

**Online Communication in the Digital Age**  
**Prof. Rashmi Gaur**  
**Department of Humanities and Social Sciences**  
**Indian Institute of Technology**  
**Lecture – 29**  
**Evolution of Social Software**

Good morning friends. In the previous module, we had looked at multimodality, hypertext and new visualization techniques focusing primarily on multimedia and hypermedia practices. In this module, we will be discussing the technological developments that have led to the socialization of the web. We will analyze most often used channels of communication like blogs, wikis, podcasting as part of the digital rhetoric and also about the new generation of the web that is known as the semantic web. The tools which are labeled as social software have introduced a new dimension to searching and other personal applications on the web. Social software or the social aspects of software become relevant because they have a major impact on how people digitally connect with other groups.

This also brings us to discuss the emergence of new phases of web known popularly as web 1.0, 2.0 and 3.0.

The key difference between web 1.0 and 2.0 is largely the internet. It is easy to get information on 1.0, but it is inconvenient to act as an author in it.

## Social Software and the Rise of Web 2.0

- Web 2.0 marks the convergence of technologies, techniques and usage patterns in the three streams of:
  - **Application Stream:** Brought the number of services on the internet.
  - **Technology Stream:** Provided the underlying infrastructure groundwork for all the advances in technology.
  - **User Participation and Socialization Stream:** Changed the way in which users (private and professional) perceive the web, interact with it and contribute by publishing their own information.
- Web 2.0 is referred to as the 'read/write' web that enables content to be read from the web and content to be written to the web.

“  
The network is opening up some amazing possibilities for us to reinvent content, reinvent collaboration  
”



Tim O'Reilly  
Source: www.wired.com

It was Tim O'Reilly who coined the term 'Web 2.0' in his seminal article "What is Web 2.0"



Web 2.0 primarily marks the convergence of technologies, techniques and usage patterns in the three streams of application, technology as well as active user participation and socialization. It has changed the way in which we perceive the web, interact with it and also contribute by publishing our own information. Web 2.0 is referred to as the read or write web that enables content to be read from the web and we can also add our own content to the websites.

Web 2.0 is designed to be behavior based that adds collaborative dimensions to social networking. They are tools designed for ease of use as well as for the rapid deployment. They are also termed as transparent technologies that have exploited second generation WWW services. In the following video, we will see a brief explanation of how the web evolved from web 1.0 to web 2.0 to have some clarity about this idea. The short video from YouTube briefly explains how web 1.0 is characterized by static web pages where the content is fixed and cannot be easily updated or modified by users. Websites were primarily one way communication platforms built for informational purposes only.

However, we find that web 2.0 has changed it to a much more interactive manner.



Source: [www.youtube.com](http://www.youtube.com)

- Evolution of Internet is divided into three phases- web 1.0, web 2.0, web 3.0 (John Markoff).
- Web 1.0 was primarily read-only.
  - It provided static content rather than dynamic or interactive HTML.
- Web 2.0 is called the Interactive and Social Web.
  - It is the web that we know and experience today.
  - It has made the web more accessible and participant-friendly.

Web 3.0 will also incorporate AI and ML technologies to increase data processing speeds and problem solving

This video talks about the evolution of internet dividing it into different phases. Web 3.0 will also incorporate AI and ML technologies to increase data processing speed and problem solving.

Let us look at the video now. The web has evolved significantly over the years and today's applications are barely recognizable from the early days. The evolution of the internet is often divided into three phases, web 1, web 2, and web 3. What is web 1? Web 1 was the first iteration of the web. Most of the participants were consumers of the content and the creators were usually the developers who created the website containing the information provided primarily in text or image format.

Web 1 lasted from about 1991 to 2004. Web 1 consisted of websites that provided static content rather than dynamic HTML. The data and content was provided by a static file system rather than a database and the website wasn't very interactive. You can think of web 1 as a read-only web. What is web 2? Most of us are primarily experiencing the web in its current form, commonly referred to as web 2.



You can think of web 2 as an interactive and social web. In the web 2 world, you don't have to be a developer to participate in the build process. Many apps are built to be easy for anyone to create. If you have an idea and want to share it with the world, you can. If you want to upload a video and have millions of people watch it, interact with it, and comment on it, that's possible.

Web 2 is really simple and because of its simplicity, more and more people around the world are becoming creators. When the video transitions to web 2.0 advancement, there is a stark difference in the visual appeal of the web. It has become dynamic and is seen used by multiple users.

In 2.0, users are not only consumers, but are also creators of content. So to say, content has become decentralized, allowing even non-technical users to easily create and publish their own content. And it is this context which is particularly useful for those scholars who are working in the field of media, communication, and also literature. It is said that we cross the threshold to the third generation web in 2007.

