

# Lecture Notes in Artificial Intelligence

2198

Subseries of Lecture Notes in Computer Science

Edited by J. G. Carbonell and J. Siekmann

Lecture Notes in Computer Science

Edited by G. Goos, J. Hartmanis, and J. van Leeuwen

**Springer**

*Berlin*

*Heidelberg*

*New York*

*Barcelona*

*Hong Kong*

*London*

*Milan*

*Paris*

*Tokyo*

Ning Zhong Yiyu Yao  
Jiming Liu Setsuo Ohsuga (Eds.)

# Web Intelligence: Research and Development

First Asia-Pacific Conference, WI 2001  
Maebashi City, Japan, October 23-26, 2001  
Proceedings



Springer

## Volume Editors

Ning Zhong

Maebashi Institute of Technology

Department of Systems and Information Engineering

460-1 Kamisadori-Cho, Maebashi-City 371-0816, Japan

E-mail: zhong@maebashi-it.ac.jp

Yiju Yao

University of Regina, Department of Computer Science

Regina, Saskatchewan, Canada S4S 0A2

E-mail: yyao@cs.uregina.ca

Jiming Liu

Hong Kong Baptist University, Department of Computer Science

224 Waterloo Road, Kowloon, Hong Kong, China

E-mail: jiming@comp.hkbu.edu.hk

Setsuo Ohsuga

Waseda University, Department of Information and Computer Science

3-4-1 Okubo Shinjuku-Ku, Tokyo 169, Japan

E-mail: ohsuga@fd.catv.ne.jp

## Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Web intelligence : research and development ; first Asia Pacific conference ;  
proceedings / WI 2001, Maebashi City, Japan, October 23 - 26, 2001.

Ning Zhong ... (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ;  
London ; Milan ; Paris ; Tokyo : Springer, 2001

(Lecture notes in computer science ; Vol. 2198 : Lecture notes in  
artificial intelligence)

ISBN 3-540-42730-9

CR Subject Classification (1998): H.3, C.2, H.4, H.5, J.1, I.2

ISBN 3-540-42730-9 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York

a member of BertelsmannSpringer Science+Business Media GmbH

<http://www.springer.de>

© Springer-Verlag Berlin Heidelberg 2001

Printed in Germany

Typesetting: Camera-ready by author, data conversion by PTP-Berlin, Stefan Sossna

Printed on acid-free paper SPIN: 10840672 06/3142 5 4 3 2 1 0

# Preface

This volume contains the papers selected for presentation at *the First Asia-Pacific Conference on Web Intelligence* (WI 2001) held in Maebashi TERRSA, Maebashi City, Japan, October 23-26, 2001. It was sponsored by ACM SIGART and Maebashi Institute of Technology, in cooperation with ACM SIGCHI, ACM SIGWEB, the Japanese Society for Artificial Intelligence (JSAI), JSAI SIGFAI, JSAI SIGKBS, and IEICE SIGKBSE. The conference was held jointly with the Second Asia-Pacific Conference on Intelligent Agent Technology (IAT 2001).

WI 2001 was the first conference on a new and emerging subfield of computer science known as *Web Intelligence*. It provided 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, we hoped to stimulate the future development of new models, new methodologies, and new tools for building a variety of embodiments of Intelligent Web Information Systems (IWIS).

In spite of its name, WI 2001 was truly an international conference that attracted 153 full-length research paper submissions from 31 countries and regions of all continents. Each submitted paper was reviewed by at least three experts on the basis of technical soundness, relevance, originality, significance, and clarity. Based on the review reports, 28 regular papers and 45 short papers were accepted for presentation and publication. Seven technical sessions were organized, namely: Web Information System Environment and Foundations; Web Human-Media Engineering; Web Information Management; Web Information Retrieval; Web Agents; Web Mining and Farming; Web-Based Applications.

