Since 1989

Happy 20th Birthday, World Wide Web


© LE FEVRE COMMUNICATIONS
The Mind Behind the Web
Tim Berners-Lee invented the World Wide Web and continues to shape its frantic evolution. He's not

© ISTOCKPHOTO.COM/ ALEX DONNA COVENEY SLOBODKIN
Facts about the Web's Creation
Everything you ever wanted to know about the Web's first days

Remembering the Day the World Wide Web Was Born
What drove Tim Berners-Lee to imagine this game-changing model for information sharing.

© ISTOCKPHOTO.COM/ ANDREY VOLODIN
The Future of Computing (Circa 1999)
M.I.T.'s Laboratory for Computer Science is developing a new infrastructure for
W3C Launches Web and Mobile Interest Group
20 August 2013 | Archive

W3C launched today a Web and Mobile Interest Group that is chartered to accelerate the development of Web technology so that it becomes a compelling platform for mobile applications and the obvious choice for cross platform development. The forum is intended to include organisations that commission such products and services, designers, developers, equipment manufacturers, tool and platform vendors, browser vendors, operators and other relevant participants in the value chain that creates and operates such products and services. Participants will focus on a wide range of sectors including retail, advertising, technology, network operators, content creation and content distribution.

The initial deliverables of the group include:

- Core Mobile Web Platform 2012 Deployment Status, which will summarize the various actions that the Interest Group is undertaking to ensure that the relevant stakeholders facilitate the deployment and adoption of the features that have been identified in the Core Mobile Web Platform 2012 report. The group will also publish new versions of the report
- Standards for Web Applications on Mobile: current state and roadmap, which will take a broader look at all the Web technologies under development that are particularly relevant to mobile devices, and tracks their status and adoption.
Don’t develop your standards at the kids’ table.

- Accredited, internationally recognized, auditable process
- Proven IPR Policy
- Liaisons with de jure bodies
- Rules that protect the level playing field

OASIS. You belong here.
Web 2.0: Where Are We?

http://youtu.be/NLiGopyXT_g

Current Anthropology
HTTP: HyperText Transfer Protocol

Client (IE, Firefox, Opera) → Request → Response
HTTP: HyperText Transfer Protocol

Request:
http://foo.bar.com/hello.htm

Connect me to:
foo.bar.com
I want file: hello.htm

Response
<html>
<body>
Hello World!
</body>
</html>
URL

- **URL**: Universal Resource Locator
  - Part of URI (Universal Resource Identifier)
  - But often used interchangeably
- **Character string to uniquely specify the location on the Web**
- **Syntax**  
  `Scheme :// domain : port / path`
  - Scheme: http, https, ftp, gopher, ...
  - Domain: ist.psu.edu
  - Port: 80 (default), 8000
  - Path: /course/2013-fall/index.html
Hyperlinks

- Hypertext: Text with a reference to other text
- Hyperlinks as the reference on the Web
  - Link any unit of information (e.g., documents) to any unit of information on the Web
- Concept suggested by V. Bush in 1945
  - “one could link any two pages of information into a trail of related information, and then scroll back and forth among pages in a trail…”
- Term coined by T. Nelson in 1965
- Independently implemented by D. Engelbart in 1966
Mark-Up

- A way to annotate text with distinctive syntax
  - Eg, TeX, LaTeX, HTML
- Structural Mark-Up: SGML
  - A notation for writing text with markup tags (<tag>)
  - Tags indicate the structure of the text
  - Tags have names and attributes
  - Tags may enclose a part of the text
  - Invented around 1970 by C. F. Goldfarb
History of HTML

- HTML: Hyper-Text Markup Language
  - Invented by Tim Berners-Lee and Robert Caillau at CERN in 1991
  - Stripped-down version of SGML
- Two important features: Hypertext & Markup
  - Can jump btw. web documents using links
  - Data in tag notations
- HTML 2.0 in 1995
- HTML 4.0 in 1997
- HTML 4.01 in 2000
- HTML 5 in 2008 (working draft)

