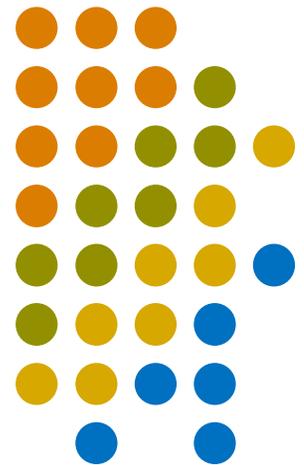




FeedFeed

A Web Service for Feed Metadata

Igor Pesenson
Advisor: Erik Wilde



Outline of presentation



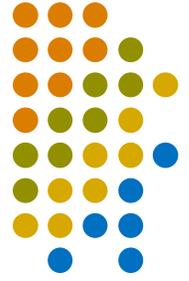
- Background
 - What are feeds
- Problem Statement
 - Why feeds
- FeedFeed Web Service
 - What it is



Bookmarks as inspiration

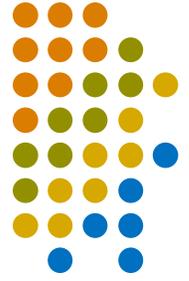


The screenshot shows a web browser window with the title "rocket1997's bookmarks on del.icio.us". The address bar contains "http://del.icio.us/rocket1997?setminposts=1". The browser's search bar shows "Google". Below the browser window, the del.icio.us profile page is visible. The profile name is "del.icio.us / rocket1997 /". There are links for "popular | recent", "login | register | help". The main content area shows "All rocket1997's items (76)" with a search bar and a "del.icio.us" dropdown menu. The page is on "page 1 of 8". The first item is titled "Disabled advocates wary of march toward open source" with a "save this" link. The text below the title reads: "BOSTON – Beleaguered Massachusetts CIO Peter Quinn has resigned, but the controversy surrounding his decision to adopt OpenDocument formats won't end with his departure. State officials haven't indicated they plan to amend Quinn's plan to adopt OpenForma". The item is dated "on nov 23, 2006". The second item is titled "» Massachusetts open source fight becomes partisan | Open Source | ZDNet.com" with a "save this" link. The text below the title reads: "Is the battle over the OpenDocument format in Massachusetts becoming a partisan battle? If it is, Microsoft is allied to the Democrats.". The item is dated "on nov 23, 2006". The third item is titled "Mass. to use Microsoft Office in ODF plan | Tech News on ZDNet" with a "save this" link. The text below the title reads: "Massachusetts will begin using OpenDocument as the default document format later this year as planned, but it will be sticking with Microsoft Office in the near term, the state's top technology executive said.". The item is dated "on nov 23, 2006". On the right side of the page, there is a "tags" section with a list of tags including "acs", "activism", "progressive", "alt.politics", "california", "cdc", "centers", "ceo", "commercial", "conferences", "cosmetics", "database", "directory", "dod", "doe", "donate", "fear", "food", "forum", "forums", "gifts", "goodcause", "government", "hazard", "health", "icon", "industry", "infoLaw", "international", "labsafety", "lbl", "library", "licence", "massachusetts", "nanoparticles", "nanoProducts", "nanotech", "nasa", "news", "ngo", "nih", "niosh", "nni", "nsa", "nsti", "ntp", "nytimes", "openSource", "oss", "papers", "pew", "plans", "policy", "publications", "publicview", "references", "research", "rice", "safety", "safety-research", "scholar", "sci.med.nutrition", "study", "thinktanks", "toprint", "toxicology", "uk", "wilsoncenter", "workshops". Below the tags section is a "tag options" section with a "view as cloud List" link.



Web feeds are...

- A way for websites to provide frequently updated content
 - BBC hourly news to your cell phone
- A way for users to pull many different information sources into one place
 - Favorite 15 news sources on your cell phone



Some feed vocabulary

- Content providers *syndicate* a feed
- Users *subscribe* to a feed
- Feeds can be *aggregated*
- Feed aggregators are called *feed readers*
- Feed *technical standards* are RSS 1.0, RSS 2.0, Atom, AtomPub, many others
- Feed *metadata* is data about the feed

Have you used iGoogle?



