was developed by NCSA and was followed by the development of Internet Explorer and Netscape [1 and 8]. In contrast, viewing documents in Hyper-G was only available through Harmony viewers [1]. In addition, Microcosm restricts the access to its data by making it accessible using proprietary viewers developed by Microcosm's team only [6]. This is because linking is done through selecting the required data and choosing to go to the destination document or choosing to create a link [9]. Although two solutions have been introduced to overcome this problem which are either adapting external viewers to be partially-aware of Microcosm's environment by modifying the underlying code or applying the concept of "clip-board links" [9], all of them weren't able to link to specific end points and having multiple interfaces is inconvenient to the users [9]. Thus, because W3 browsers are available, open sourced, can talk with any server, can display all types of documents, have consistent user interfaces and able to function in different platforms [5, 11, and 13], no entry barriers were imposed on the Web as opposed to Hyper-G and Microcosm.

2.2 Protocols
Davis H. et al. [9] explain the importance of interoperability, scalability and adaptability of links in open hypermedia systems across diverse operating systems and machines, which is also emphasized by Dexter's Hypertext Reference Model [10]. In addition, Hyper-G and the W3 have network protocols and markup languages while microcosm didn't. For the Web, the Hypertext Transfer Protocol (HTTP)'s transmission of data regardless of its format and the web servers' ability to convert the output of its programs into Hypertext Markup Language (HTML) files and vice versa without requiring storing documents in this format contributed to the W3 success [8 and 13]. Moreover, although HTML was imposed on the early Web [20 and 21], referencing documents in the Web using Uniform Resource Identifiers (URIs) irrespective of the underlying arrangement of clients and servers and the exchanged data formats allowed to link documents in legacy systems which use protocols such as FTP and WAIS Fig. [1], [8]. Thus, making the Web scalable, extendable and allows users to share many kinds of documents[9 and 20].On the other hand, an argument can be made that Hyper-G discontinued since Hyper-G's client/server communications are imposed by Hyper-G Client/Server Protocol (HG-CSP) and Hyper-G Text Format (HTF) [1 and 11]. Moreover, although the design of Hyper-G system enables its clients to retrieve information from external servers such as W3 and Gopher servers and vice versa using gateways [12 and 5], only part of the functionality provided by Hyper-G which can be accessed by W3 and Gopher clients [11]. In addition, converting Microcosm documents into other formats is complex since it involves extensive searching for every end point inside content [18]. As a result, with the absence of standards in open hypermedia, the user might face difficulties in learning the functionalities of each system and its supported data formats [18]. On the other hand, Tim Berners Lee's direction toward standardizing HTML, HTTP

and URI played a major role in making the Web a global system [19].

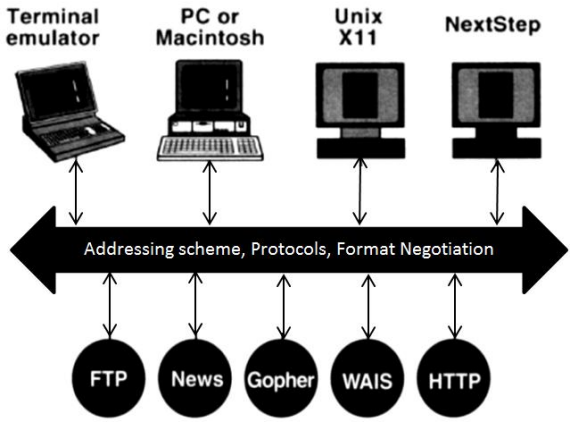

Fig.1. The Web's ability to communicate with different protocols and work on different platforms [8]