Widely implemented & supported as of now
HTML 5

- To replace: HTML 4.01, XHTML 1.0, DOM 2.0, & Web Forms 2.0, etc
- To reduce: RIA (Plug-in based Rich Internet Application) architecture usage
  - Eg, Adobe Flash, MS Silverlight
  - <audio> and <video> components
- Better error handling
  - Different web browsers give consistent results
- Inline SVG and MathML
- To remove: applet, font, frame, frameset, …
• HTML was designed to **display data** and to focus on **how data looks**.
Static vs. Dynamic HTML

- We want a piece of code that really “runs”, i.e. generate different result in different conditions
- Two kinds of scripts:
  - Client side script: running at client side, i.e. in your browser
    Eg, JavaScript
  - Server side script: running at server side, give you the result in the form of pure HTML
    Eg, PHP, Servlet
Dynamic HTML

Request:
http://foo.bar.com/date.php

Connect me to:
foo.bar.com
I want file: date.php

This is PHP code

```
<html><body>
<?echo date("l, F d Y")?>
</body></html>
```

date.php

execution

```
<html><body>
<?echo date("l, F d Y")?>
</body></html>
```

Response

```
Tuesday, October 04 2005
</body></html>
```

PHP output here.

“Explain”

This is a PHP code that outputs the current date in the format "l, F d Y".

```
<html><body>
Tuesday, October 04 2005
</body></html>
```
What does XML stand for (GRE question)?

1. X-rated Modular 3 Language
2. eXpressive Machine Learning Language
3. eXtensible UML (Unified Modeling Language) for the Web
4. eXtensible Meta Language for the Web
5. eXtensible Markup Language originating from SGML
6. None of the above
What Human Sees

About Penn State
Prospective Students
Academic Programs
Outreach Programs
Global Penn State
Campuses and Colleges
Visitors Guide
Alumni, Friends, and Giving...
What Machine Sees
Solution #1: Give Data Meaningful Tags

<School>
<Name>_solution</Name>
<TOC>…
</TOC>
<Description>…
</Description>
</School>
What Machine Sees from Solution #1

Better than before. Still NO clear and precise meanings of tags known to machines
XML

- XML is a framework for defining markup languages
- XML was designed to describe data, not format
  - Format is separately described by stylesheets
- XML separates syntax from semantics to provide a common framework for structuring information
- **NO fixed collection** of markup tags: One must define own tags, tailored for specific apps
- XML uses a schema language (eg, DTD, XML-Schema) to formally describe the data.
HTML vs. XML

**HTML**

```html
<center>
<h1>SIGMOD</h1>
<p><b><u>ACM</u> <a href="sigmod02.org">SIGMOD Conference</a>, Madison, WI, 2013</b></p>
</center>
```

**XML**

```xml
<event eID="sigmod13">
<acronym>SIGMOD</acronym>
<society>ACM</society>
<url>www.sigmod13.org</url>
<loc>
    <city>Madison</city>
    <state>WI</state>
</loc>
<year>2013</year>
</event>
```
HTML vs. XML

- Need a stylesheet to define browser presentation semantics
HTML vs. XML

- Need a stylesheet to define browser presentation semantics
XML 1.0 and 1.1

● Two current versions of XML

● XML 1.0
  ● Initially released in 1998
  ● Latest release in 2008 (5th edition)

● XML 1.1
  ● Initially released in 2004
  ● Latest release in 2006 (2nd edition)
Why is XML Important?

- Technically, ... little initially; Just old simple tree model...

- Non-technically, ...
  - Hot ($$$)
  - The standard for representation of Web information
  - The real force of XML is *generic languages* and *tools*!
  - By building on XML, you get a massive (standard) infrastructure for free
The XML Family of Specifications: The Big Picture

Last Updated: April 19, 2003

Copyright (c) 1990-2003 Kenneth D. Sall. All Rights Reserved. http://kensall.com/big-picture/
Semantic Web Video

http://youtu.be/OGg8A2zfWKg

An introduction to the Semantic Web
"The Semantic Web is an extension of the current web in which information is given well-defined meaning better enabling computers and people to work in cooperation…"

Motivation

- Current World-Wide Web is a geomantic conglomeration of information in distributed and heterogeneous environment
- Yet, the processing of such information is merely based on the syntactic matching of information
  - Eg, Web search using keyword matching
- More intelligent processing is needed with requirements:
  - Machine processing
  - Automatic processing
Current Web: Syntactic Web
Current Web: Syntactic Web

Van, Turkey

This article is about a city in Turkey. For other uses, see Van (disambiguation).

