

---

# **W3C: a Use Case for Web Technologies**

Brisbane, Australia 2009-07-02

Klaus Birkenbihl, W3C [klaus@w3.org](mailto:klaus@w3.org)



---

## The Web

After ~17 years in the wild the Web has changed and still develops in various directions:

- Web everywhere – the Mobile Web
- Web for everyone – the Accessible Web
- Web of structured information – the Semantic Web
- Web of participation and communication – Collections, Blogs and Fora
- Web as a service – Rich Web Applications
- Web for business – Web Services
- Web for media – Video and Voice on the web

*Admitted: there might be different views on this*

---

## Requirements for the Web

Some demands in order to make the Web useful and successful:

- privacy on the Web
- security and trust on the Web
- treatment of intellectual property
- content classification and rating
- accessibility of the Web
- scalability of the Web

---

## It started pretty simple

- a common syntax for documents that may contain hyperlinks (HTML)
  - *provides a base for structuring documents*
  - *a framework for applications to display document*
  - *a mechanism to link to other documents*
- a syntax for locating documents (or fragments of documents) and the option to use this syntax within HTML to build hyperlinks (URL)
  - *a base for a globally unique identification of documents*
  - *unique identification it relies on the Internet Domain Name System*
- a protocol to send request and responses between clients (usually browsers) and servers (HTTP)
  - *to make servers and clients understand each other*

Fairly simple, but most of the topics mentioned before where not yet addressed - except: it scaled!

---

## Broadening the Basis

- URL concept extended to URI and IRI
  - *a more general scheme for uniquely naming documents, real objects and beings, abstract concepts and ideas ...*
- HTML as basis is replaced by a more general concept: XML
  - *document structuring tags, can be freely chosen*
  - *can represent hierarchical databases*
  - *(apart from some syntactic restrictions of XML) HTML can be seen as a "restricted XML"*
  - *XHTML resolves this syntax problem*
- A few supporting technologies came with XML as there are:
  - *XML name spaces as a means to separate one space of tags from another*
  - *XML Schemas to allow the definition of restrictions to specify a new XML language*
  - *XQuery/Xpath as a means to navigate through or query XML documents*
  - *XML Base to cope with relative URIs/IRIs*
  - *XPointer to allow reference to XML documents and fragments*
  - *DOM an abstract Interface to process XML and related events*

## Even broader ...

- XML relies on the concept to define complex entities by composing them of simpler ones.
- RDF graphs try to define entities by describing their relationship to others.
  - An RDF graph is composed of triples of the form *subject predicate object*