Web 3.0 is known as the intelligent web as it incorporates AI and ML technologies to increase the speeds of data processing, etc. Examples can be seen in blockchains as well as in cryptocurrency payments. However, we have to admit that these terms lack in specificity. Web 2.0 is a phrase coined by O'Reilly Media in 2004 and Web 3.0 is a phrase coined by John Markoff of the New York Times in 2006. As we have discussed earlier, the essence of Web 2.0 lies in utilizing and appropriating data and data streams for deriving new information or added value. It is a service-oriented approach to create rich internet applications known primarily as RIAs. At the same time, we find that tagging, blogging, and 'wiki-ing', etc. are known as important steps into the socialization of the web and have become immensely popular amongst the people.

- Tim O' Reilly defined web 2.0 as the network that delivers software as a continually updated service that gets better as more people use it. He called this phenomena as 'perpetual beta'.
- The essence of web 2.0 is:

- Appropriating and utilizing **data and data streams** for deriving new information or added value.
- **Functionality:** Service-oriented approaches to create Rich Internet Applications (RIAs).
- Tagging, blogging and 'wiki-ing' as important steps into the **socialization** of the web.



Harnessing Collective Intelligence is the driving force of web 2.0  
Sources: www.Wikipedia.com

O'Reilly had defined Web 2.0 as a network that delivers software as a continually updated service that gets better as more people use it. He called this phenomena as perpetual beta. Wikipedia and blogs are typical examples of Web 2.0. The figure in the slide shows blogging sites as putting more emphasis on data and socialization. While Wikipedia shows advanced functionality, Web 2.0 technologies such as blogs, wikis, podcasts, and feeds are commonly referred to as social software as they are characterized by a high degree of connectivity, affording users an opportunity to collaboratively develop the web content. A term which is very popular nowadays is a social software.



- The term 'social software' came into use in 2002 and is generally attributed to the writer Clay Shirkey who simply defined it as "software that supports group interaction".\*
- It brings another dimension to searching and web-based applications by allowing users to add personal search logic to machine-based algorithms.
- Social software gets better as more people use it, that is, they can be perpetual beta sites.
  - For example, Gmail, Google Maps Google PageRank etc.
- Software as a Service (SaaS) paradigm combines community and subscription models while social software combines transactions, advertising and community models.

\*Clay Shirkey- *Social Software and the Politics of Groups*



Wikis, podcasts, blogs are web 2.0 technologies dubbed as social software  
 Source: colorlib.com, www.esquire.com, educationaltechnology.net

The term had come into use in 2002 and is normally attributed to the writer, Clay Sharkey who simply defined it as software that supports group interaction.

It has brought an altogether new dimension to searching in web-based applications by allowing us to add personal search logic to machine-based algorithms. As we have discussed earlier, social software gets better as more people use it. Software as a Service paradigm combines community and subscription models while social software combines transactions, advertising, and community models. For example, Discord, a communication platform primarily focused on gaming communities has been known to operate in a perpetual beta state. It regularly introduces new features, improvements, and changes based on user feedback, creating an evolving and dynamic user experience.

Before coming to user participation and socialization of the web, let's have an overview of the web languages which are used for web development. These languages are the backbone for creating the content and the structure of the web pages.



## Developments in Web Technology

- Webpages are commonly described using **Hypertext Markup Language (HTML)**. It uses 'markups' or 'tags' to transport information about the structure and presentation of a page.
  - Hypertext Network Analysis (HNA) is a social network analysis that looks at how websites form relations with one another via hyperlinks.
- The **Extensible Markup Language (XML)** is seen as a 'metalanguage' or a language framework used for specific applications such as news, business, biology etc.
- **Cascading Style Sheets (CSS)** can be applied with both HTML and XML to style and format the visual presentation of web pages.

XML (Extensible Markup Language)	HTML (Hypertext Markup Language)
It stores and transports data.	It displays data.
It uses user-defined tags.	It uses predefined tags.
It contains structural data.	It doesn't contain any structural data.
It can distinguish uppercase and lowercase letters (case sensitive).	It can't distinguish uppercase and lowercase letters (case insensitive).
It maintains spacing, tabs, newlines, and any other whitespace formatting.	It doesn't maintain whitespace.
It needs to have an end-tag.	It doesn't need an end-tag.
It needs structure or nesting.	It doesn't need structure.

The Table shows the differences between XML and HTML  
Source: [www.000webhost.com](http://www.000webhost.com)



There are primarily two languages known as HTML and XML. Also we would talk about CSS or Cascading Style Sheets. In the table on the right hand side, you would find that the major differences between HTML and XML have been explained.