We wish to express our gratitude to all members of the Conference Committee and the International Advisory Board for their instrumental and unflinching support. WI 2001 had a very exciting program with a number of features, ranging from technical sessions, invited talks, demos, and social programs. All of this work would not have been possible without the generous dedication of the Program Committee members and the external reviewers, of our keynote speakers, Edward A. Feigenbaum and Benjamin Wah, and invited speakers, Nick Cercone, James Hendler, W. Lewis Johnson, Riichiro Mizoguchi, Prabhakar Raghavan, and Patrick S. P. Wang, who prepared and presented very stimulating talks, and of Yiming Ye (Demos & Exhibits Chair) who solicited demo proposals and set up the program. We thank them for their strong support.

The conference Web support team at the Knowledge Information Systems Laboratory, Maebashi Institute of Technology did a terrific job of putting together and maintaining the home page for the conference as well as building a software, *cyber-chair*, which is an intelligent agent and interface among orga-

nizers, program committee members, and authors/attendees. We would like to thank Juzhen Dong, Muneaki Ohsima, and Norichika Hayazaki of the conference Web support team for their dedication and hard work.

WI 2001 could not have taken place without the great team effort of the Local Organizing Committee and the support of Maebashi Institute of Technology and Maebashi Convention Bureau. Our special thanks go to Nobuo Otani (Local Organizing Chair), Sean M. Reedy, Masaaki Sakurai, Kanehisa Sekine, and Yoshitsugu Kakemoto (the Local Organizing Committee members) for their enormous efforts in planning and arranging the logistics of the conference from registration/payment handling, venue preparation, accommodation booking, to banquet/social program organization. Our sincere gratitude goes to all of the authors who submitted papers. We are very grateful to the WI 2001 corporate sponsors: Maebashi Convention Bureau, Maebashi City Government, Gunma Prefecture Government, The Japan Research Institute, Limited, United States Air Force Office of Scientific Research, Asian Office of Aerospace Research and Development, United States Army Research Office in Far East, and Web Intelligence Laboratory, Inc. for their generous support. Last but not the least, we thank Alfred Hofmann of Springer-Verlag for his help in coordinating the publication of the proceedings.

October 2001

Ning Zhong, Yiyu Yao  
Jiming Liu, Setsuo Ohsuga

# WI 2001 Conference Committee

## *General Chairs:*

Jiming Liu  
Setsuo Ohsuga

Hong Kong Baptist University  
Waseda University, Japan

## *Program Chairs:*

Ning Zhong  
Yiyu Yao

Maebashi Institute of Technology, Japan  
University of Regina, Canada

## *Demos and Exhibits Chair:*

Yiming Ye

IBM T.J. Watson Research Center, USA

## *Local Organizing Chair:*

Nobuo Otani

Maebashi Institute of Technology, Japan

## *Advisory Board:*

Nick Cercone  
Edward A. Feigenbaum  
T.Y. Lin  
Jiming Liu  
Setsuo Ohsuga  
Ryuichi Oka  
Nobuo Otani  
Zbigniew W. Ras  
Andrzej Skowron  
Xindong Wu  
Yiyu Yao  
Philip Yu  
Ning Zhong

University of Waterloo, Canada  
Stanford University, USA  
San Jose State University, USA  
Hong Kong Baptist University  
Waseda University, Japan  
Real World Computing Partnership, Japan  
Maebashi Institute of Technology, Japan  
University of North Carolina, USA  
Warsaw University, Poland  
Colorado School of Mines, USA  
University of Regina, Canada  
IBM T.J. Watson Research Center, USA  
Maebashi Institute of Technology, Japan

*Local Organizing Committee:*

Masahiko Satori  
Tadaomi Miyazaki  
Nobuo Otani  
Sean M. Reedy  
Ning Zhong  
Masaaki Sakurai  
Toshio Kawamura  
Kanehisa Sekine  
Midori Asaka  
Yoshitsugu Kakemoto

Maebashi Institute of Technology, Japan  
Maebashi Institute of Technology, Japan  
Maebashi Institute of Technology, Japan  
Maebashi Institute of Technology, Japan  
Maebashi Institute of Technology, Japan  
Maebashi Convention Bureau, Japan  
Maebashi Convention Bureau, Japan  
Maebashi Convention Bureau, Japan  
Information Technology Agency, Japan  
Japan Research Institute, Limited, Japan



