

Information Standards

Concepts and Issues

Olle Olsson

Swedish W3C Office

Swedish Institute of Computer Science (SICS)

SU "Law and Information Technology" January 2012

© 2011 W3C



SICS – Swedish Institute of Computer Science

National research institute

- R&D in information and communication technology (ICT)

Objective:

- Conduct advanced and focused research in strategically important ICT areas



Sponsors:

TeliaSonera, Ericsson,
Saab Systems,
FMV (Defence Materiel Administration),
Green Cargo,
ABB,
Bombardier Transportation

Motivating example

Standards & compliance - example

The screenshot shows the homepage of The British Museum website. At the top left is the museum's logo. A search bar is located at the top right. Below the logo is a navigation menu with links: Home, Visiting, What's on, Explore, Research, Learning, The Museum, Join in, and Shop online. The main content area features a large image of a bronze statuette of Venus or Dione on the left. To its right are several promotional cards: 'Explore the British Museum' with a carousel of objects, 'What's on' featuring 'HADRIAN EMPIRE AND CONFLICT' (until 26 October 2008), 'STATUE PHILIA' (from 4 October 2008), and 'BABYLON MYTH AND REALITY' (from 13 November 2008). Below the statuette image is the caption 'Bronze statuette of Venus or Dione'. At the bottom right of the main content area is a yellow box containing the URL <http://www.britishmuseum.org/>.

Validating web page

Result:	8 Errors, 1 warning(s)
Address:	<input type="text" value="http://www.britishmuseum.org/"/>
Encoding:	utf-8 <input type="button" value="(detect automatically)"/>
Doctype:	XHTML 1.0 Strict <input type="button" value="(detect automatically)"/>
Root Element:	html
Root Namespace:	http://www.w3.org/1999/xhtml

- ✘ Error Line 74, Column 112: document type does not allow element "input" here; missing one of "p", "h1", "h2", "h3", "h4", "h5", "h6", "div", "pre", "address", "fieldset", "ins", "del" start-tag
- ✘ Error Line 464, Column 6: end tag for "div" omitted, but OMITTAG NO was specified.
- ✘ Error Line 467, Column 7: end tag for element "div" which is not open.
- ✘ Error Line 470, Column 7: end tag for element "form" which is not open.
- ✘ Error Line 464, Column 7: XML Parsing Error: Opening and ending tag mismatch: div line 76 and form.

.....

"Validation"

- Linguistic form
 - "X complies to Y"
 - "X conforms to Y"
- Y is a norm
 - Y is a standard
 - What is a "standard"?
- X is some object/entity/phenomenon...
 - X is an instance
- The example:
 - The web page
 - <http://www.britishmuseum.org/>
 - Does not conform to
 - XHTML 1.0 Strict

... but the page itself says that it is constructed according to this standard!

Technology standards

- Hardware
 - Example: USB (uses: memory sticks, mouse, camera, ...)
- Software
 - Example: JavaScript (uses: scripts in web browsers, ...)
- Data
 - Example: MP3 (uses: audio recording and playing, ...)

Contents

- Background
- Technologies, standards, standardisation
- Open standard
- Web standards and standardisation
- (The value of standards)
- Drawbacks/problems with standards?

- Information standards – the XML approach
- Language design – challenges
- XML standards – areas
- Bibliography

Background

ICT – what makes standards important now?

- Information and communication technology (ICT)
 - From computer centre
 - ... to desktop compute power
 - ... to hand-held
 - ... to networked society
- Trends
 - Performance evolution
 - Cost evolution
 - Accessible to non-specialists
 - The importance of information
 - Cross-sectoral
 - Globalisation



From closed to open

- Earlier:
 - Silos – (hidden problem ?) ... in those days
 - One complete supplier
 - Lock-in
 - Limited competition
- Now:
 - No fixed borders (no silos)
 - Co-operation with others
 - Many dimensions of functionality needed
 - Suppliers specialize
 - Increased lifetime and reuse
- Standards – a critical precondition
 - future safe!

Standards and societal evolution

Importance for national economy

- Enable competition
 - Push price/performance evolution
- Open up new innovation areas
 - Standards as platform
- Enlarge markets
 - Effects on volume

Standards as reusable added value

- Extend reuse of investment

As to governments:

- Establish policies for use of standards

Technologies, standards, standardisation

Standards – what?

IT standards:

- Accessible documented specifications

Types of standards:

- *De jure*: published by an officially recognised standardisation organisation – ISO, ANSI, ETSI, ...
- Consortium standards: produced within organised collaboration between a number of actors, recommended for wide use – W3C, OASIS, ...
- *De facto*: significantly broad and long-term acceptance of technology on the market – Windows XP, Linux, QWERTY keyboard, ...

Aspects

- What is standardised?
 - Scope, focus, granularity, ...
- How categorical is the standards?
 - Undefined parts; “MUST, SHOULD, MAY, ...”
- Who is responsible for the standard?
 - development, maintenance, ...
- Who is the standard targeting?
 - Suppliers, users, policy makers, ...
- What validity constraints for the standard?
 - Time and space, legal status
- What does the standard assume?
 - Other standards, policy frameworks,

Standards – about what?

- Main categories
 - Product
 - Product features, performance, compatibility, ...
 - Process
 - Requirements to be met by a process
 - Management
 - Typical “Quality management”: controlling aspects of process, organisation, procedures, resources.
- We focus on:
 - Product
 - Information Technology
 - Information/data
 - Representation formats

Actors: Standards Setting Organizations (SSO)



Successive standardisation

- Co-operation between standardisation actors
 - International => national
- MS Office Open XML Document Format
 - Microsoft => ECMA => ISO
- OOo OpenDocument Format
 - OOo => OASIS => ISO

Example: OOXML

- [Microsoft] Open Office XML (OOXML)
 - Specification of formats for MS Office applications
 - 2001 – 2002: MS implements XML-format in Office
 - 2004-05-24: EU asks MS to standardize Office formats
 - 2005-11-dd: submitted to ECMA
 - 2006-12-07: accepted as standard ECMA-376
 - 2006-12-20: submitted to ISO (fast-track)
 - Spec: 6000 pages.
 - 2007-09-04: not accepted . To be revised
 - 3522 review comments.
 - 2008-04-02: accepted as Draft standard ISO/IEC DIS 29500
 - 2008-11: published as standard ISO/IEC DIS 29500

Example: ODF

- [OpenOffice] Open Document Format (ODF)
 - Specification of formats for OpenOffice applications
 - 2000: Sun "open sources" Star Office => OpenOffice
 - 2002: OpenOffice 1.0 with XML format
 - 2002-11-dd: Sun submits "OO XML" to OASIS
 - 2005-05-01: "ODF" accepted as OASIS standard
 - 2005-11-16: submitted to ISO
 - Spec: 720 pages
 - 2006-05-03: accepted as Draft ISO/IEC standard
 - 2006-11-26: accepted as standard ISO/IEC 26300:2006

ODF & OOXML ... time line

ODF

=> OASIS

OASIS standard! => ISO

ISO standard!

