

# The Architecture of the Web

Yury Lifshits

Caltech

<http://yury.name>



Caltech IST Bunch Lunch  
May 2008

1/31

**Wikipedia:** The Web is a system of interlinked **hypertext documents** accessed via the Internet.

The Web is changing!

**Modern Version:** The Web is a system of **data, people and software** connected via the Internet and mobile networks.

2/31

## Outline

- 1 Data Markup
- 2 Linking the Web
  - Data Distribution
  - Identity Management
  - Widgets and APIs
- 3 Market Graph
- 4 Thoughts on Future

3/31

# 1

## Data Markup

4/31

## W3C Standards

- HTML, XHTML
- CSS
- XML, RDF

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.1/">
  <rdf:Description
    rdf:about="http://en.wikipedia.org/wiki/Tony_Benn">
    <dc:title>Tony Benn</dc:title>
    <dc:publisher>Wikipedia</dc:publisher>
  </rdf:Description>
</rdf:RDF>
```



5/31

## FOAF



- **Mission:** creating machine-readable pages describing people, the links between them and the things they create and do
- Started in 2000 by Libby Miller and Dan Brickley
- Tim Berners-Lee: FOAF is a basis for Global Giant Graph

6/31

## FOAF: Example

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#">
  <foaf:Person rdf:about="#JW">
    <foaf:name>Jimmy Wales</foaf:name>
    <foaf:mbox rdf:resource="mailto:jwales@bomis.com" />
    <foaf:homepage rdf:resource="http://www.jimmywales.com/" />
    <foaf:nick>Jimbo</foaf:nick>
    <foaf:depiction
      rdf:resource="http://www.jimmywales.com/aus_img_small.jpg" />
    <foaf:interest>
      <rdf:Description rdf:about="http://www.wikimedia.org"
        rdfs:label="Wikipedia" />
    </foaf:interest>
    <foaf:knows>
      <foaf:Person>
        <foaf:name>Angela Beesley</foaf:name>
      </foaf:Person>
    </foaf:knows>
  </foaf:Person>
</rdf:RDF>
```

7/31

## Microformats



- Grassroots approach to Semantic Web
- Notable microformats: hCalendar, hCard, hReview, hResume, XFN, XOXO

<http://tantek.com/presentations/2006/07/what-are-microformats/>

8/31

## Microformat Example: hCard

```
<div class="vcard">
  <div class="fn">Joe Doe</div>
  <div class="org">The Example Company</div>
  <div class="tel">604-555-1234</div>
  <a class="url" href="http://example.com/"
  >http://example.com/</a>
</div>
```

9/31

## Social Graph API

Brad Fitzpatrick presents Social Graph API:

<http://www.youtube.com/watch?v=LabCylbapuM>

10/31

# 2

## Linking the Web

11/31

## Types of Connections

- Pointers (URL, XRI & XDI, Yahoo woelD)
- Data distribution
- Identity management
- Widgets and APIs

12/31

## RSS



- Full name  
Really Simple Syndication  
RDF Site Summary
- Started in 1999 by Ramanathan V. Guha
- List of feeds: OPML
- Attention profile: APML
- Challenges: filtering, discovery, promotion

13/31

## RSS Example

```
<?xml version="1.0"?> <rss version="2.0">
<channel>
  <title>Lift Off News</title>
  <link>http://liftoff.msfc.nasa.gov/</link>
  <description>Liftoff to Space Exploration.</description>
  <language>en-us</language>
  <pubDate>Tue, 10 Jun 2003 04:00:00 GMT</pubDate>
  <lastBuildDate>Tue, 10 Jun 2003 09:41:01 GMT</lastBuildDate>
  <docs>http://blogs.law.harvard.edu/tech/rss</docs>
  <generator>Weblog Editor 2.0</generator>
  <managingEditor>editor@example.com</managingEditor>
  <webMaster>webmaster@example.com</webMaster>
  <ttl>5</ttl>

  <item>
    <title>Star City</title>
    <link>http://liftoff.msfc.nasa.gov/news/2003/news.asp</link>
    <description>How do Americans get ready to work with Russians
    aboard the International Space Station? They take a crash
    course in culture, language and protocol at Russia's
    Star City.</description>
    <pubDate>Tue, 03 Jun 2003 09:39:21 GMT</pubDate>
    <guid>http://liftoff.msfc.nasa.gov/2003/06/03.html#item573</guid>
  </item>
```

