W3C XML Query WG Status

Paul Cotton

Program Manager, XML Standards

Microsoft Canada

Organization of Presentation

◆ XML query history and QL'98

◆ XML Query WG history, goals and status

◆ XML Query Requirements

◆ XML Query Data Model

◆ XML Query Algebra

◆ Questions

W3C Process

◆ All decisions made by Advisory Committee (one AC member per company) and the W3C Director

◆ Activities contain Working Groups

◆ Activities and WGs created via a Charter

◆ WGs first produce Requirements

◆ Technical work progresses from Working Draft, to Last Call WD, to Candidate Recommendation, to Proposed Recommendation and then Recommendation

XML Query History

- Early 1998: "roll your own query language"
- XSL Working Group
  - XSLT needed syntax to select nodes
- XML Linking Working Group
  - XPointer needed syntax to select a location
- February 1999 joint meeting
  - Rapprochement on 90% of syntax
- XPath
  - W3C Recommendation in Nov 1999 with XSLT
  - http://www.w3.org/TR/xpath

XML Query History - 2

◆ Early queries facilities for SGML

◆ Academic research into semi-structured data and its operations

◆ XQL: See http://metalab.unc.edu/xql

◆ XML-QL, August, 1998

   – http://www.w3.org/TR/NOTE-xml-ql/

Query Languages Workshop '98

- W3C sponsored workshop

- Boston (USA), December 2-3, 1998

- 98 participants: W3C members, database vendors, invited experts, etc.

- 66 position papers

- See: http://www.w3.org/TandS/QL/QL98

QL'98 Candidate Requirements

◆ **QL'98 workshop summary**

– Candidate Requirements for XML Query, Paul Cotton and Ashok Malhotra, IBM

– http://www.w3.org/TandS/QL/QL98/pp/queryreq.html

◆ **See also:**

– Database Desiderata for an XML Query Language, David Maier, Oregon Graduate Institute

– http://www.w3.org/TandS/QL/QL98/pp/maier.html

W3C XML Query WG - History

- July 1999 -Working Group proposed as part of XML Activity Phase 3 rechartering

- September 1999 - WG chartered

- More than 30 W3C member companies

- Nine F2F meetings and 45 telcons so far

- Close working relationship with other W3C Working Groups (Schema, XSL, I18N)

W3C XML Query WG - Goals

- ◆ "The goal of the XML Query WG is to produce a data model for XML documents, a set of query operators on that data model, and a query language based on these query operators."

W3C XML Query WG - Status

- ◆ Jan 2000 – Requirements Working Draft
- ◆ May 2000 – XML Query Data Model WD
- ◆ May 2000 – Feedback on Schema Last Call
- ◆ August 2000 – Revised Requirements Working Draft with Use Cases
- ◆ Dec 2000 – XML Query Algebra WD
- ◆ Future public WDs every three months
- ◆ Proposed recommendation(s)

XML Query Requirements

- Usage Scenarios
- General Requirements
- XML Query Data Model Requirements
- XML Query Functionality
- Use cases for XML queries
- http://www.w3.org/TR/xmlquery-req

Usage scenarios

- ◆ Human-readable documents
- ◆ XML processing of DBMS data
- ◆ Filtering streams
- ◆ Catalog searching
- ◆ Multiple syntactic environments

Human readable documents

- ◆ Documents structured using XML
- ◆ Collections of documents
  - – Technical manuals
- ◆ Query facilities
  - – Retrieve individual documents or fragments
  - – Generate new XML output such as TOC
  - – Text search operations including ranking

XML processing of DBMS data

- ◆ DBMS data
  - – SQL databases
  - – Object-oriented databases
  - – Native XML repositories
- ◆ XML can be physical or virtual
- ◆ Queries facilities
  - – Operations on text and non-text data types
  - – Aggregations

Filtering streams

- ◆ **Messages formatted using XML**
- ◆ **Streams with multiple sources/sinks**
- ◆ **Software-based message routing**
- ◆ **Query facilities**
  - – Message selection and filtering
  - – Message transformation
  - – Text and non-text data type operations

Catalog searching

◆ Queries on catalogs that describe document servers, document types, XML schemas, etc

◆ Queries on catalog characteristics

◆ Queries on documents described by selected catalogs

Multiple syntactic environments

- ◆ XML queries must be usable in many environments

- ◆ Embedded in URLs, JSP/ASP pages

- ◆ DOM Iterators or TreeWalkers

- ◆ Usable with IETF DAV Searching and Location (DASL) or Z39.50

General Requirements

◆ Non-procedural query language

◆ XML syntax for query language but also a readable syntax

◆ Protocol independent

◆ Standard error conditions

◆ Future support for updates

XML Query Data Model

- ◆ Built on XML Infoset and XML Schema Post-Schema Validation (PSV)

- ◆ Namespace aware

- ◆ Support for XML Schema primitive and complex data types

- ◆ Support for inter- and intra- document references

XML Query Functionality

◆ Operators on all data types

◆ Text operators across element boundaries

◆ Support for hierarchy and sequence

◆ Ability to combine data from different locs

◆ Aggregation and sorting

◆ Combination of operators including queries
as operands

XML Query Functionality - 2

- ◆ Support for NULL values
- ◆ Structural preservations
- ◆ Identity preservation
- ◆ Operations on names
- ◆ Operations on "schemas"
- ◆ Extensibility
- ◆ Closure

XML Query Data Model WD

- ◆ Defines information available to a query processor
- ◆ Infoset plus the following:
  - – Support for XML Schema data types (PSV)
  - – Support for document collections
  - – Support for references
- ◆ Node-labelled tree constructor model with node identity
- ◆ Mapping from Infoset to Query Data Model defined in Annex A
- ◆ http://www.w3.org/TR/query-datamodel/

XML Query Algebra WD

- ◆ Defines operations on Query Data Model
- ◆ Simple principles, easy to use
- ◆ Firm mathematical foundation
- ◆ Many issues still open
  - References
  - Unordered data
  - Algebra subset of syntax?
- ◆ http://www.w3.org/TR/query-algebra/

XML Query Issues

- Selection of query syntax – XML or readable

- Exact use of algebra

- Support for XML Schema very important

- Relationship to XSL and XPath still under discussion

- Text operators will need I18N approval

W3C XML documents

- **XML 1.0 Second Edition**
- **Namespaces in XML**
- **XSLT** and **XPath**
- **Infoset Working Draft**
- XML Schema Candidate Rec (Oct 2000)
  - Primer: http://www.w3.org/TR/xmlschema-1/
  - Structures: http://www.w3.org/TR/xmlschema-2/
  - Datatypes: http://www.w3.org/TR/xmlschema-3/

Questions

◆ Today

◆ Later:

　　　pcotton@microsoft.com

◆ Feedback email list:

　　　www-xml-query-comments@w3.org

◆ Public email list:

　　　www-ql@w3.org