W3C Standards

Vlad Coroama
Marc Langheinrich

Seminar
Ubiquitous Information
WS 00/01
ETH Zurich
What´s coming up?

- All about the W3C
  - History, Members, Process, Domains
- XML Technology Primer
  - XML, XML Schemas, XML Linking, XML Style
- The Semantic Web
  - Metadata: RDF & Applications (CC/PP), XML Protocols & SOAP, The Web of Trust
- Summing Up
W3C History

- Founded 10/1994 by Tim Berners-Lee at MIT/LNCS
  - 4/95 INRIA hosts W3C Europe
  - 4/96 Keio University hosts W3C Asia

- Led by Jean-Francois Abramatic (Chairman) and Tim Berners-Lee (Director)

- Goals:
  - Universal Access
  - Semantic Web
  - Web of Trust

- More than
  - 20 specs in 5 years
  - 50 staff members across hosts institutions
W3C Members

- 497 members as of 01/2001
  - Companies: AOL, Appel, AT&T, Cisco, Citibank ...
  - Universities: MIT, Vrije Universiteit ...
  - Governmental Agencies: US EPA, Datenschutz-Zentrum Schleswig Holstein ...

- Yearly Membership Fee
  - US$ 50,000.- corporate members
  - US$ 5,000.- if annual revenues < US$ 50’000’000
  - US$ 5,000.- non-profit, governmental
I. All about the W3C

### Advisory Committee

- One representative from each member
  - send submission requests from their organization
  - nominate colleagues for W3C working groups
- Reviews proposals for activities, recommendations
**Submissions**

- Allows members to propose technology/ideas to W3C for consideration
- Must include IPR statement
- Reviewed by W3C Team
  - If accepted, published as **W3C Note**
W3C Recommendations

I. All about the W3C
I. All about the W3C

Working Draft (WD)
- members only
- updated often

Public Working Draft
- interim draft
- for public comment
- required: one every 3 months

Last Call Draft

Recommendation (Rec)

Proposed Recommendation (PR)

Candidate Recommendation (CR)
I. All about the W3C

W3C Recommendations

- Working Draft (WD)
  - members only
  - updated often

- Public Working Draft
  - interim draft
  - for public comment
  - required: one every 3 months

- Last Call Draft
  - WG: "We're done"
  - for public & W3C comment
  - lasts 3-4 weeks

- Recommendation (Rec)

- Proposed Recommendation (PR)

- Candidate Recommendation (CR)

Back to WD?
I. All about the W3C

W3C Recommendations

1. Working Draft (WD)
   - members only
   - updated often

2. Public Working Draft
   - interim draft
   - for public comment
   - required: one every 3 months

3. Last Call Draft
   - WG: "We’re done"
   - for public & W3C comment
   - lasts 3-4 weeks

4. Candidate Recommendation (CR)
   - WG requests implementations
   - 0-12 months (may be skipped)

5. Proposed Recommendation (PR)

6. Recommendation (Rec)

Director approves

Back to WD?

dependencies ok?

Back to WD?
I. All about the W3C

W3C Recommendations

Working Draft (WD)
- members only
- updated often

Public Working Draft
- interim draft
- for public comment
- required: one every 3 months

Last Call Draft
- WG: “We’re done”
- for public & W3C comment
- lasts 3-4 weeks

Proposed Recommendation (PR)
- for review by AC
- at least 4 weeks
- public promotion by AC reps

Candidate Recommendation (CR)
- WG requests implementations
- 0-12 months (may be skipped)

Recommendation (Rec)

Back to WD?

Director approves

dependencies ok?

