# Personal Browser

Yi-Shiou Chen, Schy Chiou, Yuan-Kai Wang and Wen-Lian Hsu

Institute of Information Science
Academia Sinica
Taipei, Taiwan, R.O.C.
fax: 886-2-2782-4814
hsu@iis.sinica.edu.tw

## 1. Introduction

Browsers [1-2] are now among the most popular software for people to surf web pages and retrieve information from Internet for their own interest. However, people usually find the information they search for are scattered in several web pages, and cannot specify which parts of contents are interesting and how these parts are presented. In other words, today's browsers are weak in personalization. To enhance the capability of personalization, browsers must have more power in managing contents of web pages. In this demo, we apply natural language processing (NLP) techniques to create a browser that can be personalized.

## 2. Personal Browser

Personal Browser allows users to record the interesting parts of web pages as chunks, which are then analyzed through NLP techniques for future information retrieval. Users can also bundle these chunks coming from different web pages, together to suit their preferences. To update information in these chunks, Personal Browser pulls relevant web pages, filter these web pages to produce the desired information according to the recorded chunks, and update the information in the original page. In Personal Browser, user can organize different information chunks in the page much like composing a personal home page.

For example, a user who is interested in Australia, Singapore and Hong Kong stock quotes finds web pages that list current prices of international stock markets. The user can copy the texts that show the three stock quotes to Personal Browser. Personal Browser then analyzes the HTML source text according to the marked text, and records the results as chunks. Applying NLP tools [3] to process the HTML source text should show that this web page is relevant to stock prices, and therefore the number after a national name, such as Australia, is a price quote, which changes

very often. The URLs, marked text, and analyzed results are all recorded in a chunk for future information retrieval.

Users can assign a priority to each information chunk. In the prioritized retrieval, Personal Browser will pull relevant web pages, applying Information retrieval techniques [4-5] to match the information chunk with the web pages to produce the desired information, and present the information in a single page.

## 3. Conclusion

Personalization is important for retrieving and managing Internet information. However, today's browsers do not provide tools for users to personalize their information any better than using bookmarks. Personal Browser allows users to manage information in the client side. Thus, users have more freedom of retrieving and filtering Internet information. Message understanding techniques are crucial for the successful implementation of Personal Browser. All kinds of IR techniques can find excellent applications in Personal Browser.

### References

[1]  Microsoft Corp., *Internet Explorer*, 1995.

[2]  Netscape Communication Corp., *Netscape Communicator*, 1994.

[3]  Anil S. Chakravarthy and Kenneth B. Hasse, *"Netserf: Using Semantic Knowledge to Find Internet Information Archives,"* in Proceedings of SIGIR'95 pp. 4-11.

[4]  Christos Faloutsos and Douglas W. Oard, *"A Survey of Information Retrieval and Filtering Methods,"* in Technical Report CS-TR-3514, University of Maryland August 1995.

[5]  Timothy A. H. Bell and Alistair Moffat, *"The Design of a High Performance Information Filtering System,"* in Proceedings of SIGIR'96, pp. 12-20.