

Introduction to Web 2.0 Data Mashups

Mustafa Jarrar

Birzeit University



Watch this lecture and download the slides



Online Courses : <http://www.jarrar.info/courses>

Outline



From Web 1.0 to Web 2.0

- Where Web 2.0 is going?
- Business/Professional Networking
- Social Networking
- Web 2.0 Technologies
- The Mashup Phenomena
- Web 2.0 challenges

Keywords: Social Community, Collaboration, Sharing, APIs, blog, Social Networking, Content syndication, Mashups, Web APIs

From Web 1.0 to Web 2.0

Social Community, Collaboration, Sharing



Where Web 2.0 is going?

most Web 2.0 sites provide APIs

Wikipedia edited in real-time by anyone

(in 252 languages, millions of articles) written collaboratively by volunteers around the world.



The image shows a screenshot of a Wikipedia article for Birzeit University. The page layout includes a sidebar on the left with navigation links, a main content area with a title and introductory text, and a right-hand sidebar with a summary table and a logo. The article text describes the university's history and its role in the Palestinian political dialogue.

Birzeit University

From Wikipedia, the free encyclopedia

Birzeit University (BU) (Arabic: **الجامعة الفلسطينية**) is a privately owned in (West Bank) Ramallah, Palestinian university. BU is among the foremost tertiary educational institutions in the Palestinian territories and has played a significant role in the Palestinian political dialogue.

Foundations of BU

The institution was originally founded in 1976 as *Madhafa* (House) as an elementary school for girls from Birzeit and the surrounding villages. It was one of the first schools in the area. In 1978, it became a co-educational secondary school, and in 1982 was renamed Birzeit Higher School.

In 1982, the name was changed to Birzeit College for (1982) for a secondary with the educational secondary level. In 1983, a freshman class was incorporated to Birzeit College, followed by a sophomore class in 1985.

Birzeit University	
Established	1976
Type	State of Palestine
President	Dr. Sabar Khan
Dean	Prof. Ibrahim Al-Jarrah
Undergraduate	1,000
Postgraduate	1,000
Location	Ramallah, West Bank
Website	www.birzeit.edu.ps



A social network for sharing photos

- 3,464 uploads in the last minute ·
- 160,129 things tagged with morning ·
- 2.3 million things geotagged this month

In 2008

Source: <http://www.flickr.com>



Upload, View, and Share Videos (and tags...)

The screenshot shows a YouTube video player with the following elements:

- Search Bar:** Located at the top left, with the YouTube logo and a search button.
- Navigation:** Tabs for Home, Videos, Channels, and Subscriptions are visible.
- Video Title:** "S.F. Breeze Youth Dance Group - Dec 2008".
- Video Player:** The main video content area showing a group of people dancing. A red circle highlights the "Share" button in the bottom left corner of the player.
- Video Description:** Located to the right of the video player, containing text about the dance group and a date. A red circle highlights the description text.
- Related Videos:** A section below the description showing thumbnails and titles of other videos.



e-Buy and e-Sell

Millions of items are listed, bought, sold, and discussed daily.



→ API

We know this! What is social?

amazon.com

Information Modeling and Relational Databases: From Conceptual Analysis to Logical Design [The Morgan Kaufmann Series in Data Management Systems] (Paperback)

Buy with confidence

20 used & new available from \$46.00

API

Upcoimg.org

A social event calendar website, uses iCalendar, supports an open API for the submission of event descriptions.



Blog (Your Own Journal)

Commentary, news, personal diaries... + comments/discussions
(+106 million blogs, as of 2007)



→ API

Business/Professional Networking

16 million registered users (2007)

60 million registered users (2007)



Social Networking

Find your friends, socialize, group, share/organize events, and more...(1.15 billion monthly users)



facebook

Email or Phone Password Log In

Facebook helps you connect and share with the people in your life.

Sign Up For Facebook
It's free and always will be.

First Name Last Name

Your Email

Re-enter Email

New Password

Birthday

Month Day Year Why do we need to provide my birthday?

Female Male

By clicking Sign Up, you agree to our Terms and that you have read our Help and Policy, including our Cookie Use.

Sign Up

API

Web 2.0 Technologies

Companies are competing in gathering content from people and offering this content back through APIs (i.e. structured data)

Content syndication

How can you deliver recent news to your users without forcing them to visit your website? → RSS, Atom

Mashups

How can you remix free content and make a new website?

