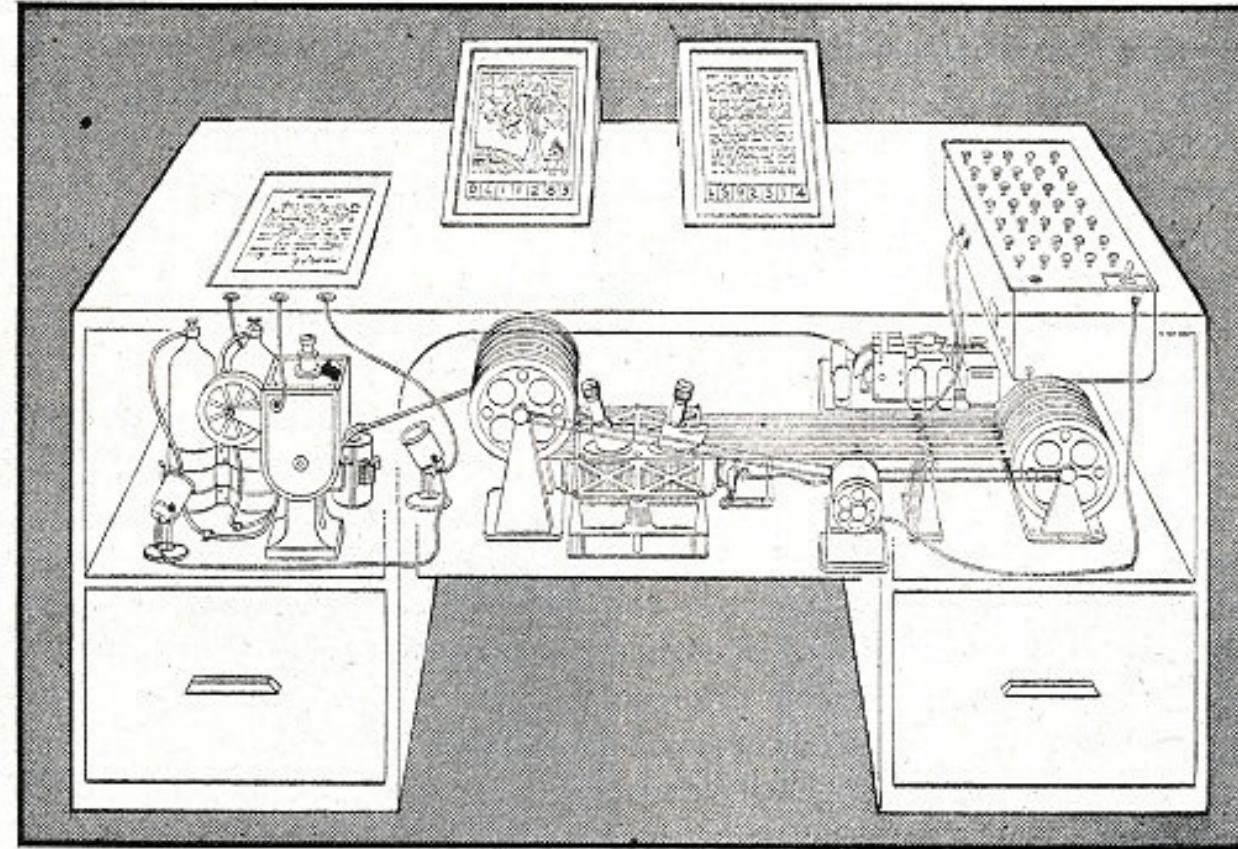# PRESENTATION ON THE WEB: HTML

INFO20002: Foundations of Informatics

March, 2016

# SOME HISTORICAL PRECEDENTS

# 1945: VANNEVAR BUSH'S MEMEX



**MEMEX** in the form of a desk would instantly bring files and material on any subject to the operator's fingertips. Slanting translucent viewing screens magnify supermicrofilm filed by code numbers. At left is a mechanism which automatically photographs longhand notes, pictures and letters, then files them in the desk for future reference.

