



# An Asian digital libraries perspective

Edward A. Fox<sup>a,\*</sup>, Elisabeth Logan<sup>b</sup>

<sup>a</sup> *Department of Computer Science, 660 McBryde Hall, Virginia Tech, Blacksburg, VA 24061, USA*

<sup>b</sup> *Division of Information Studies, Nanyang Technological University, 31 Nanyang Link, Singapore 637718, Singapore*

Available online 14 May 2004

---

## Abstract

This introduces a special issue of IP&M that covers some of the best works from the International Conference on Asian Digital Libraries, 2002. Eight papers, from eight countries, cover a broad range of topics and themes. Many present models, algorithms, systems, and/or experiments. There are approaches based on probability theory, linguistics, user-centered design, and multimedia technologies. It is hoped that readers will apply key results, and participate in future ICADL events.

© 2004 Elsevier Ltd. All rights reserved.

---

## 1. Introduction

In December 2002, the fifth in the series of International Conferences on Asian Digital Libraries was held in Singapore. The ICADL series is similar to related conference series in the Americas (i.e., the ACM/IEEE-CS Joint Conference on Digital Libraries—JCDL) and Europe (i.e., the European Conference on Digital Libraries—ECDL). Thus, ICADL broadly covers the many topics and themes that relate to “digital libraries” (DLs), and has attracted broad international attendance. Further, and of crucial importance here, many of the ICADL papers tell part of the story of important research studies.

Accordingly, after discussion with other key players on the program committee, we developed a plan to prepare a special issue of IP&M made up of extended papers selected from works in the ICADL’2002 proceedings (Lim, 2002). The eight papers selected (Bacchin, Ferro, & Melucci, 2004; Bainbridge, Dewsnip, & Witten, 2004; Feng, Jeusfeld, & Hoppenbrouwers, 2004; Renda & Straccia, 2004; Sasaki & Kiyoki, 2004; Theeramunkong & Wongtapan, 2004; Theng et al., 2004; Yeh, Ke, Yang, & Meng, 2004) are the result of this process. As can be seen from Table 1, these represent contributions from eight countries, including not only Asia but also the Americas and Europe. Table 2 shows these papers cover a broad set of topics central or closely related to “digital libraries”.

From Table 2, we note that each paper covers a particular important topic. Thus, find the central topic of each paper by looking down its column and finding the entry with “X”. Also, see the rows with bold labels to find the corresponding key topics.

---

\* Corresponding author.

E-mail address: [fox@vt.edu](mailto:fox@vt.edu) (E.A. Fox).

Table 1  
Papers (references) vs. countries

Reference	IT	JP	NL	NZ	SG	TH	TW	US
1. Bacchin et al. (2004)	x							
2. Bainbridge et al. (2004)				x				
3. Feng et al. (2004)			x					
4. Renda and Straccia (2004)	x							
5. Sasaki and Kiyoki (2004)		x						x
6. Theeramunkong and Wongtapan (2004)						x		
7. Theng et al. (2004)					x			
8. Yeh et al. (2004)							x	

Table 2  
Topics vs. Papers (references)

Topics	1	2	3	4	5	6	7	8
AI								x
Browsing			x					
Collaboration				x				
Education							x	
Experiments	x	x				x		x
GIS							x	
Handwriting						x		
HMM						x		
Images					x			
Info Retrieval	x	x			x			
<b>Knowledge</b>			X	x				x
Linguistics	x		x			x		X
LSA								x
Models	x		x	x		x		
<b>Music</b>		X						
<i>N</i> -grams		x				x		
<b>Patents</b>					X			
<b>Personalization</b>				X				
Probabilities	x					x		x
<b>Recognition</b>						X		
<b>Scenarios</b>			x				X	
Searching		x	x	x				
<b>Stemming</b>	X							
<b>Summarization</b>								X
Systems	IRON	GUIDO Greenstone		CYCLADES			GeogDL	
Text	x							x
Users			x	x			x	
Vectors	x			x		x		x

Further, from Table 2, note that there are various types of papers. Four report on experimental studies (Bacchin et al., 2004; Bainbridge et al., 2004; Theeramunkong & Wongtapan, 2004; Yeh et al., 2004). Four deal with models (Bacchin et al., 2004; Feng et al., 2004; Renda & Straccia, 2004; Theeramunkong & Wongtapan, 2004). Four describe implemented systems (Bacchin et al., 2004; Bainbridge et al., 2004; Renda & Straccia, 2004; Theng et al., 2004). Several employ probability theory (Bacchin et al., 2004; Theeramunkong & Wongtapan, 2004; Yeh et al., 2004) or knowledge-based approaches (Feng et al., 2004; Renda & Straccia, 2004; Yeh et al., 2004), while even more use vector representations (Bacchin et al., 2004;

Renda & Straccia, 2004; Theeramunkong & Wongtapan, 2004; Yeh et al., 2004). This sampling appears to be representative of current research on digital libraries.

