**LECTURE NOTES: WebTek 2018**

**Internet vs. Web**

Internet and the Web is commonly mistaken as synonymous with each other however they are two different entities.

**Internet** – global system of interconnected network of computing devices.   
 -- Use the standard Internet Protocol Suite (TCP/IP) in order to serve billions of users worldwide.

**Vinton Cerf** – Father of the Internet

**World Wide Web --** also known as **WWW, Web, W3**

A global information medium in which users can read and write via computers which are connected to the Internet.   
  
set of software services that runs/operates over the Internet.

According to W3C, the web is a universe of information that can be access through a network.

**Sir** **Tim-Berners Lee** – Father of the Web,

**World Wide Web Consortium(W3C) –** founded by Sir Tim-Berners-Lee in September 1994 at Massachusetts Institute of Technology (MIT) together with the Defense Advanced Research Projects Agency(DARPA) and the European Commission.  
 -- It is an international community where various companies work together to develop web standards and create recommendations to improve the quality of the Web.

**Sir Tim Berners Lee** -- Director of the World Wide Web Consortium(W3C).  
**Jeffrey Jaffe—**W3C’s Chief Executive Officer

THREE FUNDAMENTAL TECHNOLOGIES THAT CONTRIBUTES TO THE FOUNDATION OF THE WEB

**HTML (Hypertext Markup Language):** publishing format for the Web, has the ability to format documents and link to other documents and resources.

**URI (Uniform Resource Identifier):** address that is unique to each resources on the Web

**HTTP (Hypertext Transfer Protocol):**  Application layer used primarily to retrieve hypertext (on hypermedia) documents and resources on the Web  
 -- Jointly developed by the W3C and the IETF

**HTTP VERSION HISTORY**

**HTTP 0.9 (1991)**

**HTTP 1.0 (RFC 1945, May 1996)**

**HTTP 1.1 (RFC 2068, January 1997, RFC 2616, June 1999, RFC 7230 – 7235, June 2014)**

**HTTP 2 (RFC 7540, May 2015)**

**HTTP (HYPERMEDIA)**

**Protocol** – Set of rules need to be followed

Fundamentals

* HTTP is a stateless communication protocol that runs on top of the TCP/IP
* HTTP using TCP uses port 80 by default
* HTTPS using TCP uses port 443 by default
* HTTP resources are identifies using URIs (RFC 3986)
  + Scheme
    - HTTP
    - HTTPS
  + Authority—user information and authentication
  + Host must be identified
  + Port number
  + Resource Path – Includes actual document or resources
  + scheme specific parameters(optional)
  + URL-encoded query(optional)
  + Bookmark (fragmented identifier) – subjection of a specified document(optional)

**Cookie** (web cookie, browser cookie – refer to a small piece of data that sends by the server to the user through the browser.   
 -- very small text file

**Cache – Local storage of resources that is fetched from a server**

**HTTP RESPONSE MESSAGE**

Status Line- it shows whether the request that has been made is successful or a failure. It has 3 subcategories:

HTTP protocol version – it returns the information of the http version   
 For example , HTTP version 1.2 will return HTTP-Version = HTTP /1.2  
 Status code – it is composed of 3-digit integer where it shows the response to the requests that has been made.

Reason phrase – also called as the status text. It contains additional information and description about the status code.

Response Headers—it is where the server passes the information about the response that cannot be displayed into the status-line.

Empty Line – it indicates that the requests has already been sent.

Message Body – contains information about the request or response that has been sent.

**HTTP REQUEST METHODS – set of methods that can be done and performed for a given web resource.**

**GET: allows to retrieve a resource and be returned; involves entire query in the URL**

**HEAD: same as the get method however there is no response body included.**

**POST: puts the message in the body and used in submitting HTML form dat. It can caused a change in the server.**

**OPTION: shows the allowed actions to be performed**

**TRACE: it is used to invoke a message loop-back test to the target web resource; replicate/ echo back request message.**

**PUT: it replaces all the content within the specified web resource; generates a resource**

**DELETE: use to delete/remove a web resource**

**CONNECT: a method where the target resource is used to established a tunnel to the server.**

**METHOD PROPERTIES**

**SAFE METHODS – indicates that there is no modification/change that has been made upon requests.   
 These are read-only methods: GET and HEAD belongs to this category.**

