

# 4 Communities, the Web and Multimedia

- 4.1 Evolution of the Web
- 4.2 Social Networks and Social Media
- 4.3 Web Content Aggregation and Integration
- 4.4 Virtual Worlds in the Web

## Literature:

- T. Berners-Lee: Weaving the Web. Texere 2000
- T. O'Reilly: What is Web 2.0: Design Patterns and Business Models for the Next Generation of Software.  
*Communications & Strategies*, No. 1, p. 17, 2007.  
(<http://oreilly.com/web2/archive/what-is-web-20.html> – 2005)
- T. O'Reilly, J. Battelle: Web Squared: Web 2.0 Five Years On  
([www.web2summit.com/websquared](http://www.web2summit.com/websquared) – 2009)

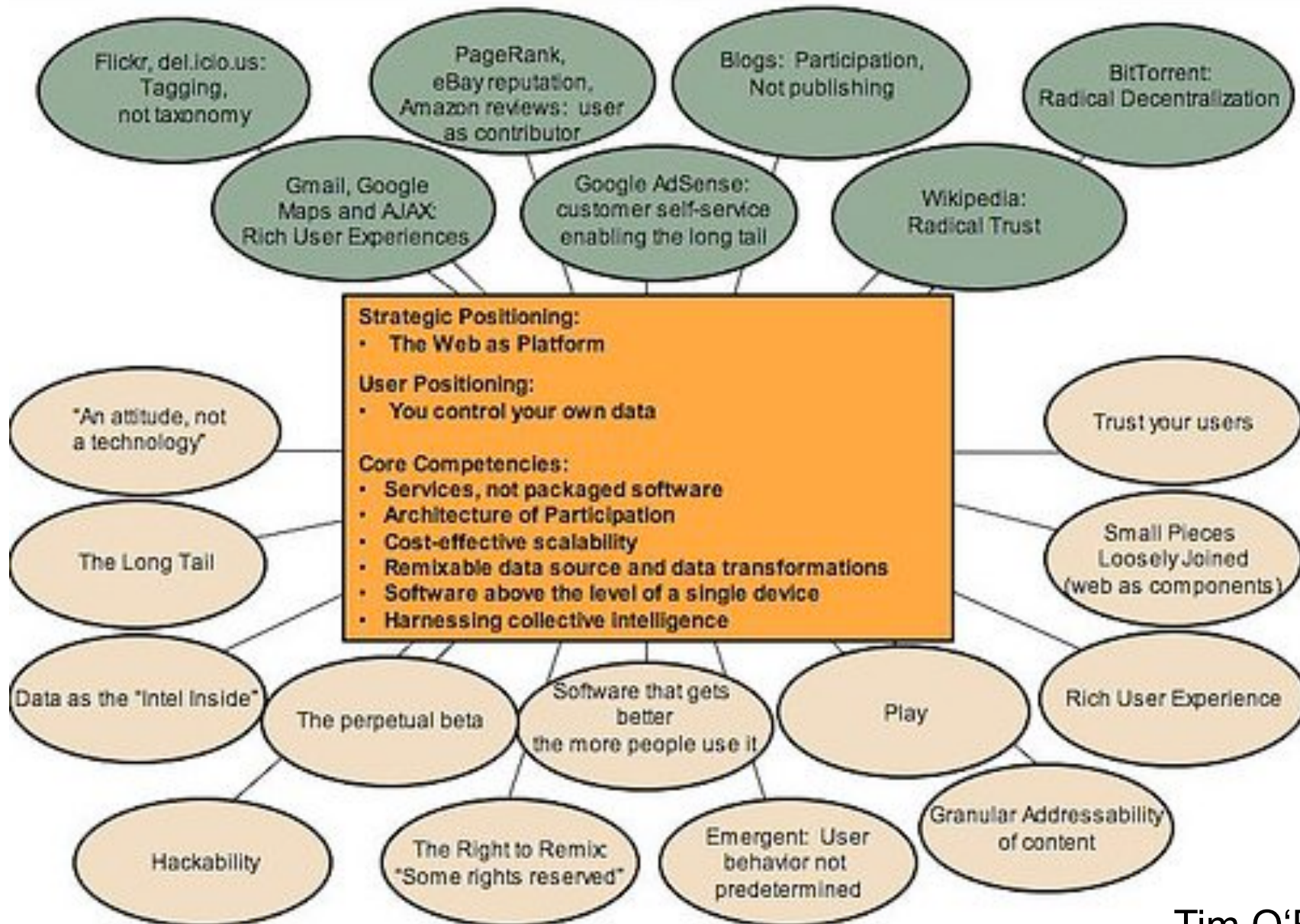
# The Key Idea of the Web

- T. Berners-Lee: “Information Management: A Proposal”.  
CERN March 1989  
“CERN is a model in miniature of the rest of the world in a few years time.”
- Disussions on Mosaic browser, 1993:  
“I ... made my now-standard case for making the Mosaic browser an editor, too. Marc [Andreessen] and Eric [Bina] explained that they had looked at that option and concluded that it was just impossible.”  
(Weaving the Web p. 77)
- Tim Berners-Lee on the future of the Web:  
“My hope and faith that we are headed somewhere stem in part from the repeatedly proven observation that people seem to be naturally built to interact with others as part of a greater system.”  
(Weaving the Web p. 223)

# Generations of the Web

- Web 0.5
  - 1988-1995
  - Only predecessors of WWW exist
- Web 1.0
  - 1996
  - Static HTML pages, few publishers - many readers
- Web 1.5
  - 1996-2001
  - Dynamic Web pages, E-Commerce
- Web 2.0
  - 2005?
  - Collaboration, communities
  - Openness, standardization, liberty

## Web 2.0 Meme Map



Tim O'Reilly

# What is the Meaning of „Web 2.0“ ?

- O'Reilly Conferences 2004/2005/2006
  - Tim O'Reilly, Dale Dougherty
  - Similar conference titles exist, e.g.: „Where 2.0“ (geospatial web)
- Basic question for the conference:
  - Which ideas have survived the burst of the dot-com bubble?
  - Creating the next wave out of the remains of the last
- „Web 2.0“ has become a „buzzword“
  - Extremely rapidly...
  - There is no agreed definition
- The conference lives on
  - Web 2.0 Summit (<http://www.web2summit.com>)



# Two Aspects of Web 2.0

- Social Aspect
  - Collaboration
  - User-Generated Content
- Technical Aspect
  - Huge bandwidth, therefore graphics, audio, pictures, videos...
  - Web browser as a universal platform for application software