2.3 Original Design Intention

The World Wide Web and Hyper-G systems were intended to work in distributed environments while Microcosm was designed as a desktop application in local area networks to allow academics to access archival files by adding links to them and not in large-scale settings [3, 4, 16 and 22]. Despite the fact that Microcosm extended to be distributed by allowing communication between the filters of each LAN, it was able to function in small scale networks only without being able to provide the functions needed by processes distributed across different platforms and also retrieving documents in W3 is faster and more efficient [16 and 20]. Moreover, Microcosm's adaptation took place in 1994 and the Web was released and started its popularity before that time [4]. Furthermore, the clients and servers of Hyper-G and W3 communicate with the assistance of TCP/IP; however, Hyper-G contrasts the Web in that each client communicates with only one local server which in turn takes the responsibility of conversation with other servers. Moreover, since local servers caches the received responses, there is the potential of dispatching old versions or having broken links when the documents residing in remote servers are updated [2]. Therefore, p- flood algorithm was successful in sending update messages to the involved servers in wide area networks and thus maintaining the integrity of links but it was associated with the scalability problem [14 and 23]. In contrast, Tim Berners Lee's simple idea of the 404 error allowed the Web to grow by not having to ensure that each anchor has a valid endpoint and thus allows authors to publish and participate freely [15 and 23].

3. ECONOMIC AND SOCIAL PERSPECTIVES

An argument can be made that commercializing and controlling Hyper-G and Microcosm were causes for their less popularity [17 and 18]. In contrast, the World Wide Web was decentralized, not controlled, free and its specifications were open and these allowed the creation of Web communities [24]. What is more, the original communities of Hyper-G and Microcosm were scientific because all of them were developed by computer scientists in universities whereas the early community of the Web wasn't because it was developed in a physics laboratory and that's why it was followed by the foundation of W3C [8 and 25]. Therefore, people had the freedom to publish whatever they want in the Web without being restricted by writing documents of high quality or by scientific evaluation of their work before publishing it [33].

4. THE FUTURE: OPEN HYPERMEDIA ON THE WEB

In spite of the success of the Web, Open Hypermedia Systems (OHSs) have sophisticated linking facilities that don't exist in the W3 such as bidirectional links, generic links and consistent linking [7]. There are several XML-based languages that can be used for the purpose of integrating OH features on the Web such as XLink [26], XPointer [27] and XPath [28]. Using these standards, an independent storage of links is possible which allows content referencing to particular areas in text, image, audio and video files [22]. Moreover, HTML hyperlinks havea one source anchor and a one destination whereas using these technologies simplifies the application of multiple sources and multiple end points (e.g. injecting them in HTML pages using XLinkProxy) [29 and 32]. Unlike the HTML, using XLink [29] will allow attaching links to read-only documents and in other cases where modifying the linkbases is cheaper than updating the documents themselves. What is more, these technologies will open the door toward automatic computation and generation of links based on different properties of linked documents [22]. Moreover, it will be possible to enrich the experience of the users by allowing them to annotate Web pages and have their own linkbases which they can update and share [30]. These annotations are supported by Open Hypermedia Interchange Format (OHIF) which is similar to Hyper-G in supporting composites and to Microcosm in supporting generic links [32]. In addition, with the assistance OH technologies and the Resource Description Framework (RDF), obtaining metadata about hyperlinks is achievable which in turn can help in deducing knowledge about how documents are interrelated to each other which could contribute to building the Semantic Web [22 and 31].

5. COMPARATIVE EVALUATION

As shown in Table 1, each of the three hypermedia systems has its own features and weaknesses. Some researchers argue that commercializing Hyper-G and Microcosm contributed to make them less popular than the Web. On the other hand, having the Web as a free and open-sourced system allowed the users to add documents and create pages without restrictions. Thus, motivating people to join this network of documents and share their information which in turn increased the value of the Web. Another reason for the increasing popularity of the Web is its simplicity. Using the Web does not require a user to take training courses or do much effort in order to learn its basics [8]. Similarly, people were able to view the data available in the Web using browsers that are free of charge as opposed to browsing Microcosm using proprietary viewers. As the Web become more valuable due to the increased volume of documents that are created in this network, more people started to contribute and take advantage of the resources available in the Web.

Table 1: A comparative evaluation

	Architecture	Viewers	License of the system	Standards and Protocols	linking	Other powerful features
Hyper-G	Client/ server	Restricted to Harmony viewers only	Commercialized	HG-CSP and HTF	Bidirectional	1. Editing the browsed content 2. Storing the links in independent linkbases 3. Using the flooding algorithm
Microcosm	Peer-to-peer	Full functionality is accessible through Microcosm viewers.	Commercialized	-	Bidirectional	1. Having generic links 2. Storing the links in independent linkbases
The Web	Client/ server	Mosaic, Internet Explorer and Netscape	Free	HTTP, HTML and URI	Unidirectional	1. Generating the 404 error whenever a document is deleted or moved. 2. Interoperability and extensibility.

