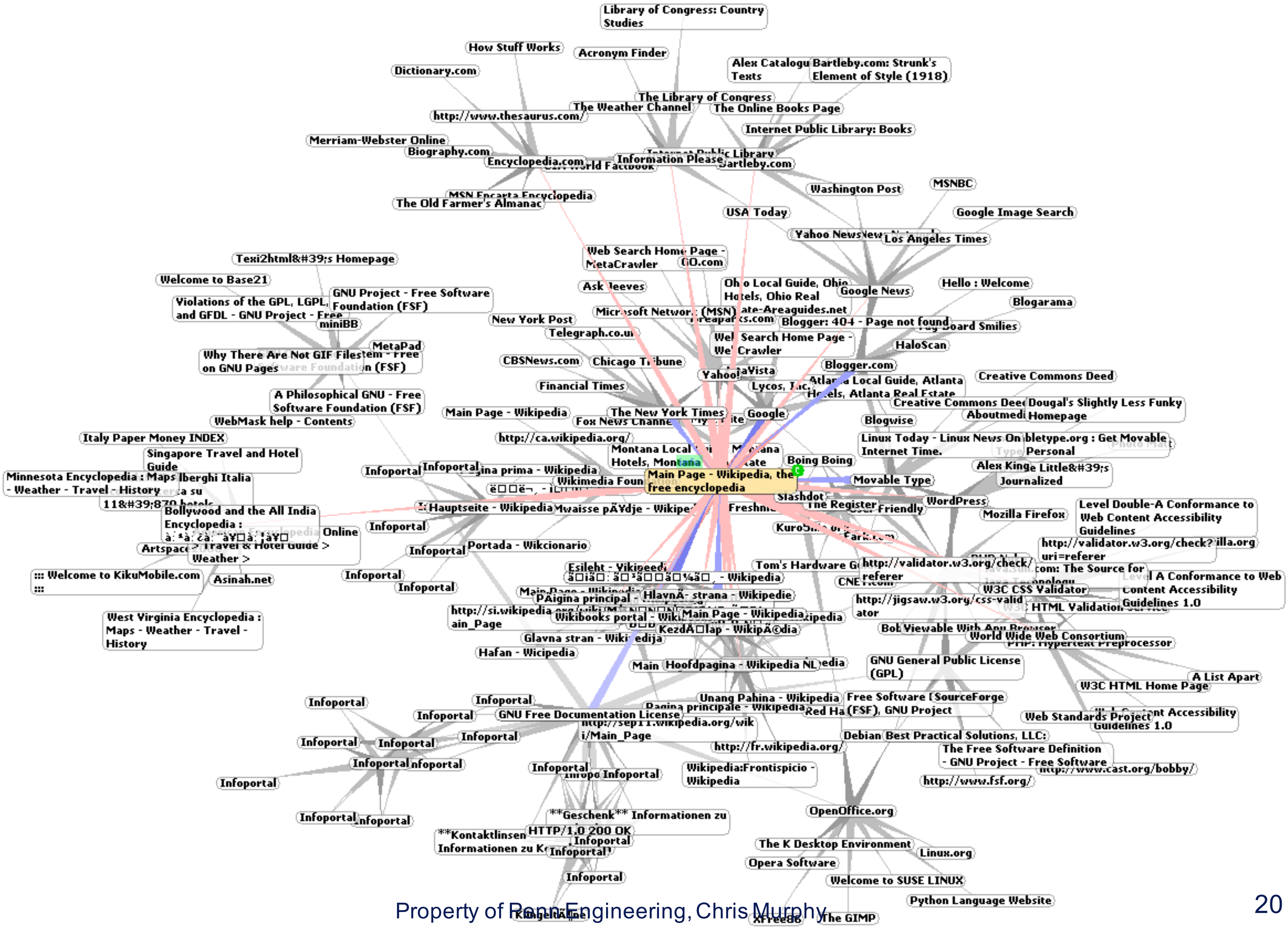
Penn Engineering

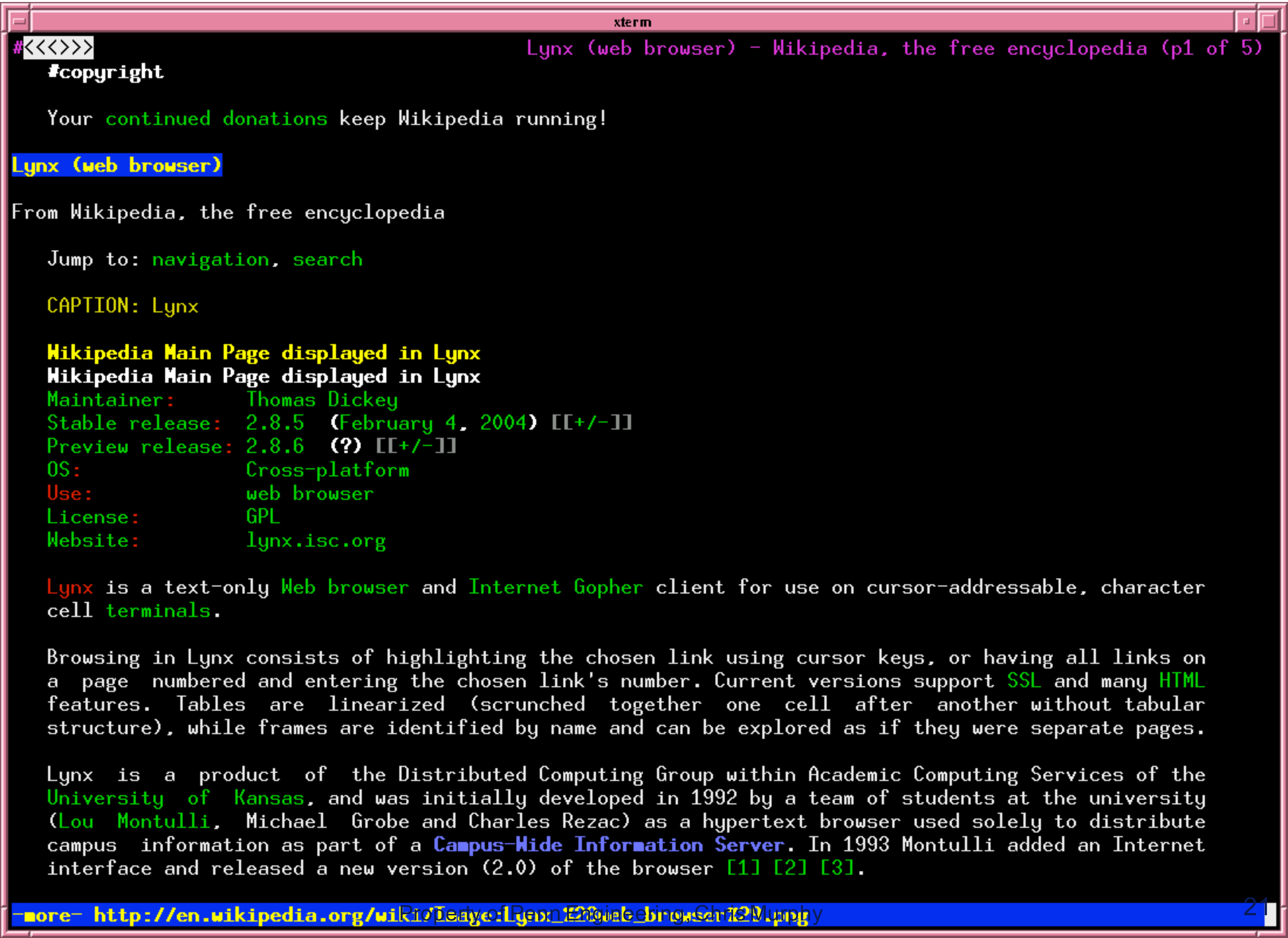
ONLINE LEARNING

# Video 1.2
# Chris Murphy

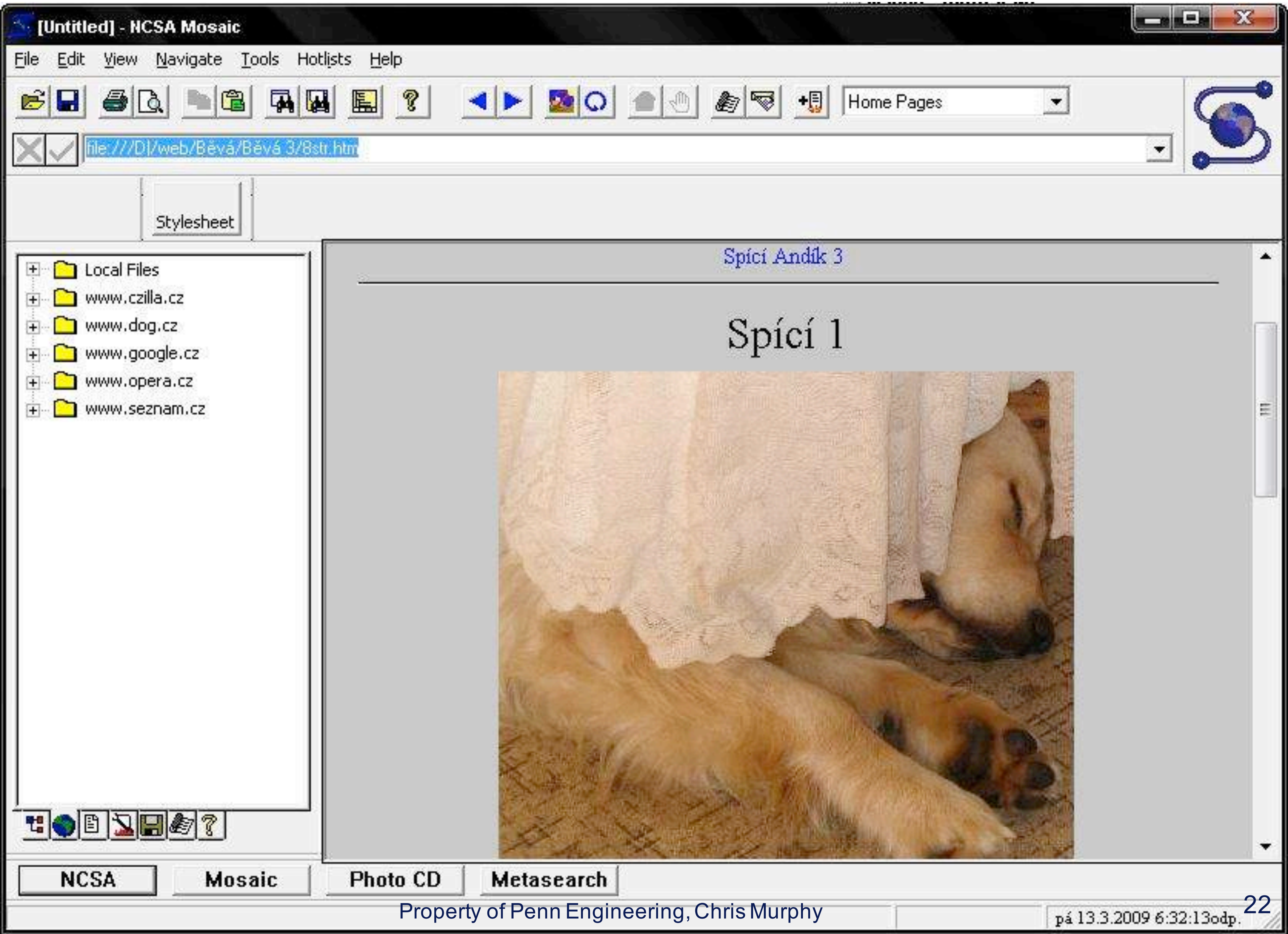
# Review

- The **Internet** is a physical network of devices
- The **World Wide Web** is an application that utilizes the Internet to allow for accessing data
- Resources on the Web have unique **URLs** that include the protocol, host name, and file/resource name


Library of Congress: Country Studies
How Stuff Works
Acronym Finder
Dictionary.com
Alex Catalogue Texts
Bartleby.com: Strunk's Element of Style (1918)
The Library of Congress
The Weather Channel
The Online Books Page
http://www.thesaurus.com/
Internet Public Library: Books
Merriam-Webster Online
Biography.com
Encyclopedia.com
Information Please
Internet Public Library
Bartleby.com
MSNBC
Washington Post
MSN Encarta Encyclopedia
The Old Farmer's Almanac
USA Today
Google Image Search
Yahoo News
Los Angeles Times
Web Search Home Page - MetaCrawler
GO.com
Texi2html&#39;s Homepage
Welcome to Base21
Ask Jeeves
Ohio Local Guide, Ohio Hotels, Ohio Real
Google News
Hello : Welcome
GNU Project - Free Software Foundation (FSF)
Violations of the GPL, LGPL, and GFDL - GNU Project - Free
Blogarama
Microsoft Network (MSN)
Areaguides.net
miniBB
New York Post
Blogger: 404 - Page not found
Board Smilies
Telegraph.co.uk
Web Search Home Page - WebCrawler
MetaPad
HaloScan
Why There Are Not GIF Files on GNU Pages
Free Software Foundation (FSF)
CBSNews.com
Chicago Tribune
Blogger.com