Web APIs (programmable web)

How can make use of the free content inside my website?

Content Syndication

Millions of online publishers including newspapers, commercial web sites and blogs now publish their latest news headlines, product offers or blog postings in standard format news feed^[1].



Content Syndication



The two main families of web syndication formats are RSS and Atom.



The image shows a screenshot of a web browser displaying the BBC News website. The address bar shows the URL http://www.bbc.co.uk/news/world_europe_070904.xml, which is circled in red. The main content area displays news items with headlines such as "French troops hit by sniper" and "Iraqi veteran returns in shock".

Overlaid on the right side of the browser window is a Notepad++ window showing the XML code of the RSS feed. The code includes the following elements:

```
<?xml version="1.0" encoding="UTF-8" ?>
<rss version="2.0" ?>
  <channel>
    <title>French troops reported by snipers</title>
    <description>Unarmed sniper has sniped British troops but claims information 'false'</description>
    <link>http://www.bbc.co.uk/news/world_europe_070904.xml</link>
    <pubDate>Wed, 21 Nov 2007 12:00:00 GMT</pubDate>
    <category>Europe</category>
    <image>http://www.bbc.co.uk/news/world_europe_070904.xml</image>
  </channel>
  <item>
    <title>Iraqi veteran home reports shocking reports</title>
    <description>About 1,000 Iraqi ex-servicemen have suffered PTSD</description>
    <link>http://www.bbc.co.uk/news/world_europe_070904.xml</link>
    <pubDate>Wed, 21 Nov 2007 12:00:00 GMT</pubDate>
    <category>Middle East</category>
    <image>http://www.bbc.co.uk/news/world_europe_070904.xml</image>
  </item>
</rss>
```

Mashups / Programmable Web

(The Puzzle of APIs)



Mashups

- A new type of **web applications**.
- Combines data from more than one source into a single integrated tool_[2].
- Creating a new and distinct web service that **was not originally provided** by either source.
- Content is typically sourced **via an API*** or a RSS Feed

Let's see some **examples..**

Mashups (Example)

Ask500People.com

Real-time mashup of votes from around the world



Mashups (Example)

Map of Forbes list of the **top 100 most powerful and best paid celebrities**. Pop-up windows for each includes related YouTube videos



Google Maps
Yahoo Geocoding
YouTube

Mashups (Example)

Pageflakes

Custom homepage builder that uses Ajax to let a user layout and customize a page with relevant information. Includes pre-built modules for Flickr, YouTube, Amazon and hundreds of other popular sites.



- Amazon S3,
- AOL Video,
- Box,
- Clickatell,
- del.icio.us,
- FedEx,
- Flickr,
- Google
- + many others

Mashups (Example)

ePartyGuide.com



Yahoo! Maps-
Yahoo! Local-
Flickr-

Mashups (Example)

Combines Google maps with real-estate databases



Google Maps-
Real-estate-

Mashups (Example)

A unified and comprehensive view of the current global state of infectious diseases and their effect on human and animal health