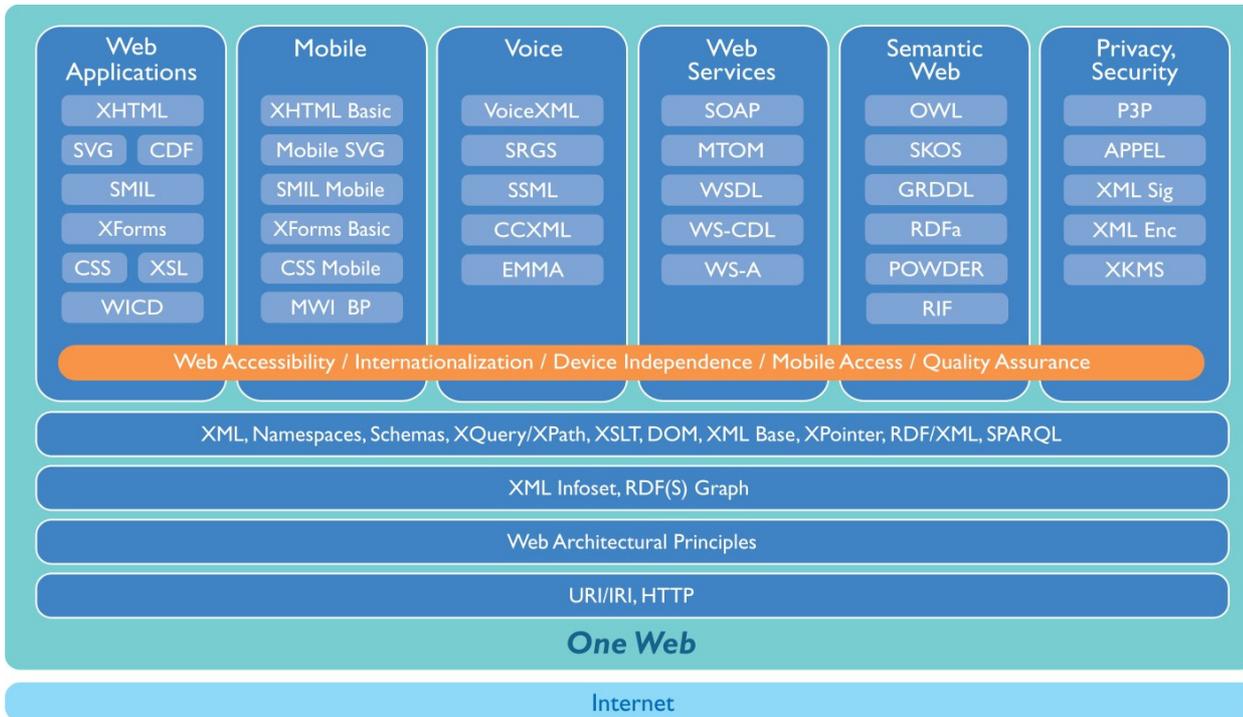
- *subject (a resource) denotes something you want to define, predicate defines the relationship (property), and object (a resource) is something for which the relationship holds*
- *Objects (and even predicates) in one triple can appear as subjects in others*
- As XML RDF comes with a set of associated technologies
  - RDFS is a Schema language that defines a vocabulary that gives meaning to some resources and properties. This opens a standardized way to express set or logic related properties. OWL supplements and extends these definitions.
  - SPARQL is a query language for RDF

---

## The universal Web

- everybody should be able to use the Web with a wide range of devices in any language he knows according to his mental capabilities
- general rules that help to achieve this:
- welcome side effects:

