

Of XML and Databases: Where's the Beef?

Michael J. Carey (*Moderator*)
IBM Almaden Research Center
650 Harry Road, K55/B1
San Jose, CA 95120-6099
carey@acm.org

ABSTRACT

This panel will examine the implications of the XML revolution, which is currently raging on the web, for database systems research and development.

Keywords

Databases, XML, semistructured data, world-wide web.

1. PANELISTS

Adam Bosworth (*Succendo*)

Bruce Lindsay (*IBM Almaden Research Center*)

Michael Stonebraker (*Informix Software*)

Dan Suciu (*AT&T Shannon Laboratory*)

Jennifer Widom (*Stanford University*)

2. OBJECTIVES

The web and XML standards are capturing the attention of database system researchers and vendors alike. To hear some tell it, it may soon be impossible to get a database research paper published if its title lacks the acronym XML, or to sell a database management system product that doesn't provide XML as its primary interface. Is all this XML interest based on hype, or is XML truly the next great wave? How will XML *really* impact the database field in the long run? Is it our community's ticket to finally impacting the web in a big way? Those are some of the questions that will hopefully be discussed in this panel session.

3. XML + DATABASES = ?

Like all panels, this one will go wherever the panelists decide to take it, but some possible questions include:

1. Does the XML revolution mean that the days of structured database systems are now numbered and that we'll soon be rid of all those pesky schemas?

Permission to make digital or hard copies of part or all of this work or personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers, or to redistribute to lists, requires prior specific permission and/or a fee.

MOD 2000, Dallas, TX USA

© ACM 2000 1-58113-218-2/00/05 ...\$5.00

2. Should we be building new database systems specifically tailored for storing and querying native XML data? If so, what should such a system look like, what are the challenges, and what sorts of applications might it be used for?
3. Since XML is a derivative of SGML, it has a document legacy. For example, it has no binary data type, which makes it less than useful in non-text environments. What should our community be recommending to W3C about ways to make XML more useful for "real data"?
4. What role will XML most likely play with respect to database-based web sites? For example, will e-businesses use XML to publish their inventories on the web, providing XML-based query access to them? Or will XML just act as the transport mechanism for remote stored procedure calls?
5. Will XML and XML query facilities supplant object-relational database systems and standards (SQL99) before they ever really have a chance to succeed? I.e., are object-relational databases already dead?
6. Given that IMS is still widely used for mission-critical data, it seems likely that SQL databases (if not SQL99 databases) will survive the XML revolution. What must be done to XML to ensure that XML data sources can interoperate with SQL data sources?
7. Will XML make it possible to someday query the web as if it were one great big database?

4. RECOMMENDED READING

- [1] T. Bray, J. Paoli, and C. Sperberg-McQueen (Editors), *Extensible Markup Language (XML) 1.0*, W3C Recommendation 10, <http://www.w3.org/TR/REC-xml>, February 1998.
- [2] D. Lomet (Editor), Special Issue on XML, *Data Engineering Bulletin*, Volume 22, No. 3, IEEE Computer Society, September 1999.
- [3] S. Abiteboul, P. Buneman, and D. Suciu, *Data on the Web: From Relations to Semistructured Data and XML*, Morgan Kaufman Publishers, San Francisco, CA, 1999.