So HTML uses markups or tags to transport information about the structure and presentation of a page. Hypertext network analysis or HNA is a social network analysis that looks at how websites form relations with one another via hyperlinks. XML is seen as a meta language or a language framework which is used for specific applications such as news, business, biology, etc. Cascading style sheets can be applied with both these languages to style and format the visual presentation of the page. The competing browsers of Mozilla, Opera, Microsoft, and others have over time to come to define their own tags like HTML to resolve issues or just in the hope of gaining competitive advantage through the support of a richer set of features.

When it comes to XML, the applications extend to web servers and web feeds.

- **Web server** is a stand-alone software component with a unique Uniform Resource Interface (URI) to communicate and exchange data over the internet.
- **Web feed** is a data format used for publishing frequently updated content to interested users. The most commonly used **formats** for web feeds are RSS (Really Simple Syndication) and Atom.
- **Really Simple Syndication (RSS)**
  - It is a web format used to publish web content like blogs, news, etc.
  - RSS aggregator can show multiple web sources in one place.
  - RSS reader checks the original sources regularly and downloads the updates.
  - It enables the readers to keep up to date with a whole series of sources of information without having to check it again.



BBC Sport offering RSS Feeds  
Source: news.bbc.co.uk



Web server is a stand alone software component with a uniform and unique resource interface known as URI to communicate and exchange data over the internet. Web feed is a data format used for publishing frequently updated content to interested users and it is this aspect which is particularly useful for scholars in our areas. The most commonly used formats for web feeds are RSS and Atom. RSS stands for Really Simple Syndication which is a web format used to publish web contents like blogs, news, etc.

RSS aggregator can show multiple web sources in one place and RSS reader checks the original sources regularly and can download the updates. In this way, it enables the reader to keep up to date with a whole series of sources of information without having to check it again repeatedly. The images given show BBC Sports RSS feeds to allow users to stay updated with their latest content. By subscribing to feeds, users can receive automatic updates whenever new articles or content are published. Once we have subscribed to the desired BBC RSS feeds, we can use a feed reader or aggregator software to manage and view the updates.

Atom was designed as an alternative and improvement to the limitations of RSS and we can also combine RSS feed with platforms like Twitter. We can tweet our headlines by



using an application called Twitter feed to push our blog's RSS feeds directly to Twitter. We will discuss more about microblogging and Twitter in the coming modules.

- **Peer to Peer (P2P)** network is a distributed network architecture in which participants (referred to as peers) share resources and directly interact with each other without the need for a centralized server or authority.
  - Unlike HTML and XML, it does not follow a client/server software paradigm that initiate conversation by sending requests to servers and accepting requests from clients.
  - It uses a decentralized framework like Bitcoin (cryptocurrency), BitTorrent Sync (Communication Protocols) etc. (This has different legal aspects).
- **Rich Internet Applications (RIA)** are web-based applications that provide enhanced functionality and user experiences similar to traditional desktop applications.
  - They utilize a combination of technologies, such as HTML, CSS, JavaScript, to deliver a rich and dynamic user interface.



P2P or Peer-to-Peer network is a distributed network architecture in which participants who are referred to as peers share resources and directly interact with each other without the need for a centralized authority or survey. Unlike the previously discussed languages, HTML and XML, it does not follow a client or server software paradigm that initiates conversation by sending requests to servers and accepting requests from clients.

It uses a decentralized framework like Bitcoin, BitTorrent Sync, etc. However, we have to be aware that in different societies and countries, there are different legal frameworks to support and put certain balances against it. Rich Internet Applications or RIA are web-based applications that provide enhanced functionality and user experiences similar to traditional desktop applications. Utilizing a combination of technologies such as HTML, CSS, JavaScript, etc., they deliver a rich and dynamic user interface.

Popular examples of RIA include streaming platforms like Netflix and Spotify, online




marketplaces like Airbnb and note-taking applications like Evernote. All these applications provide rich user interface with features like personalized recommendations, playlist creation and text formation. People have got used to the web as a medium of communication as well as a medium of socialization. Also, we use it as a business platform, discussion forum and as a constantly growing and expanding encyclopedia. Now, let us see the ways in which web has been increasingly used as a medium and also about the user participation and contribution when it comes to the socialization of the web.

## Wikis as Online Collaborative Tools

- Wiki is a webpage or a collection of pages that allows its users to add, remove or generally edit some of the available content.
- The first site to use the term was 'WikiWikiWeb' created by Howard G. Cunningham.
- The best known wiki is the collaborative encyclopedia known as Wikipedia.
- Wikis fall into three categories:

Free Wiki Services	Fee-Based Services	Self-Hosting
No cost, quick start-up, easy to monitor, minimal technical experience required, less security	Advanced management capabilities, added security, easy to add and link groups, cost involved	Allow maximum amount of control, must have own server, require end-user training



The History of Wikis Source: [www.wikipedia.com](http://www.wikipedia.com)


All of us are familiar with Wikis and we know that they are online collaborative tools. All of us are very well familiar with the best known Wiki, Wikipedia. Wikis fall into three categories. They are free wiki services, no cost, etc., quick startup and a person with minimal technical know-how can easily use and operate it.