Back to WD?
W3C Recommendations

I. All about the W3C

- Working Draft (WD)
  - members only
  - updated often

- Public Working Draft
  - interim draft
  - for public comment
  - required: one every 3 month

- Advisory Committee (AC)
  - approves

- Proposed Recommendation (PR)
  - for review by AC
  - at least 4 weeks
  - public promotion by AC reps

- Candidate Recommendation (CR)
  - WG requests implementations
  - 0-12 months (may be skipped)

- Recommendation (Rec)
  - W3C considers refined & ready for widespread deployment
  - only minor errata

- Director
  - approves

- Last Call Draft
  - WG: „We’re done”
  - for public & W3C comment
  - lasts 3-4 weeks

- Back to WD?

- W3C
  - considers refined & ready for widespread deployment
  - only minor errata

- Last Call
  - dependencies ok?

- Approves

- Back to WD?
I. All about the W3C

## W3C Domains

- **Architecture**
  - HTTP, DOM, URI, XML, XML Protocols
- **Technology & Society Domain**
  - XML Encryption, XML Signature, Privacy (P3P), Metadata, Electronic Commerce
- **User Interface**
  - HTML, Graphics, I18N, Math, Mobile, Multimedia, Style, TV/Web, VoiceBrowser
- **Web Accessibility Initiative (WAI)**
I. All about the W3C
What’s coming up?

- All about the W3C
  - History, Members, Process, Domains
- **XML Technology Primer**
  - XML, XML Schemas, XML Linking, XML Style
- The Semantic Web
  - Metadata: RDF & Applications (CC/PP), XML Protocols & SOAP, The Web of Trust
- Summing Up
The X-Files

- XML Core
  - DTD, Namespaces
- XML Linking
  - XLink, XBase, XPointer, XPath
- XML Style
  - XSL, XSLT
- XML Schema

More XML
- XML Protocol
- XML Signature
- XML Query
- XML Encryption
- XHTML
Other stuff

- Semantic Web
  - RDF
  - P3P
  - CC/PP
  - SOAP

- SyncML
- SVG / SMIL
- VoiceBrowser
XML Activity

- XML Query Working Group
- XML Schema Working Group
  - XML Schema: Primer, Structures, Datatypes
- XML Linking Working Group
  - XPointer, XLink, XML Base
- XML Core Working Group
  - Advances XML specification. XML Fragment, XInclude, XML Information Set
- Old: XML Namespaces, XML Style Sheets (XSL)
- Others: XML Protocols, XML Encryption, XML Signature

II. XML Technology Primer
XML

- markup language for documents containing structured information
- **XML is not HTML**
  - specifies neither semantics nor tag set
  - meta-language for describing markup languages
- restricted form of SGML
  - thereby usable on the Web
XML structure

- Content
- Markup
  - elements
    - attributes
  - entity references
  - comments
  - processing instructions
  - marked sections
  - document type definitions (DTDs)
XML (cont.)

- **XML Documents** can be
  - Well-formed
    - comply simplest syntactic rules
  - Valid
    - obeys the constraints of a DTD (Document Type Definition)

- **DTD**
  - context-free grammar
  - defines tag set for a specific markup vocabulary
<?xml version="1.0"?>
<oldjoke>
  <burns>Say <quote>goodnight</quote>, Gracie.</burns> <allen><quote>Goodnight, Gracie.</quote></allen>
  <applause/>
</oldjoke>

<!ELEMENT oldjoke (burns+, allen, applause?)>
<!ELEMENT burns (#PCDATA | quote)*)
<!ELEMENT allen (#PCDATA | quote)*)
<!ELEMENT quote (#PCDATA)*>
<!ELEMENT applause EMPTY>
XML Namespaces

- Name collisions for XML markup
  - single XML document contains elements and attributes defined in different vocabularies
  ⇒ every tag needs to be unique
- XML Namespace = collection of names, uniquely identified by a URI reference
  - used in XML documents as element types and attribute names

```xml
<?xml version="1.0"?>
<x xmlns:edi='http://ecommerce.org/schema'>
  <!-- the "edi" prefix is bound to http://ecommerce.org/schema for the "x" element and contents -->
</x>
```
XSL

