XML Databases
Who cares?

Mary Fernández
AT&T Labs - Research
XML Data-Exchange Illusion

XML Documents & Schema

Producer

Consumer
XML Data-Exchange Illusion

Common XML Illusion

XML Documents & Schema

Producer

Consumer

WWW2002 - Hawaii
XML Data-Exchange Illusion

Common XML Illusion

XML Documents & Schema

Legacy Database

Producer

Consumer
XML Data-Exchange Illusion

Common XML Illusion

XML Documents & Schema

Legacy Database
Producer

Consumer

WWW2002 - Hawaii
XML Data-Exchange Illusion

Common XML Illusion

Publish XML

Legacy Database

Producer

XML Documents & Schema

Consumer

WWW2002 - Hawaii
XML Data-Exchange Illusion

XML Document & Schema

Legacy Database

Producer

XML

Consumer

Persistent Storage

Publish XML

Common XML Illusion

WWW2002 - Hawaii
XML Data-Exchange Illusion

Publish XML

Legacy Database Producer

XML Documents & Schema

XML Databases Panel

Common XML Illusion

Consumer

Persistent Storage

WWW2002 - Hawaii
XML Data-Exchange Illusion

Publish XML

Legacy Database
Producer

XML Documents & Schema

Consumer

Persistent Storage

Store XML

Common XML Illusion
XML Data-Exchange Illusion

XML Documents & Schema

Common XML Illusion

Query XML

Publish XML

Store XML

Legacy Database

Producer

Consumer

Persistent Storage
What we think XML is about

- **Leverage high-performance DBMS for XML**
  Query engine, concurrency control, transaction processing, access control, ...

- **Publishing XML**
  Convert legacy data into XML
  How to specify & evaluate mapping from non-XML data to XML

- **Storing XML**
  Map XML into underlying storage system
  How to “shred” XML data into non-XML data

- **Querying XML**
  Efficient evaluation of queries over virtual XML
  How to translate queries into native query language

- **Hard, but solvable problems**
XML Data-Exchange Illusion

Understand shared XML

Move XML

XML Documents & Schema

Producer

Consumer
What XML is *REALLY* about

- **Sharing information**
  - *Understanding & processing shared data*
    - Tools for data modelling
      - Requires expert understanding of application domain
    - Tools for syntactic (representation) & semantic mappings
      - Understanding semantics of schemas
    - Old problem with no standard solutions

- **Moving data**
  - *Efficient communication & processing*
    - On-line compression & decompression
    - Querying, filtering, routing of (compressed) XML
    - Optimize filters based on known schema of XML data