There are fee-based services which require advanced management capabilities, but at the same time, they have certain added security features, but of course, certain cost is also involved. The third category is that of the self-hosting category which allows maximum amount of control and for using it, one should have one's own server. They also require end-user training. The first site to use the term as Wiki Wiki web was


created by Howard Cunningham. So Wiki is an editable web page or a collection of pages that does not require much technical competence or control over any computer language like HTML.

However, before selecting a particular Wiki, it is important to be familiar with the features that distinguish one Wiki service from another and also the prerequisites for setting up the Wiki.

- Wikis can either be internal or external.
  - External wiki is visible to anyone on the web without having to log in, while Internal wiki is visible to only those who have password access to it.
- Setting up the wiki involves three basic steps:
  - **Setting a wiki service** like WetPaint or Wikispaces.
  - **Determining the wiki's URL.**
  - **Inviting contributors.**
- Mediawikis
  - It is a free and open-source wiki software application specifically designed for collaborative editing using media elements like images.
  - One of the key aspects of MediaWiki is its ability to create an interconnected network of articles through internal links.
  - It supports the use of templates and consistent formatting across multiple pages through hyperlinked articles.



Wikipedia, WikiHow, fandom (wikia) are all examples of Mediawiki  
Sources: www.Wikipedia.com



10

Wikis can be either internal or external. For example, external Wiki is visible to everyone on the web without having to log in where internal Wiki is visible only to those who have an access through password. Setting up the Wiki involves three basic steps which include setting a Wiki service, determining the Wiki's URL and inviting contributors. Media Wikis are a free and open source Wiki software applications specifically designed for collaborative editing using media elements like images.

One of the key aspects of Media Wiki is its ability to create an interconnected network of articles through internal links. It supports the use of templates and consistent

formatting across multiple pages through hyperlinked articles. Now let us get an overview of how Wikis and Wiki clones are used in the professional and corporate world through specific instances.

## The Case of 'Twiki'

- Derivative implementations of the Wiki concept are called Wiki clones.
- 'Twiki' is a wiki clone aimed at corporate and professional use as a knowledge base for technical support.
- It helped to increase the information flow between offices and it was deployed for customer support.
- Twiki was used in Motorola Systems-on-Chip Design Technology as a team communication tool on the corporate intranet.
- The different ways in which the wiki navigated the multi-web environment are as newsgroup, meetings' calendar, issue lists, rapid production of web pages etc.



Source: <https://www.unixmen.com/install-twiki-collaboration-platform-centos-6-5/>



We will discuss the case of Twiki which is a Wiki clone aimed to corporate and professional use as a knowledge base for technical support. It has helped to increase the information flow between offices and was basically deployed for customer support.

It was used in Motorola Systems-on-Chip design technology as a team communication tool on the corporate intranet. The different ways in which the Wiki navigated the multi web environment are as news groups, meetings, calendar, issue lists, rapid production of web pages, etc. Social software like Wikis enable the creation of communication and allow people to contribute more knowledge or correct details in existing knowledge for which they are experts. Interestingly it is also for this reason that they can never be considered as a primary or a very authentic source. Wikis are collaboratively edited by multiple users which means that anyone can contribute and also can modify the content.

This lack of centralized control raises concerns about the accuracy and reliability of the



presented information and there is no guarantee that the information given on a Wiki page has been either fact checked or verified by experts in the field. Wiki pages can change frequently and the content is also continuously evolving. Maintaining quality control therefore is a challenge especially for smaller or less actively monitored Wikis. However, quality control is needed to minimize misinformation and biased content, providing users with trustworthy and credible information. But they still have a learning curve especially for users who are unfamiliar with the Wiki software or editing conventions.

User friendly interfaces and community support which is available to Wikis provide immediate feedback showing the results of edits in real time. This quick feedback loop helps newcomers gauge the impact of their contributions and learn from any mistakes they might have made. It is important to consider the two sides of the coin when it comes to Wikis. Just like Wikis, weblogs or blogs are communication forms that have contributed significantly to the read and write nature of the web. We shall have a closer look at blogs as a user generated content and how it is significantly different from Wikis.