Yahoo!
AltaVista
Creative Commons Deed
Financial Times
Lycos, Inc.
Atlanta Local Guide, Atlanta Hotels, Atlanta Real Estate
A Philosophical GNU - Free Software Foundation (FSF)
Creative Commons Deed
Dougal's Slightly Less Funky Homepage
Main Page - Wikipedia
The New York Times
Google
Aboutmedia
WebMask help - Contents
Fox News Channel
Blogwise
Italy Paper Money INDEX
http://ca.wikipedia.org/
Linux Today - Linux News On Internet Time.
Movabletype.org : Get Movable Type Personal
Singapore Travel and Hotel Guide
Montana Local Guide, Montana Hotels, Montana Real Estate
Boing Boing
Alex King
Little&#39;s Journalized
Minnesota Encyclopedia : Maps - Weather - Travel - History
Alberghi Italia
Infoportal
Infoportal
Pagina prima - Wikipedia
Main Page - Wikipedia, the free encyclopedia
Wikimedia Foundation
Movable Type
Slashdot
The Register
WordPress
User Friendly
Level Double-A Conformance to Web Content Accessibility Guidelines
Bollywood and the All India Encyclopedia :
Hauptseite - Wikipedia
Mwaisse pÃ¥dje - Wikipedia
Freshmeat
Mozilla Firefox
Online
Infoportal
KuroShin
Fark.com
http://validator.w3.org/check?uri=referer
Mozilla.org
Artspace > Travel & Hotel Guide > Weather >
Infoportal
Portada - Wikcionario
Esileht - Vikipeedia
Tom's Hardware Guide
http://validator.w3.org/check/referer
Java.sun.com: The Source for Java Technology
::: Welcome to KikuMobile.com :::
Asinah.net
Infoportal
Infoportal
CNET.com
Level A Conformance to Web Content Accessibility Guidelines 1.0
Main Page - Wikipedia
PÃ¡gina principal - Wikipedia
