Web 2.0
Lecture 3: Concepts, Statistics and Technologies

doc. Ing. Tomáš Vitvar, Ph.D.
tomas@vitvar.com • @TomasVitvar • http://vitvar.com

Czech Technical University in Prague
Faculty of Information Technologies • Software and Web Engineering • http://vitvar.com/courses/w20

Modified: Wed Feb 13 2013, 14:37:50
Humla v0.3

Overview

- Web 1.0 Timeline
- Web 2.0 Concepts

Web Basics

- Tim Berners Lee's vision:
  - ***Universal linking:*** *making a connection with any content/resource on the Web*
  - ***Separation of concerns:*** *layers divided by a standard interface, innovation on any layer may happen independently*
  - ***Web contract:*** *Web content to be independent from software that consumes it; standards-based, royalty-free technology accessible and usable by anyone, for free*
- Core technologies (global hypertext system)
  - *HTTP – protocol for client-server communication*
  - *URI – schema to identify Web's resources*
  - *HTML – language to format content on the Web*

Timeline

- **1980:** the very beginning – Tim Berners Lee
  - *ENQUIRY, a database of people and a hypertext predecesor*
  - *TBL sees problems in information presentation*
    → *how to share data among physicists with no common tools and presentation software*
- **1990:** all tools necessary for a working Web
  - *Hypertext Transfer Protocol (HTTP) 0.9*
  - *Hypertext Markup Language (HTML)*
  - *The first Web browser*
  - *The first Web server (called CERN httpd)*
  - *The first Web pages that describe the project itself*
  - *The first Web application: CERN telephone directory*
- **August 6, 1991:** Web is a publicly available service

Timeline

- **1995–1999:** Browser wars and dot-com buble
  - *Competition of browsers, custom tags, flavors of HTML*
  - *The "blink" and "marquee" era*

```
<marquee behavior="scroll" direction="left">
   marquee is a non-standard element, <blink>do not use it!</blink>
</marquee>
```

lard element, do not use it!

  - *Difficult to maintain interoperability across browsers*
  - *Web sites often sticked to IE only*
  - *dot-com companies popping up*
    - → *stock prices go up with "e-" prefix*
    - → *increasing stock prices, confidence in profit*
    - → *available venture capital*

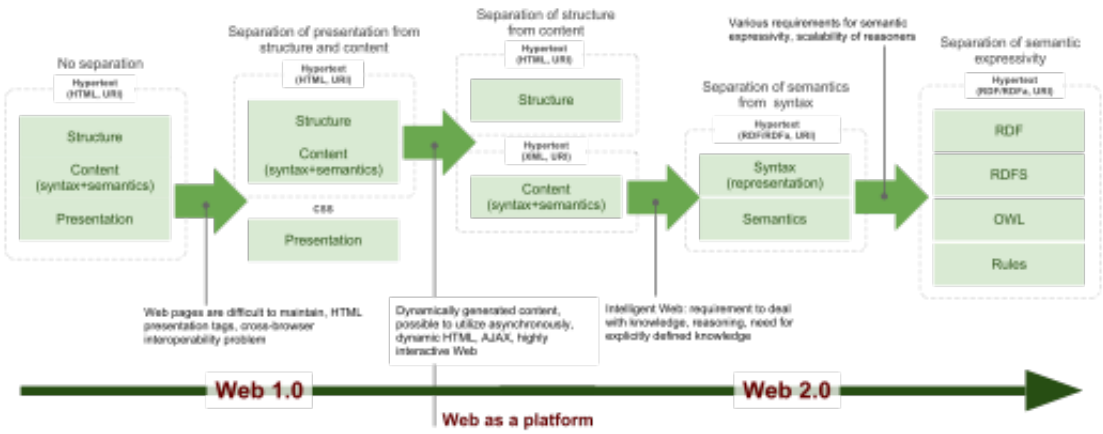
Timeline

- **1997:** HTML 3.2
  - *documenting existing authoring practices at that time*
  - *followed by HTML 4 as a baseline markup language for the Web*
- **1997–1999:** Cascading Style Sheets
  - *Separation of style rules from Web page content*
  - *Reduction of new HTML tags for a new visual effect*
    - → *presentational structures created with CSS + `<span>` and `<div>` tags*
  - *Behavior of constructs in Web pages can be customized via JavaScript and HTML Document Object Model (DOM)*
  - *Browser wars comes to the end with HTML 4*
  - *New highly interactive Web*