The screenshot shows the iGoogle homepage in a browser window. The address bar displays `http://www.google.com/ig?hl=en`. The page features a search bar with the iGoogle logo, navigation links for Web, Images, Maps, News, Shopping, and Gmail, and user account information for `ipesenson@gmail.com`. The main content area is divided into several widgets:

- ToDo:** A yellow notepad-style widget with a task list starting with "1. Type new task here".
- Weather:** A widget for Oakland, CA, showing a temperature of 75°F, "Mostly Cloudy" conditions, and a wind speed of 0 mph. It includes forecasts for "Today" (85° | 72°) and "Thu" (88° | 58°).
- Gmail:** A widget showing an inbox with 346 messages and a list of recent emails, including one from "Igor" at 9:54pm.
- Top Stories:** A widget displaying news headlines such as "Edwards Throws Support to Obama" and "House Passes Farm Bill by a Veto-Proof Margin".
- Google Calendar:** A widget showing a calendar for May 2008, with the 1st, 2nd, and 3rd highlighted.

The browser window title is "iGoogle" and the status bar at the bottom shows "Done".

Have you used MyYahoo?

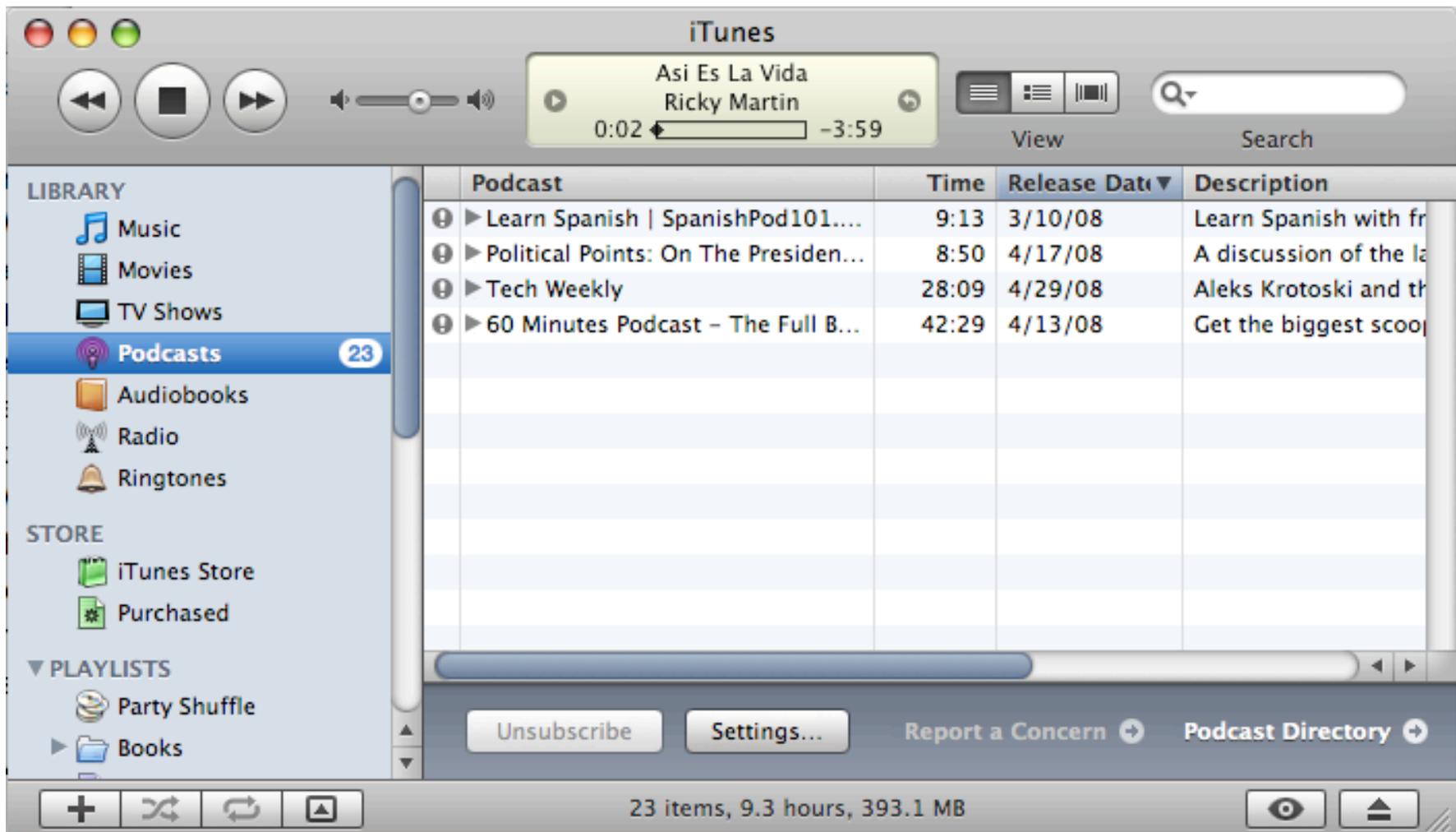


The screenshot shows a browser window titled "My Yahoo!" with the address bar displaying "http://cm.my.yahoo.com/". The page features a search bar with the "MY YAHOO! BETA" logo and navigation tabs for "Web", "Images", "Video", "Local", and "Shopping". Below the search bar, there are personalized options like "My Quicklinks", "Personalize this page", and "Share this page". The main content area is divided into several sections:

- Personal Assistant:** Includes links for Mail, Horoscope, Portfolio, Weather, Lottery, and Sports.
- Yahoo! News Full Coverage:** Lists recent news items such as "China - 25 minutes ago", "Myanmar - 25 minutes ago", and "2008 Presidential Election - 25 minutes ago".
- Financial News from MarketWatch.com:** Lists financial reports like "Rio Tinto may partner with China on ore project: report - 1 hour ago".
- Yahoo! Shopping: Popular Products & Deals:** Displays product listings for a "Garmin nuvi 350 GPS Receiver" and an "Apple iPod Nano 4G MP3 Player", both with price ranges and star ratings.

An advertisement for "THE INSPIRON™ 530" is visible on the left side, featuring a computer monitor and keyboard. The browser's status bar at the bottom shows "Done" and a yellow "P" icon.

Have you used iTunes?



Where do you find feeds?



The image shows two browser windows. The top window is the BBC News front page, with the URL `http://news.bbc.co.uk/` circled in red. The bottom window is a blog titled "grey and red, a squirrel journal", with its URL `http://greyandred.blogspot.com/` also circled in red. Red arrows point from these two circles to a magnified view of an RSS feed icon and its URL `http://news.bbc.co.uk/`.