The availability and ease of use of the browsers that allowed people to view the content published in the Web made the World Wide Web accessible and reachable which in turn enabled it to grow. However, there are powerful features in the viewers that were developed for Hyper-G and Microcosm such as the ability to edit the content presented by Harmony viewers and the dynamic generation of links in Microcosm and Hyper-G [2 and 17]. Hyper-G also had features that employed caching mechanisms to achieve efficient consumption of the network bandwidth [2].As opposed to the Web, the consistency of links is ensured in Hyper-G and Microcosm because the links are stored in separate linkbases that are updated whenever a document is created, modified or deleted. Furthermore, the interoperability of the Web contributed to its success by allowing its clients and servers to talk with a variety of different platforms seamlessly. Having a common addressing system as an essential component of the Web architecture allowed the Web to scale because people were able provide addresses and references to all the Web resources without having to pay much attention to their locations [8]. Another powerful feature that contributed to the wide adoption of the Web is retrieving all types of document formats and this feature is facilitated by the Hypertext Transfer Protocol (HTTP).Similarly, standardizing the markup language (HTML) as a language for communication and not forcing the servers to store all their documents in HTML format simplified the communication between clients and servers from all over the world [8].What is more, using the TCP/IP as a transmission protocol allowed the Web to communicate with other platforms that use the same underlying protocol. This helped in reducing the communication barriers between incompatible systems.

6. CONCLUSION

Despite the powerful features of Hyper-G and Microcosm, a closed distributed hypertext project is leading the world. This paper has analyzed the reasons which pushed the Web to be the largest distributed hypermedia system in comparison with two open hypermedia systems. Some reasons are related to the viewers, protocols, data formats of the three systems and whether they were originally designed as distributed systems or not. Others are related to their original social communities and economic factors. Because the tight-coupling of URIs with documents limits the functionality of the Web, the paper presented some of the technologies that might be considered as building blocks for open hypermedia in the Web.

Future research directions might involve implementing the features of Open Hypermedia in the Web. Allowing the users to have their own linkbases that might have different areas of interest, giving them the ability to share and modify these linkbases might be beneficial. Furthermore, the users of the Web might be given the option to annotate the Web pages that they browse. Studying the mechanisms that can improve the integrity of the links and reduce the percentage of broken links in the Web is one of the promising areas of research. Researchers also can take advantage of the features of Hyper-

G and Microcosm to improve the Web. Generic links and bidirectional links can be adopted in the World Wide Web as well. Additional areas of research include studying the factors that might positively or negatively affect the information access and retrieval in the Web.

7. REFERENCES

[1] Andrews, K., Kappe, F. and Maurer, H. 1996. "The Hyper-G network information system." J. UCS the Journal of Universal Computer Science. Springer Berlin Heidelberg, 206-220.

[2] Kappe, F. 1993. "Hyper-G: A distributed hypermedia system." Proc. INET. Vol. 93.

[3] Fountain, A. M., et al. 1990. "MICROCOSM: An Open Model for Hypermedia with Dynamic Linking."

[4] Hall, W. 2011. "The ever evolving web: the power of networks", International Journal of Communication 5 ,651-664.

[5] Andrews, K. et al. 1994. "Hyper-G: A new tool for distributed hypermedia." In International Conference on Distributed Multimedia Systems and Applications.

[6] Davis, H. C., Hall, W., Heath, Ian, Hill, Gary J. and Wilkins, Robert J. 1992. "MICROCOSM: An Open Hypermedia Environment for Information Integration. "

[7] Ashman, H., and Rosemary M. S. 1999. "Computing surveys' electronic symposium on hypertext and hypermedia: editorial." ACM Computing Surveys (CSUR) 31.4es: 1.

[8] Berners-Lee, T., Cailliau R., Luotonen, A., Nielsen, H. F., and Secret, A.. 1994"The World-Wide Web." Commun. ACM 37, 8 (August 1994), 76-82.

[9] Davis, H., et al. 1992. "Towards an integrated information environment with open hypermedia systems." Proceedings of the ACM conference on Hypertext. ACM.

