Semantic Web and Web2.0

Dr Nicholas Gibbins
Web 2.0 is the business revolution in the computer industry caused by the move to the internet as platform, and an attempt to understand the rules for success on that new platform.

Chief among those rules is this: Build applications that harness network effects to get better the more people use them

Tim O’Reilly, 2006
http://www.fotopedia.com/items/flickr-2309701375
Characteristics of Web2.0

- Enhanced user interface, experiences
  - Use of AJAX, software in the Web browser
- Linking people, forming communities
  - Social networking sites, online sharing of resources
- Organizing resources
  - Social tagging, Flickr, Yahoo Pipes, etc
- Mix and merge of the real and virtual
  - Mashups, Google Maps
Web 2.0 Meme Map

Flickr, del.icio.us: Tagging, not taxonomy
PageRank, eBay reputation, Amazon reviews: user as contributor
Blogs: Participation, Not publishing
BitTorrent: Radical Decentralization
Gmail, Google Maps and AJAX: Rich User Experiences
Google AdSense: customer self-service enabling the long tail
Wikipedia: Radical Trust

Strategic Positioning:
- The Web as Platform

User Positioning:
- You control your own data

Core Competencies:
- Services, not packaged software
- Architecture of Participation
- Cost-effective scalability
- Remixable data source and data transformations
- Software above the level of a single device
- Harnessing collective intelligence

"An attitude, not a technology"
The Long Tail
Data as the "Intel Inside"
The perpetual beta
Software that gets better the more people use it
Play
Rich User Experience
Granular Addressability of content

Hackability
The Right to Remix: "Some rights reserved"
Emergent: User behavior not predetermined
Trust your users
Small Pieces Loosely Joined (web as components)
Tagging, not taxonomy

- Resources described by bags of keywords
- “nyc, www2004, libby miller, posters”
- Vocabulary is not controlled
- Cheaper annotations – users more likely to annotate resources
Rich User Experiences

- AJAX: Asynchronous JavaScript and XML
- Googlemaps, GMail
- More responsive interfaces
- Richer interaction (beyond the click)
Radical Decentralisation

• Decentralisation of resources and processing

• Peer-to-peer computing (remember that?)

• Application components on Web

• Service orientation

• Open APIs allow service reuse in unexpected ways
  • Remixes, mashups and hackability
User as contributor

- User-written material forms basis for community websites
  - eBay feedback
  - Amazon reviews

- Rich-get-richer behaviour
  - The more people that contribute to a website, the more useful it is to you
  - Social bookmarking
  - Recommender systems
Enabling the Long Tail

• Amazon sells more minority books than mass-market bestsellers

• Small sites make up the bulk of the Web
  • DoubleClick used by large corporate sites
  • Google AdSense aimed at small sites
Participation, not publishing

- The rise of the blogosphere
- Publishing in a social context
  - Feedback and comments from readers
  - Ongoing polylogue
- Some high-profile successes
  - Blogs vs traditional news media (Iraq war, etc)
  - Grass-roots political blogging
Radical Trust

- No controls on authorship
- Everyone’s an author/editor
- “With enough eyeballs, all bugs are shallow”
- Content becomes trusted because (it is assumed that) many people have vetted it
Changing links

Linking Documents

Linking Users

Linking Communities
Social applications

• Harnessing the network effect

• Distinct functionality/Social groups/Domains
  • Wikipedia: building an open encyclopedia
  • Flickr: sharing photographs
  • Facebook: meeting new people and keeping in touch with friends
  • Last.fm: finding new music
  • Deli.cio.us: sharing bookmarks
The Personal and Social Dimensions

**Personal**

- **Blog**
  - Publish one’s own thoughts and opinions
- **Tagging**
  - Organize in one’s own way

**Social**

- **Social Network**
  - Present oneself on the Web
  - Identify common interests and topics
- **Sharing**
  - Present one’s interest and attentions
  - Users get to know each other’s interests
Facebook

- Clean interface with good usability
- Immediate results after actions
- An understanding of the audience and the domain
- The ability to refer to other members of the network.
- Vanity and the act of publishing
Facebook

- Maximise data collection by:
  - Making publishing easier
  - Adding immediate value for the users
  - Hiding the underlying system

- Now they are exposing the data
  - And making it fun....

```
SELECT name, pic FROM user WHERE uid=211031 OR uid=4801660
```
Personal Information

• You will be known for your online behaviour

• Mash-ups ...

• If you could refer to yourself using a single URI, could you more easily see where you’ve been mentioned?
Lessons Learned

• Google’s View of the Web – links as votes of confidence

• Facebook has been successful due to ease of use and immediate feedback

• Get the Data Exposed! - Complexity

• Resolvable URIs!

• Interfaces for the Semantic Web
  • mSpace
  • Tabulator
  • Dbpedia
  • Haystack
  • ...

User Interfaces
User Interfaces

Query wikipedia
This semantic database contains over 10 million statements extracted from the English Wikipedia.

Modify this query or create your own!
This knowledge base contains subject-predicate-object statements obtained from the infobox templates of the English wikipedia, such as:

DISCO

dbpedia.org
Querying Wikipedia like a Database.

OpenLink QBE
visual query builder
Encourage the deployment and publishing of Data
Tagging in context

Users
• Assign tags to resources
• Associated with other users
• Connected by the resources they tag

Resources
• Connected to other resources
• Implicit relationship through the users that tag them
Applications

YouTube
Broadcast Yourself™

flickr™

facebook

de.licio.us
social bookmarking

Technorati™

Yahoo!® Podcasts Beta

last.fm

LIVEJOURNAL

ODEO

The ESP Game
Ontologies and other animals

- Folksonomies
- Controlled Vocabularies
- Taxonomies
- Thesauri
- Ontologies

- Uncontrolled
- Controlled

- Less structure
- More structure

- Emergent
- Designed
Documents tagged with ‘sf’
(edge weights > 2)

Science Fiction

San Francisco
Users using the tag ‘sf’
Web 2.0 for the Semantic Web

- A platform with high demands:
  - Structured data
  - Information integration
  - Information discovery

- Richer user interfaces

- Domains with vibrant user communities
  (travel, web design, programming, photos, etc.)
Semantic Web for Web 2.0

• A true mashup of data from different sources (not dealing with different APIs of different sites)

• A true social network for users (not separate data being locked up in different SNS)

• A better way for users to represent themselves (not different identities in different sites)
Conclusions

• Web 2.0 needs more efficient ways to integrate data

• The Semantic Web offers ways to integrate data

• The Semantic Web and Web 2.0 open mutual opportunities