14/31

## Viral Propagation

- Digg, Yahoo! Buzz
- Del.icio.us, Stumbleupon
- Social news feeds (Facebook news, FriendFeed)

15/31

## OpenID



- Glossary: identity provider, relying party
- 2005: Started by Brad Fitzpatrick in LiveJournal project
- 2007: OpenID Foundation

16/31

## OAuth



- Open standard for [access delegation](#)
- Started in November 2006, first draft in October 2007
- Glossary: User, Consumer, Service Provider, Protected Resources
- Challenges: phishing attack, DoS attack

17/31

## Facebook Platform

- A framework for creating third-party applications that interact with core Facebook features. Launched in May 2007
- Interface (API), Query (FQL), Markup (FBML)
- [Anatomy](#): about, left nav, Facebook canvas pages, profile box, profile action links, privacy settings, news feed, alerts, message attachments, requests

18/31

## Other Application Platforms

- OpenSocial, FriendConnect, Y!OS
- Firefox extensions and userscripts (via GreaseMonkey)
- iPhone OS
- iGoogle, Netwibes, Yahoo Pipes
- Wordpress & Drupal, Salesforce.com,

19/31

## APIs

- Google AJAX Search API, Yahoo Search API, Yahoo SearchMonkey
- MySpace Data Availability, FacebookConnect
- Yelp API

20/31

# 3

## Ongoing Project: Market Graph

21/31

## Market Graph

**Market Graph** is a graph of relations between people, businesses and products.

- Transactions (I bought, I order)
- Evaluation (I like, I hate, product rating)
- Demand (wishlists, problem list)
- Authorship and employment

22/31

## Applications of Market Graph

- Recommendations
- Search
- Ad targeting
- Viral propagation
- Demand generation
- Platform for third-party applications
- Market computing

23/31

## Market Graph: Questions

- Business model and incentives
- Centralized vs. decentralized?
- Data reliability
- Publishing Market Graph data

24/31

# 4

## Thoughts on Future

25 / 31

## Emerging Technologies in the Web

- Augmented reality, Geoweb
- Vector graphics, 3D
- Internet TV, video
- Ubiquitous computing
- Semantic web
- Artificial intelligence, question answering
- Machine translation
- Intelligent agents
- New interfaces (HCI)

26 / 31

## Web OS

- Web office: Google Docs, Zoho
- Social platforms: Facebook, OpenSocial
- DesktopTwo, Netvibes, iGoogle, myYahoo
- Firefox Gran Paradiso
- CMS as Web OS?

27 / 31

## DataPortability.org



- **Participants:** Google, Facebook, Plaxo, Drupal, Netvibes, Mystrands, LinkedIn, Flickr, Six Apart, Twitter, Digg, Microsoft
- **Started:** November 2007
- **Standards involved:** APML, FOAF, hCard, OAuth, OpenID, OPML, RDF, RSS, SIOC, XFN, XRI, and XDI.
- **Next goal:** DataPortability reference design

28 / 31

## Links

Homepage <http://yury.name>

<http://businessconsumer.net/files/marketing-agenda.pdf>  
Research Agenda in Online Marketing [Working paper]

<http://yury.name/newweb.html>  
Tutorial "The New Web"

<http://yury.name/reputation.html>  
Tutorial "Reputation Systems"

<http://businessconsumer.net>  
Our research project in online marketing

29/31

## References

**Conferences:** WWW, WSDM, EC, Web 2.0 Expo, Web 2.0 Summit, The Next Web, Techcrunch50, ETech, TED

**Videos:** TED Talks (Jeff Bezos, Larry Lessig, Mena Trott, Jimmy Wales), OpenSocial Launch Event, Mark Zuckerberg (F8 Keynote)

**Blogs:** Techcrunch, ReadWriteWeb, Mashable!, Valleywag, Techmeme, VentureBeat, Greg Linden

**People:** Kevin Marks, Brad Fitzpatrick, Tantek Celik, David Recordon, Joseph Smarr, Matt Mullenweg, Mark Zuckerberg, Kevin Rose, Dries Buytaert, Chris Messina, Nova Spivack, Dick Hardt, Chris Saad

30/31

Thanks for your attention!  
Questions?

31/31