**IDEMPOTENT METHODS – even if the method is repeated multiple time, the effect is still the same (fetches the same resource)  
 Methods such as GET, HEAD, PUT AND DELETE belongs to this category.**

|  |  |  |
| --- | --- | --- |
| **Method** | **Safe** | **Idempotent** |
| **GET** |  |  |
| **HEAD** |  |  |
| **DELETE** | **X** |  |
| **PUT** | **X** |  |
| **POST** | **X** | **X** |

**CACHABLE METHODS**

**Indicate that the response to a method is allowed to be stored for future use.**

HTTP HEADERS ACCORDING TO CONTEXT (**RFCF 2616)**

**GENERAL HEADER:**

* **Cache-Control**
* **Connection**
* **Date**
* **Pragma**
* **Trailer**
* **Transfer-encoding**
* **Upgrade**
* **Via**
* **Warning**

**REQUEST HEADER:**

* **Accept**
* **Accept- charset**
* **Accept-encoding**
* **Accept-language**
* **Authorization**
* **Expect**
* **From**
* **Host**
* **If-Match**
* **If-Modified-Since**
* **If-None-Match**
* **If-Range**
* **If-Unmodified-Since**
* **Max-forward**
* **Proxy-Authorization**
* **Range**
* **Referrer**
* **User-agent**

**RESPONSE HEADER**

* **Accept-Range**
* **Age**
* **E-tag**
* **Location**
* **Proxy-Authentication**
* **Retry-after**
* **Srever**
* **Vary**
* **WWW-Authenticate**

**ENTITY HEADER**

* **Allow**
* **Content-encoding**
* **Content-language**
* **Content-Length**
* **Content-Location**
* **Content-MDS**
* **Content-Range**
* **Content- Type**
* **Expires**
* **Last Modified**

HTTP STATUS CODE

Informational (1xx)

* 100 Continue
* 101 Switching Protocols

Success (2xx)

* 200 OK
* 202 Accepted
* 203 Non-Authoritative information
* 204 No Content
* 205 Reset Content
* 206 Partial Content

Redirection (3xx)

* 300 Multiple Choices
* 301 Moved Permanently
* 302 Not Found
* 303 See Other
* 304 Not Modifies
* 305 Use Proxy
* 306 Unused
* 307 Temporary Redirect

Client Error (4xx)

* 400 Bad Request
* 401 Unauthorized
* 402 Payment Required
* 403 Forbidden
* 404 Not Found
* 405 Method not Allowed
* 406 Not Acceptable
* 407 Proxy Authentication Required
* 408 Request Timeout
* 409 Conflict
* 410 Gone
* 411 Length Required
* 412 Precondition failed
* 413 Request entity Too Large
* Request-URI Too Long
* 415 Unsupported media type
* 416 Request Range Not Satisfiable
* 417 Expectation Failed

Server Error (5xx)

* 500 Internal Server Error
* 501 Not Implemented
* 502 Bad Gateway
* 503 Service Unavailable
* 504 Gateway Timeout
* 505 HTTP Version Not Supported

**WORD VOCABULARY ACRONYM**

References:

<https://www.w3.org/WWW/>

<https://www.w3.org/Consortium/>

<http://interactivepython.org/runestone/static/webfundamentalsITUBIL103E2015Fall/WWW/history.html>

<https://developer.mozilla.org/en-US/docs/Web/HTTP/Headers>

[**https://developer.mozilla.org/en-US/docs/Web/HTTP/Cookies**](https://developer.mozilla.org/en-US/docs/Web/HTTP/Cookies)

[**https://www.tutorialspoint.com/http/http\_responses.htm**](https://www.tutorialspoint.com/http/http_responses.htm)

[**https://developer.mozilla.org/en-US/docs/Web/HTTP/Messages**](https://developer.mozilla.org/en-US/docs/Web/HTTP/Messages)

<https://www.w3.org/Protocols/rfc2616/rfc2616-sec4.html>

<https://developer.mozilla.org/en-US/docs/Web/HTTP/Methods>

<https://www.w3.org/Protocols/rfc2616/rfc2616-sec9.html>

<https://www.tutorialspoint.com/http/http_methods.htm>

<http://javarevisited.blogspot.com/2016/05/what-are-idempotent-and-safe-methods-of-HTTP-and-REST.html>

<https://www.w3.org/Protocols/rfc2616/rfc2616-sec10.html#sec10.1>

<https://tools.ietf.org/html/rfc2616>