Timeline

- **2000-2002:** move from Web content to Web apps
  - *HTML as a simple markup language evolved to:*
    - → ***HTML 4*** *– tags and attrs documents structures*
    - → ***CSS*** *– style rules for document presentation*
    - → ***DOM*** *– API to manipulate HTML in* ***JavaScript***
  - *XMLHttpRequest (XHR)*
    - → *Asynchronous JavaScript and XML (AJAX)*
  - *November 2000: eBay launches its API*
    - → *beginning of Web applications for third-party app integration*
- **2002:** beginning of Web 2.0 era
  - *After dot-com bubble burst (Fall 2001)*
  - *Web 2.0 coined by Web 2.0 conference (Tim O'Reilly)*
    - → *dot-com bubble burst was a turning point*

Web Separation of Concerns



- Evolution rather than design
  - *Information Sharing*
    - → *easy to expose/access information, standard models, semantics*
  - *Web as a platform*
    - → *reusable services, mashups, scalability and performance*

Overview

- Web 1.0 Timeline
- Web 2.0 Concepts
  - *Read-write Web*
  - *Programmable Web*
  - *Intelligent Web*
  - *Social Web*

What is Web 2.0

Web 2.0 is the same as Web 1.0, but there was a broken Web 1.5 in between.
—Tim O'Reilly

Web 1.0		Web 2.0
Britanica online	→	Wikipedia
personal websites	→	blogging
screen scraping	→	web services
content management systems	→	wikis
directories (taxonomy)	→	tagging (folksonomy)
domain name spaculation	→	search engine optimization
mailing list	→	Facebook/Twitter
desktop software	→	software as a service
initial investments	→	pay per use

Examples of Web 1.0 vs. Web 2.0 concepts and services (based on)