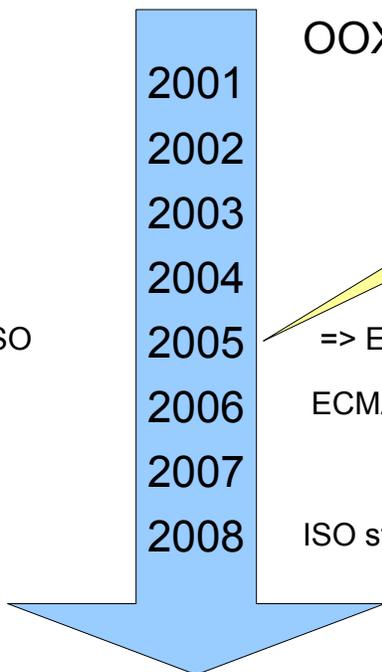
OOXML

=> ECMA

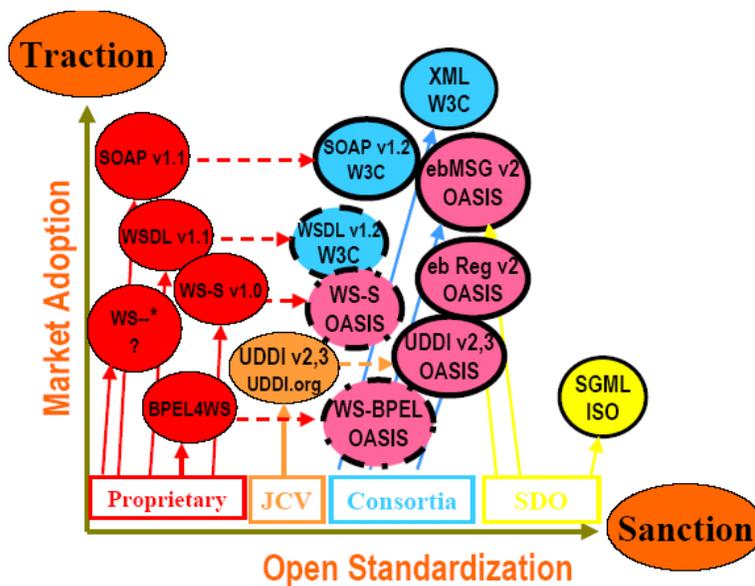
ECMA standard! => ISO

ISO standard!

State of Massachusetts Policy: use open standards



Life cycle of standards



Standards

Standard – implementation, conformance

What *is* a "standard"?

Standard = specification of

- Entities, with properties, attributes, relationships, behaviors
- constraints on props/attrs/rels/behavs

that can be fulfilled by some artefact.

The conceptual model of a standard:

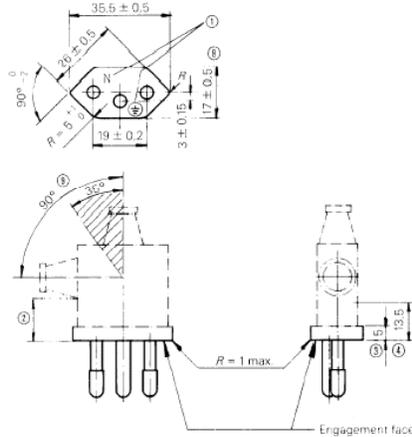
- structure of concepts

Artefact implements the standard?

- An *implementation* of the standard?
- Artefacts *conforms* to the standard?
- Artefact is *compliant* with the standard?

Example: Power plug – IEC 60906-1

The specification



An implementation



Standard conformance

Can conformance be *evaluated*?

(1) Is the specification:

- consistent?
- complete?
- unambiguous?

Cf.

- Fuzzy conditions, like "SHOULD ...", "MAY ...", etc

(2) Can one measure/evaluate:

- properties, attributes, relationships, behaviors
for a candidate implementation?

What does a standard mean?

Does it mean what I think it means?

Conceptual model of a standard:

- *Internal* concepts – artificial concepts
 - “before”/”after” for items in set, when implemented as list
- *External* concepts – representations/analogues of concepts defined elsewhere
 - “secure transmission”, “contract”, “identifier”, “transaction”

Will the standard do for me what I hoped for?

Ref:

Lundblad, N (2005) "Legal Analysis of XML-based Information Standards" in Magnusson Sjöberg, C (ed) *Legal Management of Information Systems: Incorporating Law in E-solutions* (Lund 2005)

Standards use – terms & conditions?

Using a standard – any “fine print” that I should take note of?

Standard is a specification

- For all practical purposes, a “paper” document

Remember to investigate:

- Getting access to the specification
 - Cost?
 - Legal conditions enforced?
- Implementing:
 - Licensing fee?
 - “Embedded” patents?

To think about: similarities and differences when using standards, patents, copyright as weapons in markets

Ref:

Lundblad, N op cit.

Controlling a standard

Standards as tool for competition control?

Sometimes used by sector oligopoly to reduce/eliminate competition.

May have negative effects on innovation.

But not all sorts of standards!

Open standards as an enabler.

- In contrast to closed / guarded / hidden standards.

Open standard

Open standard - statements

Why “open standard”?

Erkki Liikanen (EU Commissioner):

"Open standards are important to help create interoperable and affordable solutions for everybody. They also promote competition by setting up a technical playing field that is level to all market players. This means lower costs for enterprises and, ultimately, the consumer."

Jorma Ollila (Nokia):

"... Open standards and platforms create a foundation for success. They enable interoperability of technologies and encourage innovativeness and healthy competition, which in turn increases consumer choice and opens entirely new markets,"

Tim Berners-Lee (W3C):

"The decision to make the Web an open system was necessary for it to be universal. You can't propose that something be a universal space and at the same time keep control of it."

Open standard - definition

Definition of “open standard”?

Discussions ongoing in IGF, EC, etc.

- Open *process* ... can mean
 - Transparent process
 - Open participation
 - Technical consensus
 - etc.
- Open *results* ... can mean
 - Free and persistent specification
 - Liberal patent policy
 - Executable code
 - etc.

Standards and patents

Examples from W3C:

- P3P (Platform for Privacy Preferences)
 - Intermind participated in standardisation work.
 - Announced that they had a critical patent...
 - Other participants hesitated w.r.t. work on P3P
 - Future fees for usage?
 - Investigation started: The Intermind patent not critical
 - Result: P3P work continues
- CSS (Cascading Style Sheet)
 - Microsoft partner in work. Announced they had critical patent
 - Microsoft decided to offer patent as Royalty-Free license
- Xlink (XML Linking Language)
 - Sun had patent ... decided to offer as Royalty-Free license

Open Source, standards, patents

- Ideology underlying Open Source is in conflict with patents
- Example: W3C patent policy
 - Proposal 2001 – equally acceptable: "Reasonable And Non-Discriminatory" and "Royalty Free"
 - Open Source community protested strongly. Risks:
 - Stop using W3C standards
 - Develop alternative free standards ("balkanisation" of the web)
 - The web is taken over by commercial interests
 - Engage members of Open Source community in work
 - More attention put to requirements/needs in Open Source world

Open standards and protection

- Open Source ... objective
 - Encourage reuse and adaptation of computer software
- Open Standards ... objective
 - *Discourage* some reuse and adaptation

Open Software	Open Standards
Innovation, novelty, alternatives	Uniformity, interoperability, conformance
Darwinian unregulated evolution	Controlled and managed change
Unconstrained opportunities	No deviations

- But mutual benefits: Open Standards \Leftrightarrow Open Source

Open standards and protection

Standards

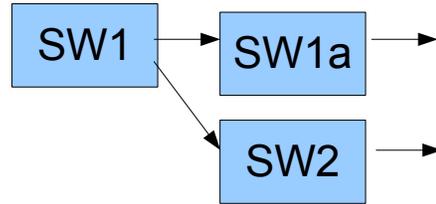
- Specifications of *what* implementations should do
- Copyright
- Derivative work
- License:
 - e.g., "Derivative works may not be created"
- Infringement – similarity
- Software is derivative?