## Blogs and Wordpress

- Blogs are expressions of personal or professional opinion or experience on which other people can at most comment.
- Blogs differ from wikis as the latter allow users to extend, modify, update or delete content.
- Blogs can be external or internal.
  - Internal blogs are hosted within a larger website or platform which is maintained by the organization that owns the website.
  - External blogs are standalone entities that operate independently from any specific website or organization for a broader audience.
- Wordpress is the open-source blogging software that is typically used as personal publishing platform.



Sources: content.time.com





Primarily, blogs are expressions of personal or professional opinion or an individual perspective on which other people can at best comment. They differ from the Wikis as the Wikis allow users to extend, modify, update or even to delete content. Blogs can also be external or internal. Internal blogs are hosted within a larger website or platform which is maintained by the organization that owns the website. External blogs are standalone entities that operate independently from any specific website or organization for a broader audience.

WordPress is the open source blogging software that is typically used as personal publishing platforms and is often used by young authors, poets and story writers, etc. The images show popular examples of blogs like HuffPost, TechCrunch and Lifehacker. The blog is a good example of how an enterprise can develop new strategies for communication. For example, HuffPost is a news and opinion website that mainly covers topics like politics, entertainment and lifestyle. TechCrunch is a leading technology blog that focuses on the latest trends, news and analysis in the tech industry.

The degree of interactivity in a blog mainly depends on the blog publishing software. Let us take WordPress as an example.

- Wordpress is a **self-hosted software** that is installed and operated on a user's own server or hosting environment, with which you edit online.
- It uses the **dynamic publishing method** which enables you to update your website's content on without manually editing HTML or other code.
- It is a **content management system** that supports blog features like ping notifications, podcasting, permanent links and multiple authors and registered visitors with the privilege to comment.
- It provides themes and plugins that allow users to customize the design, layout and functionality of their websites such as e-commerce, Search Engine Optimization (SEO), social media integration etc.



Designer services and themes on Wordpress.  
Source: [www.wordpress.com](http://www.wordpress.com)



WordPress is a self-hosted software that is installed and operated on a user's own server or hosting environment with which you edit online. It uses the dynamic publishing method which enables you to update your website's content on without manually editing HTML or any other code. It is a content management system that supports blog features like ping notifications, podcasting, permanent links and multiple authors and registered visitors with the privilege to comment.

It also provides themes and plugins that allow users to customize the design, layout and functionality of their websites such as e-commerce, search engine optimization, social media integration, etc. We will learn more about social media integration and microblogging in the coming modules. Let us now look at another communication tool of social software that is podcasts. Podcasts have been used as a means of syndicating audio material in a variety of ways and have become immensely popular with the young people. Although podcasts are restricted in their use of multimedia elements, they are widely used as a venue for reaching out to users and connecting them with the information that they might seek.

## Podcasts

- Podcasts are digital audio or video files that are available for streaming or downloading online.
- They are typically episodic in nature, with new episodes being released regularly and cover a wide range of topics, including news, storytelling, interviews, education, entertainment, and more.
- Podcast production involves:
  - ✓ Outlining the structure and flow of episodes.
  - ✓ Recording using platforms like 'Zencastr' and softwares like 'Garageband' and 'Adobe Audition'.
  - ✓ Editing in the desired audio format (like MP4).
  - ✓ Choosing podcasting hosting platforms like 'Podbean' and submitting it to platforms like 'spotify' using the RSS feed.



So, podcasts are basically digital audio or video files that are available for either streaming or downloading. They are typically episodic in nature with new episodes being released regularly and they may cover a wide range of topics including news, storytelling, interviews, education, entertainment, personal lives, etc. and many more. Podcast production involves outlining the structure and flow of episodes, recording using platforms like Zencastr and softwares like GarageBand and Adobe Audition. Editing in the desired audio format is also required and choosing podcasting hosting platforms like Podbean and submitting it to platforms like Spotify using the RSS feed.

The pictures show popular podcasting platforms targeted towards potential users covering different topics. Now let us take some examples of popular podcast as case studies to understand the modes in which discourses are conveyed through them. This includes popular podcasts like the Joe Rogan Experience, the Michelle Obama Podcast and Serial.