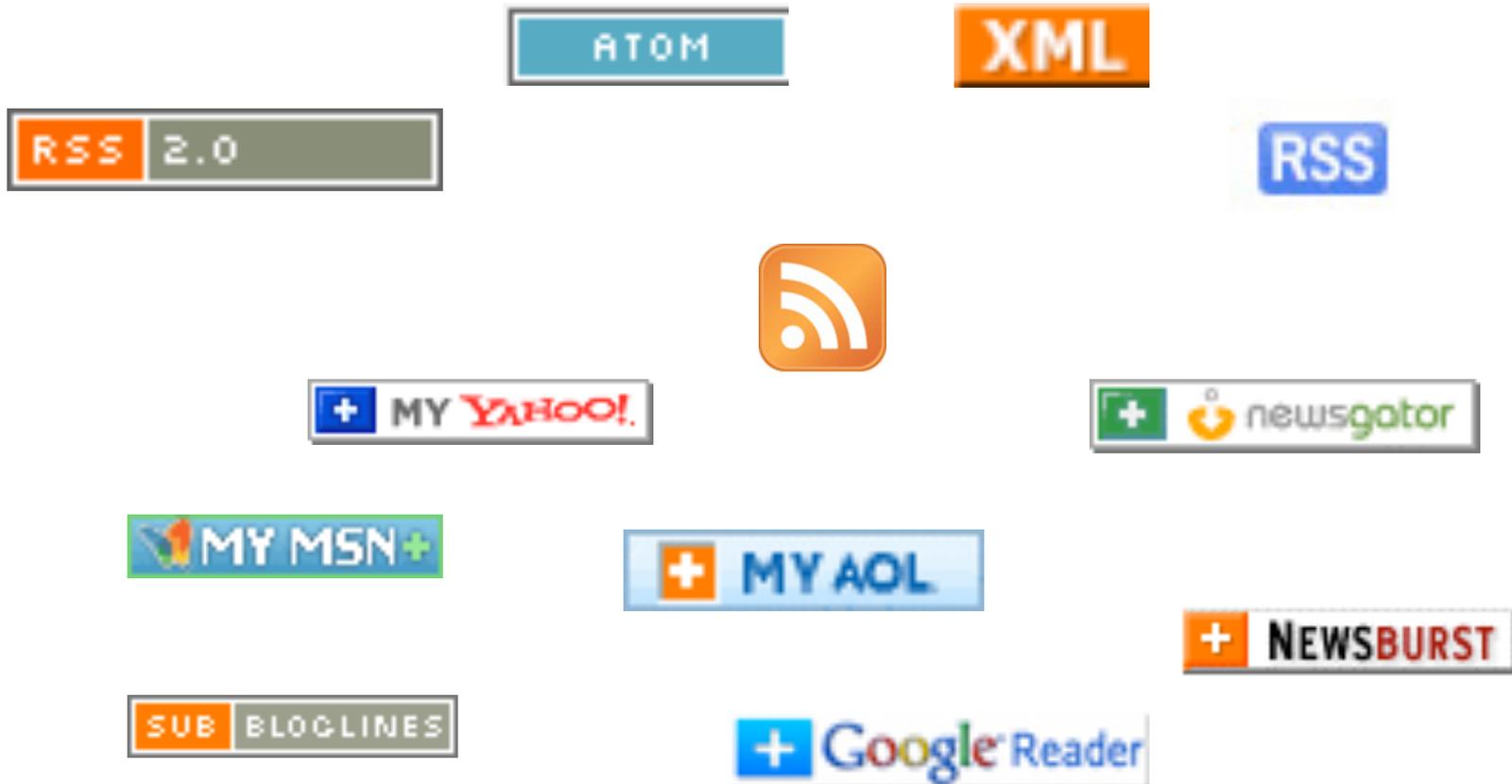
A magnified view of an RSS feed icon and its URL `http://news.bbc.co.uk/`. The icon is a white RSS symbol on an orange square background. The URL is displayed in a white box with a black border.



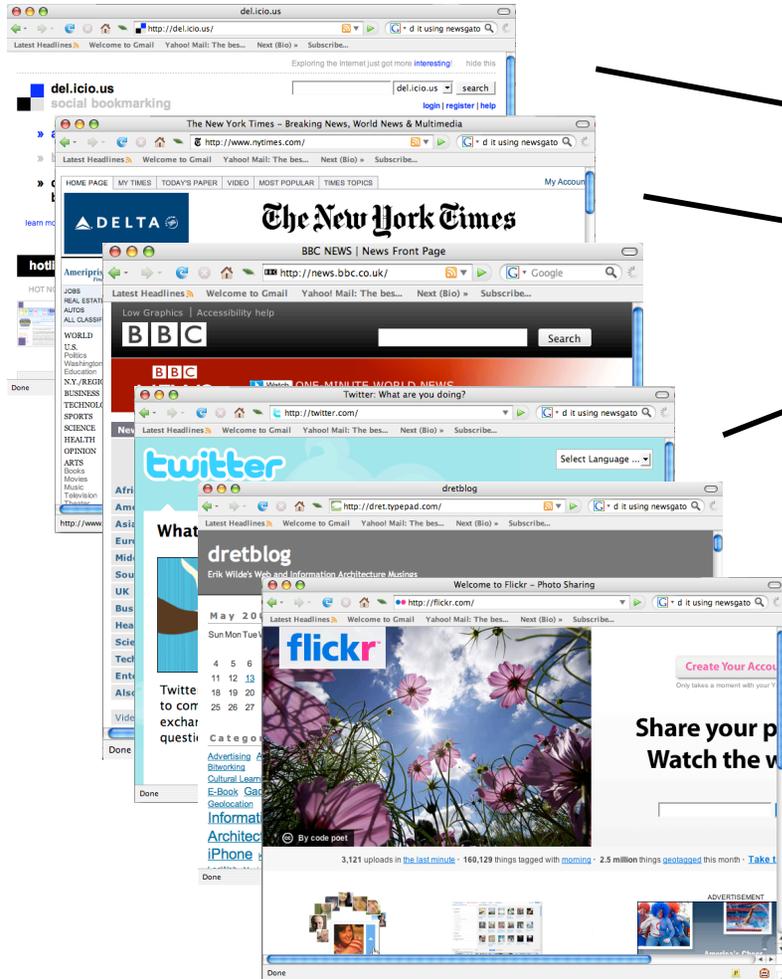
How do you know there is a feed?