Software

- Description of *how* the implementation does it
- Patent
- Derivative work
- License
- Infringement – similarity

Open standards and protection

- Standards should be exact, unique, identifiable, stable,
- Open Software should be improved
 - Forking ... good thing
- (Open) standards
 - Forking ... bad thing
- Fuzzy boundary between specification and implementation



Open standards and protection - Example

DOM – IDL

(Document Object Model - Interface Definition Language)

```
// File: dom.idl
#ifndef _DOM_IDL_
#define _DOM_IDL_
#pragma prefix "w3c.org"
module dom
{
    valuetype DOMString sequence<unsigned short>;
    typedef unsigned long long DOMTimeStamp;
    interface DocumentType;
    interface Document;
    .....
}
```

Open standards and protection - Example

XHTML – XML Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema targetNamespace="http://www.w3.org/1999/xhtml"
  xmlns:xs="http://www.w3.org/2001/XMLSchema"
  xmlns="http://www.w3.org/1999/xhtml">

  <xs:annotation>
    <xs:documentation>
      This is the XML Schema Presentation module for XHTML
      This is a REQUIRED module.
      $Id: xhtml-pres-1.xsd,v 1.1 2003/12/17 03:09:17 ahby Exp $
    </xs:documentation>
    <xs:documentation source="xhtml-copyright-1.xsd"/>
  </xs:annotation>
```

Open standards and protection

- Software can be a *derivative work* of a specification
 - Can be copyright infringement
- But Open Standards should encourage Open Source implementation
- While preventing forking of specification...
- Means available: license
- IETF: separate the specification into:
 - *Text – prohibit (meaning-changing) derivative works, and*
 - *Code – allow derivative works.*
- W3C: work on document license for HTML5
 - Prevents specification forking
 - Compatible with open source licenses (GPL, LGPL, Apache, MPL, ...)

Web standards and standardisation

Example: **World Wide Web Consortium** as standardisation initiative

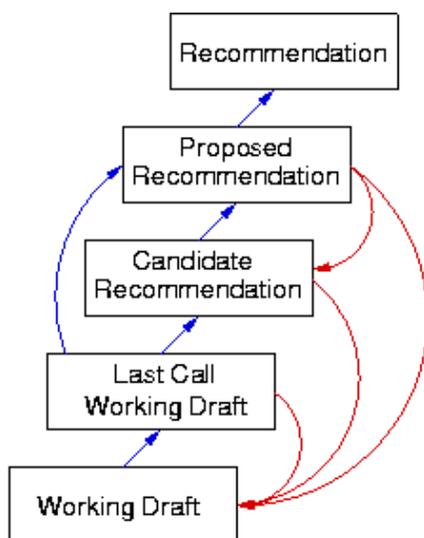
Standardisation for the web – W3C

- W3C[®] **World Wide Web Consortium** (1994-)
- Industry consortium
- Specify web technologies/standards
- Contribute to good use of standards
- Publish standards ("W3C Recommendations")
 - HTML, HTTP, XML, CSS, RDF,

W3C collaborates with standardisation initiatives

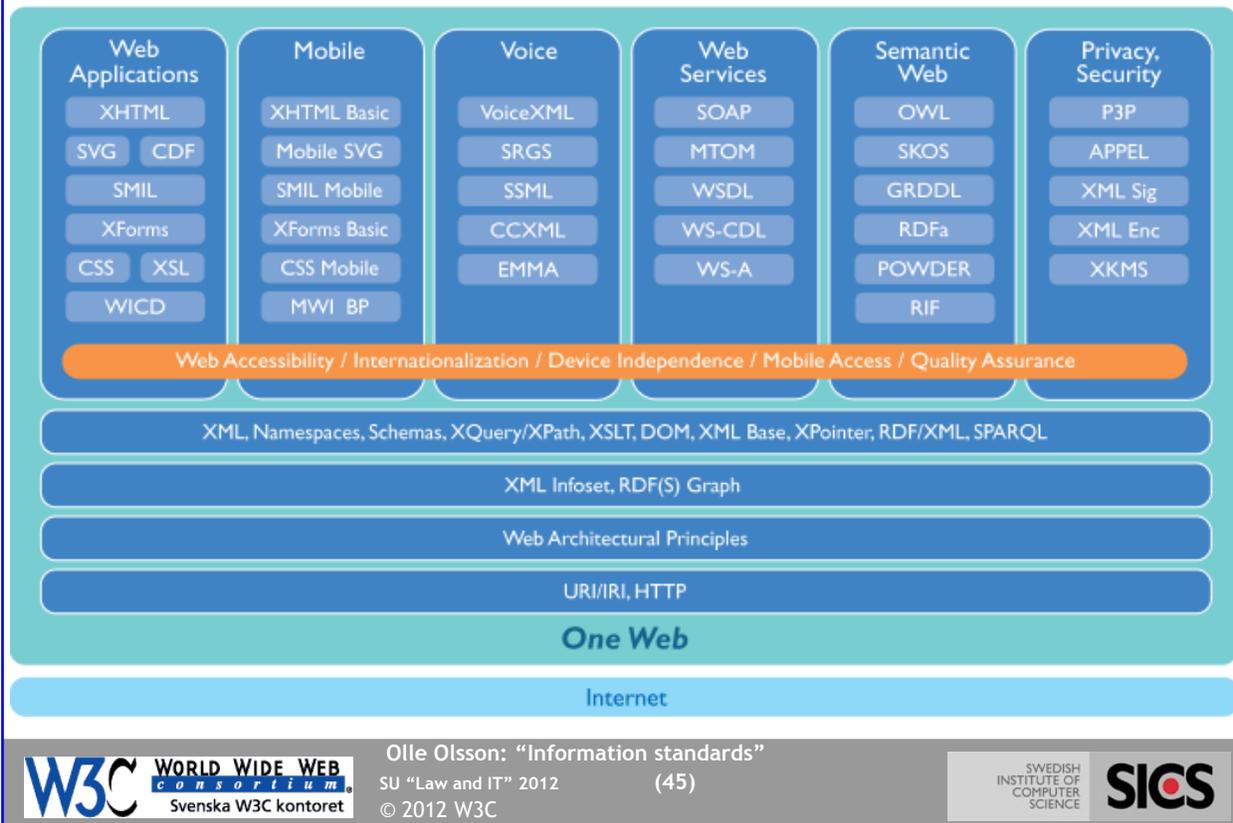
3GPP	FSTC	ITIC	OWASP
AccessBoard	GFSI	ITU	SMPTE
AILF	I3A	IW3C2	TOG
ATIA	ICANN	JIS	Unicode
Apache	ICC	Liberty Alliance	UN/CEFACT
BSI	IEEE	MPIC	Unicode
CEN	IETF	NIST	VoiceXML
CESI	IGF	OASIS	WAB-Cluster
DATSCG	IGF-DCOS	OGF	WASP
DCMI	IMS	OMA	Web3D
Daisy	INCITS	OMG	WS-I
EuroAccessibility	IPTC	OeBF	
ETSI	ISO	Open GIS Consortium	

Life cycle for W3C standardisation process



1. *Members propose work to be done*
2. *Advisory Council supports proposal*