			<p>Sources: <a href="http://www.powerfulJRE.com">www.powerfulJRE.com</a>, <a href="http://www.spotify.com">www.spotify.com</a>, <a href="http://www.serialpodcast.org">www.serialpodcast.org</a></p>
<ul style="list-style-type: none"> <li>• His audience is largely young, male audiences that advertisers covet.</li> <li>• His guests are usually provocative that appeal to young people who do not support the mainstream politics.</li> <li>• His style is a blend of comedy, controversy and ideological whiplash.</li> <li>• He has become a mass cultural product with over 11 million listeners per episode.</li> </ul>	<ul style="list-style-type: none"> <li>• It is an investigative journalism podcast.</li> <li>• The narrative consisted of exploring various aspects, witness testimonies, and evidences.</li> <li>• Ethics of crime podcasts can be questionable as they appear to promote social justice but may also utilize media tactics.</li> </ul>	<ul style="list-style-type: none"> <li>• Michelle Obama interviews celebrities, close friends and family members.</li> <li>• There is a characteristic combination of warmth and reason.</li> <li>• She creates a sense of closeness with her audience through her skills as a public figure.</li> </ul>	
 IIT ROORKEE  NPTEL ONLINE CERTIFICATION COURSE		15	

In the Michelle Obama Podcast, we find that there is a characteristic combination of warmth and reason. She interviews celebrities, close friends and family members and creates a sense of closeness with them providing a certain warmth and charm during her podcasts.

Serial is an investigative journalism podcast and as we can easily imagine the narrative consists of exploring various aspects like witness testimony and evidence etc. However, ethics of crime podcasts can sometimes be questionable as though they promote social justice, they also utilize different media tactics to gain popularity. The Joe Rogan Experience is a blend of comedy, controversy and ideological whiplash. His audience is chiefly young male audience that advertisers covet. His guests are usually provocative, they appeal to the young people who do not support the mainstream politics.


He has become a mass cultural product with over 11 million listeners per episode. What we have to understand here is that most podcasts employ the business model of finding an ideological space connecting via cross-promotion and guest selection with similar shows and allowing the algorithms of social media, finding the target audience is important to drive traffic their way. Another social software that has introduced new

dimensions to searching and other applications on the web are tags and social bookmarking.


It had an interesting impact on online search and let us see how.

## Tagging

- Tagging refers to the practice of assigning descriptive keywords or labels (tags) to web pages, online content, or elements within a webpage like photos or links.
- Tagging can be categorized into two practices based on the user incentives:
  - Organizational: Used as an alternative to structured filing for personal purposes using common tags created by others.
  - Social: Expresses the communicative nature of tagging where the tags are by-products of opinion, self-expression, play and competition.
- There are various dimensions to tagging. For example,
  - Tagging rights: Self-tagging, permission-based, free-for-all.
  - Type of object: Textual (like blog posts), non-textual (like images).



Sources: [www.techcrunch.com](http://www.techcrunch.com)



16

Tagging refers to the practice of assigning descriptive keywords or labels to web page, online content or elements within a web page like photos or links. Tagging can be categorized into two practices based on the user incentive. They can be organizational or social. The names or the titles are self-explanatory. Organizational are used as an alternative to structured filing for personal purposes using common tags created by others.

Social expresses the communicative nature of tagging where the tags are byproducts of opinion, self-expression, play and competition. There are various dimensions to tagging. For example, the tagging rights can be permission based, free for all or self-tagging. The types of objects can be either textual or non-textual. Other dimensions include tagging support which decides who can view the tags associated with the resource and also type of connectivity that is whether to link resources or to connect socially.





However, in recent years the popularity of Flickr has declined and we find that social media platforms like Instagram have grown in popularity. As they have better streamlined photo sharing experiences. Now let us talk about social bookmarking and folksonomies.

## Social Bookmarking and Folksonomies

- Social bookmarking allows users to save, organize, and share bookmarks or web links to various online resources, such as websites, articles, blog posts, images etc.
  - Classifies content by means of user-created tags as well as AI algorithms.
- The integration of social bookmarking and tagging, in turn, led to the idea of 'folksonomy'.
  - It is a user-generated taxonomy admitting an unrestricted vocabulary for tagging objects.
  - They characterize a bottom-up consensus on the basis that a system of largely agreed upon terms evolves from collaborative tagging, which coexists next to personal and minority categories.
  - Another defining characteristic is 'flat-name space', causing it to develop into a non-hierarchical and non-exclusive collection of tags.



Social Bookmarking Sites  
Source: [www.onaplatterofgold.com](http://www.onaplatterofgold.com)

Social bookmarking allows users to save, organize and share bookmarks or web links to various online resources such as websites, article, blog post, images etc.

It also classifies content by means of user created tags as well as AI algorithms. The integration of social bookmarking and tagging in turn led to the idea of folksonomy. Folksonomy is a user generated taxonomy, an arrangement of information into groups admitting an unrestricted vocabulary for tagging objects. They characterize a bottom-up consensus on the basis that a system of largely agreed upon terms evolve from collaborative tagging which co-exist next to personal and minority categories. Another defining characteristic is flat-name space causing it to develop into a non-hierarchical and non-exclusive collection of tags. Folksonomies represent an important development



in the direction of creating a common understanding agreement and maybe even trust within communities on the web.

