Web–Based Personal Digital Library

Sheng-Uei Guan
National University of Singapore, Singapore

INTRODUCTION

Memex became an influential ideal and was hailed as the inspiration for hypertext and other new ways for information retrieval and organization. To deal with the explosion of scientific information, Bush's proposal for Memex focused on the problems of "locating relevant information in the published record and recording how that information was intellectually connected" (Bush, 1945). The expansion of information will be more and more serious as the Internet grows and most future computers are network-based. Thus, researchers have to consider the problems of "locating relevant published records in the published machines and recording how those records are intellectually connected".

This article proposes a mechanism that provides a new service paradigm for a network-based personal computer to browse, search, retrieve, organise, share, and publish information on the Web. Information in a machine should be made shareable with other machines, and other machines could be part of a machine's "memory extender".

As collections of relevant information either for public use or for personal use in a Web-based personal computer look like a digital library to its owner, this mechanism is called a Personal Digital Library (PDL). PDL will realise the selections of books, PDLs, and related persons through intellectual associations. As a server, PDL can store information simply and efficiently for easy retrieval and search, and provide intelligent supports for users (clients) to browse and find information. On the other hand, as a client, PDL can concurrently browse and automatically retrieve information from different PDLs (servers), and find a related person to communicate with during browsing. In addition, other PDLs can serve as memory extensions to overcome storage limitation in the local computer and avoid information duplication on the Web. PDL has potential uses in many areas such as personal use, education, commerce, finance, and entertainment. Personal users can employ any PDL with an Internet connection to manipulate their distributed information from anywhere around the world.

This article introduces PDL design and prototyping. The prototype implemented shows that the PDL concept is feasible under existing technologies. Although PDL is designed to manage personal information collections in a network-based personal computer, it will be fruitful if the design or ideas presented in this article can stimulate further development in the Digital Library research.

BACKGROUND

Memex

Bush described Memex as "a sort of mechanised private file and library" and as "a device in which an individual stores his books, records, and communications, and which is mechanised so that it may be consulted with exceeding speed and flexibility". Memex would store information on microfilms, which would be kept on a user's desk. Memex would have a scanner to enable the user to input new material, and it would also allow him to create hand-written marginal notes and comments.

In addition to the establishment of individual links, Bush also wanted Memex to support the building of trails through the material in the form of a set of links that would combine information of relevance for a specific topic. Bush emphasised (1967), that the mechanisation of "selection by association" would bring about a successful personal machine that would allow a human being to "think creatively and wisely, unencumbered by unworthy tasks," and that would allow people to "face an increasingly complex existence with hope".

The proposed PDL is a network-based Memex, which will be embedded in a personal computer. Information including bookmarks in a PDL can be made shareable with other PDLs. A Web-based "memory extender" is formed by linking PDLs together. Another essential feature of PDL is to allow people to locate and browse a certain or similar resource quickly and easily through intelligent associations.

Digital Library

Although PDL is not a particular digital library (DL) project, it has very similar goals as most DLs. In terms of service and resource provision in a DL, a number of service levels should be considered (Ormes & McClure, 1997): no services or resources provided, resource provision, self-assisted services, interactive services, video-on-demand services (Lin & Guan, 1996), and knowledge-based services.

PDL still needs to focus on the Internet, resource sharing and providing access to more information. However, there are some reasons that PDL can reach the top level. First, it is fully self-assisted because any PDL uses the same way to organise its collections, and has the same user interface,

therefore, users do not need to be assisted. Second, as to interactive service, a PDL user might discuss the use of a particular book or PDL with other users, authors, or PDL owners. Third, PDL has knowledge-based services. Examples include regularly checking changes of the favourite books or PDLs and then automatically informing the owner; providing access details and statistics of each book; and automatically forwarding a query to other PDLs. Due to such similarities with digital libraries, this proposed mechanism is called the Personal Digital Library.

In the following, we introduce two DL research projects (Cousins, 1996; Fox, 2003; Frew, Freeston, Freiras, Hill, Janee, Lovette, Nideffer, Smith, & Zheng, 2000; Guan, Yu, & Yang, 1998; Hassan & Paepcke, 1997; Kamiya, Röscheisen, & Winograd, 1996; Liu, Maly, Zubair, & Nelson, 2001; Paepcke, Cousins, Hector, Hassan, Ketchpel, Röscheisen, & Winograd, 1996; Wilensky, 1995, 1996) and some recent work on personalized bookmark services (Chen, Chen, & Sun, 2002; Kanawati & Malek, 2000, 2002; Li, Vu, Agrawal, Hara, & Takano, 1999; Yamada & Nagino, 1999).