3. Working Draft:
 - Technical proposal to be reviewed
4. Candidate Recommendation
 - Proposal that can be validated via implementations
5. Proposed Recommendation
 - Reviewed and validated proposal fulfilling requirements
6. Recommendation
 - Accepted as web standard

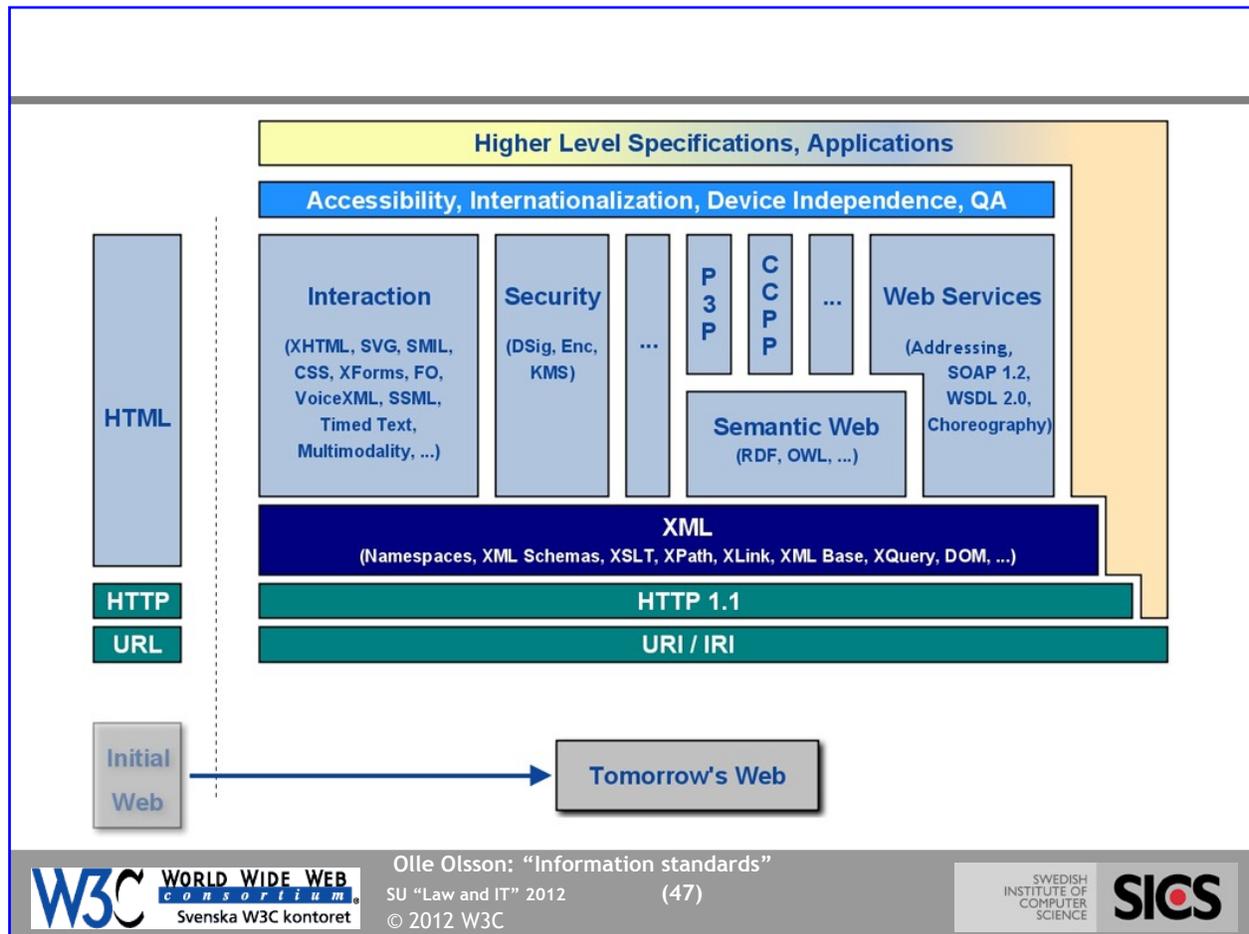
Standardised web technologies



Development – early years

1996	1997	1998	1999	2000	2001	2002	2003	2004
PNG	PICSRules	XHTML 1.0	CSS 1	ATAG 1.0	MathML 2.0	XML Signature	SVG 1.1	CC/PP
PICS 1.1		MathML 1.0	Namespaces	XML 1.0	Canonical XML	P3P 1.0	SVG Mobile	DOM 3
CSS1		CSS 2	WebCGM	DOM 2	XML Schemas	XML Canonicalization	XPointer	Infoset(2nd)
		SMIL 1.0	RDF (Old Version)	(X)HTML	Ruby	XPath Filter	SOAP 1.2	Namespaces 1.1
		DOM 1	WCAG 1.0		XLink 1.0	Decrypt Transform	XForms 1.0	XML 1.0 (3rd)
			Style Sheets PI		XML Base	XML Encryption	XML Events	XML 1.1
			MathML 1.01		SMIL 2.0	UAAG1.0	MathML 2.0	OWL
			XPath 1.0		SMIL Animation		PNG (2nd)	RDF
			XSLT 1.0		SVG 1.0			Voice Framework
					XSL 1.0			
					Infoset			
					WebCGM			

Date: 25-Sep-2004



Examples: development time

XML

- WD 14-Nov-96
- WD 31-Mar-97
- WD 30-Jun-97
- WD 07-Aug-97
- WD 17-Nov-1997
- PR 8-Dec-1997
- Rec 10-Feb-1998

Xforms 1.0

- Extensible Forms Description Language (XFDL) 4.0 Proposal submitted 2-Sep-1998
- XML Forms Architecture (XFA) Proposal submitted 14-Jun-1999
- WD 06-Apr-2000
- WD 15-Aug-2000
- WD 19-Dec-2000
- WD 16-Feb-2001
- WD 08-Jun-2001
- WD 28-Aug-2001
- WD 07-Dec-2001
- WD 18-Jan-2002
- WD 21-Aug-2002
- CR 12-Nov-2002
- PR 01-Aug-2003
- Rec 14-Oct-2003

XForms 1.1

- WD 15-Nov-2004
- WD 09-Dec-2005
- WD 14-Jul-2006
- WD 03-Nov-2006
- WD 12-Dec-2006
- WD 22-Feb-2007
- CR 29-Nov-2007
- PR 18-Aug-2009
- Rec 20-Oct-2009

WD – Working Draft
CR – Candidate Recommendation
PR – Proposed Recommendation
Rec – Recommendation

W3C Patent Policy

- Standards should not depend on patented technologies
- Objective:
 - “In order to promote the widest adoption of Web standards, W3C seeks to issue Recommendations that can be implemented on a Royalty-Free (RF) basis. Subject to the conditions of this policy, W3C will not approve a Recommendation if it is aware that Essential Claims exist which are not available on Royalty-Free terms.”
- Exceptions may be acceptable

The value of standards

How standards landscape changes over time

- The landscape of standards evolves
- To use standards in the best way, one should know in what ways things typically change
- ... useful to have a model of the universe of standards
 - as an ecosystem

Supplier – why standards?

Driving forces for suppliers:

- Broaden customer base
- Provide "pluggable" technologies
- In practice "outsourcing" of platforms / components
- Standards-based products extended with "features"