# Comparison by Examples

## Web 1.0

- DoubleClick
- Ofoto
- Akamai
- mp3.com
- Britannica Online
- Personal homepage
- Personal bookmarks
- Content mgmt
- Taxonomy

## Web 2.0

- Google AdSense
- Flickr
- BitTorrent
- Napster
- Wikipedia
- Blogging
- Del.icio.us
- Wikis
- Folksonomy

# Web 2.0 Principles

- Web 2.0 is about harnessing collective intelligence!
- Reach out to the entire Web, to the edges and not just the center, to the long tail and not just the head.
- The service automatically gets better the more people use it.
- Network effects from user contributions are the key to market domination in the Web 2.0 era.
- Web 2.0 companies build value as a side-effect of the ordinary use of their application.
- The race is on to own certain classes of core data (e.g. location, identity, calendaring, product identifiers)
- Consequences for software development:
  - Software will cease to perform unless it is maintained on a daily basis
  - Users must be treated as co-developers
  - Lightweight programming models and loose coupling are needed
  - Design for remixability

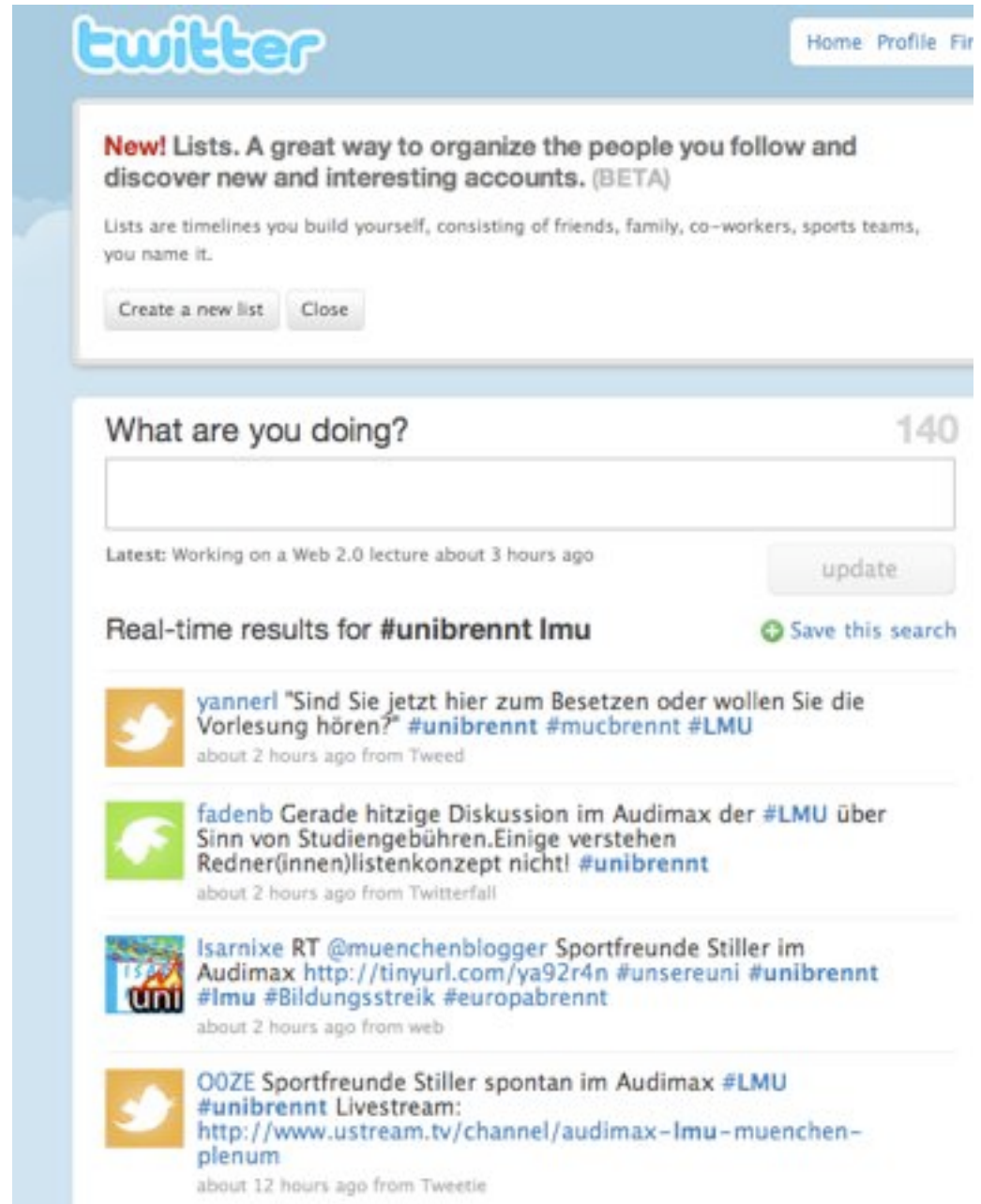


# Web 2.0 Five Years On: Web Squared

- Smartphones and other recent/upcoming developments:
  - Collective intelligence no longer being driven solely by humans but, increasingly, by sensors.
  - Talking to the Web becomes a reality (Google Mobile App on iPhone)
  - Information shadows, Internet of Things: Web meets World
  - Automatic geo-tagging of pictures by GPS built into cameras
  - Face recognition built into photo archiving software
  - Object recognition via smartphone camera
  - Infinite Images (Adobe MAX 2008, Shai Avidan, see <http://www.youtube.com/watch?v=QxNx2OyeCHA>)
- A key competency of the Web 2.0 era is discovering implied metadata, and then building a database to capture that metadata and/or foster an ecosystem around it.
- Systematic identities/primary keys are being replaced by clever recognition mechanisms (cf. CDDB/Gracenote signature of CD)

# Real Time Media

- Current main example: Twitter
  - Search, analytics and social networks built around real-time communication
  - Triggering of innovations:
    - » Example:  
Shorthand URLs
- ***Everything in the world is now real time.***
  - Technically, it is no problem notifying the manufacturer *immediately* if a certain type of shoe is not selling at a certain shop.
  - "Houdini" system used by Obama campaigners



# 4 Communities, the Web and Multimedia

4.1 Evolution of the Web

4.2 Social Networks and Social Media

4.3 Web Content Aggregation and Integration

4.4 Virtual Worlds in the Web

Literature:

Eigner/Leitner/Nausner/Schneider: Online-Communities, Weblogs  
und die soziale Rückeroberung des Netzes,  
Nausner & Nausner 2003

C. Stöcker: Die Zeit der Kopfjäger,  
Spiegel-online.de, 1. August 2006

# Social Networks, First Generation



Hauptmenü

Anmeldung

Log in

- College traditions:
  - Class listings, alumni listings, freshman listings
  - Personal profiles
  - Keeping in touch (classmate reunions)
- 1995: Classmates.com
  - German version: stayfriends.de
- Many similar platforms
  - E.g. facebook.com

Leistungen	Basis	Gold
Klassenfoto einstellen und Mitschüler markieren	✓	✓
Aktives Auffinden selbst festlegen	✓	✓
Klassentreffen melden	✓	✓
Alle Klassentreffeninformationen einsehen		✓
Mit dem Organisator des Klassentreffens in Kontakt treten		✓
Mailingliste für Klassentreffen verwalten		✓
»Eigene Seite« anlegen und Profil erstellen	✓	✓
Fotos auf »Eigene Seite« hochladen	✓	✓
»Eigene Seite« von anderen komplett einsehen		✓
Fotos ansehen		✓
Nachrichten hinterlegen	✓	✓
Nachrichten lesen und beantworten		✓
Kontaktliste anlegen	✓	✓
Ein Poesiealbum führen	✓	✓
Jedes Poesiealbum lesen		✓

# Social Networks, Second Generation

- From documentation of existing relationships to creation of new relationships
- General platform for self presentation
  - Easy way to personal homepage
- Examples:
  - MySpace.com
  - Friendster.com
  - StudiVZ.de
  - Xing.com (OpenBC)
- Establishment of “friend” link by mutual agreement
- Tracing of social network
  - 2nd degree contacts
  - Former colleagues



Holtzbrinck zahlt Millionenpreis für StudiVZ  
03. Jan 2007 17:52, ergänzt 18:54



StudiVZ-Website  
Foto: StudiVZ

Quelle: netzeitung.de

Der Verlagskonzern Holtzbrinck hat das Internetportal StudiVZ vollständig übernommen. Mehr als 50 Millionen Euro zahlte das Unternehmen für die Kontaktbörse.

# Example: facebook.com (1)

- History:
  - Mark Zuckerberg and friends, Harvard, October 2003:
    - » Facemash: Comparing student photos
  - Mark Zuckerberg, February 2004: "The Facebook" for Harvard students
  - Stepwise expansion to other universities, colleges and high schools
  - Sept, 2006: Open to everybody of age 13 and up
- Popularity:
  - Leading social networking site (2009, according to *comScore*)
  - 500 million active users (July 2010)
  - Has overtaken MySpace.com in popularity by April 2008
  - Alexa.com traffic rank:
    - » Sept. 2006: Traffic rank 60
    - » Nov. 2010: Traffic rank 2 (rank 2 in Germany, rank 1 in 6 countries)
  - Still high user growth (getting slower)



Country	Percent of Site Traffic
United States	26.3%
India	6.5%
Germany	4.3%
France	4.0%
United Kingdom	3.9%
Italy	3.8%
Mexico	2.9%
Indonesia	2.7%
Spain	2.4%
Canada	2.4%

alexa.com

Total active users<sup>[N 1]</sup> (in millions)

Date ▲	Users ▶	Days later ▶	Monthly growth <sup>[N 2]</sup> ▶
August 26, 2008	100 <sup>[36]</sup>	1,665	178.38%
April 8, 2009	200 <sup>[37]</sup>	225	13.33%
September 15, 2009	300 <sup>[38]</sup>	150	10%
February 5, 2010	400 <sup>[39]</sup>	143	6.99%
July 21, 2010	500 <sup>[6]</sup>	166	4.52%
—	600	118 (ongoing)	—

Wikipedia

# Example: facebook.com (2)

- Features:
  - Wall, Photos, Pokes, Status, Newsfeed, Notes (blogging), Gifts, Marketplace, Events, Video, ...
    - » More than 30 billion pieces of content shared each month
- Facebook Platform:
  - May 2007: Software development platform (PHP5, JavaScript, ActionScript, ...)
  - Nov. 2007: 7,000 applications  
Nov. 2010: 550,000 active applications
- Financial side:
  - 2007: Microsoft buys 1.6% share for \$240 million
  - September 2009: First time positive cash flow
  - Value estimate 2010: \$41 billion

## Facebook Value Tops EBay, Becomes No. 3 U.S. Web Company

By Brian Womack - Nov 15, 2010 6:01 AM GMT+0100

<http://www.facebook.com/press/info.php?statistics>  
<http://www.bloomberg.com/news/2010-11-15/>

# Multimedia and Social Networks

- Obvious parts of homepage:
  - Personal photograph
  - Private pictures
  - Background music
- Copyright for all uploaded content owned by the user
  - User-generated (multimedia) content
- Legal trading of public domain music and videos
  - MySpace Music



• Völlig kostenlos  
• Lade MP3s hoch -- nicht komprimieren!  
• Hol dir deinen eigenen Band-URL

**Verbreite deine Musik bei 148 Millionen von Fans!**

Web | MySpace | Leute | Musik | Music Videos | Blogs | Videos | Film ▶

Search powered by Google™

Musiker – HIER **KOSTENLOS** ANMELDEN!



# Social Media

- Media consisting of user-generated (multimedia) content
- Classics: Flickr.com (photo), YouTube.com (video)
- Tagging
  - By originator
  - By others
  - Folksonomy
  - Tag clouds
- Comments
  - Discussion
  - Feedback
- Ratings
- Automation
  - Most recent, most popular

All time most popular tags

amsterdam animal animals april architecture art australia baby barcelona  
beach berlin bird birthday black blackandwhite blue boston bridge building bw  
california cameraphone camping canada car cat cats chicago  
china christmas church city clouds color colorado concert day dc dog dogs england  
europe family festival fireworks florida flower flowers food france  
friends fun garden geotagged germany girl graduation graffiti green hawaii  
holiday home honeymoon house india ireland italy japan july june kids lake  
landscape light london losangeles macro march may me mexico moblog  
mountains museum music nature new newyork newyorkcity newzealand night  
nyc ocean orange oregon paris park party people phone photo pink portrait  
red reflection river roadtrip rock rome sanfrancisco school scotland sea seattle sign  
sky snow spain spring street summer sun sunset taiwan texas thailand  
tokyo toronto travel tree trees trip uk unfound urban usa vacation  
vancouver washington water wedding white winter yellow zoo