HlavnÃ strana - Wikipedie
W3C CSS Validator
http://jigsaw.w3.org/css-validator
W3C HTML Validation Service
West Virginia Encyclopedia : Maps - Weather - Travel - History
http://si.wikipedia.org/wiki/Main_Page
Wikibooks portal
Main Page - Wikipedia
KezdÅ‘lap - WikipÃ©dia
Bob
Viewable With Any Browser
World Wide Web Consortium
Glavna stran - Wikipedija
PHP: Hypertext Preprocessor
Hafan - Wicipedia
Main
Hoofdpagina - Wikipedia NL
GNU General Public License (GPL)
A List Apart
W3C HTML Home Page
Infoportal
Infoportal
Unang Pahina - Wikipedia
Pagina principale - Wikipedia
Free Software Foundation (FSF), GNU Project
SourceForge
Red Hat
Infoportal
GNU Free Documentation License
http://sep11.wikipedia.org/wiki/Main_Page
Web Standards Project
Web Content Accessibility Guidelines 1.0
Infoportal
Infoportal
Infoportal
Debian
Best Practical Solutions, LLC:
http://fr.wikipedia.org/
The Free Software Definition - GNU Project - Free Software
Infoportal
Infoportal
Infoportal
Infoportal
Wikipedia:Frontispicio - Wikipedia
http://www.cast.org/bobby/
http://www.fsf.org/
Infoportal
**Geschenk** Informationen zu
OpenOffice.org
Infoportal
Infoportal
**Kontaktlinsen** Informationen zu K
HTTP/1.0 200 OK
Infoportal
Infoportal
The K Desktop Environment
Linux.org
Opera Software
Infoportal
Welcome to SUSE LINUX
Python Language Website
XFree86
The GIMP



```
#<<<>>>                                   Lynx (web browser) - Wikipedia, the free encyclopedia (p1 of 5)
   #copyright

   Your continued donations keep Wikipedia running!