The U.C. Berkeley Digital Library Project (Wilensky, 1995)

The goal of this project is to develop technologies that support "work-centred" digital library services, oriented to address the mission of work groups. Research areas include automated indexing, intelligent retrieval and search processes; database technology to support digital library applications; new approaches to document analysis; and data compression and communication tools for remote browsing.

PDL is a work-centred digital information system because the services provided by PDL also concentrate on the user needs. PDL improves the way to organise information for easy retrieval and search. PDL users can browse and even use different types of collections (books) not only from the local PDL but also from external PDLs. In addition, PDL provides users with a universal, customisable interface to perform information organising, browsing, retrieving, searching, and publishing.

The Stanford Digital Library Project (Stanford Group, 1995)

This project focuses on integration and interoperation (Hassan & Paepck, 1997; Paepcke et al., 1996). Its research areas include information sharing and communication models (Kamiya et al., 1996), client information interfaces (Cousins, 1996), and information finding services (Guan, Yu, & Yang, 1998).

The Integrated Digital Library will create a shared environment that links everything from personal information collections, to collections found today in conventional libraries, to large data collections shared by scientists. PDL is an integrated information system designed for network-based personal computers. Information including bookmarks in a PDL can be made shareable with other PDLs. Since bookmarks become shared resources, the information collections from other PDLs can be linked together with the local collections.

PowerBookmarks (Li et al., 1999)

PowerBookmarks (Li et al., 1999) provides personalized organization and management of bookmarks by combining the database with Web technologies. It can achieve advanced query, classification, and navigation functions and classifies the bookmarks of all users.

Personal Web Map (PWM) (Yamada & Nagino, 1999)

SeiJi Yamada and Norikatsu Nagino (1999) proposed a database named Personal Web Map (PWM) and developed the Anytime-Control algorithm to let users control their own Web map construction. PWM can help users to gather relevant information in the WWW to a small database for convenient retrieval. It will be interesting to see how their work can be extended for group work.

Semantic Web Personal Agents

Subhash, Kunjithapatham, Sheshagiri, Finin, Joshi, Peng, and Cost (2002) described the semantic Web as a vision to simplify and improve knowledge reuse on the Web. It uses software agents to collect, process, and exchange information. The PDL differs from the semantic Web personal agent in the sense that the PDL is a Memex-based approach as opposed to an agent based approach.

DESCRIPTION OF PERSONAL DIGITAL LIBRARY

Architecture

Figure 1 shows a typical PDL architecture and information flow. A PDL consists of the following components:

- **Library Explorer:** An integrated tool for information organising, browsing, retrieving, publishing, and representing. It is also a control panel to manipulate other mechanisms and application tools, and can give assistance to visitors during navigation through a series of views and facilities. Thus, the Library Explorer is a service centre in a PDL.

Related Content

Technologies for Information Access and Knowledge Management

Thomas Mandl (2009). *Encyclopedia of Information Science and Technology, Second Edition (pp. 3680-3685).*

www.irma-international.org/chapter/technologies-information-access-knowledge-management/14125

The eAssessment of Key Competences and their Relationship with Academic Performance

Fernando Martinez-Abad, Patricia Torrijos-Finciasand María José Rodríguez-Conde (2016). *Journal of Information Technology Research (pp. 16-27).*

www.irma-international.org/article/the-eassessment-of-key-competences-and-their-relationship-with-academic-performance/172089

IS Management and Success of an Italian Fashion Shoe Company

Donatella Sciuto, Giocomo Ravarine Aurelio Buonanno, Marco Tagliavine, Paolo Faverioand Joan Shroyer-Keno (2003). *Annals of Cases on Information Technology: Volume 5 (pp. 563-584).*

www.irma-international.org/chapter/management-success-italian-fashion-shoe/44565

On-Line Analytical Processing at Washtenaw Mortgage

John H. Heinrichsand William J. Doll (2001). *Annals of Cases on Information Technology: Applications and Management in Organizations (pp. 196-216).*

www.irma-international.org/article/line-analytical-processing-washtenaw-mortgage/44616

Investigating the Needs, Capabilities and Decision Making Mechanisms in Digital Preservation: Insights from a Multiple Case Study

Daniel Burdaand Frank Teuteberg (2013). *Information Resources Management Journal (pp. 17-39).*

www.irma-international.org/article/investigating-the-needs-capabilities-and-decision-making-mechanisms-in-digital-preservation/80181