# Technology

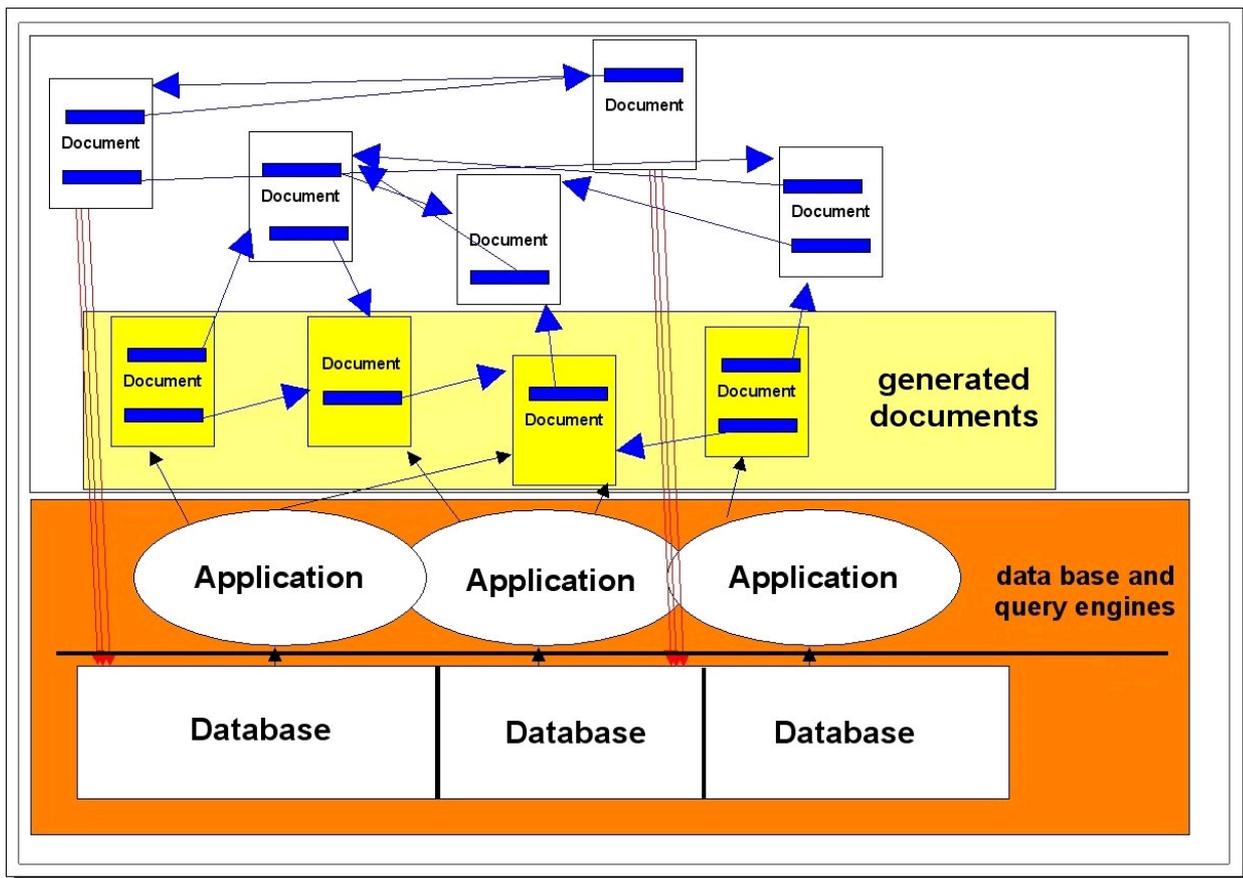


We cannot cover all. More on <http://www.ict-media.de/webtech.html>

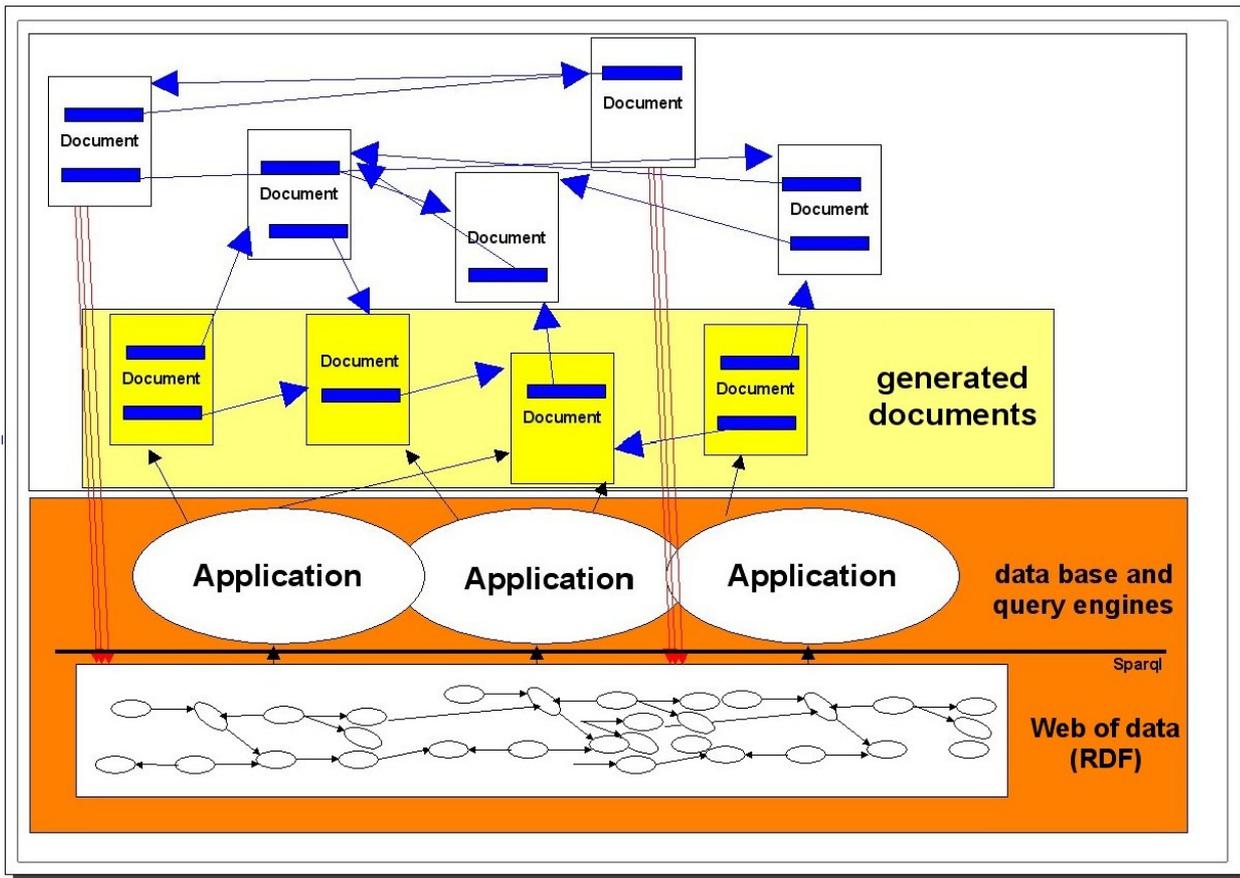
# Web applications

- plain HTML is all you need to build a Web application
- on the other end we see fully fledged applications for:
  - *office work or picture processing.*
  - *Video, audio*
  - *sophisticated graphics*
  - *presentational software for mathematics or chemistry*
- Plugins and Interpreters for programming languages allow nearly any content or processing that works on the desktop. Using technologies that are
  - *standardized*
  - *proprietary*
  - *Open Source*
  - *locked*
- some important standards
  - *HTML ... 2 are emerging actually*
    - the main focus of XHTML 2 is to better use the advantages of XML within HTML
    - HTML 5 is more on advancing the existing HTML (HTML 4 or XHTML 1) in a compatible way
  - *CSS to control presentation in a structured way*
  - *Xforms as a better way to exchange information with servers*
  - *SVG for vector based two dimensional graphics*
  - *compound document standards (WICD, CDF...) to allow better integration of parts of a document using different technologies*
- The Web – especially when combined with powerful programming languages and interfaces – can act as a platform for sophisticated applications.
  - *Games, maps, image processing or complete office suites...*
  - *using technologies like Compound Document Formats, XML, the DOM (and DOM events), Xforms, XHTML, SVG and some more*
- SMIL to build presentations with media from various sources in a synchronized manner

# The web today



# Semantic Web



---