[10] Halasz, F., et al. 1994. "The Dexter hypertext reference model." Communications of the ACM 37.2: 30-39.

[11] Ehmayer, G., Kappel, G., and Reich, S. 1997. "Connecting databases to the Web: A taxonomy of gateways." Database and Expert Systems Applications. Springer Berlin Heidelberg.

[12] Bouvin, N. O. 1999. "Unifying strategies for Web augmentation." Proceedings of the tenth ACM Conference on Hypertext and hypermedia: returning to our diverse roots: returning to our diverse roots. ACM.

[13] Berners-Lee, T. and Cailliau, R., 1992. "The World-Wide Web", in proceedings of theconference"Computing in High Energy Physics", Annecy, France. Yellow report, CERN, Geneva, Switzerland.

[14] Kappe, F. 1996. "A scalable architecture for maintaining referential integrity in distributed information systems." J. UCS the Journal of Universal Computer Science. Springer Berlin Heidelberg. 84-104.

[15] Davis, H. C. 1999. "Hypertext link integrity." ACM Computing Surveys (CSUR)31.4es: 28.

[16] Gary, H. and Hall, W. 1994. "Extending the microcosm model to a distributed environment." Proceedings of the 1994 ACM European conference on Hypermedia technology. ACM.

[17] Hall, W., Gary, H., and Davis, H. 1993. "The microcosm link service." Fifth ACM Conference on Hypertext.

[18] Anderson, K. M. 1997. "Integrating open hypermedia systems with the World Wide Web." Proceedings of the eighth ACM conference on Hypertext.

[19] Halsey, B., and Anderson, K. M. 2000. "XLink and open hypermedia systems: a preliminary investigation." Proceedings of the eleventh ACM on Hypertext and hypermedia.

[20] Goose, S., et al. 1996 "An open framework for integrating widely distributed hypermedia resources." Multimedia Computing and Systems, Proceedings of the Third IEEE International Conference.

[21] Berners-Lee T. and Connolly D., Hypertext Markup Language 2.0, September 22, (1995) http://www.w3.org/MarkUp/html-spec/html-spec_4.html#SEC4

[22] Carr, L., Hall, W., and Roure, D. D. 1999. "The evolution of hypertext link services." ACM Computing Surveys (CSUR) 31.4es: 9.

[23] Lawrence, S., et al. 2001. "Persistence of web references in scientific research." Computer 34.2: 26-31.

[24] Gibson, D., Kleinberg, J., and Raghavan, P. 1998. "Inferring web communities from link topology." Proceedings of the ninth ACM conference on Hypertext and hypermedia: links, objects, time and space---structure in hypermedia systems: links, objects, time and space---structure in hypermedia systems.

[25] Leiner, B. M., et al. 2009. "A brief history of the Internet." ACM SIGCOMM Computer Communication Review 39.5: 22-31.

[26] DeRose, S., Orchard, D., and Maler, E., 1999. "XML Linking Language (XLink)", http://www.w3.org/TR/xlink/

[27] DeRose, S., Daniel, R., Grosso, P., Maler, E., Marsh,J., Walsh, N., 2002. "XML Pointer Language (XPointer)" (2002) http://www.w3.org/TR/xptr/

[28] Clark, J., DeRose,S. 1999. "XML Path Language (XPath)" ,http://www.w3.org/TR/xpath/

[29] Ciancarini, P., et al. 2002. "XLinkProxy: external linkbases with XLink."Proceedings of the 2002 ACM symposium on Document Engineering.

[30] Bouvin, N. O., et al. 2002. "Fluid annotations through open hypermedia: using and extending emerging web standards." Proceedings of the 11th international conference on World Wide Web.

[31] Grønbæk, K., Sloth, L., and Bouvin, N. O. 2000. "Open hypermedia as user controlled meta data for the Web." Computer Networks 33.1: 553-566.

[32] Christensen, B. G., Hansen, F. A., and Bouvin, N. O. 2003. "Xspect: bridging open hypermedia and XLink." Proceedings of the 12th international conference on World Wide Web.

[33] Donato, F. D. 2005. "Designing a Semantic Web Path to e-Science."*SWAP*.