Lynx (web browser)

From Wikipedia, the free encyclopedia

   Jump to: navigation, search

   CAPTION: Lynx

   Wikipedia Main Page displayed in Lynx
   Wikipedia Main Page displayed in Lynx
   Maintainer:      Thomas Dickey
   Stable release:  2.8.5  (February 4, 2004) [[+/-]]
   Preview release: 2.8.6  (?) [[+/-]]
   OS:              Cross-platform
   Use:             web browser
   License:         GPL
   Website:         lynx.isc.org

   Lynx is a text-only Web browser and Internet Gopher client for use on cursor-addressable, character
   cell terminals.

   Browsing in Lynx consists of highlighting the chosen link using cursor keys, or having all links on
   a  page  numbered and entering the chosen link's number. Current versions support SSL and many HTML
   features.  Tables  are  linearized  (scrunched  together  one  cell  after  another without tabular
   structure), while frames are identified by name and can be explored as if they were separate pages.

   Lynx  is  a  product  of  the Distributed Computing Group within Academic Computing Services of the
   University  of  Kansas, and was initially developed in 1992 by a team of students at the university
   (Lou  Montulli,  Michael  Grobe and Charles Rezac) as a hypertext browser used solely to distribute
   campus  information as part of a Campus-Wide Information Server. In 1993 Montulli added an Internet
   interface and released a new version (2.0) of the browser [1] [2] [3].

-more- http://en.wikipedia.org/wiki/Image:Lynx_%28web_browser%29.png
```