- Stability – investing in product offerings
- Etc.

"Enrol and lock-in customers"

"... standardization benefits entrants, complementors, and consumers, but may hold little interest for dominant incumbents." (Shapiro & Varian)

Suppliers: participate in standardisation – why?

- Influence standards
 - For own benefit
- Influence standardisation process
 - Take advantage of time-wise effects
- Create ecosystem as means of competition
 - Cooperation with other standardisation participants
- Observe / close study of technology field
 - “insider”, what other participants know/do
- Add strength to standardisation work
 - Create expectations
- Guarantee own products future safe
 - Risk management (bet on the right things)
 - Early standards conformance

Industry and standardisation



Computer Sweden

Sök i arkivet

UTFÄRDAR EGNA RIKTLINJER | 2008-09-25 14:14

IBM hotar att lämna standardorgan

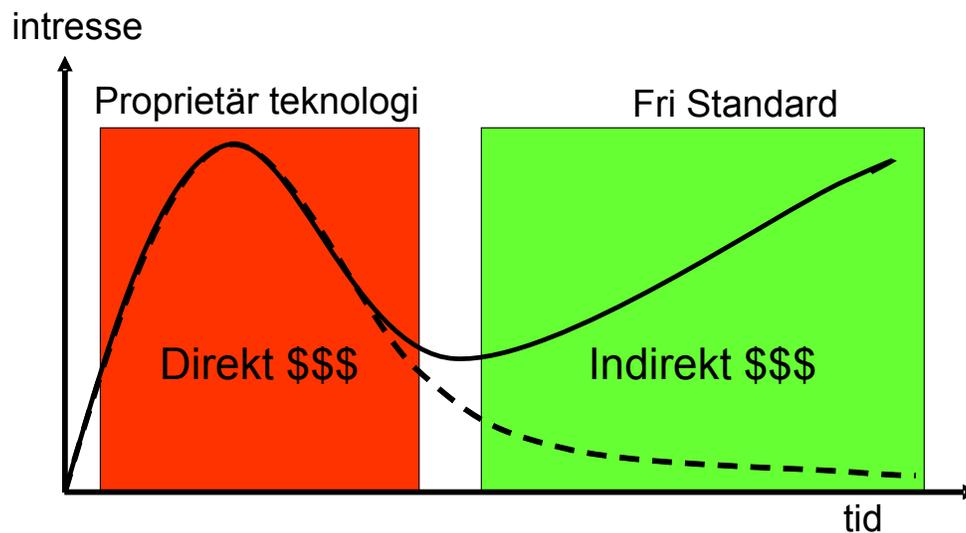


Av [Martin Wallström](#) |

Affärer & företag IBM hotar nu att lämna ett antal ledande standardorgan. Skälet är att företaget anser att den process som ligger bakom framröstandet av standarder inte är tillräckligt rättvis.

ANNONS

Supplier -- profitability?



Users – why standards?

Driving forces for suppliers:

- Avoid odd solutions
- Prolong effective life time of investment
- Secure access to competence
- Quality assurance
- Increase probability that **some** supplier exists
- Increased vendor independence
- Possibility to have **several** providers

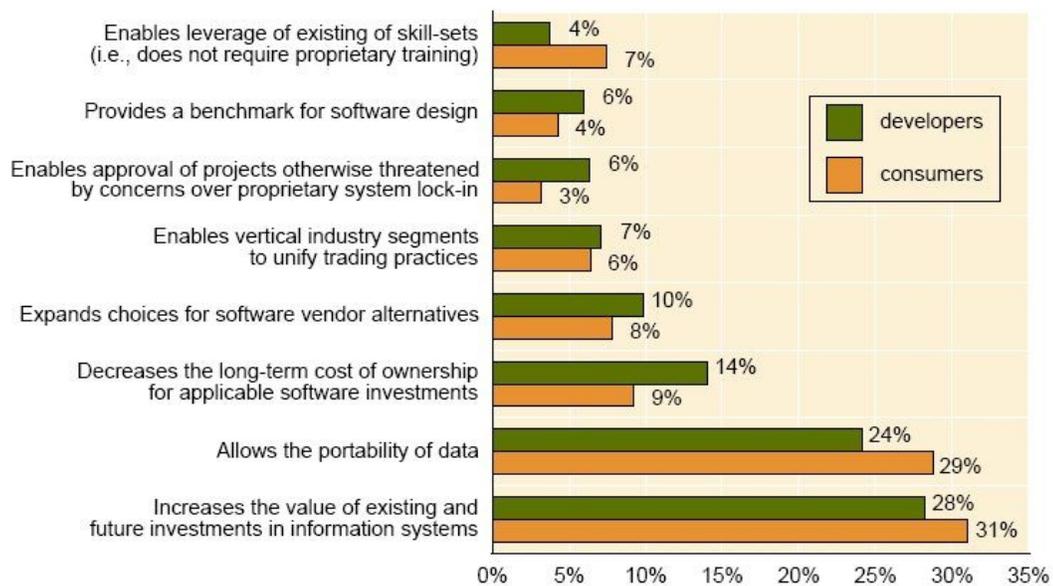
"Long-term effective investments"

Users: participate in standardisation work – why?

- Argue for users' needs, priorities, preferences, ...
- Act as counter balance for suppliers

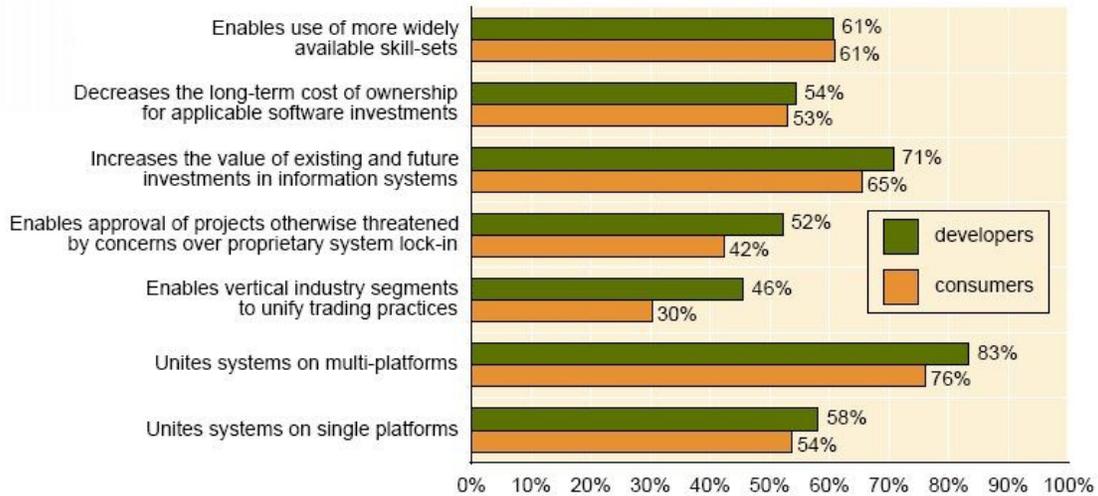
Stated advantages

Single most important advantage of general standard



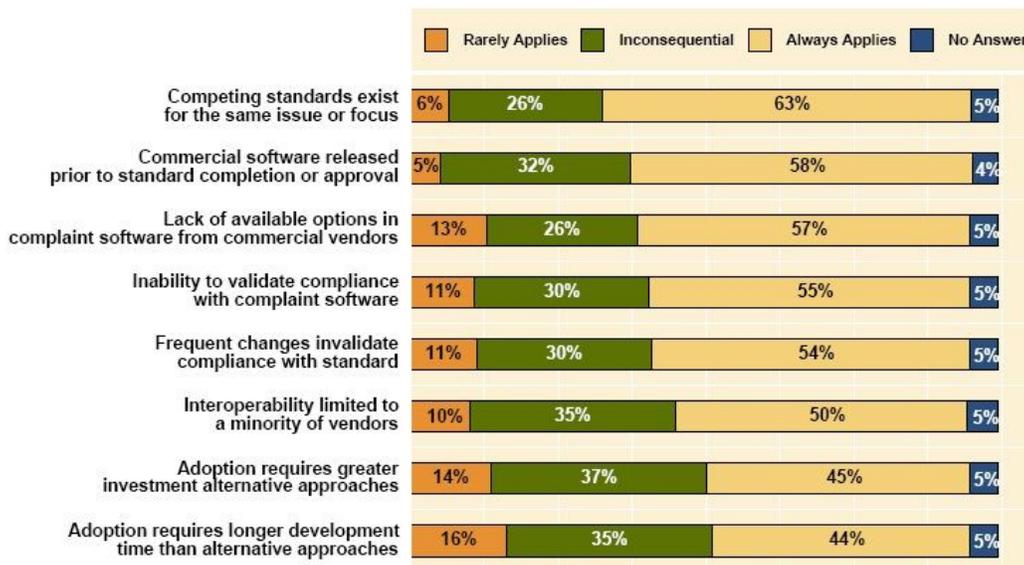
Stated advantages

Single most important advantage of *open* standard

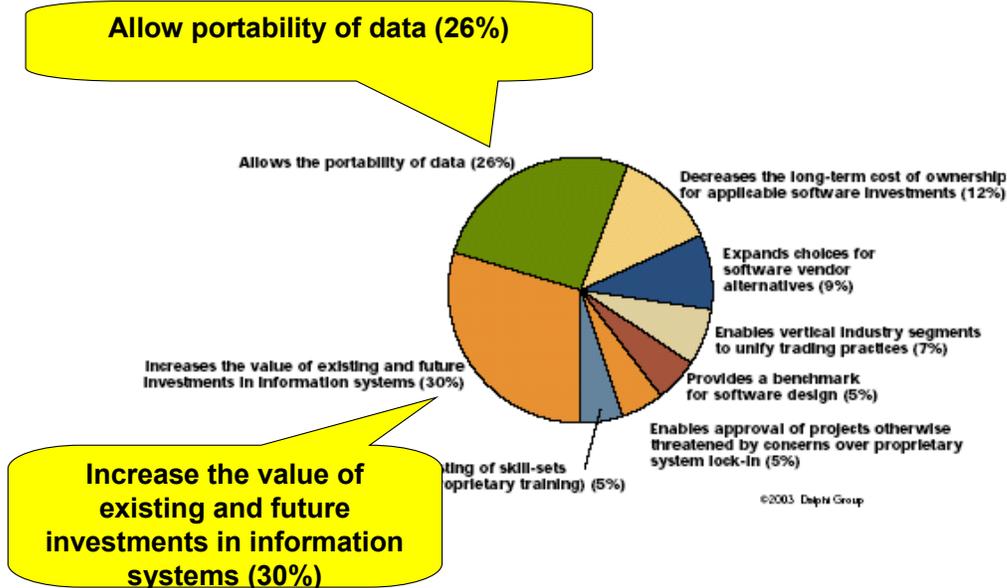


Stated risks

Drawbacks and threats to standards



Users' advantages



Conclusion about value of standards

- It is profitable to use **vendor independent** standards – especially web standards:
 - Today
 - Decreased costs for development and maintenance
 - Improved interoperability
 - Ensuring quality
 - Tomorrow
 - Increased vendor independence
 - Decreased costs for migration and rejuvenation
 - Improved support for heterogeneous environment
 - Simplified approach to service-oriented infrastructures

Why standards? Well, because ...

- Improve market
 - Foster international trade
 - Increased market size
 - Lower barriers to entry
 - Increased competition
 - Diffuse new technologies
- Decrease sector barriers
 - Improved compatibility, interoperability, ...