## Program Committee

Sarabjot Singh Anand	MINEit Software Limited, USA
Hendrik Blockeel	Katholieke Universiteit Leuven, Belgium
Peter Bollmann-Sdorra	Technische Universität Berlin, Germany
Cory Butz	University of Ottawa, Canada
Keith Chan	Hong Kong Polytechnic University
Hsinchun Chen	University of Arizona, USA
Ming-Syan Chen	National Taiwan University
Jingde Cheng	Saitama University, Japan
David Cheung	Hong Kong University
Robert Cooley	University of Minnesota, USA
Stefan Decker	Stanford University, USA
Liya Ding	National University of Singapore
Dieter Fensel	Vrije Universiteit Amsterdam
Benjamin Groszof	Massachusetts Institute of Technology, USA
Jiawei Han	Simon Fraser University, Canada
James Hendler	DARPA/ISO, USA
Bernardo A. Huberman	Xerox Palo Alto Research Center
W. Lewis Johnson	University of South California, USA
Tomonari Kamba	NEC Human Media Research Lab., Japan
Yasuhiko Kitamura	Osaka City University, Japan
Ramamohanarao Kotagiri	University of Melbourne, Australia
Bing Liu	National University of Singapore
Chunnian Liu	Beijing Polytechnic University, China
Jiming Liu	Hong Kong Baptist University
Brien R. Maguire	University of Regina, Canada
Akira Namatame	National Defense Academy, Japan
Jian-Yun Nie	Université de Montréal, Canada
H-O Nyongesa	Sheffield Hallam University, UK
Yukio Ohsawa	University of Tsukuba, Japan
Terry R. Payne	Carnegie Mellon University, USA
Gregory Piatetsky-Shapiro	Knowlegde Stream, USA
Mohamed Quafafou	University of Nantes, France
Vijay V. Reghavan	University of SW Louisiana, USA
Qiang Shen	University of Edinburgh, UK
Timothy K. Shih	Tamkang University, Taiwan
Myra Spiliopoulou	University of Magdeburg, Germany
Jaideep Srivastava	University of Minnesota, USA
Yasuyuki Sumi	ATR Laboratory Japan
Einoshin Suzuki	Yokohama National University, Japan
Roman W. Swiniarski	San Diego State University, USA
Atsuhiko Takasu	National Institute of Informatics, Japan

Pierre Tchounikine  
Hiroshi Tsukimoto  
Shusaku Tsumoto  
Gottfried Vossen  
Lipo Wang

Takashi Washio  
Michael S.K. Wong  
Graham Williams  
Seiji Yamada  
Yoneo Yano  
Yiyu Yao  
Yiming Ye

Tetuya Yoshida  
Ning Zhong

Lizhu Zhou  
Wojciech Ziarko

University of Maine, France  
Toshiba Corporation, Japan  
Shimane Medical University, Japan  
University of Munster, Germany  
Nanyang Technology University,  
Singapore  
Osaka University, Japan  
University of Regina, Canada  
CSIRO, Australia  
Tokyo Institute of Technology  
Tokushima University, Japan  
University of Regina, Canada  
IBM T. J. Watson Research Center,  
USA  
Osaka University, Japan  
Maebashi Institute of Technology,  
Japan  
Tsinghua University, China  
University of Regina, Canada

# Table of Contents

## Introduction

Web Intelligence (WI) (Research Challenges and Trends in the New Information Age) . . . . .	1
<i>Y.Y. Yao, N. Zhong, J. Liu, S. Ohsuga</i>	

## Invited Talks

Knowledge Is Power: The Semantic Web Vision . . . . .	18
<i>J. Hendler, E.A. Feigenbaum</i>	
From Computational Intelligence to Web Intelligence: An Ensemble from Potpourri . . . . .	30
<i>N. Cercone</i>	
Pedagogical Agents for Web-Based Learning . . . . .	43
<i>W.L. Johnson</i>	
Ontological Engineering: Foundation of the Next Generation Knowledge Processing . . . . .	44
<i>R. Mizoguchi</i>	
Social Networks on the Web and in the Enterprise . . . . .	58
<i>P. Raghavan</i>	
3D Object Recognition and Visualization on the Web . . . . .	61
<i>P.S.P. Wang</i>	