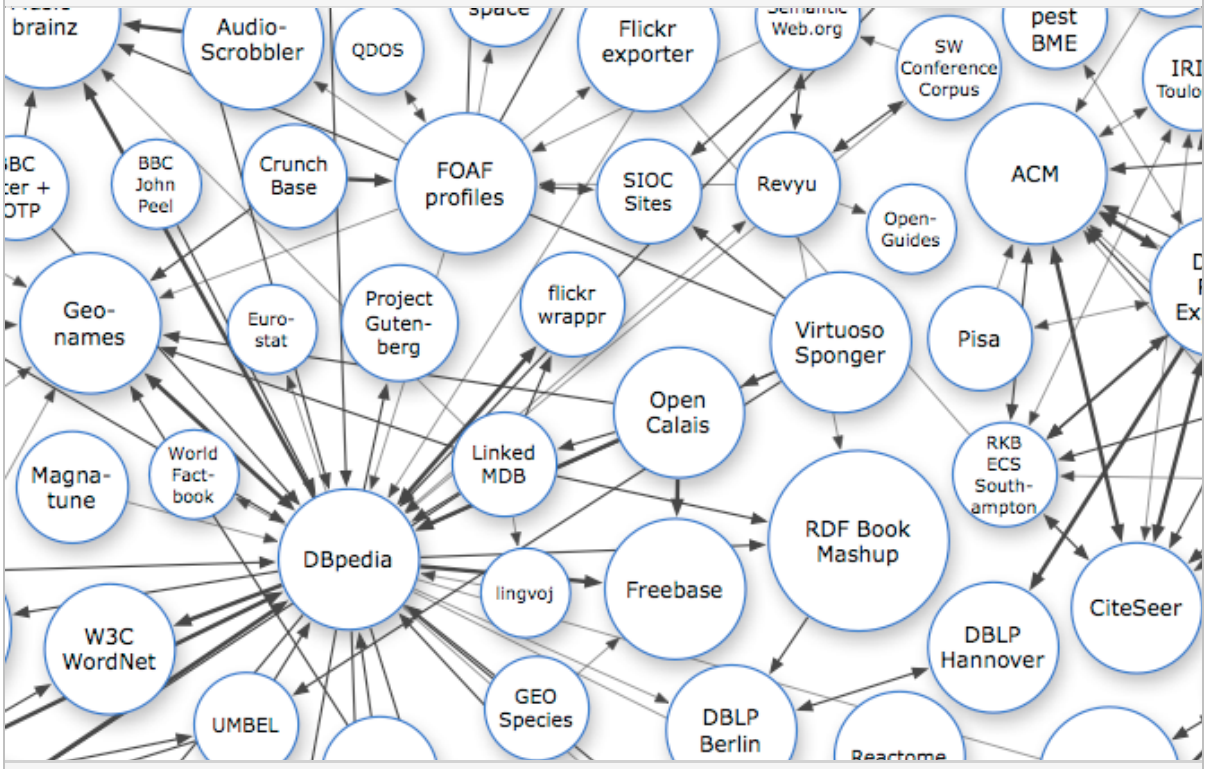
Overview

- Web 1.0 Timeline
- Web 2.0 Concepts
  - *Read-write Web*
  - *Programmable Web*
  - *Intelligent Web*
  - *Social Web*

Read-write Web

- Initial idea of TBL
- Consumers become producers: **prosumers**
  - *participation, contributions by users*
  - *users add value, and they are often the value*
    → *Twitter with 150M users has a value of ~$10B*
- Best example: **Wikipedia**
  - *and many others: Flickr, del.icio.us, Digg, etc.*
  - *but also personal blogs*
  - *as well as Amazon (satisfaction, ranks of books)*
- Users/their knowledge – basis for Web intelligence
  - *Collective intelligence*
    → *mining and utilization of contributed data*

Linked Data

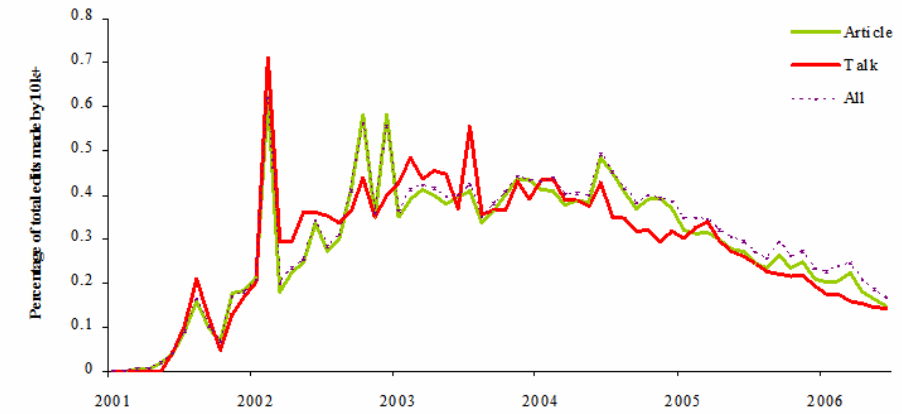


Read-write Web Technologies

- Wiki
  - *Easy, versioned content management*
  - *simplified markup for non-IT people*
- Tagging, folksonomies
  - *Easy creation of metadata describing content*
- Syndication, Atom/RSS, AtomPub, trackback
  - *enabling streams of thought, dialogues*
- New License Models
  - *Creative Commons*