- language for expressing stylesheets, consisting of
  1. language for transforming XML documents (XSLT), and
  2. an XML vocabulary for specifying formatting semantics

II. XML Technology Primer

![Diagram of XSL Two Processes: Transformation & Formatting]

XSL Transform
Result Tree (element and attribute nodes)

Result XML tree is the result of XSLT processing.
XSLT

- language for transforming XML documents into other XML documents
- Part of XSL (which is XSLT + XML vocabulary for formatting)
- A transformation expressed in XSLT describes rules for transforming a source tree into a result tree
- The transformation is achieved by
  - associating patterns with templates.
  - Matching patterns against elements in the source tree.
  - Instantiating a template to create part of the result tree.
  - Separating the result tree from the source tree
II. XML Technology Primer

XML Tech Tree

- XML
- SGML
- HTML
- SOAP
- XHTML
- DTD
- PICS
- XForms
- XMLQuery
- P3P 1.0
- XLink
- XPath
- XPointer
- RDF
- RDF Schema
- XML Signature
- Canonical XML
- XSL/T
- XML Schema
- P3P 2.0
- CC/PP
- XHTML
- SOAP
- XML
- Web heute
- Web morgen
XLink

- XML application
- Defines additional attributes, rules, for linking two or more XML resources
- **HTML Link:** `<a href="foo.xml">bar</a>`
- **XLink:** `<my:baz xlink:type="locator" xlink:label="buzz" xlink:href="foo.xml">bar</my:baz> <my:gogo xlink:type="arc" xlink:to="buzz"/>
- Use XPointer or XPath for fine-grained linking!
XPath

- Common syntax and semantics for XSLT and XPointer
- Addresses parts of XML document
- Provides basic facilities to manipulate strings, numbers and booleans
- Can also be used for matching parts of XML document (used XSLT)
XPointer

- language to be used as the basis for a fragment identifier
- Based on XPath
- Adds arbitrary referencing inside XML documents (e.g. for continuous selection with a mouse, which crosses element boundaries)
- DOES NOT USE XML itself, but rather URI structure (since it will be embedded in e.g. href attributes)

```xml
<button xlink:type="simple" xlink:href="#xpointer(here()/ancestor::slide[1]/preceding::slide[1])"> Previous </button>
```
XML Schema

- XML language for describing and constraining the content of XML documents
- Def. „Schema“: enumeration, structure and definition of terms used to make (metadata) assertions.
- XML Spec defines valid, well-formed XML Syntax. XML Schema adds (few) semantics.
DTD drawbacks

- **DTD**
  - does not support data types beyond character data
    
    `<year>Hello world!</year>`
  - no support for Namespaces
  - is (only) CF-grammar
  - (DTD != XML) => XML technologies (DOM, SAX) cannot parse and expose DTD
XML Schema improvements

- New built-in data types, based on SQL and Java data types
  - also User-defined data types
- Explicit support for namespaces
- Schema constructs can be imported from existing schemas
- Elements can inherit content and attributes of other elements through refinement
What’s coming up?

- All about the W3C
  - History, Members, Process, Domains
- XML Technology Primer
  - XML, XML Schemas, XML Linking, XML Style
- The Semantic Web
  - Metadata: RDF & Applications (CC/PP), XML Protocols & SOAP, The Web of Trust
- Summing Up
The Semantic Web

- **Tim Berners-Lee’s Vision**
  - “... the Semantic Web approach develops languages for expressing information in a machine processable form.”

- **Tools**
  - Universal Addressing Scheme (URIs)
  - Universal Data Format (XML)
  - Ontologies (RDF, RDF Schemas)
  - Object Protocols, Communication (SOAP)
Why RDF?

- „Resource Description Framework“
  - Describing Resources („Things“)
  - Metadata - Data about Data
- Metadata helps us finding things
  - Yellow Pages
  - Library Card Catalog
- RDF is Metadata for the Web
  - Structure instead of Brute-force Text Indexing or Manual Directories
Why not use XML?