Enough symbols to confuse everyone...



Publishers syndicate you aggregate



...anywhere

Example Feed Reader



The screenshot shows the Google Reader interface in a browser window. The address bar displays the URL <http://www.google.com/reader/view/#stream/user/>. The page title is "Google Reader (1000+)".

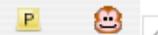
The interface includes a navigation bar with links for "Gmail", "Calendar", "Documents", "Photos", "Reader", "Web", and "more". The user's email address "ipesenson@gmail.com" is visible, along with links for "Offline", "Settings", "My Account", "Help", and "Sign out".

The main content area is titled "Google Reader" and features a sidebar on the left with navigation options: "Home", "All items (1000+)", "Starred items", "Trends", "Your stuff", and "Friends' shared items (26)". Below these are buttons for "Add subscription" and "Discover", and a "Refresh" button.

The main feed is titled "News" and shows a list of items. The first item is "Amazon dilemma - Logging both" from "BBC News | News Front Pag" at 9:55 PM. Other items include "Janish wins it in debut on walk-off", "Lebanon Reverses Decisions", "Edwards Endorses Obama for", "Palestinians mark 'catastrophe'", "Yankee Hitters Give Mussina Just", "Flores continues to haunt Mets in", "UN chief to send envoy to Burma", "US to Continue Burma Aid", "Judge Allows Suit Against", "'American Idol' Sets its Top Two -", "Sabathia dominant in complete", "China scales up rescue", "Winding Mountain Road", "California could become the", "Bacon Painting Auctioned for", "Council approves drug policy", and "John Edwards Endorses Barack".

At the bottom of the feed, there are buttons for "Previous item" and "Next item", and a note that there are "more than 60 items".

Done



How to get a feed



The screenshot shows the BBC News website with a red circle around the RSS icon in the browser's address bar. A red arrow points from this icon to a second screenshot of the same page, where a red circle highlights the 'Subscribe to this feed using' dropdown menu, which is set to 'Google'. Another red arrow points from this dropdown to the iGoogle interface on the right.

BBC NEWS | News Front Page

http://news.bbc.co.uk/

Latest Headlines Welcome to Gmail Yahoo! Mail: The bes... (Bio) Subscribe...

Low Graphics Accessibility help

BBC NEWS Watch ONE-MINUTE WORLD NEWS

News Front Page Page last updated at 19:28 GMT, Wednesday, 14 May 2008 20:28 UK

LATEST: Pakistan's army swaps prisoners with the Taliban.

China quake toll close to 15,000

Nearly 15,000 people died and 25,000 are still trapped in the earthquake that hit China, official media say.

- Quake survivors rally together
- Close-up man of area

UN raises Burma cyclone estimate

The UN sharply increases to 2.5m its estimate of those severely affected by Cyclone Nargis in Burma.

BBC News | News Front Page | World Edition

Subscribe to this feed using **Google**

Always use Google to subscribe to feeds

Subscribe Now

BBC News | News Front Page | World Edition

Visit BBC News for up-to-the-minute news, breaking news, video, audio and feature stories. BBC News provides trusted World and UK news as well as local and regional perspectives. Also entertainment, business, science, technology and health news.

[China quake toll close to 15,000](#)

Nearly 15,000 people died and 25,000 are still trapped in the earthquake that hit China, official media say.