Van (Armenian: Վան) is a city in eastern Turkey and the seat of Van Province, and is located on the eastern shore of Lake Van. The city's official population in 2009 was 360,810, but many estimates put this as much higher with a 1996 estimate stating 500,000 and former Mayor Burhan Yengin quoted as estimating the population at approximately 600,000.

Contents [hide]

1 History
   1.1 Urartu
   1.2 From the Orontids to the Kingdom of Armenia
   1.3 The Byzantines and the kingdom of Vaspurakan
   1.4 The Seljuk Empire
   1.5 The Ottoman Empire
     1.5.1 City life
     1.5.2 Demographics
   1.6 World War I and the Armenian Genocide
   1.7 Turkish War of Independence and Republic

2 Van today
3 Cuisine

Cars > Vans

Used Vans for Sale

2004 Ford Econoline 350 for $12,775

2004 Ford Econoline, 83,384 miles, White
Color-keyed engine console cover w/dual stowage, (4) cup holders, Black plastic stepwell pads,...

Posted by Haldeman Ford with Oodle Pro

Hamilton, NJ · 10 hours ago

2009 Chevrolet Express 2500 WORK VAN for $15,619

2009 Chevrolet Express, 34,127 miles, White
Audio system, radio provisions only. Includes alarm warning chimes, radio wiring harness, radio...

Posted by Archer Kia Volkswagen Automotive with Oodle Pro
Q: Search Chevrolet Express Van with less than 10K miles → Currently, hard to answer this query!
Current Web: Syntactic Web w. Links
Current Web: Syntactic Web

- Markup consists of:
  - Rendering information (e.g., font size and colour)
  - Hyper-links to related content
- Semantic content is accessible to humans but not (easily) to computers…
What Human Sees

About Penn State
Prospective Students
Academic Programs
Outreach Programs
Global Penn State
Campuses and Colleges
Visitors Guide
Alumni, Friends, and Giving
What Machine Sees
Solution #1: XML w. Meaningful Tags

```xml
<School>
  <Name>解决方案
  </Name>
  <TOC>目录
  </TOC>
  <Description>描述
  </Description>
</School>
```
What Machine Sees from Solution #1

Better than before.
Still NO clear and precise meanings of tags known to machines
What is Needed? → Semantic Web

- **External agreement** on meaning of annotations
  - E.g., Dublin Core (DC) agree on the meaning of a set of annotation tags for “documents”
  - But, limited number of things can be expressed

- **Use Ontologies** to specify meaning of annotations
  - Ontologies provide a vocabulary of terms
  - New terms can be formed by combining existing ones
  - Meaning (i.e., *semantics*) of such terms is formally specified
  - Can also specify *relationships* between terms in multiple ontologies
Solution #2: XML w. Semantic Tags

<School>
<Name>_solution</Name>
<TOC>
...content...
</TOC>
>Description>
...content...
</Description>
</School>
What Machine Sees from Solution #2

- **Official title of an institution**: `<School>`
- **Same meaning as `<desc>` from The namespace: http://foo.com/myown**: `<TOC>`
- **Clear and precise meanings of tags known to machines**: `<Description>`
Reference

- XML Cover Page: www.oasis-open.org/cover/
- XML Tutorial:
  - www.w3schools.com/xml/default.asp
  - www.brics.dk/~amoeller/XML
  - www.xml101.com/xml/default.asp
- RDF Primer
- Introduction to the Semantic Web and RDF, A.M. Kuchling
  - http://www.amk.ca/talks/2004-12-02/
- Tutorial on the W3C OWL