## How do we get there

- Stop: Mash-ups/Ajax do it already!
  - *API based, often one API per source*
  - *APIs mostly proprietary*
  - *quite some effort in implementing*
  - *dependency from API provider*
  - *instrumental to show how powerful a Web of data can be*
- Microformats:
  - *standard to represent data*
  - *could be used by mash-up*
  - *not very flexible, rather static*
- standards to resolve this:
  - *RDF family (RDFS, OWL) to represent data*
  - *SPARQL to access data*
  - *RDFa allows to include RDF statements in an XHTML document*
  - *SKOS provides a data model for sharing and linking knowledge organization systems as application of RDF*

---

## Web Services

- Web Services is not about browsing
  - *interoperation of software applications that run on computers interconnected by a network*
  - *especially in e-business*
  - *pass data from one business step to the next ones*
  - *automate parts of the business process*
  - *lot of concepts and buzzwords (EAI, ESB, MOM, SOA...)*
- technologies
  - *SOAP specifies the format of messages that can be used to communicate with a service*
  - *WS-Addressing describes how to address Web Services and messages*
  - *WSDL is a language to describe a service by the kind of messages it will receive or send and its functionality. There is also a standard for Semantic Annotations to WSDL (SAWSDL) which allows to use ontologies to describe the semantics of Web Services.*
  - *a means to specify policies for the use of a service can be specified and attached to a service by using the Web Services Policy specifications.*
  - *Semantic Annotations for WSDL allow to include semantics into service descriptions*