[UN raises Burma cyclone estimate](#)

The UN sharply increases to 2.5m its estimate of those severely affected by Cyclone Nargis in Burma.

[New deadline for Zimbabwe's vote](#)

Zimbabwe's government extends by 90 days the deadline a presidential run-off vote must be held.

[Estonian cyber defence hub set up](#)

Nato backs a new cyber defence centre in Estonia, which last year blamed Russia for targeting its internet structure.

The screenshot shows the iGoogle interface with a red circle around the 'Google' dropdown menu in the 'Subscribe to this feed using' section. A red arrow points from this dropdown to the 'Gmail' widget, which shows an 'Inbox (321)' and a 'Compose' button.

iGoogle

http://www.google.com/ig

Latest Headlines Welcome to Gmail Yahoo! Mail: The bes... Next (Bio) Subscribe...

Web Images Maps News Shopping Gmail more ipesenson@gmail.com | Classic

iGoogle

Google Search I'm Feeling Lucky

Home Add a tab Get artist themes Select theme Add stuff

ToDo

Gmail **Inbox (321)** Hide preview Compose Mail

Raymond - [Jan 12]

Daniela - [maste 11:00am

Michael, Igor (2) 10:58am

Apple Develop 10:25am

Igor, Drew (2) - E 9:52am

Top Stories

[Obama Strikes Back With Negative Ad](#)

New York Times - [all 1992 related](#)

[Palfrey Suicide Notes Are Released](#)

Washington Post - [all 504 related](#)

[Myanmar cyclone kills 4000](#)

State of Technology



- Feeds are becoming widely used
 - News sites, blogs, podcasts
- Originally ambiguity with technology standards but now more stable
 - RSS, Atom, AtomPub

Jane the Feed User



- Where
 - work, home, mobile
- What
 - news, blogs, audio
- Why
 - fun, work, contacts
- How much
 - 100s of feeds

Jane has problems



- Difficult to share feeds with people
- A pain to subscribe to feeds from three different places - home, work, mobile



What does Jane need?



- A way to organize the many feeds
- A way to create, organize and share feeds with friends
- A way to manage multiple subscriptions

How do we help Jane?



www.jafs.com



FeedFeed Service

Jafs.com

A feed service for Jane



Latest Headlines Welcome to Gmail Yahoo! Mail: The bes... Next (Bio) » Subscribe...

FeedFeed subscriptions of igor:

Add a feed:

[This American Life Updates](#) Add a tag: bioproject boring cell cool
crazy forjim fun jimsnewtag
Presented by Chicago Public Radio, "This American Life" is an award-winning radio program as well as jimtag nyc toread work
a television program on the Showtime network. Hosted by Ira Glass, we feature first-person stories
and short fiction pieces that are touching, funny, and

to: toread cell fun boring crazy fun fun fun bioproject toread toread nyc work work cell jimtag
jimsnewtag jimsnewtag fun nyc fun bioproject toread crazy cool boring work forjim nyc work
forjim cell fun work fun work crazy boring fun cell fun fun boring

The Onion Radio News is a daily podcast featuring a short news clip from The Onion's award-winning
24-hour radio news network. Read The Onion at www.theonion.com.

to: toread cell fun boring crazy fun fun fun bioproject toread toread nyc work work cell jimtag
jimsnewtag jimsnewtag fun nyc fun bioproject toread crazy cool boring work forjim nyc work
forjim cell fun work fun work crazy boring fun cell fun fun boring

Done

Did we help Jane?



The screenshot shows a Mozilla Firefox browser window displaying a FeedFeed page. The address bar shows the URL `http://groups.ischool.berkeley.edu/feedfeed/jafs/getfeedfeed.php`. The page title is "FeedFeed subscriptions of igor:". The page content includes three feed entries: "This American Life Updates", "Atom", and "The Onion Radio News". Each entry has a description and a list of tags. Red annotations highlight specific elements: a large bracket on the left groups the feed entries under the label "Jane Feeds"; the word "Organize" is centered above the page with arrows pointing to the "Add a tag:" input fields; the word "Browse" is on the right with an arrow pointing to the vertical scrollbar; the word "Share" is on the right with an arrow pointing to the "export" button; and the word "Subscribe" is on the right with an arrow pointing to the "Feed" button. Red circles highlight the "Add a tag:" input fields, the tag lists for the first two feeds, and the "export" button.