Bush, Vannevar. "As We May Think". The Atlantic. N.p., 1945. Retrieved from http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/. "Memex conceptual sketch". Retrieved from http://www.computerhistory.org/revolution/the-web/20/370/2111.

# 1965: TED NELSON'S HYPER-TEXT

*"Mr. Nelson pointed out that we often do not think in linear sequences but rather in 'swirls' and in footnotes. He introduced the concept of the hyper-text, which would be a more flexible, more generalized, non-linear presentation of material on a particular subject."*

Wedeles, Lauren. "Prof. Nelson Talk Analyzes P.R.I.D.E." (1965, Feb 3). Retrieved from http://faculty.vassar.edu/mijoyce/MiscNews_Feb65.html.
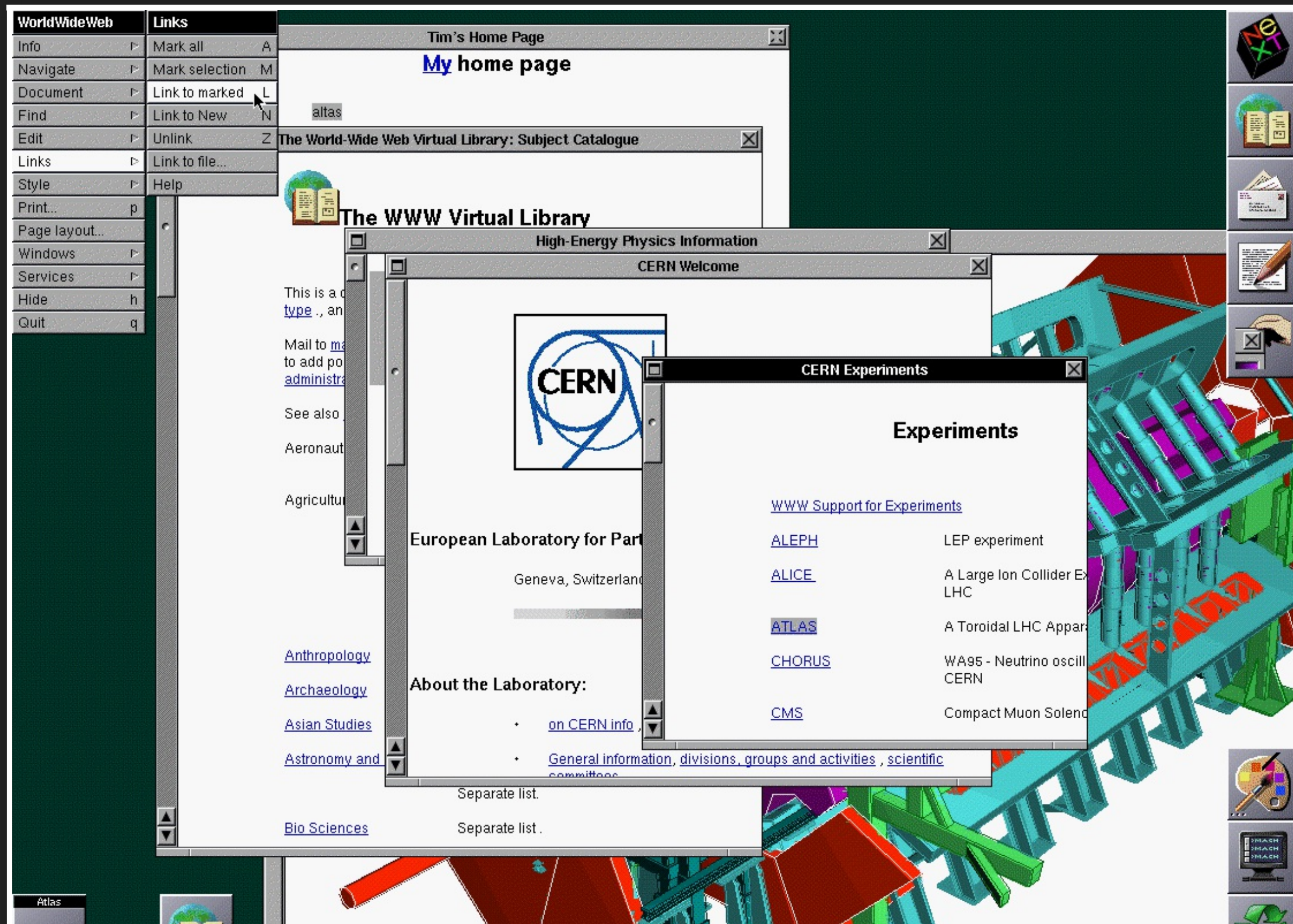
# 1980: TIM BERNERS-LEE'S ENQUIRE