---

## Mobile Web

- devices with limited capabilities (e.g. mobile phones)
  - *connectivity*
  - *bandwidth*
  - *processing power*
  - *display size and sound quality*
- limited capabilities are addressed by
  - *down sizing of existing standards to match the constraints of mobile devices*
  - *the development of guidelines especially for authors (but for authoring-tool manufacturers as well) on how to make sure that a Web site is mobile friendly*
  - *a way for servers to learn about the capabilities of an attached client*
- technologies:
  - *downsized standards XHTML Basic, SVG Tiny, SVG Basic, SML mobile, Xforms Basic and CSS mobile*
  - *DDR simple API describes an interface that servers can use to query device description*

---

## Voice: the phone on the Web

- designing voice dialogs
  - *VoiceXML* is about defining frameworks for dialogs.
  - *speech oriented standards to support the use of VoiceXML:*
    - The *SSML* is an *XML* based language to assist the generation of synthetic speech.
    - *SRGS* allows the specification of grammars for recognition of speech.
- *CCXML* provides support for telephony call control and can be used in conjunction with dialog systems such as *VoiceXML*.
- *EMMA* is an *XML* language to describe input received on several channels (voice, handwriting, keyboard...) simultaneously.

---

# Privacy and security

## ■ issues

- *while using the Web we disclose a lot of our personal data, our interests and preferences*
- *can we be sure that the Web site that pretends to be the home page of our bank is really our bank?*
- *can we securely transfer confidential data over the Web?*
- *can we securely identify ourselves on the web?*

## ■ technologie that help

- *P3P is a standard for specifying a machine readable description of a privacy policy*
- *TLS/SSL resolves much of the authentication, authorithation, confidentiality issues on the Internet transport layer*
- *XML Encryption, XML Signature and XML Key Management: 3 related technologies that allow to apply well known cryptographic methods on fragments of XML documents*

---

## The Web live

- here we are and ...
  - *technologies and standards are a means to an end*
  - *the list is far from complete*
  - *available bandwidth and compute power are steadily increasing*
  - *Web turned from geek's toy to a powerful infrastructure*
  - *formed by millions of individuals, business people, journalists, artists, scientists...*
  - *Google, Yahoo, Flickr, Ebay, Youtube, Myspace, Wikipedia, Mozilla are only a few popular projects that live in this one space*
- here we go
  - *a lot of open questions that wait for ideas, technical solutions, standards or a social agreement*
  - *in parallel to the work on millions of Web sites, Blogs, Wikis, shops and fancy applications there is ongoing work on better standards and technologies*
  - *W3C provides a framework where vendors, authors and users cooperate to improve the supporting frame of the Web*

## W3C

- W3C <http://www.w3.org/> the home of (X)HTML, XML, CSS, RDF, the Web and Semantic Web ...
- ~400 members (the usual suspects, SMEs, users, grassroots ...) <http://www.w3.org/Consortium/Member/List>
- **invited experts and volunteers**
- 65 groups doing the work <http://www.w3.org/Consortium/activities>
- 17 world offices all over the world <http://www.w3.org/Consortium/Offices/>
- a team of ~60 individuals - working and living distributed around the globe - coordinated by 3 hosts
  - MIT, US <http://www.csail.mit.edu/>
  - ERCIM, Europe <http://www.ercim.org/>
  - Keio University, Japan <http://www.keio.ac.jp/>
- director: WWW inventor Tim Berners Lee <http://www.w3.org/People/Berners-Lee/>



---

## Making Standards at W3C

- a very open and transparent process to form consensus <http://www.w3.org/Consortium/Process/>
- responsive to the public
- a transparent patent policy <http://www.w3.org/2004/02/05-patentsummary.html> that protects IPRs and promotes proliferation of standards
- standards are available for free (might be essential for their success!) - find all of them here: <http://www.w3.org/TR/> (and use and implement!)
- any other info please read: About W3C <http://www.w3.org/Consortium/>

# A bit of History



# WORLD WIDE WEB CONSORTIUM Tenth Anniversary

## Pre-W3C Web and Internet Background

- 1945: Vannevar Bush article in Atlantic Monthly describes a photo-electrical-mechanical device called a Memex, for memory extension, which could make and follow links between documents on microfiche.
- 1960: J.C.R. Licklider publishes "Man-Computer Symbiosis."
- 1962: Douglas Englebart publishes "Augmenting Human Intellect: A Conceptual Framework."
- 1965: Ted Nelson coins the term "Hypertext" in "A File Structure for the Complex, the Changing, and the Indeterminate." 20th National Conference, New York, Association for Computing Machinery.
- 1968: Douglas Englebart demonstrates Online System (NLS).