## Web Information System Environment and Foundations

A Web Proxy Cache Coherency and Replacement Approach . . . . .	75
<i>J. Aguilar, E. Leiss</i>	
Content Request Markup Language (CRML): A Distributed Framework for XML-Based Content Publishing . . . . .	85
<i>C.-H. Chiu, K.-C. Liang, S.-M. Yuan</i>	
A Rough Set-Aided System for Sorting WWW Bookmarks . . . . .	95
<i>R. Jensen, Q. Shen</i>	
Average-Clicks: A New Measure of Distance on the World Wide Web . . . . .	106
<i>Y. Matsuo, Y. Ohsawa, M. Ishizuka</i>	
Autonomy Oriented Load Balancing in Proxy Cache Servers . . . . .	115
<i>K.C. Tsui, J. Liu, H.L. Liu</i>	

Emerging Topic Tracking System ..... 125  
*K.K. Bun, M. Ishizuka*

On Axiomatizing Probabilistic Conditional Independencies  
in Bayesian Networks ..... 131  
*C.J. Butz*

Dynamic Expert Group Models for Recommender Systems ..... 136  
*D. Kim, S.W. Kim*

The ABC's of Online Community ..... 141  
*R. McArthur, P. Bruza*

Sufficient Conditions for Well-Behaved Adaptive Hypermedia Systems ... 148  
*H. Wu, P. De Bra*

## **Web Human-Media Engineering**

Towards Formal Specification of Client-Server Interactions  
for a Wide Range of Internet Applications ..... 153  
*V. Doubrovski*

Collecting, Visualizing, and Exchanging Personal Interests  
and Experiences in Communities ..... 163  
*Y. Sumi, K. Mase*

Audio Content Description in Sound Databases ..... 175  
*A.A. Wiczorkowska, Z.W. Raś*

Intelligent Interfaces for Distributed Web-Based Product  
and Service Configuration ..... 184  
*L. Ardissono, A. Felfernig, G. Friedrich, D. Jannach, R. Schäfer,  
M. Zanker*

Using Networked Workshop System to Enhance Creative Design ..... 189  
*S.S.J. Lin, E.Z.-F. Liu, M.C. Cheng, S.-M. Yuan*

Discovering Seeds of New Interest Spread from Premature  
Pages Cited by Multiple Communities ..... 195  
*N. Matsumura, Y. Ohsawa, M. Ishizuka*

Personalized Web Knowledge Management ..... 200  
*K. Takeda, H. Nomiyama*

## **Web Information Management**

Event and Rule Services for Achieving a Web-Based  
Knowledge Network ..... 205  
*M. Lee, S.Y.W. Su, H. Lam*

Knowledge-Based Validation, Aggregation, and Visualization of Meta-data: Analyzing a Web-Based Information System . . . . .	217
<i>H. Stuckenschmidt, F. van Harmelen</i>	
Online Handwritten Signature Verification for Electronic Commerce over the Internet . . . . .	227
<i>W.S. Wijesoma, K.W. Yue, K.L. Chien, T.K. Chow</i>	
A Data Model for XML Databases . . . . .	237
<i>V. Wuwongse, K. Akama, C. Anutariya, E. Nantajeewarawat</i>	
Conference Information Management System: Towards a Personal Assistant System . . . . .	247
<i>T. Mine, M. Amamiya, T. Mitamura</i>	
Automatic Intelligence Gathering from the Web: A Case Study in Container Traffic . . . . .	254
<i>J. Perdigao, A. Garg, T. Barbas, S. Scheer, G. Mastrangelo, G. Rubino</i>	
The Work Concept RBAC Model for the Access Control of the Distributed Web Server Environment . . . . .	262
<i>W.B. Shim, S. Park</i>	
A New Conceptual Graph Generated Algorithm for Semi-structured Databases . . . . .	267
<i>K.F. Wong, Y.F. Su, D. Yang, S. Tang</i>	
<b>Web Information Retrieval</b>	
A Contextual Term Suggestion Mechanism for Interactive Web Search . . . .	272
<i>C.-K. Huang, Y.-J. Oyang, L.-F. Chien</i>	
3DGML: A 3-Dimensional Graphic Information Retrieval System . . . . .	282
<i>J.H. Hwang, K.H. Lee, S. Hwang</i>	
An Evolutionary Approach to Automatic Web Page Categorization and Updating . . . . .	292
<i>V. Loia, P. Luongo</i>	
Automatic Web-Page Classification by Using Machine Learning Methods . .	303
<i>M. Tsukada, T. Washio, H. Motoda</i>	
A Theory and Approach to Improving Relevance Ranking in Web Retrieval . . . . .	314
<i>Z.W. Wang, R.B. Maguire</i>	
A Fast Image-Gathering System on the World-Wide Web Using a PC Cluster . . . . .	324
<i>K. Yanai, M. Shindo, K. Noshita</i>	