**Documentation of the RPC project (concept)**

Most of the documentation is available on VMS, with the two principle manuals being stored in the CERNDOC system.

1) includes: The VAX/NOTES conference VXCERN::RPC
2) includes: Test and Example suite
3) includes: RPC BUG LISTS
4) includes: RPC System: Implementation Guide
Information for maintenance, porting, etc.
5) includes: Suggested Development Strategy for RPC Applications
6) includes: "Notes on RPC", Draft 1, 20 feb 86
7) includes: "Notes on Proposed RPC Development" 18 Feb 86
8) includes: RPC User Manual
How to build and run a distributed system.
9) includes: Draft Specifications and Implementation Notes
10) includes: The RPC HELP facility
11) describes: THE REMOTE PROCEDURE CALL PROJECT in DD/OC

**Help Display Select Back Quit Mark Goto_mark Link Add Edit**

Berners-Lee, Tim. "Enquire Manual - In HyperText" (1980, October) Retrieved from https://www.w3.org/People/Berners-Lee/EnquireManual.htm.

# 1990: TIM BERNERS-LEE'S WORLDWIDEWEB

# HYPERTEXT MARKUP LANGUAGE (HTML)

# HELLO WORLD (HTML5)

```html
<!DOCTYPE html>
<html>
    <head>
        <meta charset="utf-8">
        <title>Hello world!</title>
    </head>
    <body>
        <h1>Hello world!</h1>
        <p>This is the basic structure of an HTML document.</p>
    </body>
</html>
```

# THREE BROAD CATEGORIES OF BODY MARKUP

# THREE BROAD CATEGORIES OF BODY MARKUP

- Structural - Defines document and layout structure

# THREE BROAD CATEGORIES OF BODY MARKUP

- Structural - Defines document and layout structure
- Stylistic - Modifies the appearance of text

# THREE BROAD CATEGORIES OF BODY MARKUP

- Structural - Defines document and layout structure
- Stylistic - Modifies the appearance of text
- Semantic - Identifies specific types of information

## STRUCTURAL MARKUP

- Headings: `<h1>`, `<h2>`, `<h3>`, `<h4>`, `<h5>`, `<h6>`
- Paragraphs: `<p>`
- Lists: `<dl>`, `<ol>`, `<ul>`
- Block and inline content: `<div>` and `<span>`

## STYLISTIC MARKUP

- Bold: `<em>`, `<strong>`
- Deleted: `<del>`
- Italic: `<i>`
- Preformatted: `<pre>`
- Subscript: `<sub>`
- Superscript: `<sup>`

# SEMANTIC MARKUP

- Links: <a>
- Tables: <table>
- Forms: <form>
- Images: <img>
- Others: <address>, <code>

"Lesser known semantic elements". Retrieved from https://www.w3.org/wiki/Lesser_-_known_semantic_elements

```html
<address>
    <span>Mark Norman Francis</span>,
    <span class="tel">1-800-555-4865</span>
</address>
```