Tag cloud from Flickr

# Giving, Exchanging, Buying

- Elementary way of exchanging goods
  - Reciprocal needs
- Buying:
  - Exchange is eased by money
- Media products may be different!
  - Artist has a need for being recognized, getting feedback
  - Consumer has a need for being entertained, informed
  - Reciprocal needs exist
- Amateur content producers
  - Do not in the first place expect revenue
  - “Giving” instead of exchanging
  - See Open Source software
- Global medium is more than its parts
  - Automatic creation of new valuable content by aggregation, filtering
  - “Wisdom of the crowd”

# Social Bookmarks

- Organizing Web content:
  - Hierarchical directories, taxonomy:
    - » Gopher, Yahoo
  - Personal bookmarks
    - » Retrieval problem, metadata
  - Sharing platform for links to information in the Web
  - Classics: Del.icio.us, digg.com, spurl.net, furl.net (with archive)
- Tagging:
  - Adds a semantic dimension to Web search
- Browser extensions for keeping personal bookmarks on server
- Digg effect:
  - Small web sites becoming quickly popular – overload situations
- Overview paper (found via del.icio.us...):
  - <http://www.dlib.org/dlib/april05/hammond/04hammond.html>

# Weblog, Blog

- Definition: A collection of chronologically (backwards) ordered, regularly added contributions to an umbrella topic.  
(adapted from Ebner/Baumann/Krcmar)
  - Traditionally: Contributions comment on one specific hyperlink
- Technical view:
  - Simple content management system
- History:
  - First online diary by Simon Gisler 1994 (according to Wikipedia)
  - John Barger 1997: Term “Weblog”; Term “Blog” since 1999 (Peter Merholz)
  - Huge popularity since 2002
- Platforms: e.g. wordpress.org, blogger.com
  - Links point to individual contribution and are permanent (“permalinks”)
- Variants by media type:
  - Vlog, linklog, photoblog, moblog
- Problematic issues:
  - Borderline between advertisement, propaganda, free speech
  - Law violations, offensive statements

# Reading and Writing

- Traditional cultural techniques: writing and reading
  - Mostly separated activities
- Blogging:
  - Writing as a continuation of reading
  - High motivation to start writing, by external trigger (commented link)
  - Often very short texts
  - Reading as a continuation of writing (e.g. reading comments)
  - Continuous process of sign production
- A new cultural technique?
  - Reading-writing-reading-writing-...

From Eigner et al p. 122

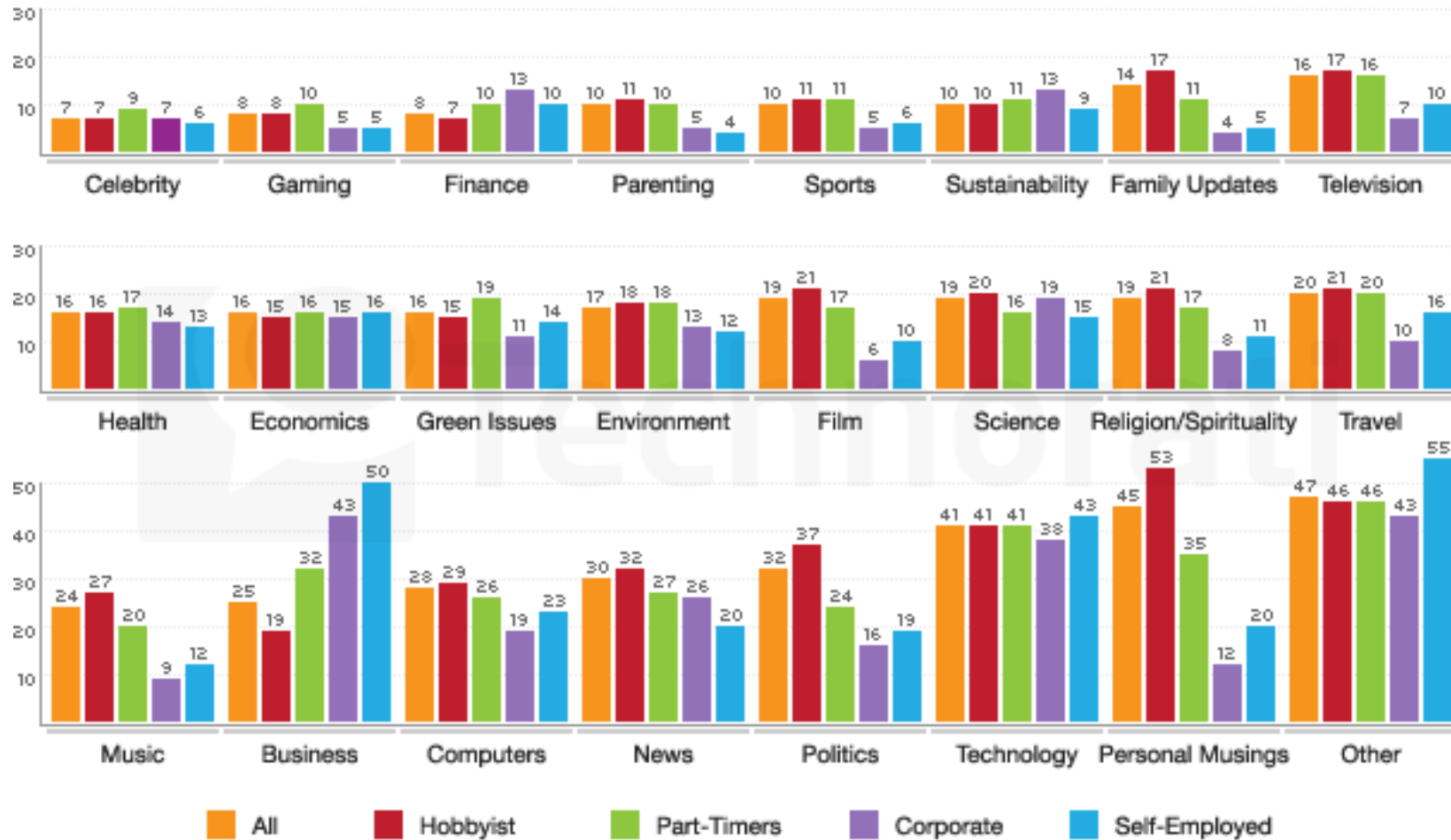
# Blog as an “Oscillation Medium”

