

# TR-102

## MASTERING THE SEMANTIC WEB

### DAY-5

#### ❖ Evolution from Web 1.0 to Web 3.0

The evolution of the web can be divided into three major phases: Web 1.0, Web 2.0, and Web 3.0. Each phase represents a significant shift in the way the internet is used and experienced.

##### ➤ Web 1.0

**Time Period:** Early 1990s to early 2000s

**Characteristics:**

- **Static Pages:** Content was mostly static and displayed in a read-only format. Websites were collections of static pages with fixed content.
- **HTML and HTTP:** Web pages were primarily built using HTML and accessed through the HTTP protocol.
- **Limited Interaction:** User interaction was minimal, typically limited to navigating through hyperlinks.
- **Information-Centric:** Websites were focused on delivering information rather than providing interactive experiences.
- **Personal Websites:** The web was dominated by personal websites and basic web portals.
- **Read-Only:** Users were consumers of content rather than contributors.

**Examples:** Early Yahoo, GeoCities, and static company websites.

##### ➤ Web 2.0

**Time Period:** Early 2000s to present

**Characteristics:**

- **Dynamic Content:** Websites became more dynamic, with content generated in real-time based on user interactions.
- **User-Generated Content:** Platforms encouraged users to create and share their own content, leading to the rise of blogs, social media, and wikis.
- **Interactivity:** Enhanced interactivity through technologies like JavaScript, AJAX, and rich media.
- **Social Media:** Social networking sites like Facebook, Twitter, and YouTube became central to the web experience.

- **Collaboration:** Emphasis on collaboration and sharing, with platforms like Wikipedia and Google Docs.
- **Read-Write Web:** Users could both consume and contribute content.

**Examples:** Facebook, YouTube, Twitter, Wikipedia, and WordPress.

## ➤ **Web 3.0**

**Time Period:** Emerging (Late 2010s and beyond)

### **Characteristics:**

- **Semantic Web:** The web is more intelligent, capable of understanding and interpreting data in a way that is meaningful to both humans and machines. Data is linked and contextualized.
- **Artificial Intelligence:** Integration of AI and machine learning to provide personalized experiences and intelligent data processing.
- **Decentralization:** Increased use of blockchain technology for decentralized applications and data storage.
- **Interoperability:** Seamless integration and interoperability between different platforms and devices.
- **Enhanced Privacy and Security:** Focus on user privacy and security through encryption and decentralized networks.
- **Ubiquitous Connectivity:** Enhanced connectivity through the Internet of Things (IoT), enabling devices to communicate and share data.

**Examples:** Decentralized applications (dApps), blockchain platforms like Ethereum, AI-driven platforms, and IoT applications.

## ❖ **Differences between Web 1.0 to Web 3.0**

Feature	Web 1.0	Web 2.0	Web 3.0
<b>Content</b>	Static	Dynamic	Semantic and contextual
<b>User Interaction</b>	Limited	Extensive (social media, blogs)	Intelligent (AI, machine learning)
<b>Data Control</b>	Centralized	Centralized with user contribution	Decentralized (blockchain)
<b>Technology</b>	HTML, HTTP	AJAX, JavaScript, Web 2.0 tools	AI, blockchain, IoT
<b>User Role</b>	Consumer	Consumer and contributor	Active participant with control over data
<b>Focus</b>	Information delivery	Social interaction and collaboration	Intelligence, decentralization, privacy
<b>Examples</b>	Early Yahoo, GeoCities	Facebook, YouTube, Wikipedia	Ethereum, decentralized apps (dApps)

## ❖ URL, URI and URN

### ➤ URL (Uniform Resource Locator)

**Definition:** A URL is a specific type of URI that provides a means to access a resource on the internet by specifying its location.

**Structure:**

- **Scheme:** Protocol used to access the resource (e.g., http, https, ftp).
- **Host:** Domain name or IP address of the server.
- **Path:** Specific location of the resource on the server.
- **Query:** Optional parameters to pass to the server.
- **Fragment:** Optional reference to a specific part of the resource.