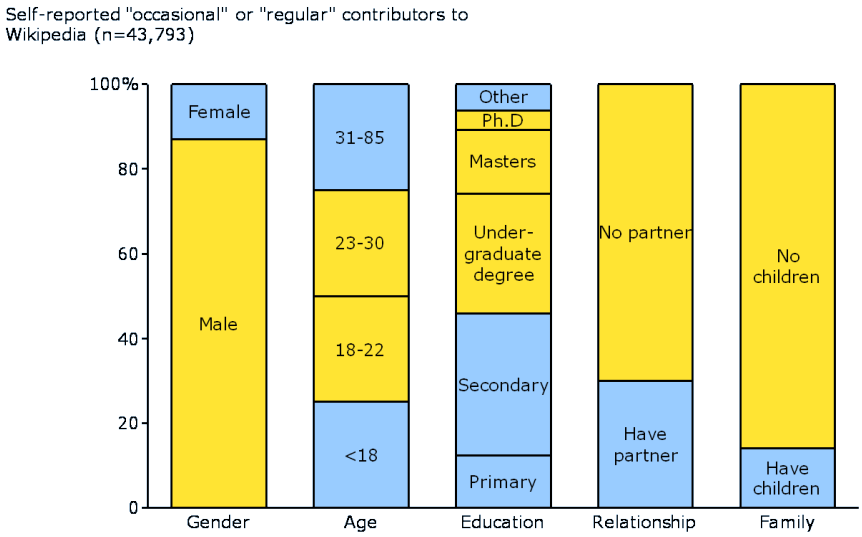
How Many Producers

- Wikipedia
  - *~1% of Wikipedia users were responsible for 50% of Wikipedia edits (2003/2004)*



Participation of users with 10k+ edits, source Long Tail of User Participation in Wikipedia, Augmented Social Cognition Research Blog from PARC.

Demography of Wikipedia Editors



Note: Data for age category also includes respondents who were not contributors but who did read Wikipedia. Average age for contributors is 26.8 (vs. 25.3 for readers). "Regular" contributors include authors, editors, and administrators. "Occasional" contributors include readers who occasionally contribute as authors or editors.
Source: "Wikipedia Survey – First Results," UNU-MERIT, April 2009

Demography of Wikipedia Editors, source: Wikimedia Strategic Planning

Overview

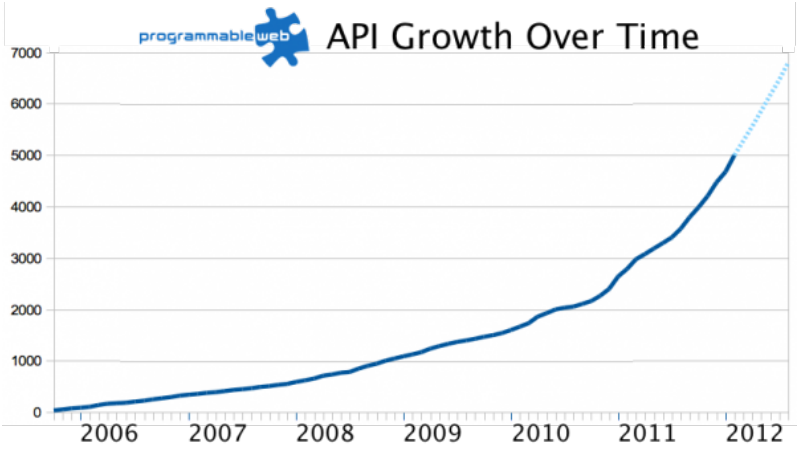
- Web 1.0 Timeline
- Web 2.0 Concepts
  - *Read-write Web*
  - *Programmable Web*
  - *Intelligent Web*
  - *Social Web*

Programmable Web

- Web as a platform
- APIs replace screen scraping
  - *Access to information and **functionality***
  - *RESTful web services, HTTP as a basis*
- Data in structured and standardized form
  - *XML and JSON for data exchange*
  - *RSS/Atom Feeds*
  - *Microformats*
  - *Domain-specific formats, Linked Data*
- Mashups, Aggregation
  - *Yahoo! pipes, Google AppsScript*
  - *AJAX (XMLHttpRequest), JSONP*

How Many Web APIs?

- Web APIs Growth
  - *8 years to get to 1,000 APIs in 2008*
  - *18 months to get to 2,500 APIs in 2010*
  - *18 months to get to **5,000 APIs in 2012**.*



API Growth over Time, source: ProgrammableWeb.com

API Billionaires Club

- Web API Traffic
  - *New highly dependable back-end technologies to serve increasing traffic*

twitter	13 billion API calls / day *(May 2011)*
Google	5 billion API calls / day *(April 2010)*
facebook	5 billion API calls / day *(October 2009)*
NETFLIX	10 billion API calls / month *(January 2011)*
ebay	8 billion API calls / month *(Q3 2009)*
bing	3 billion API calls / month *(March 2009)*
npr	1.1 billion API-delivered stories / month *(March 2010)*
salesforce.com	Over 50% of all traffic via API *(March 2008)*
amazon web services	Over 260 billion objects stored in S3 *(January 2011)*

