



# XML Databases

Access to XML documents (II)

Benjamin Piwowarski



# Object Model

*“A description of the structure of the objects in a system including their identity, relationships to other objects, attributes, and operations”*

(Computer Science Glossary)



# Object Model II (+ pragmatic)

*“Formal description of a programmable object, which can consist of one or more classes, events, functions, interfaces, methods, namespaces, objects, and properties”*

(Microsoft)



# Object Model III (how-to)

- 1) Identify the objects
- 2) Identify attributes
- 3) Identify associations



# DOM

- Document Object Model
- Platform- and language-neutral interface (like SAX)
- Dynamic access to the content/structure of documents
- Representation of XML documents with an object-oriented model
- 3 (4) levels / Level 0 = pre-W3C epoch



# DOM level 1

- Core (XML) + HTML model
- Access to document
  - Composed of nodes (12 types): Document, Element, PI, Comment, DocType, Text, Attr, CDATA, EntityRef, DocumentFragment



# DOM level 2

- Core + HTML
- Filtered views
- Document Traversal & Range
- Events
- Refined OM
  - XML Namespace support
  - ex: `document.getElementById(id)`



# DOM level 3 (not finished)

- 6 components

- 1) Core
- 2) Load and Save
- 3) Validation
- 4) XPath
- 5) Views and Formatting
- 6) Requirements (Bare minimum)





# XML Databases

XML technologies, XML applications



# XML related technologies

- XML applications
  - X(HT)ML + CSS
  - SVG
  - MathML
- XML Technology
  - XLink
  - XInclude



# X(HT)ML

## ■ Problems with HTML

- ☐ Used before it was defined
- ☐ Not extensible
- ☐ SGML (no namespaces, complexity)
- ☐ Both presentation and structuring tags

# XHTML: The idea

- XHTML is an XML application (-> allow to mix applications, eg MathML, CML, Xlink, etc.)
- Separate (as of XHTML 1.1)
  - The structure + content (XHTML nodes + text)
  - The presentation (CSS)
- XHTML 1.1: modularity (core, applet, text extension, forms, table, image, ... = 21 modules in total)