**Example:**

`https://www.example.com/path/to/resource?query=example#fragment`

**Components:**

- **Scheme:** https
- **Host:** www.example.com
- **Path:** /path/to/resource
- **Query:** query=example
- **Fragment:** fragment

**Use Cases:** Accessing web pages, files, images, videos, etc., over the internet.

### ➤ URI (Uniform Resource Identifier)

**Definition:** A URI is a string of characters used to identify a resource either by location, name, or both. It is a broader category that includes both URLs and URNs.

**Structure:**

- Can be a URL, URN, or both.
- Identifies a resource in a generic way, without necessarily locating it.

**Example:**

`urn:isbn:0451450523`

This is a URI, specifically a URN, which identifies a book by its ISBN.

**Components:**

- **Scheme:** urn

- **Namespace Identifier (NID):** isbn
- **Namespace Specific String (NSS):** 0451450523

**Use Cases:** Identifying resources without necessarily providing a way to access them directly.

## ➤ **URN (Uniform Resource Name)**

**Definition:** A URN is a specific type of URI that names a resource in a persistent way, independent of its location or how it can be accessed.

### **Structure:**

- Typically includes a namespace identifier (NID) and a namespace-specific string (NSS).
- Does not imply the availability or accessibility of the resource.

### **Example:**

urn:ietf:rfc:2141

### **Components:**

- **Scheme:** urn
- **Namespace Identifier (NID):** ietf
- **Namespace Specific String (NSS):** rfc:2141

**Use Cases:** Persistent naming of resources such as books (ISBN), documents (RFC), and other entities.

## ➤ **Key Differences**

Aspect	URL	URI	URN
<b>Purpose</b>	Locates a resource	Identifies a resource	Names a resource
<b>Structure</b>	Includes scheme, host, path, query, fragment	Includes scheme, can be a URL or URN	Includes scheme, NID, NSS
<b>Access</b>	Specifies how to access a resource	Generic identification	No implication of access method
<b>Examples</b>	https://www.example.com	https://www.example.com, urn:isbn:0451450523	urn:isbn:0451450523

- **URI** is the most general term, encompassing both URLs and URNs.
- **URL** is a type of URI that not only identifies a resource but also provides a means to locate and access it.
- **URN** is a type of URI that provides a persistent, location-independent identifier for a resource, focusing on its identity rather than its accessibility.

## ➤ **IRI (Internationalized Resource Identifier)**

**Definition:** An IRI is an extended version of a URI that supports a wider range of characters, including those from non-Latin scripts, allowing for internationalized resource identifiers.

**Structure:**

- Similar to URIs, but can include Unicode characters beyond the ASCII set.
- IRIs can be converted to URIs using a process called "percent-encoding" for compatibility with systems that only support URIs.

**Example:**

`https://example.com/こんにちは`

**Use Cases:** Identifying resources on the web using characters from various languages and scripts, making the web more accessible internationally.

## ➤ **IETF (Internet Engineering Task Force)**

**Definition:** The IETF is an open standards organization that develops and promotes voluntary internet standards, particularly the standards that comprise the Internet protocol suite (TCP/IP).

**Responsibilities:**

- Development of internet protocols and standards.
- Publication of technical documents called RFCs (Request for Comments) that specify protocols, procedures, and best practices for the internet.

## ➤ **ISBN (International Standard Book Number)**

**Definition:** An ISBN is a unique identifier for books, intended to simplify the distribution and purchase of books by providing a standardized identification system.

**Structure:**

- Consists of 13 digits (formerly 10 digits) divided into five parts: a prefix (978 or 979), registration group, registrant, publication, and check digit.

**Example:**

`ISBN 978-3-16-148410-0`

## ➤ **ISSN (International Standard Serial Number)**

**Definition:** An ISSN is an eight-digit identifier used to uniquely identify serial publications, such as magazines, journals, newspapers, and other periodicals.

**Structure:**

- Composed of eight digits, typically split into two groups of four, separated by a hyphen.

**Example:**

ISSN 1234-5678

## ❖ **W3C (World Wide Consortium)**

**Definition:** The W3C is an international community that develops open standards to ensure the long-term growth of the web. Founded in 1994 by Tim Berners-Lee, the W3C's mission is to lead the web to its full potential by creating protocols and guidelines that ensure the web remains open, accessible, and interoperable.