Let us now consider some examples of web service and platforms that fall under search engine optimization or SEO services. We will look at the examples of web services like Delicious, Plurk and Pinterest.

**del.icio.us**

- First web service to integrate combination of social bookmarking and tagging.
- It offered 'stacks' feature that allows users to group related bookmarks into collections.
- It had the ability to follow other users and browse their bookmarks.

**Pinterest**

- Pinterest users create 'boards' which they can save and organize by topic, theme or interest.
- It can be a powerful tool for driving traffic to other websites by clicking on the links in their boards.
- It provides tools to track the performance of their 'pins' and boards.

**PLURK**

- Plurk is a social networking and microblogging platform that combines elements of social bookmarking.
- It has timeline-based interface that displays users' posts in a horizontal timeline.
- It allows users to organize their content using tags.

[www.mashable.com](http://www.mashable.com), [www.pinterest.com](http://www.pinterest.com),  
[www.commonswikimedia.org](http://www.commonswikimedia.org)

19

So the details of these services are mentioned on the slides since they are very popular sites and almost everybody is aware of them, I will skip the details. If there are any further questions about it, it may be asked during the discussion session. Many other websites like Twitter and Reddit offer approaches under the social bookmarking umbrella to curate, tag and share content with the added benefit of search engine optimization.

Data and data streams are one of the core features of the web today and this is mainly because it shows the monetary and commercialized side of the media materialized in the form of electronic commerce or e-commerce which we have already discussed in the previous module. Data is generated automatically in many applications and is collected

on purpose by applications such as search engines or satellites, sometimes even by several private players including the dark web. This brings us to the topic of data ownership and of course several other related problems which we shall be discussing in the next module.

## Data Ownership

- Data can be human-generated where the data is collected by humans and processed by machines or it can be machine-generated where data often arises by itself and is simply collected and stored on the servers of a company or a provider.
- Data ownership and collection is used for creating or adding strategic value. This includes:
  - Creation strategies for creating and owning data that is difficult to re-create.
  - Access strategies provide access to data that was previously difficult to find or get to.
- Data ownership implies responsibilities like security management, data quality and management of business rules.



When we talk of data ownership, we find that data can be human generated where the data is collected or created by humans and uploaded and processed by machines or it can be machine generated where data arises by itself and is simply collected and stored on the service of a company or a provider. Data ownership and collection is used for creating or adding strategic value which includes creation strategies as well as access strategies for creating and owning data and also to provide access to data that was previously difficult to find.

Data ownership implies responsibilities too like security management, data quality and management of business rules as well as following certain ethical considerations. For example, in a system that contains sensitive data like confidential information or corporate intelligence, it is necessary to define a security and authorization policy for its enforcement. The pictures shown depict Google Street View as a service that enables



creation strategies by owning data that is difficult to recreate. This is because of the extensive geographic coverage, technological infrastructure and privacy as well as legal considerations. For a search to be perfect, it needs to be ubiquitous that is it should search everywhere on the web taking also the personal intentions and preferences into account and this is the point where the semantic web enters the picture.

## The Semantic Web

- The semantic web is an extension of the current web that aims to enable machines to understand and process the meaning (semantics) of information on the web.
- Semantic web starts with RDF (Resource Description Framework), a standard data model that provides a flexible way to describe resources, their properties, and relationships in a machine-readable format.
- In the context of semantic web, ontology refers to a formal representation of knowledge that defines the concepts and enables machines to understand and reason about the meaning of data and information on the web.
  - They are represented using Ontology languages like Web Ontology Language (OWL).



Source: [www.inf.ed.ac.uk](http://www.inf.ed.ac.uk)



swajani



So, what is the semantic web? It is an extension of the current web that aims to enable machines to understand and process the meaning or the semantics of information on the web. Semantic web starts with RDF that is Resource Description Framework, a standard data model that provides a flexible way to describe the sources, their properties and relationships in a machine readable format. In the context of semantic web, ontology refers to a formal representation of knowledge that defines the concepts and enables machines to understand and reason about the meaning of data and information on the web. Let us listen to a video about the semantic web to have a clearer idea about it before we get into its details.

The video is a one minute series on YouTube whose content mainly focuses on the web



3.0, Metaverse, blockchain and machine learning. This video explains the goal and benefits of semantic web.



Source: <https://youtu.be/EC46HwEgVp0>



Swajani



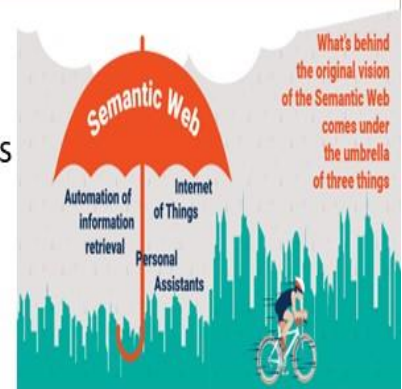
The semantic web is a vision about how to make the world wide web even better by giving software programs machine readable metadata and data about the information and data that is already on the internet. Computers would be able to use the semantic web, which would add more information about what is on the web, in the same way humans use information to make sense of things and achieve our goals. The goal of the semantic web is for computers to be able to manipulate information on our behalf.