# BLOCK-LEVEL AND INLINE ELEMENTS

- By default, block-level elements begin on newlines whereas inline elements do not
- Block-level elements may contain inline elements and other block-level elements
- Inline elements may contain only data and other inline elements

"Block-level elements". Retrieved from https://developer.mozilla.org/en/docs/Web/HTML/Block-level_elements. "Inline elements". Retrieved from https://developer.mozilla.org/en-US/docs/Web/HTML/Inline_elements

# ANCHOR

Provides a one-way link from one document to another

```html
<a href="page.html">Page</a>
```

# UNIFORM RESOURCE LOCATOR (URL)

A formal representation of the location and means of access for a resource on the Internet.

Berners-Lee, Tim. "Uniform Resource Locators (URL)" Retrieved from https://www.w3.org/Addressing/URL/url-spec.txt.

scheme://user:password@host:port/path?query#fragment

http://www.example.com/index.html

http://www.example.com

http://www.example.com:8080

http://www.example.com/path/to/image.jpg

```
http://www.example.com/search?
query=cats&sortby=cuteness
```

http://www.example.com/page.html#section1

# ABSOLUTE VS RELATIVE URLS

Absolute URL to a file in the top level directory

Absolute URL to a file in the top level directory

- `http://www.example.com/page.html`

Absolute URL to a file in the top level directory

- `http://www.example.com/page.html`
- `/page.html`

Absolute URL to a file in a subdirectory

# Absolute URL to a file in a subdirectory

- `http://www.example.com/subdir/page.html`

# Absolute URL to a file in a subdirectory

- `http://www.example.com/subdir/page.html`
- `/subdir/page.html`

Relative URL to a file in the current directory

# Relative URL to a file in the current directory

- `page.html`

# Relative URL to a file in the current directory

- `page.html`
- `./page.html`

Relative URL to a file in a subdirectory of the current directory

Relative URL to a file in a subdirectory of the current directory

- `subdir/page.html`

# Relative URL to a file in a subdirectory of the current directory

- `subdir/page.html`
- `subdir/subdir2/page.html`

Relative URL to a file in the parent directory

# Relative URL to a file in the parent directory

- `../page.html`

# Relative URL to a file in the parent directory

- `../page.html`
- `../../page.html`

# HTML ESCAPES

# YOU MUST ALWAYS ESCAPE

- Less than (<) - `&lt;`
- Greater than (>) - `&gt;`
- Ampersand (&) - `&amp;`

"Using escape characters in markup and CSS". Retrieved from https://www.w3.org/International/questions/qa-escapes.

# FORMS

```
<form action="http://localhost/cgi-bin/env.py" method="post">
    <input type="text" name="foo" />
    <input type="submit" />
</form>
```

Submit

# FORM ATTRIBUTES

# FORM ATTRIBUTES

- Action is the URL of a program that processes the form input

# FORM ATTRIBUTES

- Action is the URL of a program that processes the form input
- Method must be one of POST or GET

# FORM ATTRIBUTES

- Action is the URL of a program that processes the form input
- Method must be one of POST or GET
  - An idempotent request is one that does not produce "side effects".

# FORM ATTRIBUTES

- Action is the URL of a program that processes the form input
- Method must be one of POST or GET
  - An idempotent request is one that does not produce "side effects".
  - GET requests are expected to be idempotent whereas POST requests are expected to produce side effects

```python
# A Python CGI program that prints out its environment variables,
# and the contents of stdin.
import os
import sys

print 'Content-Type: text/plain\n'
for var in os.environ:
    print '%s = %s' % (var, os.environ[var])
print 'stdin = %s' % sys.stdin.read()
```

```
HTTP_REFERER = http://localhost/~bjpop/forms-lecture/simpleform.html
    SERVER_SOFTWARE = Apache/2.2.8 (Unix) mod_ssl/2.2.8 OpenSSL/0.9.7l DAV/2
    SCRIPT_NAME = /~bjpop/env.py
    SERVER_SIGNATURE =
    REQUEST_METHOD = POST
    SERVER_PROTOCOL = HTTP/1.1
    QUERY_STRING =
    PATH = /usr/bin:/bin:/usr/sbin:/sbin
    CONTENT_LENGTH = 27
    HTTP_USER_AGENT = Mozilla/5.0 (Macintosh; U; Intel Mac OS X 10_5_4; en-us) ...
    HTTP_CONNECTION = keep-alive
    SERVER_NAME = localhost
    REMOTE_ADDR = ::1
    SERVER_PORT = 80
    SERVER_ADDR = ::1
    SCRIPT_FILENAME = /Users/bjpop/Sites/env.py
```