- User/usage support
 - Set limits for safety protection

Why standards? Well, because ... /2

- Innovation
 - Create forces that move innovation to new areas
- etc

When to standardise

- Reactive standardisation
 - Some market and need exist
 - Some technology is a clear winner
 - Has broad usage
 - Might be a basis for a standard with a future
 - “Rubber-stamp” what is already seen
- Proactive standardisation
 - A need for a standardised solution
 - No clear “winner” seen
 - Preconditions look good
 - Foreseen result differs from what we have
 - (Enough support)

Drawbacks/problems with standards?

Standards – a competitive field

All standards are not equal

- Relevant or irrelevant?
- Alive or archaic?
- Better or worse?
- Popular or marginal?

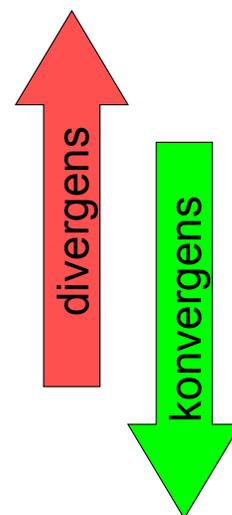
Evolutionary landscape

- “Survival of the fittest”
- Internal battles
- Qualitative changes in surrounding world

Standards – multiplicity

”stacken av standarder”:

- verksamhet
- affärsprocessmodeller
 - datalager
 - distribuerad bearbetning
 - meddelandestruktur
 - katalogstrukturer
- infrastruktur
- datakodning
 - protokoll för transport
 - nätverksprotokoll
 - fysisk konnektivitet



Standardisation negative for innovation?

- Standards “freeze” aspects on technology
 - Prevents new ways of thinking?
 - Negative effects on innovation?
- Standards move attention and releases resources
 - To new implementations of standards
 - To new neighbouring areas
 - “upwards in the stack”
- Natural selection ... in the long term perspective
 - Disruptive technologies ... radical changes of the landscape

Standardisation – bad compromises?

“Engineering”

- Make decisions about alternatives
- “trade-offs”
- Useful and rational results

Standardisation

- Make decisions about alternatives
- “trade-offs”
- Useful and rational results



“Not optimal for any specific case, but useful and valuable for most”

Extended standards?

”There's a sordid history in the technology world of everybody trying to get a little leverage over somebody else by developing **proprietary extensions** or **vendor-specific add-ons** to the core technology.

In general, those have been bad, because they **don't end up being extendible** over time and that costs companies like us a lot of money.”

CIO of a Fortune 100 corporation

Information standards – the XML approach

The basic concepts

- **Mark-up language**
 - Special annotations are introduced in a text
- **XML** (eXtensible Markup Language)
 - Set of rules for XML-based markup languages
- **XML-based** markup language
 - Set of rules for a markup with some intended use

- **XML**
 - The meta language for markup languages
 - The tool for designers of markup languages

XML application - example