- Traditional web sites (including online versions of traditional media):
  - Closed content
  - Links mostly internal to web site
  - Plus a few “related links”
- Bookmark collections:
  - Completely open content
  - Only reference to outer location
- Oscillation media:
  - Both closed and open
  - Blog comments on a link and contains external links
  - Reader is “oscillating” between open and closed reading
    - » Shall I follow the link?
    - » Shall I read on?
  - Hypertextuality as a media creation force

From Eigner et al p. 119

# Diversity of Blog Topics

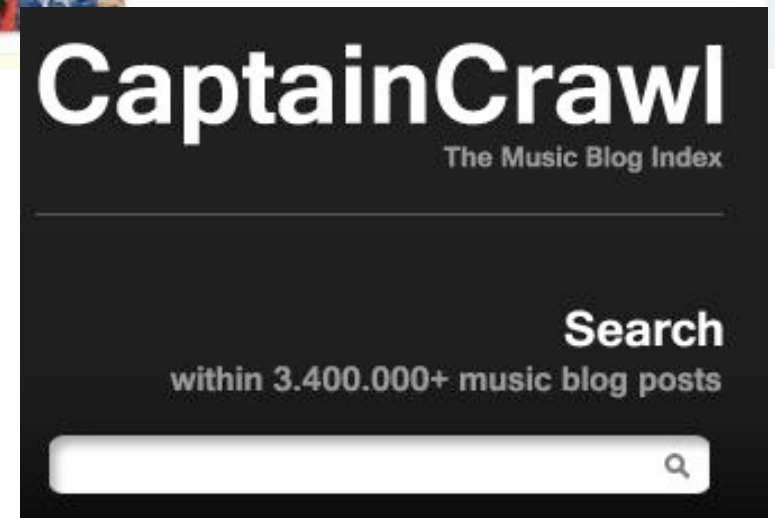
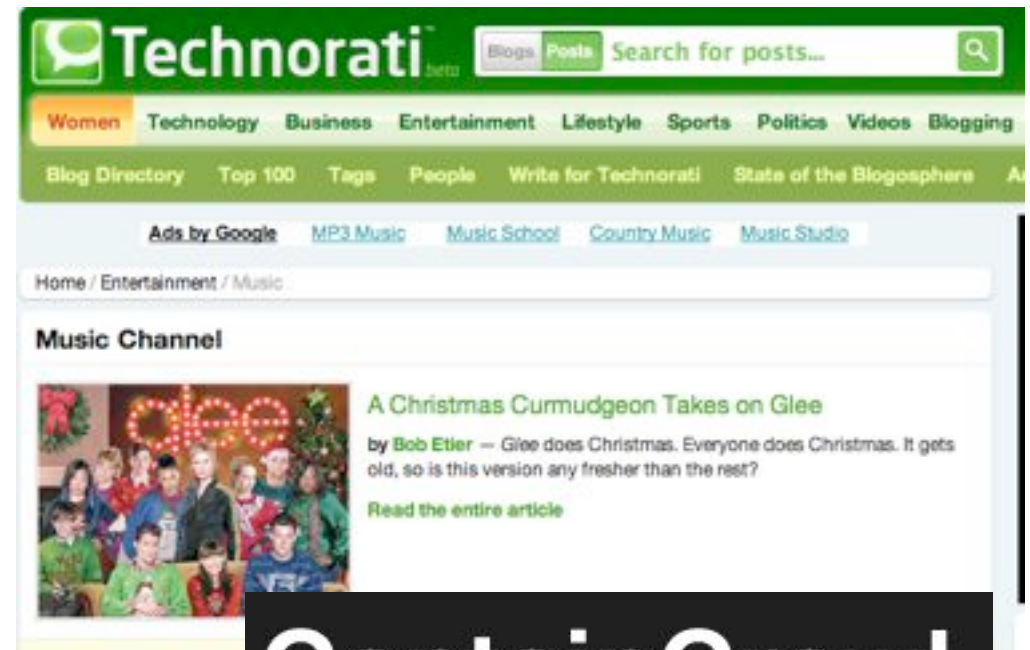
What topics do you blog about?



Technorati.com State of the Blogosphere 2010

# Blog Search

- Blog search engine
  - Combining information from many blogs
  - Including tagging, rating etc.
  - Examples: Technorati.com, blogsearch.google.com
  - Being expanded towards multimedia
    - » Counting links from blogs to music albums, videos, movies etc.



## The Hype Machine follows music blog discussions

Every day, **thousands of people** around the world write about music they love — and it all ends up here. |



# Microblogging

Anarchaia

Experimental, impressionistic sub-paragraph  
tumblin' (think obstsalat)

April 2005: Term "tumblelog"

- Tumblelog:
  - Relatively unstructured "stream of consciousness"
  - Small bits of information and media
- Simplified blogging platforms
  - Tumblr (2006)
  - Twitter (2006/7)
- Microblog:
  - Brief updates (text or small media units), published on the Web
  - Submitted through various means, also from mobile devices
- Microblogs built into social Web platforms
  - "Status Update" on Facebook



tumblr.com

# Podcasting

Suchbegriff > English | Übersicht | Abo | Impressum | Kontakt

Die Bundeskanzlerin

Angela Merkel  
Bundeskanzlerin Angela Merkel

START AKTUELL ANGELA MERKEL KANZLERAMT UNTERWEGS

PODCASTS

Anzahl der Einträge: 166 1-25 | 26-50 | 51-75 | 76-100 | 101-125 >>

	Datum	Titel
	17.10.2009	Investitionen in Kultur sind Investitionen in die Zukunft
	10.10.2009	Buchmesse unter wunderbaren Vorzeichen

- “History”:
  - Discussed since 2000, massive use since 2003
  - iPod & Broadcasting
  - Word of the year of the New Oxford American Dictionary 2005
- Media file distributed by subscription (paid or unpaid)
  - Playback on computers or mobile devices
  - Mainly audio, partly video information, may be any file technically
  - Automation of download by “feeds” (RSS or Atom)
- Often **User-Generated Content (UGC)**:
  - Amateur podcasts
  - Production of audio podcasts has minimal hardware/software requirements

# Corporate Blogs and Podcasting

- Companies use blogs and podcasts for:
  - General information on company
  - Brand formation, general public relations
  - Topic blogs
  - Campaign blogs
  - Knowledge distribution and customer service
  - Internal information channels (intranet blogs)
    - » Executive blog, team blog
- Problematic issues (for the company):
  - Negative image campaigns (e.g. attac)
  - Confidential or problematic issues discussed openly in internal blogs (e.g. cases of mobbing)
  - Danger of emotional escalations