Google News-
ProMED-
World Health
Organization-

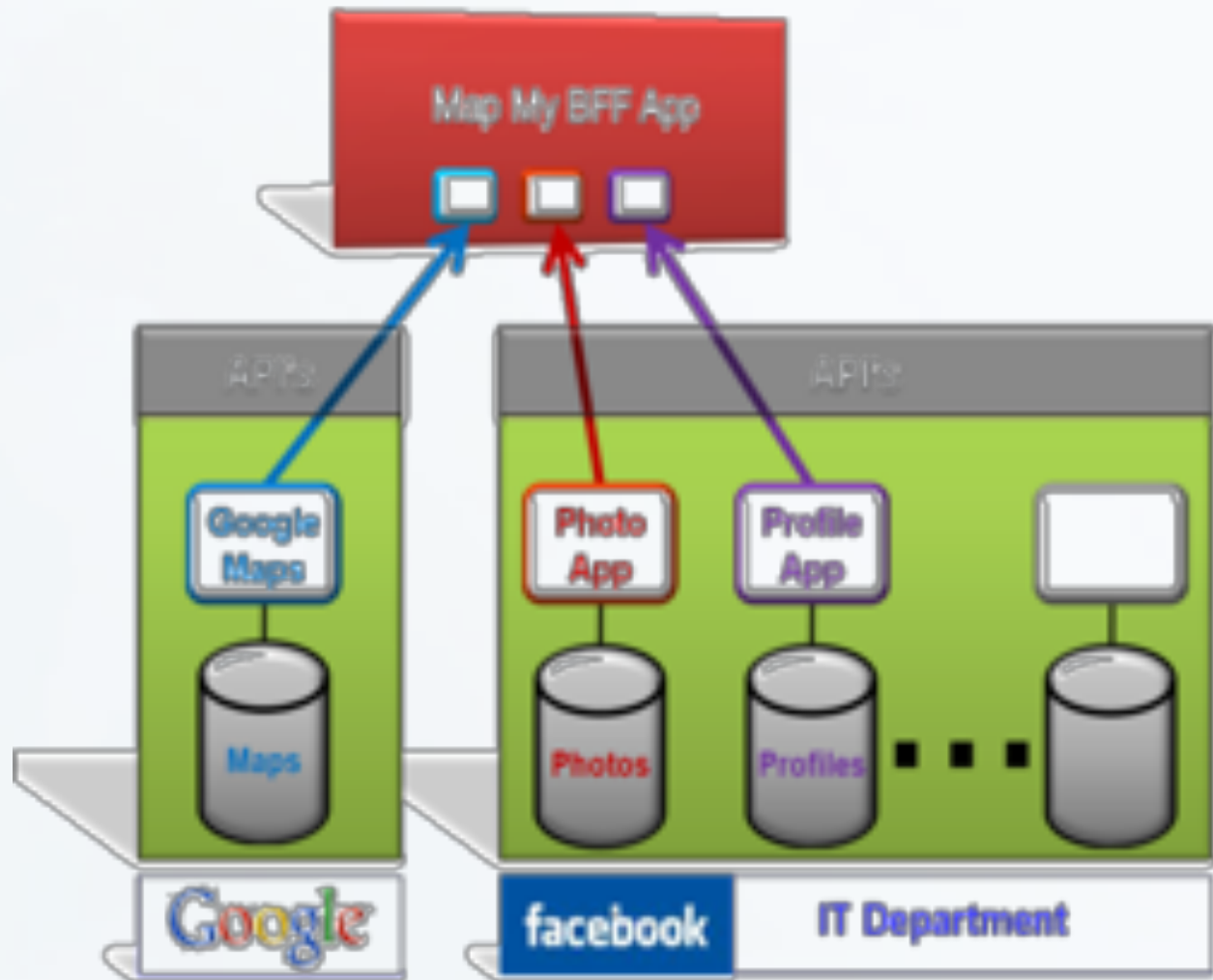
Mashups (Example)

A visually rich application that combines the *Amazon API* to show the relationship between movies, bands, actors, etc.



Amazon API-
You Tube-
Others-

Be Innovative... (The Puzzle of APIs)



Find popular APIs



The screenshot shows the ProgrammableWeb.com website. The logo "ProgrammableWeb" is circled in red. The page is titled "Web 2.0 API Listing" and displays a table of APIs. The table has columns for API name, description, category, and methods. The APIs listed include Google Maps, Flickr, Amazon, YouTube, Microsoft, eBay, Yahoo Maps, and others.

API	Description	Category	Methods
Google Maps	Mapping service	Mapping	120
Flickr	Photo sharing service	Photo	100
Amazon	Online retailer	Shopping	129
YouTube	Video sharing and streaming	Video	100
Microsoft	Mapping service	Mapping	147
eBay	Online auction marketplace	Shopping	104
Yahoo Maps	Mapping service	Mapping	110
LinkedIn	Web 2.0 social networking	Networking	119
Microsoft	Local bookmarking	Bookmarking	61
Google Search	Search service	Search	88
Yahoo Search	Search service	Search	66
Yahoo	Search service	Mapping	66
Search	Search service	Search	66
Search	Search service	Search	66

Find popular APIs here



Also, How-To start a mashup

(API_x + API_y = Money)

... Be innovative to know x and y ;-)

➔ See the “Mashup Market”

The Mashup Phenomena



→ From Web of Documents (Hypertext) into Web of Data

In other words,

The **Web is becoming a Database**, where
a data source is seen as a table, and
a mashup is seen a query over multiple tables/sources.

→ **Data Web**

Web 2.0 challenges

Suppose you want to build a mashup:

Build up your publications “publications webpage”, combining all data from all sources/APIs?