### ➤ **Key functions and goals of W3C**

- **Standards Development:** The W3C develops web standards through a collaborative process involving member organizations, full-time staff, and the public.
- **Ensuring Interoperability:** W3C standards are designed to ensure that different web technologies can work together seamlessly.
- **Promoting Accessibility:** The W3C is committed to making the web accessible to all people, including those with disabilities.
- **Encouraging Innovation:** By providing a stable, open platform, the W3C encourages the development of new and innovative web technologies.

### ➤ **Key W3C Standards and Technologies**

- **HTML (HyperText Markup Language):** The standard markup language for creating web pages and web applications.
- **CSS (Cascading Style Sheets):** A style sheet language used for describing the presentation of a document written in HTML or XML.
- **XML (eXtensible Markup Language):** A markup language that defines rules for encoding documents in a format that is both human-readable and machine-readable.
- **SVG (Scalable Vector Graphics):** A language for describing two-dimensional graphics in XML.

## ❖ HTTP vs HTTPS

HTTP (HyperText Transfer Protocol) and HTTPS (HyperText Transfer Protocol Secure) are protocols used for transmitting data over the web. While both are fundamental to the functioning of the internet, they differ significantly in terms of security and data integrity.

### ➤ HTTP (HyperText Transfer Protocol)

**Definition:** HTTP is the foundational protocol used for transmitting data on the web. It defines how messages are formatted and transmitted, and how web servers and browsers should respond to various commands.

#### **Characteristics:**

- **Insecure:** Data transmitted over HTTP is not encrypted, making it vulnerable to interception and eavesdropping.
- **Port:** Uses port 80 by default.
- **Speed:** Slightly faster due to the absence of encryption overhead, but the difference is minimal with modern computing power.
- **Usage:** Commonly used for public websites where security is not a primary concern, but this usage is declining in favor of HTTPS.

#### **Example:**

`http://www.example.com`

### ➤ HTTPS (HyperText Transfer Protocol Secure)

**Definition:** HTTPS is an extension of HTTP that uses encryption to secure data transmitted between the client and server. It ensures data integrity, confidentiality, and authenticity.

#### **Characteristics:**

- **Secure:** Data is encrypted using SSL/TLS (Secure Sockets Layer/Transport Layer Security), protecting it from interception and tampering.
- **Port:** Uses port 443 by default.
- **Encryption:** Employs SSL/TLS protocols to encrypt the data transmitted, ensuring that sensitive information such as login credentials and personal data are secure.
- **Authentication:** Provides server authentication through SSL/TLS certificates, ensuring users are communicating with the legitimate server.
- **Data Integrity:** Ensures that data is not altered during transmission.
- **SEO Advantage:** Search engines like Google prioritize HTTPS websites in search rankings.
- **Browser Indicators:** Modern web browsers display visual indicators (such as a padlock icon) in the address bar to show that the connection is secure.

**Example:**

https://www.example.com

**➤ Key differences**

Feature	HTTP	HTTPS
<b>Security</b>	No encryption; data can be intercepted	Encrypted data; secure from eavesdropping
<b>Port</b>	Uses port 80	Uses port 443
<b>Encryption</b>	None	SSL/TLS encryption
<b>Authentication</b>	None	Server authentication via certificates
<b>Data Integrity</b>	Data can be altered during transmission	Ensures data integrity
<b>SEO Impact</b>	No SEO advantage	SEO ranking boost
<b>Browser Indicator</b>	None or "Not Secure" label in some browsers	Padlock icon or "Secure" label in browsers

**❖ SSL/TLS**

SSL (Secure Sockets Layer) and its successor TLS (Transport Layer Security) are cryptographic protocols designed to provide secure communication over a computer network. They are commonly used to secure data transfer over the Internet, ensuring privacy, data integrity, and authentication between client-server applications.