Jane Feeds

Organize

Browse

Share
Subscribe

export

Jane is happy...

What about her company?



- MegaHealth Inc.
 - Who
 - 5000 doctors, 400,000 patients
 - What
 - benefits, procedural, studies results
 - How
 - flyers, email... wants to use feeds
- Classic problem
 - Vast amount of info centrally produced
 - Highly segmented audiences



What does MegaHealth need?



- A way to organize the information
- A way to create, organize and share information
- A way to manage multiple subscriptions

How do we help MegaHealth?

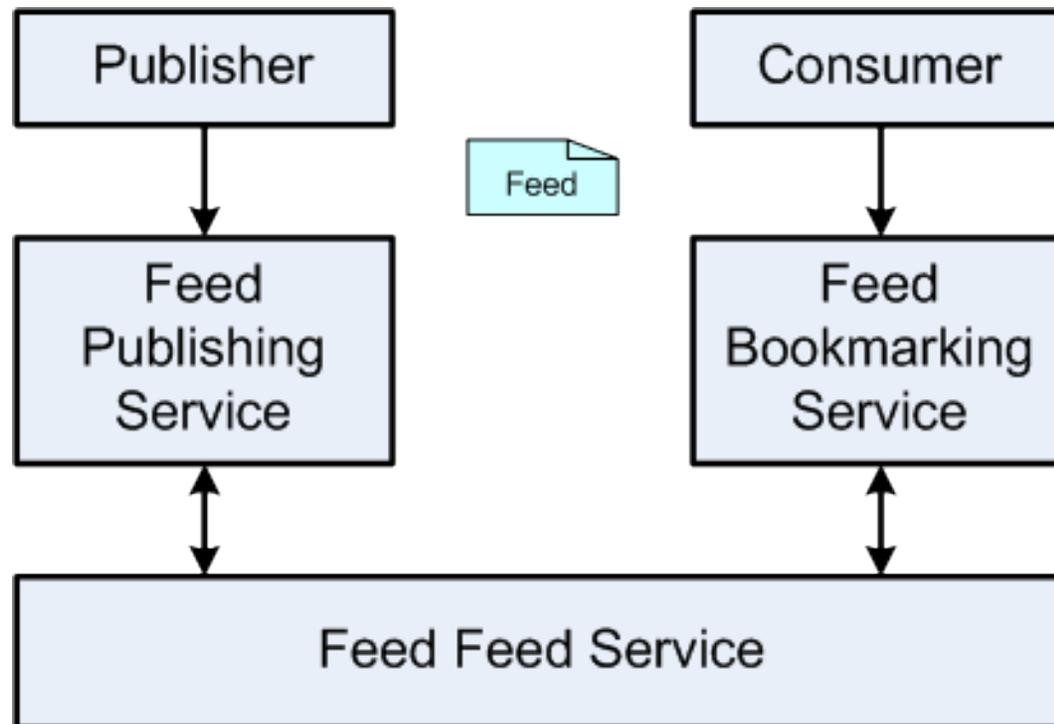


www.hafs.com



FeedFeed Service

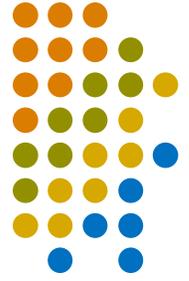
Architecture of FeedFeed





Requirements

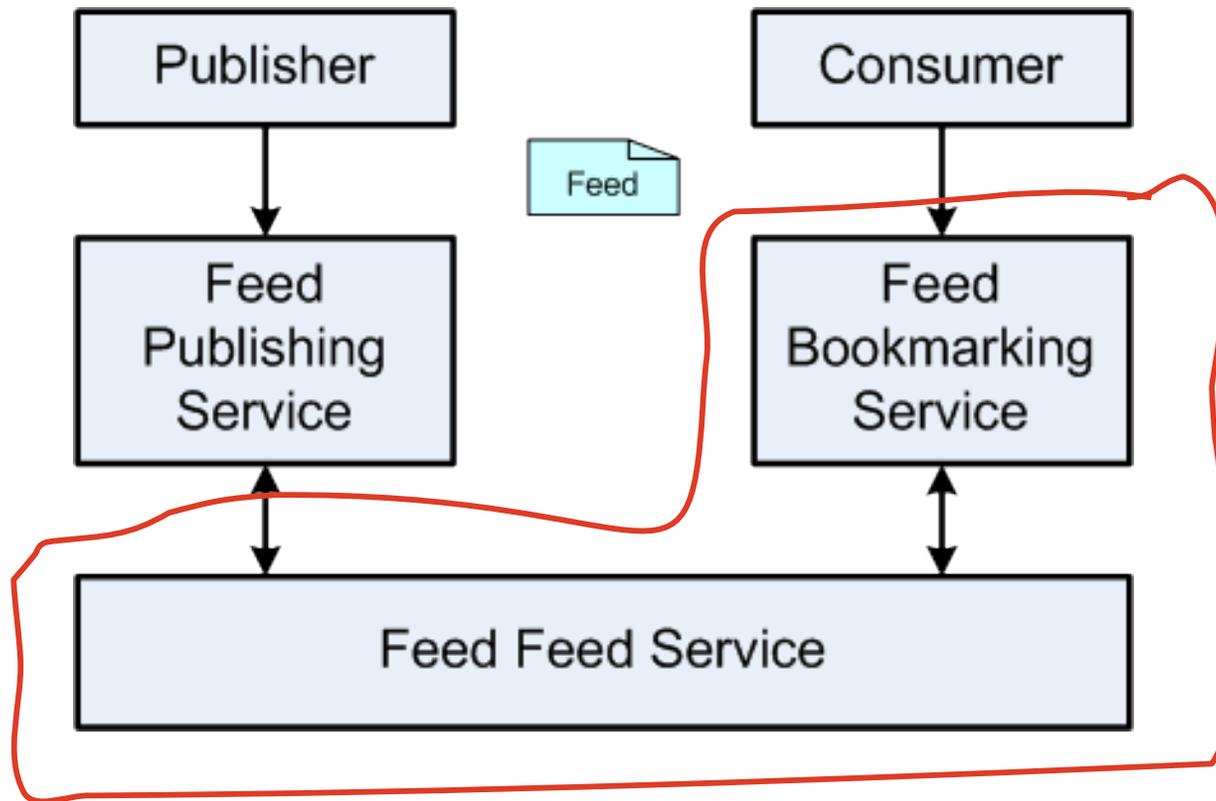
- Abstracted
 - One approach for different problems
- Broad access to data
 - People adding and modify different data