- How to link data items cross sources?
- How to answer structured queries (over different APIs) such as “Give me all people who works-at Birzeit and previously lived in Belgium”?
- How to know the meaning/type of data Items? *Bessan is a person name, or a city name?*

➔ Linking data items cross sources, such that it becomes easy to integrate search, and retrieve (use RDF).

From Web 2.0 to Web 3.0

Next lecture

References

- [1] Mustafa Jarrar, Anton Deik: The Graph Signature: A Scalable Query Optimization Index for RDF Graph Databases Using Bisimulation and Trace Equivalence Summarization. International Journal on Semantic Web and Information Systems, 11(2), 36-65,. April-June 2015
- [2] Mustafa Jarrar, Anton Deik, Bilal Faraj: Ontology-Based Data And Process Governance Framework -The Case Of E-Government Interoperability In Palestine . In pre-proceedings of the IFIP International Symposium on Data-Driven Process Discovery and Analysis (SIMPDA'11). Pages(83-98). 2011.
- [3] Mustafa Jarrar and Marios D. Dikaiakos: A Query Formulation Language for the Data Web. The IEEE Transactions on Knowledge and Data Engineering. IEEE Computer Society. Pages(783-798). Volume 24, Number 4, April 2012
- [4] Paolo Ceravolo, Chengfei Liu, Mustafa Jarrar, Kai-Uwe Sattler: Special Issue on Querying the Data Web -Novel techniques for querying structured data on the web. The World Wide Web Journal. Volume(14), Issue (5-6). Springer. August 2011. ISSN:1573-1413.
- [5] Anton Deik, Bilal Faraj, Ala Hawash, Mustafa Jarrar: Towards Query Optimization for the Data Web - Two Disk-Based algorithms: Trace Equivalence and Bisimilarity. Proceedings of the 3rd Palestinian International Conference on Computer and Information Technology (PICCIT 2010). 2010.
- [6] Mustafa Jarrar, Marios D. Dikaiakos: Querying the Data Web: the MashQL Approach. IEEE Internet Computing. Volume 14, No. 3. Pages (58-670). IEEE Computer Society, ISSN 1089-7801. May 2010.
- [7] Mustafa Jarrar, Marios D. Dikaiakos: Querying the Data Web: the MashQL Approach. IEEE Internet Computing. Volume 14, No. 3. Pages (58-670). IEEE Computer Society, ISSN 1089-7801. May 2010. Mustafa Jarrar and Marios D. Dikaiakos: A Data Mashup Language for the Data Web. Proceedings of LDOW, WWW'09. ACM. ISSN 1613-0073. (2009).
- [8] Mustafa Jarrar and Marios D. Dikaiakos: MashQL: a query-by-diagram topping SPARQL -Towards Semantic Data Mashups. Proceedings of ONISW'08, part of the ACM CiKM conference. ACM. pages (89-96) ISBN 9781605582559.(2008).
- [9] Mustafa Jarrar: Towards methodological principles for ontology engineering. PhD Thesis. Vrije Universiteit Brussel. (May 2005)
- [10] Mustafa Jarrar, Luk Vervenne, Diana Maynard: HR-Semantics Roadmap- The Semantic challenges and opportunities in the Human Resources domain . Technical Report. The Ontology Outreach Advisory, Belgium. (OOA-HR/2007-08-20/v025). August 2007
- [11] Lyndon Nixon, Malgorzata Mochol, Mustafa Jarrar, Stamatia Dasiopoulou, Vasileios Papastathis, and Yiannis Kompatsiaris: Prototypical business use cases. Deliverable D1.1.2 (WP1.1), The Knowledge Web Network of Excellence (NoE) IST-2004-507482, Luxemburg. January 2005.
- [12] Peter Spyns, Daniel Oberle, Raphael Volz, Jijuan Zheng, Mustafa Jarrar, York Sure, Rudi Studer, and Robert Meersman: OntoWeb- a Semantic Web Community Portal. Proceedings of the 4th International Conference on Practical Aspects of Knowledge Management (PAKM 2002). Pages (189-200). LNCS 2569, Springer. ISBN: 3540003142. December 2002.
- [13] "Web Syndication." Wikipedia. Wikimedia Foundation. Web. 16 Feb. 2015. <http://en.wikipedia.org/wiki/Web_syndication>.
- [14] Dengel, Andreas. KI 2008, Advances in Artificial Intelligence 31st Annual German Conference on AI, KI 2008, Kaiserslautern, Germany, September 23-26, 2008 : Proceedings. Berlin: Springer, 2008. Print.