**➤ Key Features and Functions**

- **Encryption:** SSL/TLS protocols encrypt data transmitted between clients (e.g., web browsers) and servers. This encryption ensures that even if intercepted, the data cannot be read without the decryption key.
- **Data Integrity:** SSL/TLS protocols use message integrity checks to ensure that transmitted data is not tampered with during transmission. This prevents attackers from altering the data en route.
- **Authentication:** SSL/TLS protocols enable server authentication, verifying that clients are communicating with the intended server and not an impostor. This is achieved through digital certificates issued by trusted Certificate Authorities (CAs).
- **Compatibility:** SSL/TLS is widely supported by web browsers, email clients, and other applications, making it a versatile choice for securing various types of online communication.



## ➤ SSL vs TLS

- **SSL (Secure Sockets Layer):** Developed by Netscape in the mid-1990s, SSL was the original protocol designed to secure online transactions. Versions include SSL 2.0, SSL 3.0.
- **TLS (Transport Layer Security):** TLS is the successor to SSL, designed to address vulnerabilities found in earlier versions of SSL. Versions include TLS 1.0, TLS 1.1, TLS 1.2, and TLS 1.3 (the latest version as of now).

## ❖ Internet vs www (World Wide Web)

While the terms "Internet" and "World Wide Web" (WWW or the Web) are often used interchangeably, they refer to different concepts.

### ➤ The Internet

**Definition:** The Internet is a vast global network of interconnected computers and other devices. It is the infrastructure that enables communication and data transfer between devices around the world.

#### Characteristics:

- **Infrastructure:** The Internet consists of physical elements like servers, routers, cables, satellites, and wireless systems.
- **Protocols:** It uses a suite of protocols known as TCP/IP (Transmission Control Protocol/Internet Protocol) to transmit data.
- **Services:** Supports a variety of services including email (SMTP), file transfer (FTP), online gaming, video streaming, and, notably, the World Wide Web.
- **Connectivity:** It connects millions of private, public, academic, business, and government networks.

#### Example Services:

- **Email:** Exchange of messages via Simple Mail Transfer Protocol (SMTP).
- **File Transfer:** Transferring files using protocols like FTP (File Transfer Protocol).
- **VoIP (Voice over Internet Protocol):** Voice communication over the Internet, e.g., Skype.
- **Streaming:** Real-time transmission of audio and video, e.g., Netflix, YouTube

### ➤ Key differences

Aspect	Internet	World Wide Web
<b>Definition</b>	Global network of interconnected devices	System of interlinked hypertext documents
<b>Function</b>	Infrastructure for global communication	Service running on the Internet for accessing web content

Aspect	Internet	World Wide Web
<b>Protocols</b>	TCP/IP, SMTP, FTP, etc.	HTTP/HTTPS
<b>Components</b>	Routers, servers, cables, satellites	Web browsers, web servers, web pages
<b>Services</b>	Email, VoIP, streaming, file transfer	Websites, search engines, social media
<b>Usage</b>	Enabling data exchange and communication	Accessing and sharing information on the web

## ❖ Some Chrome extensions

### ➤ AbBlock Plus

**Function:** Adblock Plus is a popular ad-blocking extension that blocks intrusive ads on websites, making browsing faster and less cluttered.

**Features:**

- **Blocking Ads:** Filters out banners, pop-ups, and video ads.
- **Customizable:** Allows users to whitelist specific sites or customize filters.
- **Privacy:** Helps protect privacy by blocking tracking scripts.

### ➤ LastPass

**Function:** LastPass is a password manager that securely stores and autofills usernames and passwords for websites.

**Features:**

- **Password Management:** Stores passwords, usernames, and other credentials securely.
- **Autofill:** Automatically fills login credentials on websites.
- **Security:** Generates strong passwords and encrypts data locally before syncing.

### ➤ ColorZilla

**Function:** ColorZilla is a color picker and gradient generator extension.

**Features:**

- **Color Picker:** Allows users to pick any color from a webpage.
- **Gradient Generator:** Generates CSS gradients visually.
- **Color Palette:** Saves picked colors to a palette for future use.

## ➤ **WhatFont**

**Function:** WhatFont identifies fonts used on web pages.

**Features:**

- **Font Identification:** Shows the font family, size, weight, and style used on text elements.
- **Characteristics:** Displays additional details like line height and color.
- **Browser Compatibility:** Works across various browsers, including Chrome.

These extensions enhance functionality and productivity while browsing, providing tools for ad management, password security, color management, and font identification.