Abstract
The Extensible Markup Language (XML) is becoming the dominant standard for exchanging data over World Wide Web. Due to its flexibility, XML is rapidly emerging as the de facto standard for exchanging and querying documents on the Web to the next generation web applications including electronic commerce and intelligent web searching. XML documents can be stored and queried by using specialized semistructure repositories. While this is a viable approach to handle XML data, these specialized systems are yet premature since the semistructure data research itself is still young. Furthermore, it does not allow us to use the state-of-the-art relational database technologies available today that have been developed and tuned for last decades. Not only that, large volumes of data available today generally exist in relational database systems. Therefore, efficient storage and retrieval of native XML data and existing relational data seamlessly is becoming important. In fact, all major commercial relational vendors are working hard to support such capabilities and to build an infrastructure to enable them. However, they face the following three challenges: (1) how to represent XML data in relational model, (2) how to support XML query's processing over XML data stored in relational databases, and (3) how to publish existing relational data to XML format. We will discuss current state of the art technologies for these challenges and present the future research issues.

Surajit Chaudhuri is a senior researcher and manager of the Data Management, Exploration and Mining Group at Microsoft Research. He has worked extensively in the area of self-tuning database technology, query processing, data warehousing and data mining on SQL systems. He has published many papers in leading database conferences and journals. His work on self-tuning database technology and data mining has been incorporated in the Microsoft SQL Server product. Surajit has been a member of program committees of leading database and data mining conferences. In 1999, he was the co-chair of the ACM International Conference on Knowledge Discovery and Data Mining and co-chair of the industrial track of the ACM SIGMOD conference. Surajit did his Ph.D. from Stanford University and B.Tech from Indian Institute of Technology (Kharagpur, India). Prior to joining Microsoft Research, he was a member of the research staff at Hewlett-Packard Laboratories (Palo Alto) from 1992-1995.

Kyuseok Shim is an Assistant Professor at Seoul National University in Korea. Previously, he was an Assistant Professor at KAIST in Korea and a Member of Technical Staff at Bell Laboratories. He was one of the key contributors to the Serendip data mining project in Bell Laboratories. Before that, he worked for Quest Data Mining project at IBM Almaden Research Center and contributed to IBM Intelligent Miner for Data. He received B.S. degree in Electrical Engineering from Seoul National University in 1986, and the MS and Ph.D. degrees in Computer Science from University of Maryland, College Park in 1988 and 1993, respectively. Kyuseok has been working in the area of data mining and databases. He has published several research papers in prestigious conferences and journals. He has served as a program committee member on ACM SIGKDD, ACM SIGMOD, ICDE, PAKDD, and VLDB conferences. He is also currently on the Editorial Board of VLDB and KAIS journals.