- Easy development platform
 - Accepted technical standards
 - Simple model

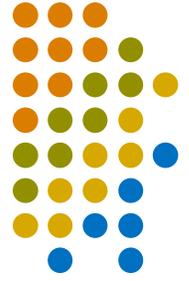


Standards and Tools

- Data format
 - Accept all feed formats, serve Atom
- Application Programming Interface
 - Atom Publishing Protocol
- Query Support
 - Feed Item Query Language
- Tools
 - Python, CherryPy, MySQL, Dejavu

Progress Made





Some Lessons Learned

- Many opportunities
 - Even advanced users not happy with existing support in the domain
- There are few generic AtomPub standard implementations
 - Amplee framework does not abstract out the data model
- AtomPub standard rules!

Future work

- Implement the full AtomPub API
- Build out the consumer interface
- Create the publisher interface



A Big Thanks to:

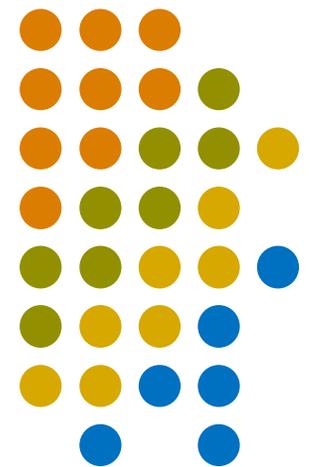
Erik Wilde

Jim Miller

Chris Volz

Kevin Heard

Yiming Liu



Questions?