- There’s more than one way to do it (in XML):
  - `<car color="red" />`
  - `<car><color>red</color></car>`
  - `<car color="#cc" />`<br>      `<color id="cc" shade="red" />`

- The RDF way of things
  - *Resource*: car
  - *Property*: color
  - *Statement*: red

- RDF: `<Subject> has <Predicate><Object>"}
RDF Core Concepts

III. The Semantic Web

Resource
- referenced by URI
- grouping possible
  - Bags (unordered)
  - Sequences
  - Alternatives

Property
- Uses XML namespaces
- Implicitly referenced by URI

Statement
- Application of Property with Value
- Value can be
  - Strings, or
  - Other Resources
RDF/XML

- RDF = Structured graphs (ER-Model)
- RDF/XML = serialized RDF
  - Other serializations possible, e.g. SOAP
- RDF/XML
  - doesn’t look different than XML - it is XML,
  - but with particular data model and
  - predefined set of element types
The students in course 6.001 are Amy, Tim, John, Mary, and Sue

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:s="http://mycollege.edu/schema/students">
  <rdf:Description about="http://mycollege.edu/courses/6.001">
    <s:students>
      <rdf:Bag>
        <rdf:li resource="http://mycollege.edu/students/Amy"/>
        <rdf:li resource="http://mycollege.edu/students/Tim"/>
        <rdf:li resource="http://mycollege.edu/students/John"/>
      </rdf:Bag>
    </s:students>
  </rdf:Description>
</rdf:RDF>
```
RDF/XML Syntax

- `<rdf:RDF>`
- `<rdf:Description>`
  - ID=string
  - about=#URI
  - aboutEach)#URI
  - aboutEachPrefix=string
- `<x:property> ... </x:property>`
- `<x:property resource=#URI />`
- `<rdf:Bag>`
- `<rdf:Seq>`
- `<rdf:Alt>`
- `<rdf:li>`
RDF Schema

- Schema definition language
- Basic type system

Written in RDF
Example: Schema for RDF Schema

III. The Semantic Web
RDF Schema Example

III. The Semantic Web

```xml
<rdf:Description ID="MotorVehicle">
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="http://www.w3.org/2000/01/rdf-schema#Resource"/>
</rdf:Description>

<rdf:Description ID="PassengerVehicle">
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdf:Description>

<rdf:Description ID="Van">
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
</rdf:Description>

<rdf:Description ID="MiniVan">
  <rdf:type resource="http://www.w3.org/2000/01/rdf-schema#Class"/>
  <rdfs:subClassOf rdf:resource="#MotorVehicle"/>
  <rdfs:subClassOf rdf:resource="#Van"/>
  <rdfs:subClassOf rdf:resource="#PassengerVehicle"/>
</rdf:Description>
```
CC/PP
- Composite Capabilities/Preferences Profile
- part of W3C Mobile Access Activity
- „user preferences and device capabilities“

Device Profiles
a device profile lists the (display) abilities of a particular device

Document Profiles
a document exists in different variants, each including a document profile, describing the browser support it needs to display it

Negotiate Correct Content for Device
CC/PP

- **Idea**
  - Device sends Pointer to Device Profile along with request
  - Server replies with best matching document
  - Device Profiles written in RDF
  - UAProf Specification (WAP Forum) defines Client Capabilities for Mobile Phones

- **Example:**