MELISSA: Mobile Electronic LSA Internet Server Search Agent . . . . .	335
<i>H.D. Brian, M.H. Garzon</i>	
Construction of a Fuzzy Multilingual Thesaurus and Its Application to Cross-Lingual Text Retrieval . . . . .	340
<i>R. Chau, C.-H. Yeh</i>	
Declustering Web Content Indices for Parallel Information Retrieval . . . . .	346
<i>Y. Chung, H.-C. Kwon, S.-H. Chung, K.R. Ryu</i>	
Indexing a Web Site to Highlight Its Content . . . . .	351
<i>E. Desmontils, C. Jacquin</i>	
Using Implicit Relevance Feedback in a Web Search Assistant . . . . .	356
<i>M. Fasli, U. Kruschwitz</i>	
The Development and Evaluation of an Integrated Imagery Access Engine . . . . .	361
<i>T. Fukumoto, K. Akahori</i>	
Conceptual Information Extraction with Link-Based Search . . . . .	367
<i>K.-J. Kim, S.-B. Cho</i>	
World Wide Web – A Multilingual Language Resource . . . . .	373
<i>F. Li, H. Sheng, W. Weisweber</i>	
Query by History Tree Manipulation . . . . .	379
<i>T. Sakairi, H. Nomiyama</i>	
Web-Based Information Retrieval Using Agent and Ontology . . . . .	384
<i>K.M. Sim, P.T. Wong</i>	
Content-Based Sound Retrieval for Web Application . . . . .	389
<i>C. Wan, M. Liu, L. Wang</i>	
<b>Web Agents</b>	
Collaborative Filtering Using Principal Component Analysis and Fuzzy Clustering . . . . .	394
<i>K. Honda, N. Sugiura, H. Ichihashi, S. Araki</i>	
<i>i</i> JADE IWSshopper: A New Age of Intelligent Mobile Web Shopping System Based on Fuzzy-Neuro Agent Technology . . . . .	403
<i>R.S.T. Lee</i>	
Wireless Agent Guidance of Remote Mobile Robots: Rough Integral Approach to Sensor Signal Analysis . . . . .	413
<i>J.F. Peters, S. Ramanna, A. Skowron, M. Borkowski</i>	

Ontology-Based Information Gathering Agents .....	423
<i>Y.-J. Chen, V.-W. Soo</i>	
An Effective Conversational Agent with User Modeling Based on Bayesian Network .....	428
<i>S.-I. Lee, C. Sung, S.-B. Cho</i>	
Information Fusion for Intelligent Agent-Based Information Gathering ....	433
<i>Y. Li</i>	
An Adaptive Recommendation System with a Coordinator Agent .....	438
<i>M. Lim, J. Kim</i>	
Interactive Web Page Filtering with Relational Learning .....	443
<i>M. Okabe, S. Yamada</i>	
A Fuzzy Rule-Based Agent for Web Retrieval-Filtering .....	448
<i>S. Vrettos, A. Stafylopatis</i>	
<b>Web Mining and Farming</b>	
Implementation Issues and Paradigms of Visual KDD Systems .....	454
<i>J. Han, N. Cercone</i>	
Re-engineering Approach to Build Domain Ontologies .....	464
<i>A. Kayed, R.M. Colomb</i>	
Discovery of Emerging Topics between Communities on WWW .....	473
<i>N. Matsumura, Y. Ohsawa, M. Ishizuka</i>	
Mining Web Logs to Improve Web Caching and Prefetching .....	483
<i>Q. Yang, H.H. Zhang, I.T.Y. Li, Y. Lu</i>	
Mining Crawled Data and Visualizing Discovered Knowledge .....	493
<i>V. Dubois, M. Quafafou, B. Habegger</i>	
Categorizing Visitors Dynamically by Fast and Robust Clustering of Access Logs .....	498
<i>V. Estivill-Castro, J. Yang</i>	
Online Learning for Web Query Generation: Finding Documents Matching a Minority Concept on the Web .....	508
<i>R. Ghani, R. Jones, D. Mladenic</i>	
A Formal Ontology Discovery from Web Documents .....	514
<i>N. Ogata</i>	
The Variable Precision Rough Set Model for Web Usage Mining .....	520
<i>V.U. Maheswari, A. Siromoney, K.M. Mehata</i>	