Peter Wolff: Die Macht der Blogs, Datakontext 2006

# Context-Sensitive Advertisement

- Important source of revenue in Web 2.0 sites
  - Advertisement precisely targeted at customer
- Market leader: Google AdSense
  - Ad server operated by Google
  - Websites register with Google
    - » Advertisement placed based on analysis of content of page to be shown (Javascript)
    - » Generate revenue per click or per thousand impressions
  - Selection among relevant ads and order of ads by real-time auction
    - » Ads creating highest revenue are shown
    - » Using bid price of advertised and quality score of the ad (e.g. Click-Through-Rate)
    - » Paid price may be lower than the bid (minimal price to keep position on the list)
  - Advertisers arrange fixed budgets in advance
- See: <http://www.google.com/adwords/displaynetwork/control-your-costs/pricing.html>

# Lead User Innovation Process with Online Communities

- Traditional Lead User Innovation Process:
  - Workshops with selected users
  - Bring together product developers and end users
  - Discuss ideas for new products and product improvements
  - Create new ideas through creativity techniques
- Closed Online Communities:
  - Lead user community organized through web-based platform
  - Specific software tools in addition to community-platform functions
- Example:
  - TMG München projects with Baluff and Webasto
  - Ideas for automotive products created and tested with large group of Chinese participants

# Innovation Community: innocentive



Search Challenges:

Search

• Register • Log In • Help • Contact Us

**INNOCENTIVE**  
Maximize your RETURN ON INNOVATION

[Learn More](#)

Interested in having a problem solved?

Contact Us

- Products
- Seekers
- Solvers
- Challenges
- FAQ

*"The ability to pick and choose your customized solution from a number of applicants ensures you get the best of the best."*  
- Mark Bent, CEO, SunNight Solar

[Solution Seekers](#)

[Problem Solvers](#)

### Featured Challenge

**Novel Approaches to Protecting Maize from Insect Damage**

Challenge Reward: **\$20,000 USD**      Challenge Type: **Theoretical-IP Transfer**      INNOCENTIVE **8836928**

The Seeker is looking for novel approaches to protecting maize from insect damage. This Challenge requires only a written proposal.

# 4 Communities, the Web and Multimedia

4.1 Evolution of the Web

4.2 Social Networks and Social Media

4.3 Web Content Aggregation and Integration

4.4 Virtual Worlds in the Web

Literature:

R. Yee: Pro Web 2.0 Mashups, Remixing Data and Web Services,  
Apress 2008

# Content Aggregation, Indexing, Curation

- Content aggregation:
  - Combination of content on specific topics from various sources
  - Essential steps:  
Finding content (indexing), filtering content, publishing combined content
- Indexing:
  - May address different types of source (web, blogs, local sources)
- Filtering:
  - Automatic filtering
  - Careful manual selection of content: *Curation*
  - Community-based automatic filtering (voting, tagging, counts)
- Publishing
  - User-configurable content (personal news)
  - Syndication to other online media
  - Basis for non-online media (e.g. print media)

---

<http://emediavitals.com/article/1005/how-publishers-curate-world-content>



# Content Indexing + Automatic Filtering: Examples

The screenshot shows the Google News interface. At the top, there is a search bar with the text "Google news" and two buttons: "Search News" and "Search the Web". Below the search bar, there is a link for "Advanced news search". On the left side, there is a "Top Stories" section with a list of topics: TSA, Roy Halladay, Initial public offering, Kate Middleton, Geno Auriemma, Mitch McConnell, Ronni Chasen, Viktor Bout, iTunes, and Cholera. The main content area shows a search for "iTunes". Below the search bar, there is a button "Add to my personalized news page". The main article is titled "The Beatles' Abbey Road leads to top 20 on iTunes" with a star icon. The article is from "Vancouver Sun - 50 minutes ago". The text of the article reads: "The now-settled battle between the Beatles and iTunes isn't the first time the band has fought to protect its profits. LOS ANGELES - The Beatles soared up ...". There are several links below the article: "Video: BEATLES AND iTUNES STRIKE DEAL" (with a YouTube icon), "Beatles storm iTunes charts on first day of sale" (with a Guardian icon), "Techtree.com - NEWS.com.au - Los Angeles Times - Wikipedia: Apple Corps v Apple Computer", "all 3,636 news articles >", and "AAPL" (with a stock icon).

The screenshot shows the banner for the HPANA website. The banner features the text "THE HARRY POTTER AUTOMATIC NEWS AGGREGATOR" at the top. Below this, there is a large image with the word "HPANA" in the center. The image shows several characters from the Harry Potter series, including Sirius Black, Hermione Granger, Harry Potter, and Albus Dumbledore. Below the image, there are social media icons for Google+, MySpace, and RSS. To the right of the icons, there is a link "Translate to: Español · Français · Português · Italiano · Deutsch". At the bottom of the banner, there is a navigation menu with links: "Home", "News Browser", "Community", "Shopping", "Members", and "About". Below the navigation menu, there are links: "Front Page", "HP Fan Trips", and "MugglePost.com".

hpana.com

# News Curation Tools and Services: Examples

Solutions About **LOUD3R** News Contact

## LOUD3R gives publishers and brands powerful content curation tools

**Why do I need LOUD3R?**

In this era of information overload, publishers and marketers are looking for an efficient, low cost way to pull in quality content that keeps their sites fresh, authoritative and engaging. But adding content isn't always scalable. The bill adds up quickly—with extra editorial staff, IT support, licensing, production—making it difficult to compete at web speed. Meanwhile, finding a way to integrate social media isn't always obvious or easy, and consumer expectations for real-time content, updated 24/7, are growing every day.

**LOUD3R helps Publishers & Marketers:**

- 1. Quickly add content areas to sites, apps and campaigns**
  - Build out full topic sites as sections of your existing site
  - Launch completely new standalone sites, apps and social media profiles
  - Create content-rich email newsletter campaigns
- 2. Complement in-house editorial**
  - Add engaging 3rd party content adjacent to your in-house editorial or products
- 3. Leverage social media for real-time publishing**
  - Finds you the highest value tweets for every

**Glaci3r - Stream Manager**

**The LOUD3R Curation Platform**

A powerful, flexible admin tool that helps you and your team curate and publish the great content LOUD3R delivers. Adjust filters so content matches your brand. Automate or control what goes live and how it's featured. Add

loud3r.com

daylife.com

**Publishers: Outpublish your competitors, don't outspend them.**

**daylife**

Quick and easy editorial consoles

Search news & the largest editorial photo archive on the web

Search

Trending: Michael Vick Kate Middleton Charles Rangel Mark Zuckerberg

Learn more Publisher Login

# Video Curation and Aggregation: Example

The screenshot displays the magnify.net website interface. At the top left is the logo "magnify.net" with the tagline "The Realtime Video Curation Engine". To the right, it says "video powered by Akamai" and "Solutions". A navigation menu includes "Support", "F", and "NY Mag Search".