Spící Andík 3

Spící 1

# Popular Browsers

# What is a Web Browser?

- **Browser**: software that is used to access and display Web content, and to navigate across the Web

- **Main Components of the Browser**
  - Rendering Engine (HTML/CSS) – responsible for static content presentation, formatting, and layout
  - JavaScript Engine (JavaScript) – responsible for creating and modifying dynamic content and appearance

# How Does a Web Browser Work?

- Browser and the World Wide Web utilize **Hypertext Transfer Protocol (HTTP)** to transfer documents

# How Does a Web Browser Work?

- Browser and the World Wide Web utilize **Hypertext Transfer Protocol (HTTP)** to transfer documents

**Client**

# How Does a Web Browser Work?

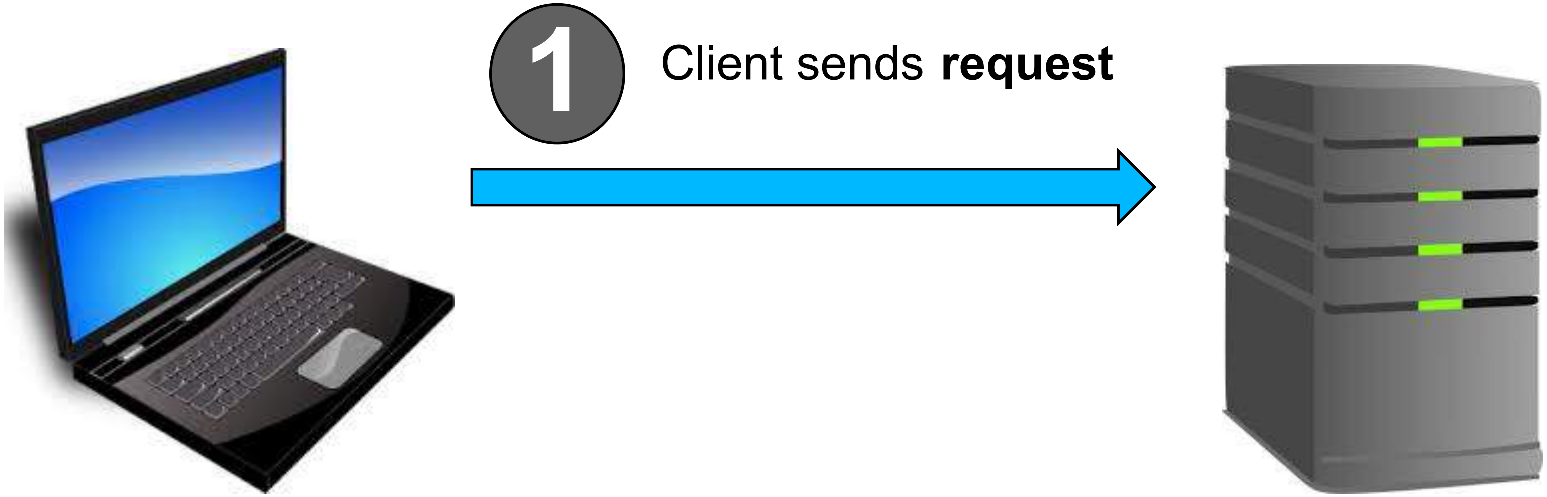
- Browser and the World Wide Web utilize **Hypertext Transfer Protocol (HTTP)** to transfer documents

**Client**

**Server**

# How Does a Web Browser Work?

- Browser and the World Wide Web utilize **Hypertext Transfer Protocol (HTTP)** to transfer documents



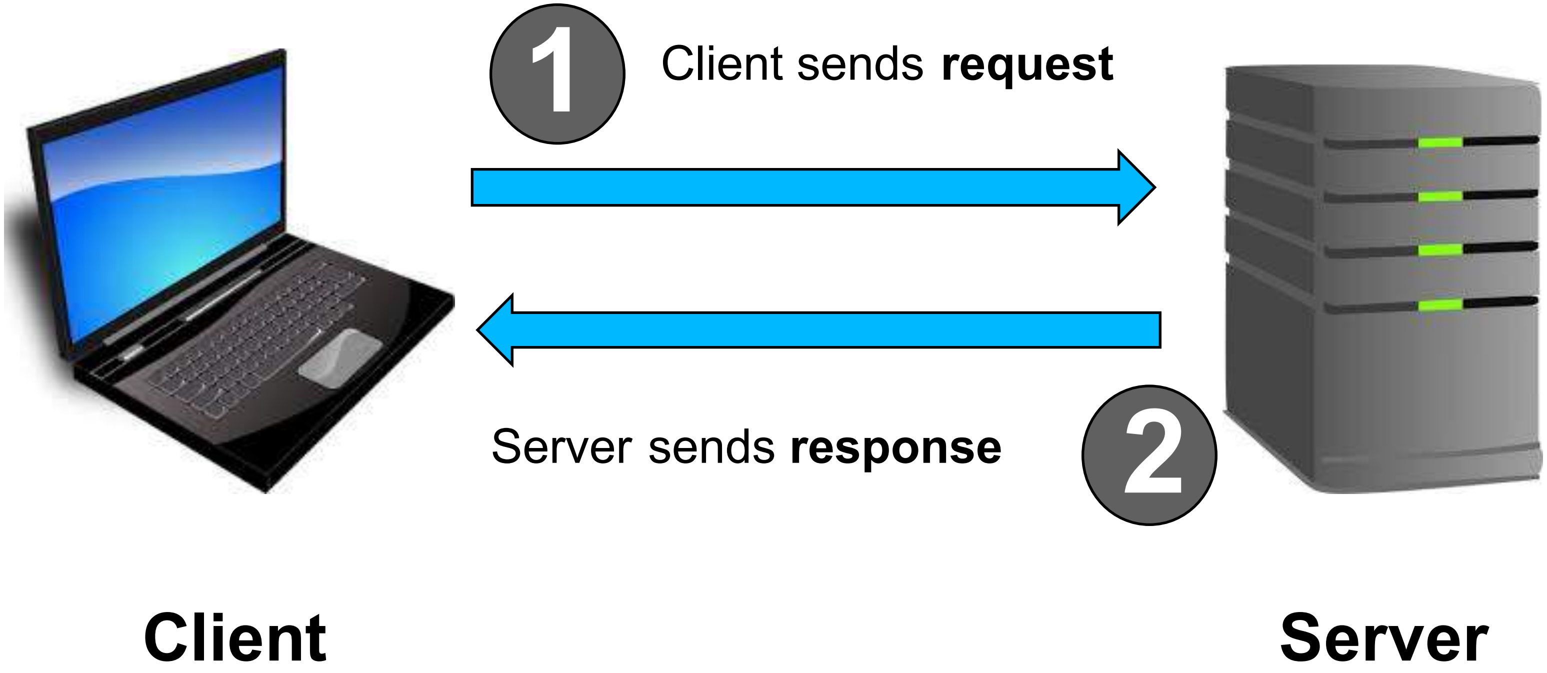
1 Client sends **request**

**Client**

**Server**

# How Does a Web Browser Work?

- Browser and the World Wide Web utilize **Hypertext Transfer Protocol (HTTP)** to transfer documents