- 1980: While consulting for CERN, Tim Berners-Lee writes a program, "Enquire-Within-Upon-Everything", which allows links to be made between arbitrary nodes.

- May: Tim Berners-Lee publishes version 2 of "Information Management: A proposal."
- End 1990: Development begins for first browser (called "WorldWideWeb"), editor, server, and line-mode browser. Culminates in first Web client-server communication over Internet in December 1990.

- Dec: First Web server outside of Europe set up at Stanford University.

- Dec: Hypertext '91 Conference in San Antonio, Texas (USA). Tim Berners-Lee paper on Web only accepted as poster session.

- 1984: Paul Mockapetris introduces Domain Name System (DNS).

- Mar: Tim Berners-Lee circulates "Information Management: A Proposal" for comments at CERN.

- 1969: Advanced Research Projects Agency commissions ARPANET to conduct research on networking.
- 1971: Ray Tomlinson of BBN creates email program to send messages across a distributed network.
- 1972: Tomlinson expands program to ARPANET users, using the "@" sign as part of the address.
- 1974: Vint Cerf and Bob Kahn publish, "A Protocol for Packet Network Interconnection", which specifies in detail the design of a Transmission Control Protocol (TCP).
- 1978: Part of TCP published separately as the Internet Protocol (IP).

## W3C

- Mark Andreessen and colleagues leave NCSA to form Mosaic Communications Corp., which later became Netscape.
- Traditional dial-up systems (CompuServe, AOL, Prodigy) begin to provide Internet access.

- Feb: Tim Berners-Lee meets Michael Dertouzos in Zurich to discuss possibility of starting new organization at MIT.
- Apr: Alan Kotok, then at DEC, visits CERN to discuss creation of Consortium.
- 1 Oct: W3C created.

- Apr: INRIA becomes W3C Host in Europe.
- Jun: W3C holds first Workshop, on Content Rating; leads to PICS.

pre 1979

pre 1989

1989

1990

1991

1992

1993

- Jan: Number of browsers increases; includes Mida, Erwise, Viola, and Samba.

- Mar: NCSA releases first alpha version of Mosaic for X Windows.

- Apr: CERN agrees to allow anyone to use Web protocol and code royalty free.

- Jun: Dale Dougherty of O'Reilly hosts WWW Wizards Workshop in Cambridge, Massachusetts, USA.

- Nov: At a Newcastle, U.K. conference, Tim Berners-Lee discusses the future of the Web with MIT's David Gifford, who suggests that Tim contact Michael Dertouzos.

1994

1994

1995

## Starting year of current W3C Activities

- Graphics
- Hypertext Markup Language (HTML)
- Style

## Web conference

- ★ Geneva (Jun)
- ★ Chicago (Oct)

- ★ Darmstadt (Apr)
- ★ Boston (Dec)



## Web servers



## Big picture

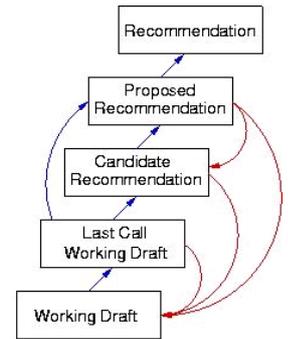
---

## Some simple rules - great impact

- everything is on the web
  - *if its not on the Web it does not exist!*
  - *paperless*
  - *history - archiving and versioning by CVS!*
  - *proof of concept*
  - *gives us feedback (about use)*
- eat your own dogfood!
  - *use standards for your work wherever you can*
  - *be your own usecase*
  - *store your ideas on the server - there you can refer to them*
- conform to standards - no cheap excuses please!
  - *credibility*
  - *no tools???*
- document your communication - use mailinglists and logged IRC
- be open to new applications of standards
  - *RSS feeds*
  - *WIKIs*
  - *Blog!*
  - *use Semantic Web applications*
  - *...*
- go public - listen to the public - response to the public
  - *its not necessary to be a member to submit ideas (though it helps)*
  - *member or not - you are invited to contribute*

