# Web Intelligence (WI)

Ning Zhong
Maebashi Inst. Tech.

Jiming Liu
Hong Kong Baptist U.

Y.Y. Yao
U. Regina

Setsuo Ohsuga
Waseda U.

## 1 Introduction

The 21st century is the age of Internet and World Wide Web. The Web revolutionizes the way we gather, process, and use information. At the same time, it also redefines the meanings and processes of business, commerce, marketing, finance, publishing, education, research, development, as well as other aspects of our daily life. The revolution is just beginning. Although individual Web-based information systems are constantly being deployed, advanced issues and techniques for developing and for benefiting from Web intelligence still remain to be systematically studied.

This position paper defines a new research field, namely *Web Intelligence* (WI for short) by giving a complete picture of WI related topics for systematic study on advanced Web technology and developing Web-based intelligent information systems. Roughly speaking, WI exploits AI and advanced information technology on the Web and Internet. It is the key and the most urgent research field of IT for business intelligence.

## 2 WI Related Topics

What are issues and research topics on WI? In order to study advanced Web technology systematically, and develop advanced Web-based intelligent information systems, we give an overview of WI related topics as shown in Figure 1 and list several major subtopics in each topic below.

- Web Human-Media Engineering:
  the art of Web page design, multimedia information representation, multimedia information processing, visualization of Web information, and Web-based human computer interface.

- Web Information Management:
  data quality management, information transformation, Internet and Web-based data management, multidimensional Web databases and OLAP (on-line analytical processing), multimedia information management, new data models for the Web, object oriented Web information management, personalized information management, semi-structured data management, use and management of metadata, Web knowledge management, Web page automatic generation and updating, as well as Web security, integrity, privacy and trust.

- Web Information Retrieval:
  approximate retrieval, conceptual information extraction, image retrieval, multi-linguistic information retrieval, multimedia retrieval, new retrieval models, ontology-based information retrieval, as well as automatic Web content cataloging and indexing.

- Web Agents:
  dynamics of information sources, e-mail filtering, e-mail semi-automatic reply, global information collecting, information filtering, navigation guides, recommender systems, remembrance agents, reputation mechanisms, resource intermediary and coordination mechanisms, as well as Web-based cooperative problem solving.

- Web Mining and Farming:
  data mining and knowledge discovery, hypertext analysis and transformation, learning user profiles, multimedia data mining, regularities in Web surfing and Internet congestion, text mining, Web-based ontology engineering, Web-based reverse engineering, Web farming, Web-log mining, and Web warehousing.

- Web Information System Environment and Foundations:
  competitive dynamics of Web sites, emerging Web technology, network community formation and support, new Web information description and query languages, theories of small world Web, Web information system development tools, and Web protocols.

- Web-Based Applications:
  business intelligence, computational societies and markets, conversational systems, customer relationship management (CRM), direct marketing, electronic commerce and electronic business, electronic library, information markets, price dynamics and pricing algorithms, measuring and analyzing Web merchandising, Web-based decision support systems, Web-based

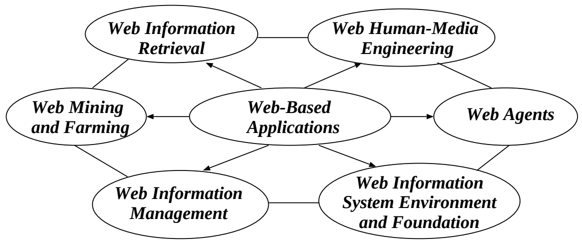


**Figure 1. A schematic diagram of WI related topics**

distributed information systems, Web-based electronic data interchange (EDI), Web marketing, and Web publishing.

## 3 WI Related Case Studies

WI presents an excellent opportunity as well as challenge for the research and development of new generation of information processing technology, as well as for exploiting business intelligence. Specifically, e-commerce activity that involves the end user is undergoing a significant revolution [8]. The ability to track users' browsing behavior down to individual mouse clicks has brought the vendor and end customer closer than ever before. It is now possible for a vendor to personalize his product message for individual customers at a massive scale. This is called *Targeted Marketing*.

Hackathorn proposed Web farming that is the systematic refining of information resources on the Web for business intelligence [3].

Abiteboul et al. systematically investigated the data on the Web and the features of semistructured data [1].

Zhong, Yao et al. proposed a way of mining peculiar data and peculiarity rules that can be used for Weblog mining [11]. They proposed ways for targeted marketing by mining classification rules and market value functions [15, 10]. They are also working on text mining on the Web including automatic construction of ontology, e-mail filtering systems, and Web-based e-business systems [14, 16].