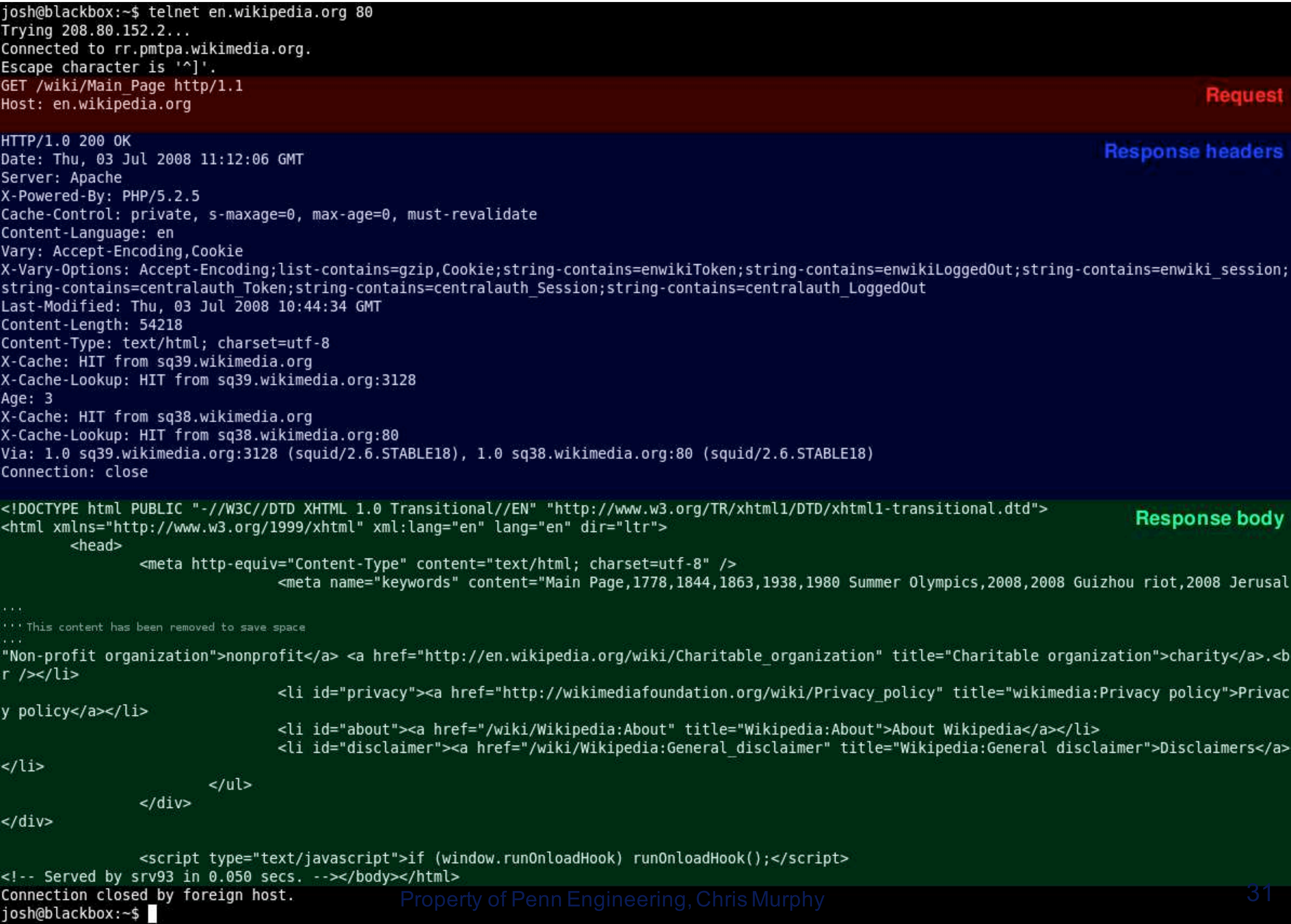


# HTTP Overview

- HTTP is a plain-text, human-readable protocol used for exchanging data on the Web