Supporting Cooperative Consensus Formation via Ontologies ..... 525  
*K. Sumi, R. Mizoguchi*

**Web-Based Applications**

Web-Based Intelligent Call Center for an Intensive Care Unit ..... 530  
*K. Han, D. Lee*

Electronic Homework on the WWW ..... 540  
*C. Liu, L. Zheng, J. Ji, C. Yang, J. Li, W. Yang*

The Shopping Gate - Enabling Role- and Preference-Specific  
 e-Commerce Shopping Experiences ..... 549  
*M. Stolze, M. Ströbel*

Building Reusable and Adaptable Web-Based Courses ..... 562  
*P. Forcheri, M.T. Molfino, S. Moretti, A. Quarati*

Group Learning Support System for Software Engineering Education –  
 Web-Based Collaboration Support between the Teacher Side  
 and the Student Groups – ..... 568  
*A. Hazeyama, A. Nakako, S. Nakajima, K. Osada*

The Intelligent Electronic Shopping System Based on  
 Bayesian Customer Modeling ..... 574  
*J. Ji, L. Zheng, C. Liu*

Leveraging a Web-Aware Self-Organization Map Tool  
 for Clustering and Visualization ..... 579  
*S.-T. Li*

Experiencing NetPeas: Another Way of Learning ..... 584  
*E.Z.-F. Liu, S.S.J. Lin, S.-M. Yuan*

ITMS: Individualized Teaching Material System – Web-Based Exploratory  
 Learning Support System by Adaptive Knowledge Integration – ..... 589  
*H. Mitsuhashi, Y. Ochi, Y. Yano*

An Intelligent Sales Assistant for Configurable Products ..... 596  
*M. Molina*

Reconciling of Disagreeing Data in Web-Based Distributed Systems  
 Using Consensus Methods ..... 601  
*N.T. Nguyen*

Web Based Digital Resource Library Tracing Author’s Quotation ..... 606  
*Y. Ochi, Y. Yano, R. Wakita*

**Author Index** ..... 613



# Web Intelligence (WI)

## Research Challenges and Trends in the New Information Age

Y.Y. Yao<sup>1</sup>, Ning Zhong<sup>2</sup>, Jiming Liu<sup>3</sup>, and Setsuo Ohsuga<sup>4</sup>

<sup>1</sup> Department of Computer Science, University of Regina  
Regina, Saskatchewan, Canada, S4S 0A2  
yyao@cs.uregina.ca

<sup>2</sup> Department of Information Engineering, Maebashi Institute of Technology  
Japan  
zhong@maebashi-it.ac.jp

<sup>3</sup> Department of Computer Science, Hong Kong Baptist University  
Hong Kong  
jiming@Comp.HKBU.Edu.HK

<sup>4</sup> Department of Information and Computer Science, Waseda University  
Japan

**Abstract.** This paper is about a new research field called *Web Intelligence* (WI for short). We try to explain the needs for coining the term as a sub-discipline of computer science for systematic studies on advanced Web related theories and technologies, as well as the design and implementation of Intelligent Web Information Systems (IWIS). Background information and related topics are discussed in an attempt to demonstrate why we consider WI to be a subject worthy of study and, at the same time, to establish a starting point for the further development of WI.