```
<?xml version="1.0"?>
<rss version="2.0">
  <channel>
    <title>Example Channel</title>
    <link>http://example.com/</link>
    <description>My example channel</description>
    <item>
      <title>News for September the Second</title>
      <link>http://example.com/2002/09/01</link>
      <description>other things happened today</description>
    </item>
    <item>
      <title>News for September the First</title>
      <link>http://example.com/2002/09/02</link>
    </item>
  </channel>
</rss>
```

XML application - text

```
<?xml version="1.0"?>
<rss version="2.0">
  <channel>
    <title>Example Channel</title>
    <link>http://example.com/</link>
    <description>My example channel</description>
    <item>
      <title>News for September the Second</title>
      <link>http://example.com/2002/09/01</link>
      <description>other things happened today</description>
    </item>
    <item>
      <title>News for September the First</title>
      <link>http://example.com/2002/09/02</link>
    </item>
  </channel>
</rss>
```

XML application - elements

```
<?xml version="1.0"?>
<rss version="2.0">
  <channel>
    <title>Example Channel</title>
    <link>http://example.com/</link>
    <description>My example channel</description>
    <item>
      <title>News for September the Second</title>
      <link>http://example.com/2002/09/01</link>
      <description>other things happened today</description>
    </item>
    <item>
      <title>News for September the First</title>
      <link>http://example.com/2002/09/02</link>
    </item>
  </channel>
</rss>
```

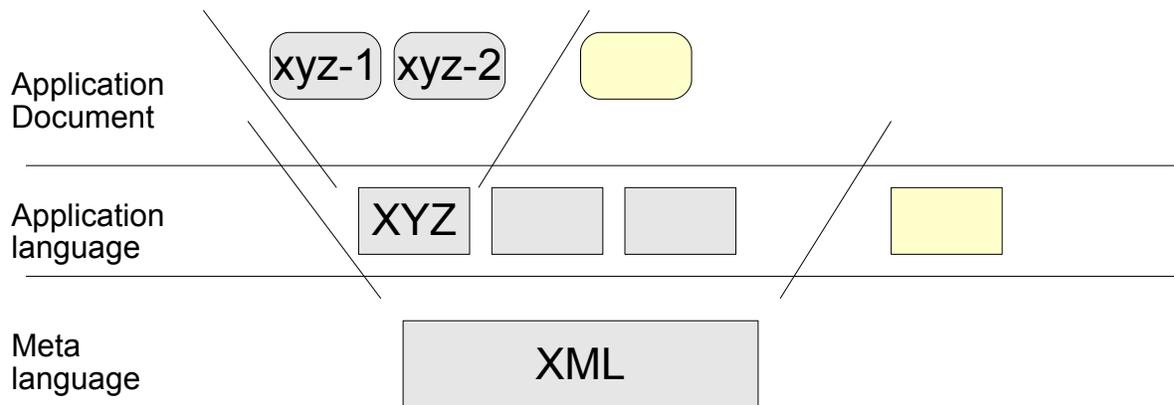
XML application - attributes

```
<?xml version="1.0"?>
<rss version="2.0">
  <channel>
    <title>Example Channel</title>
    <link>http://example.com/</link>
    <description>My example channel</description>
    <item>
      <title>News for September the Second</title>
      <link>http://example.com/2002/09/01</link>
      <description>other things happened today</description>
    </item>
    <item>
      <title>News for September the First</title>
      <link>http://example.com/2002/09/02</link>
    </item>
  </channel>
</rss>
```

Defining new languages

- XML-based language
 - Is an application of XML
 - Looks like XML: **<foo bar="6">Abc <fum>def</fum></foo>**
 - What elements and what attributes?
 - How can they be mixed?
 - What texts? What attribute values?
- Given a defined XML-based language XYZ
 - Documents expressed in XYZ
 - Meaning as intended by definition of XYZ

Being of a type ... being well-defined ...



Being of a type ... checking

A: For an application markup language XYZ

- Is it an XML-based language?

B: For an application document

- Is it an XYZ type of document?

For B:

- Traditional linguistic approach:
 - Grammar for language ... parse application document
- Specific XML approach:
 - Schema for language ... analyse application document

Schemas and schema processing

- XML application document
 - Annotated text, linear
 - Represents hierarchy of elements
 - Tree structure
- An XML Schema definition
 - Defines permissible tree structures
 - What types of elements may contain what other types of elements, in what order and what attributes
- PSVI – Post Schema Validation Infoset
 - Default values, ...
- An application document xyz-1 **conforms to** a schema XYZ:
 - XYZ **validates** xyz-1

XML format – example document

```
<shipTo country="US">
  <name>Alice Smith</name>
  <street>123 Maple Street</street>
  <city>Mill Valley</city>
  <state>CA</state>
  <zip>90952</zip>
</shipTo>
```

XML Schema – example definition