The main content area features a large heading "Content" and the text "Make, Gather, Curate Magnify Multi-Source Curation". Below this, it states "Build a curated collection of high-quality content, fast, brand-safe. And it drives traffic." To the right of this text are three overlapping boxes: "Your Pro Content" (with a camera icon), "User Submitted" (with a smartphone icon), and "Web Curated" (with logos for Dailymotion, YouTube, metacafe, hulu, and blip.tv).

On the right side of the page, there is a "New York VIDEO" logo and a navigation bar with categories: "NEWS & FEATURES", "RESTAURANTS", "BARS", "ENTERTAINMENT", and "FASH". Below this is a "Most Recent Videos" section with two video thumbnails. The first thumbnail is titled "Robin Sparkles: Two Beavers Are Better Than One" and the second is "Late Night: Kathy Griffin Spoofs Justin Bieber Hitting Puberty".

At the bottom of the screenshot, the URL "video.nymag.com" is visible.

# Print Media from Curated Online Content



---

*In May 2010, we conducted a two-day media experiment. 8,000 people signed up, 1,500 submissions came in, 35 editors selected 70 pieces to fill a 60-page magazine.*

*People liked it. We broke our distributor's sales records, received positive reviews in The New York Times, PBS, and the Village Voice, and won a Knight-Batten Award for Innovation in Journalism. Here, we present selected work from the print edition of Issue Zero.*

[48hrmag.com](http://48hrmag.com)

# Content Sources

- Data feeds (XML files)
  - RSS (Really simple syndication)
    - » Channels and items
  - Atom, Atom Syndication Format (ASF)
    - » Successor for RSS
    - » IETF Standard
  - Proprietary file formats
- Database access
  - Often databases specific for application domain (e.g. in a company)
- Public Web Services
  - Access to information provided by large Web sites
  - E.g. Amazon, Google Maps
  - See later for details

# Mashup

- Application integrating diverse Web content seamlessly
- Presentation screen and layout:
  - May be based on existing Web site
  - May be created specifically
- General architectural principle:
  - Web sites provide program access (API) over the Internet (Web Services)
  - Several Web Services are contacted and results are evaluated
- Basic alternatives:
  - Client-side mashup
  - Server-side mashup (more frequent)
- Various technologies for transmission/invocation:
  - REST
  - SOAP
  - XML-RPC

# Example: Alkemis Local

- local.alkemis.com
- NYC live traffic cams
- Yahoo Traffic News
- Flickr images
- Del.icio.us links
- A9 Blockview photos
- APIs:
  - [Amazon A9](#)
  - [OpenSearch](#)
  - [del.icio.us](#)
  - [Flickr](#)
  - [Google Maps](#)
  - [Yahoo Traffic](#)
- See [programmableweb.com](#)



# Web Service APIs

- Example: Flickr API
- Existing methods are grouped in packages
- For each method, allowed parameters are defined
- Often a registration key is required which has to be obtained from Web service provider

## flickr.photos.getInfo

Get information about a photo. The calling user must have permission to view the photo.

### Authentication

This method does not require authentication.

### Arguments

**api\_key** (Required)

Your API application key. [See here](#) for more details.

**photo\_id** (Required)

The id of the photo to get information for.

**secret** (Optional)

The secret for the photo. If the correct secret is passed then permissions checking is skipped. This enables the 'sharing' of individual photos by passing around the id and secret.

## people

- [flickr.people.findByEmail](#)
- [flickr.people.findByUsername](#)
- [flickr.people.getInfo](#)
- [flickr.people.getPublicGroups](#)
- [flickr.people.getPublicPhotos](#)
- [flickr.people.getUploadStatus](#)

## photos

- [flickr.photos.addTags](#)
- [flickr.photos.delete](#)
- [flickr.photos.getAllContexts](#)
- [flickr.photos.getContactsPhotos](#)
- [flickr.photos.getContactsPublicPhotos](#)
- [flickr.photos.getContext](#)
- [flickr.photos.getCounts](#)
- [flickr.photos.getExif](#)
- [flickr.photos.getFavorites](#)
- [flickr.photos.getInfo](#)
- [flickr.photos.getNotInSet](#)
- [flickr.photos.getPerms](#)
- [flickr.photos.getRecent](#)
- [flickr.photos.getSizes](#)
- [flickr.photos.getUntagged](#)
- [flickr.photos.getWithGeoData](#)
- [flickr.photos.getWithoutGeoData](#)
- [flickr.photos.recentlyUpdated](#)
- [flickr.photos.removeTag](#)
- [flickr.photos.search](#)
- [flickr.photos.setContentType](#)
- [flickr.photos.setDates](#)
- [flickr.photos.setMeta](#)
- [flickr.photos.setPerms](#) (excerpt)



# REST (Representational State Transfer)

- REST is one of many possible methods to call a Web Service API
- History:
  - Roy Fielding 2000, Ph.D. thesis
    - » Analyzes and generalizes architecture of the Web
- Main features which made the Web architecture successful:
  - Identification of resources (in most cases by URIs)
  - Manipulation of resources through these representations
  - Stateless operation of server (regarding application state)
  - Hypermedia as base engine
- Applying REST to Web Services:
  - All resources on the server are identified by URI strings
    - » API method plus parameters coded in URI
  - Client uses only standard HTTP methods, mainly GET
  - Response contains clear metadata about the used language and an information body containing hyperlinks (to further resources)
- Atom feeds also provide a REST-based API

# Example: REST Request/XML Response

- Request:

```
http://api.flickr.com/services/rest/  
?method=flickr.photos.search&api_key=8c...93  
&tags=puppy&per_page=3
```

- Response:

```
<?xml version="1.0" encoding="utf-8" ?>  
<rsp stat="ok">  
<photos page="1" pages="276125" perpage="3"  
total="828375">  
  <photo id="41150XXXX20" owner="41905YYY@N03"  
    secret="13a...1c" server="2638" farm="3"  
    title="MY PHOTO TITLE!" ispublic="1"  
    isfriend="0" isfamily="0" />  
  <photo id="4116JJJ47" owner="225GGG@N08"  
    ... />  
  <photo id="4176GGG653" owner="45HHHH06@N00"  
    ... />  
</photos>  
</rsp>
```