Liu et al. are working on e-commerce agents [7]. Liu and Zhong are also working on Web agents and KDDA (Knowledge Discovery and Data Mining Agents) [5, 6].

## 4 WI Conferences

We initiated a new high-quality, high-impact biennial conference series, namely the Asia-Pacific Conference on Web Intelligence (WI). The first meeting in this new series, WI'2001, will be held in Maebashi City, Japan, October 23-26, 2001 (http://kis.maebashi-it.ac.jp/wi2001).

WI'2001 is an international forum for researchers and practitioners to present the state-of-the-art in the development of Web intelligence, to examine performance characteristics of various approaches in Web-based intelligent information technology, and to cross-fertilize ideas on the development of Web-based intelligent information systems among different domains. By idea-sharing and discussions on the underlying foundations and the enabling technologies of Web intelligence, WI'2001 is expected to stimulate the future development of new models, new methodologies, and new tools for building a variety of embodiments of Web-based intelligent information systems.

## References

[1] Abiteboul, S., Buneman, P., and Suciu, D. *Data on the Web*, Morgan Kaufmann (2000).

[2] Dong, J.Z., Zhong, N., and Ohsuga, S. "Probabilistic Rough Induction: The GDT-RS Methodology and Algorithms", Z.W. Ras and A. Skowron (eds.) *Foundations of Intelligent Systems*. LNAI 1609, Springer-Verlag (1999) 621-629.

[3] Hackathorn, R.D. *Web Farming for the Data Warehouse*, Morgan Kaufmann (2000).

[4] Ling, C.X. and Li, C. "Data Mining for Direct Marketing: problems and solutions", *Proceedings of KDD'98* (1998) 73-79.

[5] Liu, J. and Zhong, N. (eds.) *INTELLIGENT AGENT TECHNOLOGY: Systems, Methodologies, and Tools*, World Scientific (1999).

[6] Liu, J., Zhong, N., Tang, Y.Y., and Patrick S.P. Wang (eds.) *Agent Engineering*, World Scientific (2001).

[7] Liu, J. and Ye Y. (eds.) *Advances in E-commerce Agents: Brokerage, Negotiation, Security, and Mobility*, Springer-Verlag (2001).

[8] Srivastava, J. et al. "Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data", SIGKDD Explorations, Newsletter of SIGKDD, Vol.1 Issue 2 (2000) 12-23.

[9] Yao, Y.Y. and Zhong, N. "Mining Market Value Functions for Targeted Marketing" (manuscript).

[10] Yao, Y.Y. and Zhong, N. "Granular Computing using Information Tables", Lin, T.Y., Yao, Y.Y., and Zadeh, L.A. (eds.) *Granular Computing and Data Mining* In Studies in Fuzziness and Soft Computing series, Physica-Verlag (2000).

[11] Zhong, N., Yao, Y.Y., and Ohsuga, S. "Peculiarity Oriented Multi-Database Mining", J. Zytkow and Jan Rauch (eds.) *Principles of Data Mining and Knowledge Discovery*. Lecture Notes in Artificial Intelligence 1704, Springer-Verlag (1999) 136-146.

[12] Zhong, N., Skowron, A., and Ohsuga, S. (eds.) *New Directions in Rough Sets, Data Mining, and Granular-Soft Computing*, LNAI 1711, Springer-Verlag (1999).

[13] Zhong, N. and Zhou, L. (eds.) *Methodologies for Knowledge Discovery and Data Mining*, LNAI 1574, Springer-Verlag (1999).

[14] Zhong, N., Yao, Y.Y., and Kakemoto, Y. "Automatic Construction of Ontology from Text Databases", *Proc. Second International Conference on DATA MINING (Data Mining'2000)*, Wessex Press (2000).

[15] Zhong, N., Dong, J.Z., and Ohsuga, S. "Rule Discovery by Soft Induction Techniques", in Special Issue on Rough-Neuro Computing, *Neurocomputing, An International Journal*, Elsevier Science Publishers (2000).

[16] Zhong, N. "A Study on E-mail Filtering by Uncertainty Sampling and Relation Learning", Technical Report, Yamaguchi University (2000).

[17] Zhong, N. *Knowledge Discovery and Data Mining*, in the Encyclopedia of Microcomputers, Volume 27 (Supplement 6) Marcel Dekker (May of 2001).