- Initially developed by Tim Berners-Lee at CERN in 1989

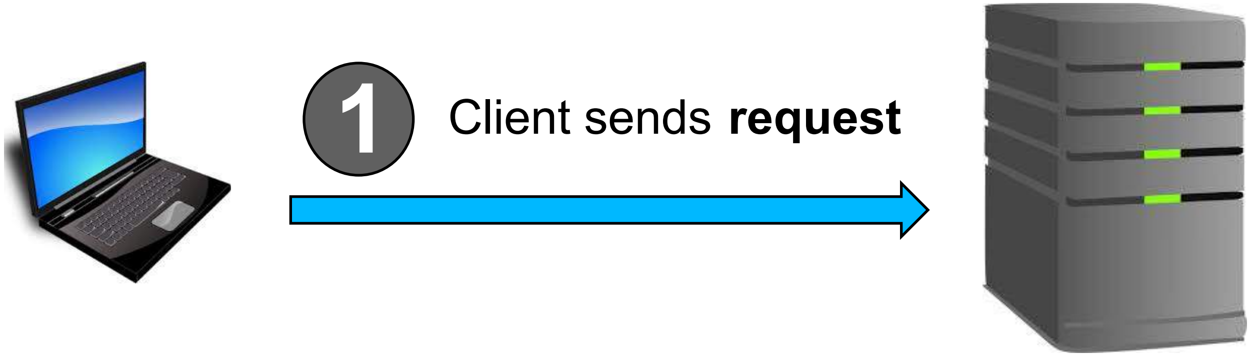
- Based on client-server model:
  - Client sends **request** for resource, possibly including information about the client
  - Server sends **response**, including header (status information) and requested resource



```
josh@blackbox:~$ telnet en.wikipedia.org 80
Trying 208.80.152.2...
Connected to rr.pmtpa.wikimedia.org.
Escape character is '^]'.
GET /wiki/Main_Page http/1.1
Host: en.wikipedia.org

HTTP/1.0 200 OK
Date: Thu, 03 Jul 2008 11:12:06 GMT
Server: Apache
X-Powered-By: PHP/5.2.5
Cache-Control: private, s-maxage=0, max-age=0, must-revalidate
Content-Language: en
Vary: Accept-Encoding,Cookie
X-Vary-Options: Accept-Encoding;list-contains=gzip,Cookie;string-contains=enwikiToken;string-contains=enwikiLoggedOut;string-contains=enwiki_session;
string-contains=centralauth_Token;string-contains=centralauth_Session;string-contains=centralauth_LoggedOut
Last-Modified: Thu, 03 Jul 2008 10:44:34 GMT
Content-Length: 54218
Content-Type: text/html; charset=utf-8
X-Cache: HIT from sq39.wikimedia.org
X-Cache-Lookup: HIT from sq39.wikimedia.org:3128
Age: 3
X-Cache: HIT from sq38.wikimedia.org
X-Cache-Lookup: HIT from sq38.wikimedia.org:80
Via: 1.0 sq39.wikimedia.org:3128 (squid/2.6.STABLE18), 1.0 sq38.wikimedia.org:80 (squid/2.6.STABLE18)
Connection: close

<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Transitional//EN" "http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en" lang="en" dir="ltr">
        <head>
                <meta http-equiv="Content-Type" content="text/html; charset=utf-8" />
                                <meta name="keywords" content="Main Page,1778,1844,1863,1938,1980 Summer Olympics,2008,2008 Guizhou riot,2008 Jerusal
'''
''' This content has been removed to save space
'''
"Non-profit organization">nonprofit</a> <a href="http://en.wikipedia.org/wiki/Charitable_organization" title="Charitable organization">charity</a>.<b
r /></li>
                                <li id="privacy"><a href="http://wikimediafoundation.org/wiki/Privacy_policy" title="wikimedia:Privacy policy">Privac
y policy</a></li>
                                <li id="about"><a href="/wiki/Wikipedia:About" title="Wikipedia:About">About Wikipedia</a></li>
                                <li id="disclaimer"><a href="/wiki/Wikipedia:General_disclaimer" title="Wikipedia:General disclaimer">Disclaimers</a>
</li>
                        </ul>
                </div>
</div>

                <script type="text/javascript">if (window.runOnloadHook) runOnloadHook();</script>
<!-- Served by srv93 in 0.050 secs. --></body></html>
Connection closed by foreign host.
josh@blackbox:~$ 
```

# Anatomy of an HTTP Request

- First line of request will always be a **verb** followed by an **argument**
  - **GET** – retrieve resource
  - **HEAD** – retrieve only headers (information about the resource)
  - **POST** – create resource (usually used in form submission context)
- Next comes the protocol (usually HTTP/1.1)
- Optionally include other information about the request and/or the client