# SOAP and XML-RPC

- Remote procedure call (RPC):
  - Technology to execute a procedure (method) with certain parameter values on a different (remote) computer
  - Various technologies exist (e.g. CORBA, DCOM) outside the Web area
- Web Service Invocation:
  - Invoking a Web Service using Web standards
- SOAP (earlier acronym: Simple Object Access Protocol)
  - XML-based syntax for messaging between applications
  - Independent of transport protocol
  - Web Services are a special application of SOAP
  - W3C standard
- XML-RPC:
  - Similar to SOAP (somehow its predecessor)
  - Transport protocol is HTTP
  - Simpler but limited in functionality

# SOAP Example

- From Flickr.com:

```
<s:Envelope
  xmlns:s=http://www.w3.org/2003/05/soap-envelope
  xmlns:xsi=http://www.w3.org/1999/XMLSchema-instance
  xmlns:xsd="http://www.w3.org/1999/XMLSchema">
  <s:Body>
    <x:FlickrRequest xmlns:x="urn:flickr">
      <method>flickr.test.echo</method>
      <name>value</name>
    </x:FlickrRequest>
  </s:Body>
</s:Envelope>
```

- SOAP makes use of XML namespaces
- Relatively high organizational overhead
- Compare equivalent REST request format

```
http://api.flickr.com/services/rest/
?method=flickr.test.echo&name=value
```

# XML-RPC Example

- From Flickr.com:

```
<methodCall>
  <methodName>flickr.test.echo</methodName>
  <params>
    <param>
      <value>
        <struct>
          <member>
            <name>name</name>
            <value><string>value</string></value>
          </member>
        </struct>
      </value>
    </param>
  </params>
</methodCall>
```

- Simple structure, deep nesting, also large overhead

# Conceptual Difference REST vs. SOAP/XML-RPC

- Tradeoff between
  - diversity of method names and
  - complexity of parameter structure
- Simple classical example
  - Special method name: fib
    - » Call: fib(13)
    - » SOAP style
  - Universal method name: exec
    - » Call: exec(fib, 13)
    - » REST style (GET is universal method name)
- Programs as data structures
  - Universal interpreter (compare Turing machine)
  - Basic idea of all current computer technology

# API Toolkits

- Requests are constructed and responses are evaluated in scripts
  - Mostly server-side scripts, e.g. PHP
  - Constructing a request in PHP:  
`$content = file_get_content($url);`
  - Evaluating the response:  
XML parsing is standard part of PHP since version 5
- Simplifying development for specific API:  
*API toolkits*
  - Example: `phpflickr.com`
  - "Wrapper" around API functions and invocation
  - Direct PHP call to required functionality
  - Response processed and data array returned
    - » Example functions:  
`people_findByUsername(), getPhotos()`

Selected(!) API toolkits  
for Flickr

- ActionScript**
  - [flickr api \(docs\)](#)
  - [Flashr](#)
  - [Flickr API Interfaces REST](#)
  - [as3 flickr lib](#)
- C**
  - [Flickcurl](#)
- Cold Fusion**
  - [CFlickr](#)
- Common Lisp**
  - [Clickr](#)
- cUrl**
  - [Curlr](#)
- Delphi**
  - [dFlickr](#)
- Java**
  - [flickrj](#)
  - [jickr](#)
- .NET**
  - [Flickr.NET](#)
- Objective-C**
  - [ObjectiveFlickr](#)
- Perl**
  - [Flickr::API 0.03](#)
  - [Flickr::Upload 1.06](#)
- PHP**
  - [PEAR::Flickr\\_API](#)
  - [phpFlickr](#)
- PHP5**
  - [Phlickr](#)
- Python**
  - [Beej's Python Flickr API](#)
  - [flickr.py](#)

# Yahoo Pipes

- Example of a tool (Web application itself) for *data mashup* development:
  - Interactive feed aggregator and manipulator
- Graphical environment to
  - Fetch data from source
  - Extract data
  - Apply filters
  - Apply simple programming tools

we make flickr not art\*

Layout Expand All Collapse All

Sources

User inputs

Operators

- For Each: Annotate
- For Each: Replace
- Count
- Filter
- Truncate
- Content Analysis
- Sort
- BabelFish
- Location Extractor
- Union
- Unique

Url

String

Date

My pipes

Fetch

URL

http://www.we-make-money

Union

Content Analysis

For Each: Replace

Replace each item in input feed with al

Flickr

Find 2

images of heise.de

[pipes.yahoo.com](http://pipes.yahoo.com)



# Screenscraping

- Technically the following is possible ("*Screenscraping*"):
  - Send HTTP request from server script to a Web site (even if it does not offer a Web Service API)
  - Analyse the returned HTML code
  - Proceed depending on the result
- The script simulates a human person using a Web browser
  - "Web Robot"
  - Frequently used by search engines
- Most Web site providers do not agree with automated access
  - Dangerous in particular in the area of authentication
  - Recommendation:  
Check Terms of Use carefully, or better refrain from Screenscraping

# 4 Communities, the Web and Multimedia

4.1 Evolution of the Web

4.2 Social Networks and Social Media

4.3 Web Content Aggregation and Integration

4.4 Virtual Worlds in the Web

# Virtual Worlds

- Online communities and online games are merging
  - Example World of Warcraft
- Non-Game online communities with virtual world
  - Old idea, see
    - » Gibson: Neuromancer
    - » Stephenson: Snow Crash
  - Was tried several times, but this time a bit more successful...
- Secondlife.com
  - Created and run by Linden Labs
  - Sixteen million accounts (2009), twenty thousand concurrent users
  - Full virtual environment, avatars, extensive creative tools
- Many simpler virtual worlds
  - E.g. [www.habbo.de](http://www.habbo.de)
    - » Virtual hotel for kids

# Example: Habbo.de



# Second Life



Linden  
Gallery  
Of Resident Art

<http://video.google.com/videoplay?docid=-5182759758975402950>

# Second Life and Business

- Large companies are/were using Second Life
  - For meetings, conferences, customer care
  - As sales channel



☆ AP Photo by IBM Corp. - 1 month ago

In this screen grab provided by IBM Corp., one of their virtual islands which will open to the public next week is shown. |