# Outline of the process

- Finding new Ideas
  - make a (members) Submission <http://www.w3.org/Submission/>
  - organize/go to a Workshop <http://www.w3.org/2003/08/Workshops/>
  - set-up an Incubator Group to develop ideas <http://www.w3.org/2005/Incubator/>
- Getting things started
  - team drafts a charter for a group <http://lists.w3.org/Archives/Public/public-new-work/>
  - identifying team-contact and chair
  - get the "go" from the membership
- working along
  - publish working drafts from the working group
  - discuss with members and the public
  - promote to (candidate/proposed) recommendation (aka standard)



## Some of our tools

- work is done in meetings (either face to face or - mostly - on phone) working group tools:
  - <http://www.w3.org/2004/12/wg-tools>
  - logged IRC (RFC1459 <http://www.ietf.org/rfc/rfc1459.txt>) for minuting
  - Semantic Web Based Tool "Zakim" meeting and phonebridge management <http://www.w3.org/2001/12/zakim-irc-bot.html>
  - an "RRSAgent" agent to draft meeting minutes <http://www.w3.org/2002/03/RRSAgent>
  - Action tracking, Issue tracking
- Outside meetings working groups use
  - public and members only mailing list (100s) <http://lists.w3.org/Archives/Public/>
  - Wikis <http://en.wikipedia.org/wiki/Wiki>
  - Blogs <http://en.wikipedia.org/wiki/Blog>
  - IRC (RFC1459 <http://www.ietf.org/rfc/rfc1459.txt>) for instant messaging
  - action tracking <http://www.w3.org/2005/06/tracker/>, issue tracking ... <http://esw.w3.org/topic/TrackingIssues>
  - WBS: Web-Based Straw-poll and balloting system <http://www.w3.org/2002/09/wbs/1/>
  - ...
- publishing about the results disseminating the results
  - "pubrule checker" <http://www.w3.org/2005/07/pubrules>
- Quality - life after rec
  - validators <http://www.w3.org/QA/Tools/#validators>
  - test suites z.B. <http://www.w3.org/Graphics/SVG/Test/20061213/> oder <http://www.w3.org/Style/CSS/#browsers>
  - tutorials <http://www.w3.org/2002/03/tutorials>
  - lists of implementations (see respective working group or activity pages)
  - translations database (SW based) <http://www.w3.org/Consortium/Translation/>
  - talks database (SW based) <http://www.w3.org/Talks/>
  - ...

---

## Summary

- could not mention all
  - *maybe I forgot some important*
  - *e.g. Slidy <http://www.w3.org/Talks/Tools/Slidy/> the tool that I use for my slides*
- technology makes life easier
- we reach the community (4Mio hits/day on <http://www.w3.org/>)
- people can access
  - *fast*
  - *search*
  - *lots of additional information*
- **because everthing is on the Web**
- **many tools are available open source**

---

## See more

- Most links are embedded in the slides of this talk: <http://www.w3.org/2009/Talks/0702Brisbane-KB/>
- Article on Web technologies: <http://www.ict-media.de/webtech.html>
- About W3C: <http://www.w3.org/Consortium/>
- All W3C technical reports (including recommendations aka standards): <http://www.w3.org/TR/>
- W3C Process: <http://www.w3.org/Consortium/Process/>
- Talks database: <http://www.w3.org/2007/11/Talks/search/>
- Translations database: <http://www.w3.org/Consortium/Translation/>
- W3C Home page (leading to much more stuff as press, mailing list, people, groups, news, Offices ...): <http://www.w3.org/>