```xml
<ccpp:component>
  <Description about="http://www.example.com/TerminalHardware">
    <type resource="http://www.example.com/Schema#HardwarePlatform"/>
    <uaprof:ScreenSize>640x400</uaprof:ScreenSize>
  </Description>
</ccpp:component>
```
Needed: Infrastructure

- Vocabularies
  - CC/PP, P3P, PICS
- Query Language
  - RDF Query efforts
- Data Storage
  - Rdfdb, Redland
- Characterization
  - How much do I know?
XML Protocols Activity

- XML Protocol Activity
  - Since May 2000

- Deliverables
  - An envelope to encapsulate XML data for transfer
  - an operating system-neutral convention for the content of the envelope when used for RPC
  - A mechanism to serialize data based on XML Schema datatypes
  - a non-exclusive mechanism layered on HTTP transport

- Starting Point: SOAP/1.1
SOAP 1.1

- Microsoft, IBM, et al.
  - "mechanism for exchanging structured and typed information between peers in a distributed environment using XML"
  - http://msdn.microsoft.com/soap/
  - part of Microsoft’s .NET framework
POST /StockQuote HTTP/1.1
Host: www.stockquoteserver.com
Content-Type: text/xml; charset="utf-8"
Content-Length: nnnn
SOAPAction: "Some-URI"

<SOAP-ENV:Envelope
 SOAP-ENV:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/">
 <SOAP-ENV:Body>
   <m:GetLastTradePrice xmlns:m="Some-URI">
     <symbol>DIS</symbol>
   </m:GetLastTradePrice>
 </SOAP-ENV:Body>
</SOAP-ENV:Envelope>

SOAP 1.1 Message Embedded in HTTP Request

- SOAP Envelope Framework
  - what is in a message
  - who should deal with it
  - whether it is optional or mandatory
  - error handling („faults“)
SOAP Serialization

- Defines XML Schema for
  - Simple Types
    - Strings, Integer, Floats
    - Enumeration
    - Byte Arrays
  - Compound Types
    - Structs
    - Arrays

SOAP 1.1 Array Type Example

```xml
<SOAP-ENC:Array SOAP-ENC:arrayType="xsd:string[,][4]">
    <SOAP-ENC:Array SOAP-ENC:position="[2]" SOAP-ENC:arrayType="xsd:string[10,10]">
        <item SOAP-ENC:position="[2,2]">Third row, third col</item>
        <item SOAP-ENC:position="[7,2]">Eighth row, third col</item>
    </SOAP-ENC:Array>
</SOAP-ENC:Array>
```
More non-W3C Stuff

- ebXML (e-business XML, „SOAP+“)
  - www.ebxml.org
- UDDI (Description & Discovery)
  - www.uddi.org
- WSDL (Web Services Descr. Language)
  - http://msdn.microsoft.com/xml/general/wSDL.asp
  - integrates with SOAP
- CORBA/SOAP
  - OMG request for proposals

III. The Semantic Web
No more time for:

- **XML Signature**
  - Canonical XML, Requirements (since 06/1999)
- **XML Encryption**
  - Goal: Encrypting Parts of XML (since 01/2001)
- **XML Query**
  - Requirements, Algebra & Data Model done
- **XForms**
  - replaces HTML/XHTML forms
- **XHTML**
  - XHTML 1.1 done
  - now working on modularizing, XHTML 2.0
What’s coming up?

- All about the W3C
  - History, Members, Process, Domains
- XML Technology Primer
  - XML, XML Schemas, XML Linking, XML Style
- The Semantic Web
  - Metadata: RDF & Applications (CC/PP), XML Protocols & SOAP, The Web of Trust
- Summing Up
Take Home Message

- **W3C**
  - almost 500 members, more than 1000 participants in Working Groups

- **The Semantic Web**
  - XML as universal exchange language
  - RDF as (weak) semantics
  - SOAP et al as lightweight CORBA
  - Description and Discovery Standards emerging

- **We’re just getting started!!**
Building the Web of Trust

IV. Summing Up

- RDF Schema
- RDF syntax in XML
- Resource Description Framework: basic ER-like model
- Namespaces
- XML-Schema
- XML - Structured documents
- Universal Resource Identifiers
- Unicode
- Web of Trust
- Proof
- KR rules
- Logic
- Ontology support
- KR data
- Dublin Core element set
- Moz
- P3P

© 9/2000 Tim Berners-Lee