## 1 Introduction

With the rapid growth of Internet and World Wide Web (WWW), we have now entered into a new information age. The Web provides a total new media for communication, which goes far beyond the traditional communication medias, such as radio, telephone and television. The Web has significant impacts on both academic research and ordinary daily life. It revolutionizes the way in which information is gathered, stored, processed, presented, shared, and used. The Web offers new opportunities and challenges for many areas, such as business, commerce, marketing, finance, publishing, education, research and development. For computer scientists, the Web introduces many new research topics and provides a new platform to reconsider old problems. It might be high time to create a new sub-discipline of computer science covering theories and technologies related to the Web. *Web Intelligence* is our proposal for this purpose.

The authors of this paper conceived *Web Intelligence* (WI for short) in late 1999. We felt that although a number of conferences and journals publish or

cover Web or Internet related topics, there was no conference and journal devoted to intelligence aspects in the design and implementation of Web information systems. We suspected that there exists a need for a conference devoted to Web Intelligence. At the 24th Annual International Computer Software and Applications Conference (IEEE COMPSAC) in 2000, we first introduced Web Intelligence and formally announced the new *Web Intelligence* conference in a position paper at a Panel on Data Mining and Web Information Systems [49]. We are impressed by the quick and vast responses, as well as kind support, from research community and reputable publishers.

The main objective of this paper is to formally initiate a sub-discipline of computer science by coining the term Web Intelligence, into which Web related research can be fitted. It is more a proposal and an appeal for the creation of WI on its own rights, rather than a precise definition of what is exactly WI. We are more concerned with the necessity and benefits of WI, as well as research topics of WI. It is our intention to create further discussion and critical examination of WI among researchers working on Web related topics.

The rest of the paper is organized as follows. In Section 2, we provide a definition of Web Intelligence. In Section 3, we argue that it is necessary and beneficiary to have a new sub-discipline of computer science labelled by WI. In Section 4, we present an overview of Artificial Intelligence and show its relevance to WI. In Section 5, we provide a list of topics of WI. In Section 6, we discuss trends and challenges of WI related research and development. Section 7 is devoted to intelligent Web Agents (WA). Finally, Section 8 introduces the Web Intelligence conference, and Section 9 gives conclusion, respectively.

## 2 What Is Web Intelligence?

At this very early stage, we are not sure if a formal definition of Web Intelligence is useful or desirable. Nevertheless, we suggest the following definition:

“Web Intelligence (WI) exploits Artificial Intelligence (AI) and advanced Information Technology (IT) on the Web and Internet.”

This definition has the following implications. The basis of WI is AI and IT. The “I” happens to be shared by both “AI” and “IT”, although with different meanings in them, and “W” defines the platform on which WI research is carried out. The goal of WI is the joint goals of AI and IT on the new platform of the Web. That is, WI applies AI and IT for the design and implementation of Intelligent Web Information Systems (IWIS). An IWIS should be able to perform functions normally associated with human intelligence, such as reasoning, learning, and self improvement.

There perhaps might not be a standard and non-controversial definition of WI, as the case that there is no standard definition of AI. One may argued that our definition of WI focuses more on the software aspects of the Web. It is not our intention to exclude any research topic using the proposed definition. The term, Web Intelligence, should be considered as an umbrella or a label of a new

branch of research centered on the Web. Our definition simply states the scopes and goals of WI. This allows us to include any theories and technologies that either fall in the scopes or aim at the same goals. To complement the formal definition, we try to make the picture clearer by listing topics to be covered by WI.

WI will be an ever-changing research branch. It will be evolving with development of the Web as new media for information gathering, storage, processing, delivery and utilization. It is our expectation that WI will be evolved into an inseparable research branch of computer science. Although no one can predict the future in detail and without uncertainty, it is clear that WI would have huge impacts on the application of computers, which in turn will effect our everyday lives.

### 3 Motivations and Justifications for WI

The introduction of Web Intelligence (WI) can be motivated and justified from both academic and industrial perspectives.