# FOO.PY

```python
import cgi

print 'Content-Type: text/html'
print

store = cgi.FieldStorage()

print '>html<>head<>/head<>body<>h1<'

if 'foo' in store:
    print store.getvalue('foo')

print '>/h1<>/body<>/html<'
```

# FOO.PY

```python
import cgi

print 'Content-Type: text/html'
print

store = cgi.FieldStorage()

print '>html<>head<>/head<>body<>h1<'

if 'foo' in store:
    print cgi.escape(store.getvalue('foo'))

print '>/h1<>/body<>/html<'
```

# INPUT CONTROLS

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from
https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)
- checkboxes

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)
- checkboxes
- radio buttons

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)
- checkboxes
- radio buttons
- submit, reset buttons

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from
https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)
- checkboxes
- radio buttons
- submit, reset buttons
- buttons (other than submit and reset)

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from https://www.w3.org/TR/html5/forms.html#forms.

# INPUT CONTROLS

- text input boxes (single line and multi-line)
- checkboxes
- radio buttons
- submit, reset buttons
- buttons (other than submit and reset)
- hidden fields

Forms in HTML5, A vocabulary and associated APIs for HTML and XHTML, W3C Recommendation 28 October 2014. Retrieved from
https://www.w3.org/TR/html5/forms.html#forms.

# A FORM WITH A FEW DIFFERENT CONTROLS

```html
<form method="get" action="http://localhost/cgi-bin/env.py">
    Enter something:
    <input type="text" name="mytextbox" value="a default value" />
    <br />
    <input type="checkbox" name="mycheckbox1" />
    <input type="checkbox" name="mycheckbox2" checked="checked" />
    <br />
    <input type="radio" name="myradio" value="value1" checked="checked" />
    <input type="radio" name="myradio" value="value2" />
    <input type="radio" name="myradio" value="value3" />
    <br />
    <input type="submit" />
    <input type="reset" />
</form>
```

Enter something: a default value

☐ ☑

◉ ○ ○

Submit  Reset

- A single HTML document can contain multiple forms, and hence multiple submit buttons.

- A single HTML document can contain multiple forms, and hence multiple submit buttons.
- However, forms cannot be nested. That is, one form cannot be inside another.

- A single HTML document can contain multiple forms, and hence multiple submit buttons.
- However, forms cannot be nested. That is, one form cannot be inside another.
- It is possible to write the CGI script to produce a form as output, and also act as the recipient of the form data (by being the "action" of the form).

- Accessibility - Using descriptive tags to aid screen-readers and audio browsers

- Accessibility - Using descriptive tags to aid screen-readers and audio browsers
- Search-engine optimisation (SEO) - Bumping your site to the top of the Google listing …

- Accessibility - Using descriptive tags to aid screen-readers and audio browsers
- Search-engine optimisation (SEO) - Bumping your site to the top of the Google listing ...
- "Semantic web" - All about having descriptive tags in everything so that it can be more efficiently mined. controversial idea ...

# REFERENCES

- HTML5. A vocabulary and associated APIs for HTML and XHTML. W3C Recommendation 28 October 2014
- MDN HTML Reference
- http://www.html5rocks.com/en/

# CHECK YOUR WORK!

https://validator.w3.org/