API Billionaires Club, source: ProgrammableWeb.com

Overview

- Web 1.0 Timeline
- Web 2.0 Concepts
  - *Read-write Web*
  - *Programmable Web*
  - *Intelligent Web*
  - *Social Web*

Intelligent Web

- Semantic Web
  - *Separation of semantics from syntax*
  - *Variety of levels for semantic expressivity*
  - *Reasoning support*
- Collective Intelligence
  - *Utilization of human-knowledge for benefits of users*
  - *Ranking of products/services, collaborative filtering*
- Human Intelligence
  - *Utilization of crowds to do computation or perform tasks*
  - *Amazon Mechanical Turk, reCAPTCHA*

Overview

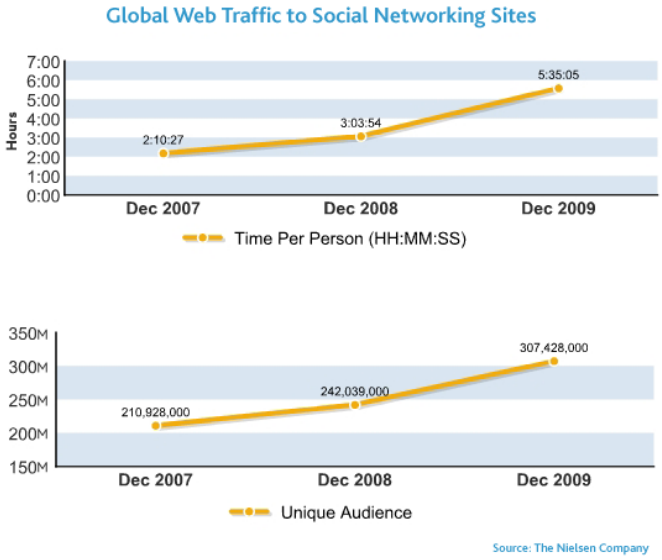
- Web 1.0 Timeline
- Web 2.0 Concepts
  - *Read-write Web*
  - *Programmable Web*
  - *Intelligent Web*
  - *Social Web*

Social Web

- Web Science
  - *need to study the Web in a wider context*
  - *inter-disciplinary themes involving sociology, economics, and computer science fields*
  - *Social web in the core of Web science*
- Data formats
  - *OpenGraph protocol, FOAF, SIOC*
- Social Network Analysis
  - *Structure*
  - *Behavior*
  - *Visualization*

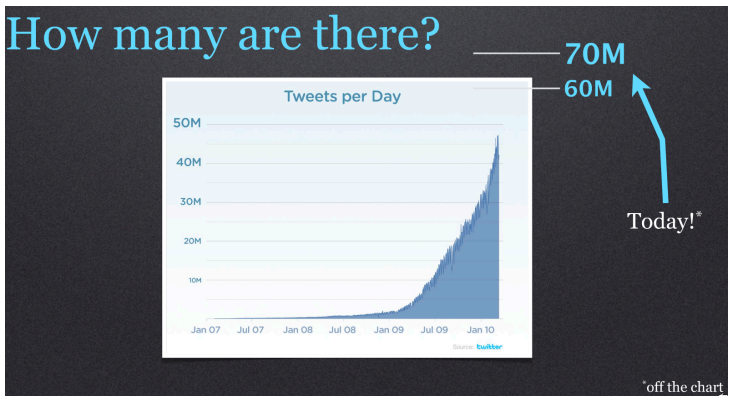
Social Web Stats: Time

- Time spent by people
  - *time spent by people increases by 80% every year*
  - *700B minutes per month on Facebook*



Social Web Stats: Data

- Data and User Base Sizes
  - *300M users on Twitter, 800M on Facebook (as of 2011)*
  - *Twitter: 300M tweets per day, 50GB of tweets alone*
  - *75% of all Twitter traffic through API!*
  - *Twitter serves ~800 tweets per second*



Source: Twitter By Numbers, Raffi Krikorian's talk at UC Berkley, September, 2010.