Tim Berners-Lee came up with the idea for the semantic web and he says that the word schematic means that a machine can do something with the data. In this way the semantic web can be seen as a knowledge graph that is made up of connected, linked data and intelligent content. This allows machines to understand and process content, metadata and other types of information at a large scale. Customer experiences will get better and easier thanks to the semantic web. This is because the web will be able to understand and show itself in the best way possible for each customer's unique needs.

Semantic web standards could be the key to a big change in the way the web becomes smarter. This means that the content we post online could be shown in a way that machines can understand, connect and combine.

## The Semantic Web

- Semantic Web enables web content to be linked in a way that connects related information across various websites and data sources.
- This structured data representation make it easier for machines to understand the context and meaning of information.
- The primary goal of the Semantic Web is to enable machines to process and understand the content on the web in a way that goes beyond simple keyword-based searches.



Source: [www.ontotext.com](http://www.ontotext.com)




The semantic web enables web content to be linked in a way that connects related information across various websites and data resources. This structured data representation makes it easier for machines to understand the context and meaning of information. The primary goal of the semantic web is to enable machines to process and understand the content of the web in a way that goes beyond simple keyword based searches. The video clarifies that semantic web can be understood as a smarter and better organized version of the web.

Just like we use tags and categories to organize things in real life, the semantic web uses special codes and functions to organize information on web pages. It helps search engines and AI assistants to better understand the content and connections between different pieces of information. Widespread adoption of the semantic web has been slower than initially envisioned and its full realization still remains a work in progress.




There are also search engines that use semantic web technologies to improve search results and provide more contextually relevant information like ‘Swoogle’ and ‘Watson Discovery’.

However, there is a gap between semantic ontologies and Web 2.0 developments like blogs and web pages. Let us now look at the ways to overcome this and some of its applications.

- Semantic web acts as an “upper” logical layer that provides the semantics while web 2.0 acts as a “lower” physical layer that provides the syntax and technology.
- ‘Micro-formats’ and ‘collaborative tagging’ are concepts that aim to bridge the gap between the two.
  - Micro-formats embed structured data within HTML to identify content with semantic meaning in webpages. They aim to balance human-readable content and machine-readable data.
  - Collaborative tagging is also a method of categorizing and organizing information on the web through user-generated tags.
- The BBC created a Semantic Web-based website for the 2010 FIFA World Cup. The basis of the system was an ontology that described how world cup facts related to each other.
  - The publishing platform had both manual and automated tagging features.



BBC World Cup Website Showcased Semantic Technologies  
Sources: readwrite.com

24

Semantic web acts as an upper logical layer that provides the semantics while Web 2.0 acts as a lower physical layer that provides the syntax and technology.

Microformats and collaborative tagging are concepts that aim to bridge the gap between the two. We have given the example of how BBC has created a semantic web based website for the 2010 FIFA World Cup and this example can be seen on the right hand side of the slide. It had incorporated semantic technologies to gather and present information from various sources in a unified manner allowing users to explore team statistics etc. in a better structured way. Similarly, in the case of blogs, microformats can even introduce structure to an otherwise unstructured text.

## Conclusion

- Social Software covers the aspect of user-generated content as it occurs in blogs or wikis, in tagging as well as in social bookmarking.
- Socialization also encompasses the software and the application that the software is providing.
- Search engine providers offer services for a personalized search that aims at improving search results by taking the empirical information about the user's search behavior into account.
- Paradigms like Software as a Service (SaaS), social software, data ownership combine community models with the business models to accumulate profit.

So, we can see that the social software covers the aspect of user generated content as it occurs in blogs or wikis in tagging as well as in social bookmarking.

Socialization also encompasses the software and the application that the software is providing. Search engine providers offer services for a personalized search that aims at improving search results by taking the empirical information about the user search behavior into account. And paradigms like SAS social software combine community models with the business models to accumulate profit. The contemporary modes of artificial intelligence like the chatbots are also an evolved form of social software. So, in this module we have looked at the evolution of social software. We will discuss about the more contemporary modes of social network and their enhancement in the coming module.

In the next module, we will focus on the central role of artificial intelligence as part of communication technologies. We will look at important breakthroughs, perspectives, case studies, robotic communications and critical aspects related to it. Thank you.