## 2. Sequence and highlights

In order to aid the reader of this special section, we have ordered the papers (Bacchin et al., 2004; Bainbridge et al., 2004; Feng et al., 2004; Renda & Straccia, 2004; Sasaki & Kiyoki, 2004; Theeramunkong & Wongtapan, 2004; Theng et al., 2004; Yeh et al., 2004) in a way that we hope will be of greatest benefit. While readers are free to pick and choose, and to jump directly to works of special interest, it is worth noting that this particular sequencing (see below) might be useful.

The first paper in the special section (Renda & Straccia, 2004) illustrates how DLs can broadly serve user needs, by considering communities. Personalization may be achieved best by studying the common interests of collaborative communities. A helpful model, involving actors, data items, collections, and folders, can support recommendation and collaboration in addition to searching, and has been built into the CYC-LADES system.

The second paper (Theng et al., 2004) offers guidance on the design of DLs, by engaging representative users in the development of scenarios as well as claims analysis. Applied to support education, the GeogDL integrates textual and geographic information, with guidance of students aged 13–15.

The third paper (Bainbridge et al., 2004) surveys and demonstrates improvements to algorithms for searching in music DLs, and their integration in the GUIDO and Greenstone systems. Continuing this multimedia theme, the fourth paper (Sasaki & Kiyoki, 2004) surveys key approaches and systems for content-based image retrieval. Its perspective brings in the legal and intellectual property rights aspects of DLs, since it addresses which types of processes can/have been patented.

The fifth paper (Yeh et al., 2004) gives experimental results comparing several advanced approaches to text summarization, making use of the corpus characteristics that devolve from DLs. We can gain insight from the comparison of: (a) using linguistic and heuristic knowledge combined with training through genetic algorithms, versus (b) latent semantic analysis integrated with Salton et al.'s text relationship maps.

The sixth paper (Feng et al., 2004) broadens our view of DLs from searching and browsing, i.e., tactical cognition support functions, to strategic cognition support functions—supported by a model employing a knowledge subspace to supplement the usual document subspace.

In the seventh paper (Bacchin et al., 2004), on the other hand, probabilities based on local and global data allow language-independent stemming, improving upon expensive linguistic knowledge-based methods, as can be seen in the IRON system.

The final paper (Theeramunkong & Wongtapan, 2004) has particular import for Asia, where many languages are still in need of extensive study. Experimental results show that recognition of handwriting in Thai can be supported through both  $n$ -gram and hidden Markov model (HMM) methods, especially coupled with multi-directional island-based projection.

## 3. Conclusion

In this special issue, solid studies are presented that extend early papers appearing in the ICADL'2002 proceedings (Lim, 2002). These papers illustrate the broad range of topics and types of works that are presented at international DL conferences. It is hoped that the findings reported will further advance the field, and that interested readers will plan to attend upcoming conferences (e.g., ICADL 2004 in Shanghai).

## References

- Bacchin, M., Ferro, N., & Melucci, M. (2004). A probabilistic model for stemmer generation. *IP&M*, this issue.
- Bainbridge, D., Dewsnip, M., & Witten, I. H. (2004). Searching digital music libraries. *IP&M*, this issue.
- Feng, L., Jeusfeld, M. A., & Hoppenbrouwers, J. (2004). Beyond information searching and browsing: acquiring knowledge from digital libraries. *IP&M*, this issue.
- Lim, E.-P., Foo, S., Khoo, C. S. G., Chen, H., Fox, E. A., Urs, S. R., & Thanos, C. (2002). Digital Libraries: People, knowledge, and technology. In *Lecture Notes in Computer Science, Proceedings 5th international conference on Asian digital libraries, ICADL 2002, Singapore, December 11–14*. Springer, ISBN 3-540-00261-8.
- Renda, M. E., & Straccia, U. (2004). A personalized collaborative digital library environment: a model and an application. *IP&M*, this issue.
- Sasaki, H., & Kiyoki, Y. (2004). A formulation of the U.S. practice for patenting content-based retrieval processes in digital libraries. *IP&M*, this issue.
- Theeramunkong, T., & Wongtapan, C. (2004). Off-line isolated handwritten Thai OCR using island-based projection with  $n$ -gram model and hidden Markov models. *IP&M*, this issue.
- Theng, Y.-L., Goh, D. H.-L., Lim, E.-P., Liu, Z., Yin, M., Pang, N. L.-S., & Wong, P. B.-B. (2004). Applying scenario-based design and claims analysis to the design of a digital library of geography examination resources. *IP&M*, this issue.
- Yeh, J.-Y., Ke, H.-R., Yang, W.-P., & Meng, I.-H. (2004). Text summarization using a trainable summarizer and Latent Semantic Analysis. *IP&M*, this issue.