```
<xsd:element name="shipTo" type="USAddress"/>

<xsd:complexType name="USAddress" >
  <xsd:sequence>
    <xsd:element name="name" type="xsd:string"/>
    <xsd:element name="street" type="xsd:string"/>
    <xsd:element name="city" type="xsd:string"/>
    <xsd:element name="state" type="xsd:string"/>
    <xsd:element name="zip" type="xsd:decimal"/>
  </xsd:sequence>
  <xsd:attribute name="country" type="xsd:NMTOKEN"
    fixed="US"/>
</xsd:complexType>
```

What it basically is about

- Designing a language
 - ... to express certain facts
 - ... embedded in XML syntax
 - ... understandable by others
 - ... good "citizen" in the community of standards
-
- Who designs?
 - When?
 - How?
 - How evaluate?

Language design – challenges

XML-based language – design challenges

- What elements?
 - What names?
 - What attributes? What value types?
 - What names?
 - What structural hierarchy?
 - What constraints on structure?
 - What reuse of element/attributes from other languages?
 - Embedding other languages in this language?
 - Embedding this language in other languages?
 -
- When is it a good language design?*

Weakness in specification – formal

- Does it cover the real needs?
 - Enough expressibility
- Is it of usable size?
 - Total size
 - Modularization
 - Profiles
- Can it be extended?
 - Evolution, new versions
- Does it build on strong standards?
 - Foundation building blocks ...

Weakness in specification – formal/2

- Is the textual specification consistent?
 - Are there statements that are in conflict with each other?
- Is the textual specification complete?
 - Are all important cases covered?
- Is the textual specification deliberately vague?
 - Does it use terms like "SHOULD", "SHOULD NOT", "MAY", ...?

Weakness in specification - pragmatical

- Does it embed patented technologies?
- Is it a free standard?
- Is it an open standard?
- Is there “authoritative” informative material describing *intended use* of the standard?

Weakness in use of specification

- Does use conform to specification?
 - Do concrete instances comply to specification?
- Does used software respect the specification?
 - Is the software implementation “compliant” to specification
- Is it used according to intended uses (use cases)?
 - Is this a targeted use?

Most critical perspective

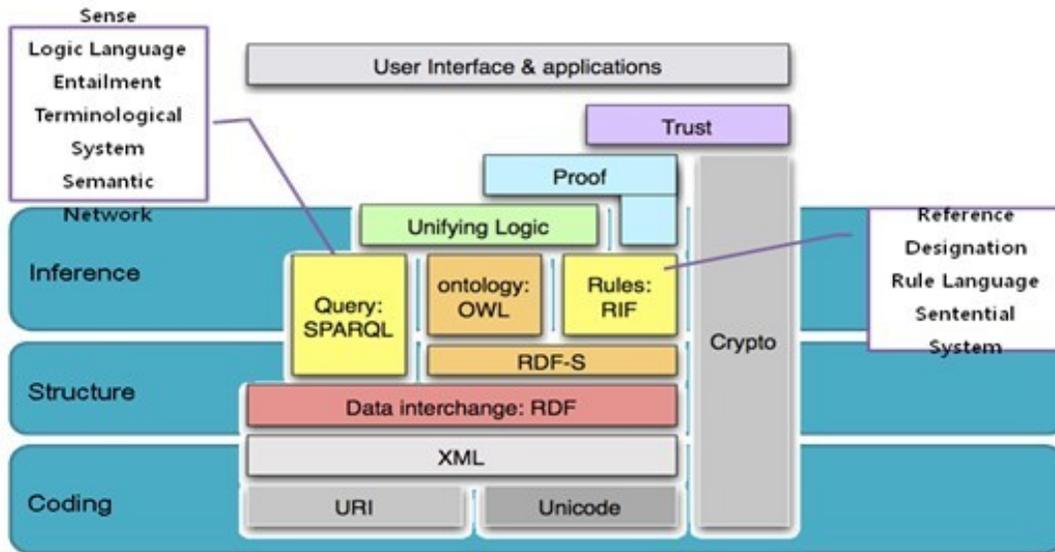
- What is the conceptual model underlying a language?
 - Entities, relationships, properties, ...
- Does my domain correspond to the language's conceptual model?
- Are the conceptual models of two languages compatible?
 - If not, then full interoperability may be impossible

Challenge: Sets of standards

- Interoperability among standards
- Building-block reuse of standards

- Can they be combined?

Technology stack



Profiles – hypothetical example

Web Services Interoperability (WS-I) Basic Profile 1.0:	Options:	Total: 268 435 456
1. Simple Object Access Protocol (SOAP) 1.1	1 2 3 4	
2. Extensible Markup Language (XML) 1.0 (Second Edition)	1 2 3 4	
3. Hypertext Transfer Protocol -- HTTP/1.1	1 2 3 4	
4. HTTP State Management Mechanism	1 2 3 4	
5. Web Services Description Language (WSDL) 1.1	1 2 3 4	
6. XML Schema Part 1: Structures	1 2 3 4	
7. XML Schema Part 2: Datatypes	1 2 3 4	
8. UDDI Version 2.04 API Specification	1 2 3 4	
9. UDDI Version 2.03 Data Structure Reference	1 2 3 4	
10. UDDI Version 2 XML Schema	1 2 3 4	
11. RFC2818: HTTP Over TLS	1 2 3 4	
12. RFC2246: The TLS Protocol Version 1.0	1 2 3 4	
13. The SSL Protocol Version 3.0	1 2 3 4	
14. RFC2459: Internet X.509 PKI Certificate and CRL Profile	1 2 3 4	

XML-based languages

Areas and standards

Standards & use areas

- XBRL – eXtensible Business Reporting Language
 - language for the electronic communication of business and financial data
 - U.S. Securities and Exchange Commission (SEC): companies to submit financial reports in XBRL.
- P3P – Platform for Privacy Preferences Project
 - Websites can express their privacy practices in a standard format that can be retrieved automatically and interpreted easily by user agents
- EDRM – Electronic Discovery Reference Model
 - discovery in civil litigation which deals with information in electronic format (Electronically Stored Information, ESI).

Standards & use areas/2

- MetaLex – CEN Workshop on an Open XML Interchange Format for Legal and Legislative Resources
 - aims to standardize the way in which sources of law and references to sources of law are to be represented in XML.
- Crown XML Schema for Legislation
 - full and comprehensive encoding for all United Kingdom primary and secondary legislation.
- See also SDU BWB, LexDania, ...

Standards & use areas/3

- Digital rights management
 - XrML (eXtensible Rights Markup Language); describes rights, fees and conditions together with message integrity and entity authentication information
 - CcREL (Creative Commons Rights Expression Language) ; language for descriptive metadata to be appended to media that is licensed under any of the Creative Commons licenses.

Standards & use areas/4

- Access control
 - XACML (eXtensible Access Control Markup Language); access control policy language and a processing model, describing how to interpret the policies
- Security
 - SAML (Security Assertion Markup Language) ; standard for exchanging authentication and authorization data between security domains

Other emerging areas

- Public sector information
 - GovML (Governmental Markup Language) – an XML vocabulary to support the delivery of content and services to citizens (businesses) in terms of life-events (business episodes)
- Contracts
 - cf eContracts in LegalXML

Bibliography

Bibliography (selected)

- Hebner, R. E. (1999), "Standards and Trade – Who Really Cares". In *Technology, Standards, and the Standardization Process – Their Impact on Innovation and International Competitiveness*, Stanford University, Palo Alto, California, U.S.A., 1999, pp. 4-16.
- Shapiro, S. and Varian, H.R. (2002), "The Art of Standards Wars". In Raghu Garud, Arun Kumaraswamy, Richard N. Langlois (eds): *Managing in the modular age: architectures, networks, and organization*, Wiley 2002, pp 247-269
- Meeks, G. and Swann, G.M.P. (2009), "Accounting Standards and the Economics of Standards", *Accounting and Business Research*, Vol.39(3), pp. 191-210.
- Swann, G.M.P. (2000), "Final Report for Standards and Technical Regulations Directorate Department of Trade and Industry"; <http://www.berr.gov.uk/files/file11312.pdf>

- Lee, N. (2006), "Patented Standards and the Tragedy of Anti-Commons". *Teollisoikeudellisia Kirjoituksia*, 2006.
- Zhu, H., Fu, L. (2009) "Towards Quality of Data Standards: Empirical Findings from XBRL, *30th International Conference on Information System (ICIS'09)*, December 15-18, Phoenix, AZ, USA.
- Merges, R.P. and Kuhn J.M. (2009), "An Estoppel Doctrine for Patented Standards", *California Law Review*, February 2009, Vol. 97, No. 1
- Geradin, D. and Rato, M. P.L. (2006), "Article 82, IP Rights and Industry Standards: In Reply". Available at SSRN: <http://ssrn.com/abstract=950215>
- DeNardis, Laura and Tam, Eric (2009), "Open Documents and Democracy: A Political Basis for Open Document Standards". *The Indian Journal of Law and Technology*, Volume 5, 2009, p 39-78

- Simcoe, T.S. (2008), "*Standard Setting Committees*". Available at SSRN: <http://ssrn.com/abstract=899595>
- Baird, S. (2007), "Government at the Standard Bazaar". *Stanford Law & Policy Review*, Vol. 18, No. 35, 2007.
- Baird, S. (2009), "Government Role and the Interoperability Ecosystem". *Journal of Law and Policy for the Information Society*, Vol. 5, No. 2, p. 219, Summer 2009.
- Rosen, L. (2009) "Defining Open Standards". In *The Standards Edge: Unifier or Divider?*, Sheridan Books (to appear)
- Rosen, Lawrence (2011), "Implementing Open Standards in Open Source", *Software Tech Journal*, Vol. 14, Number 1 ("DoD and Open Source Software", February 2011)
https://journal.thedacs.com/stn_view.php?stn_id=56&article_id=178
- Camp, L. J. and Vincent, C., (2004) "*Setting Standards: Looking to the Internet for Models of Governance*". Available at SSRN: <http://ssrn.com/abstract=615201>

---end---