Two features of the Web make it a useful and unique platform for computer applications and research, the size and complexity. The Web contains a huge amount of interconnected Web documents known as Web pages. For example, the popular search engine Google claims that it can search 1,346,966,000 pages as of February 2001. The sheer size of the Web leads to difficulties in the storage, management, and efficient and effective retrieval of Web documents. The complexity of the Web, in terms of connectivity and diversity of Web documents, forces us to reconsider many existing information systems, as well as theories, methodologies and technologies underlying those systems. One has to deal with a heterogeneous collection of structured, unstructured, semi-structured, inter-related, and distributed Web documents consisting of texts, images and sounds, instead of homogeneous collection of structured and unrelated objects. The latter is the subject of study of many conventional information systems, such as databases, information retrieval, and multi-media systems. To accommodate the needs of the Web, one needs to study issues on the design and implementation of the Web-based information systems by combining and extending results from existing intelligent information systems. Existing theories and technologies need to be modified or enhanced to deal with complexity of the Web. Although individual Web-based information systems are constantly being deployed, advanced issues and techniques for developing and for benefiting from the Web remain to be systematically studied. The challenges brought by the Web to computer scientists may justify the creation of the new sub-discipline, WI, for carrying out Web-related research.

The Web increases the availability and accessibility of information to a much larger community than any other computer applications. The introduction of Personal Computers (PCs) brought the computational power to ordinary people. It is the Web that delivers more effectively information to everyone at finger tips. The Web, no doubt, offers a new means for sharing and transmitting in-

formation unmatched by other media. The revolution started by the Web is just beginning. New business opportunities, such as e-commerce, e-banking, and e-publication, will increase with the maturity of the Web. It can hardly over-emphasize more impacts of the Web on the business and industrial world. The creation of a new sub-discipline devoted to Web related research and applications might have a significant value in the future.

The needs for WI may be further illustrated by the current fast growing research and industrial activities centered on it. We searched the Web by using the keyword “Web Intelligence” through several search engines in February 2001. The results are summarized in Table 1.

**Table 1.** A Statistics on WI

Search Engine	Number of hits
Lycos ( <a href="http://search.lycos.com/">http://search.lycos.com/</a> )	1,102,279
Google ( <a href="http://www.google.com/">http://www.google.com/</a> )	1,080,000
Excite ( <a href="http://www.excite.com">http://www.excite.com</a> )	223,825
AltaVista ( <a href="http://www.AltaVista.com/">http://www.AltaVista.com/</a> )	1,271
Netscape ( <a href="http://Netscape.com/">http://Netscape.com/</a> )	77
Yahoo ( <a href="http://www.yahoo.com/">http://www.yahoo.com/</a> )	74
LookSmart ( <a href="http://www.looksmart.com/">http://www.looksmart.com/</a> )	62

There are some interesting observations from the search results. The Web pages returned by most search engines contain both keywords “Web” and “Intelligence”, although they may not appear as a phrase in many pages. The co-occurrences of the two keywords show their strong association. This provides a piece of convincing empirical evidence supporting WI. The identification of this association may lead to the recognition of the importance of WI. We also used advanced search option of Google to search for the exact phrase “Web Intelligence”. We obtained 3,660 hits. We found that many companies concentrate on WI to provide intelligent solutions to business in the new Web-based information age. In fact, the majority of the top 40 pages returned by Google is industry related. For comparison, we search ResearchIndex (the NECI Scientific Literature Digital Library, <http://citeseer.nj.nec.com/cs>) containing an extreme large collection of scientific papers on-line. We found only one paper contains the phrase “Web Intelligence”. A further search of “Web” and “Intelligence” within two words results in 12 documents. They deal with topics such as Web browser intelligence, artificial intelligence for Web search, and Internet marketing intelligence through Web log mining. We also used “Web Intelligence” to query Ask Jeeves (<http://www.ask.com/>) and obtained related topics, such as intelligent Web systems, Web artificial intelligence, Web business intelligence, intelligent Web agents, intelligent Web robots, intelligent user interfaces, and Web user interfaces. Those topics clearly fit the proposed research areas of WI. From the