1

Client sends **request**

# HTTP Request Example

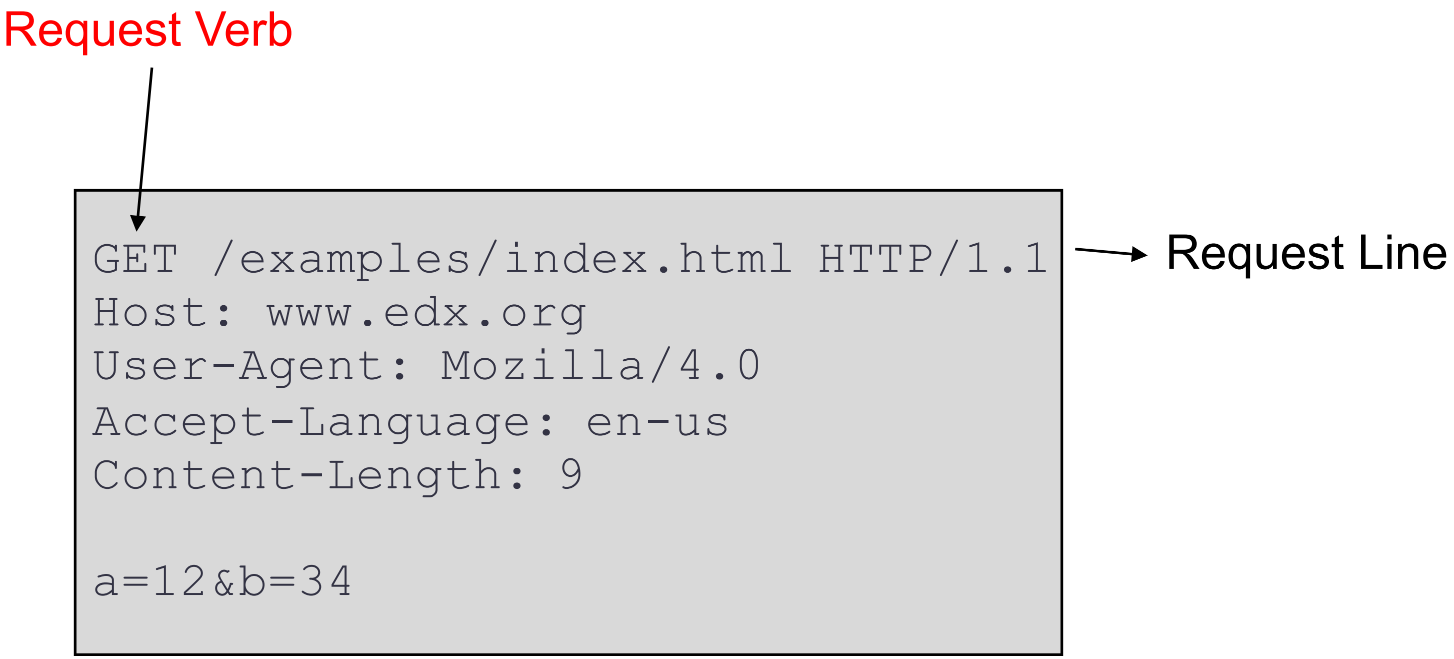
```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

# HTTP Request Example

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

→ Request Line

# HTTP Request Example

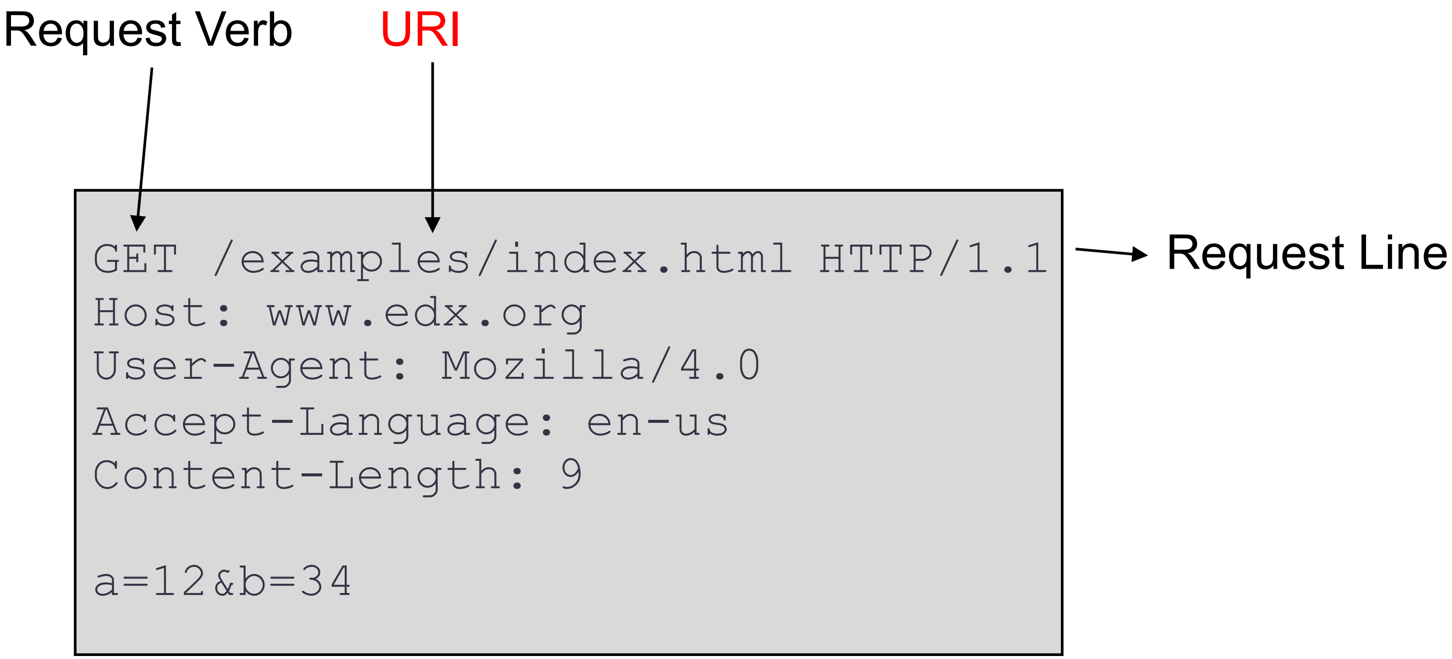


Request Verb

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

Request Line

# HTTP Request Example



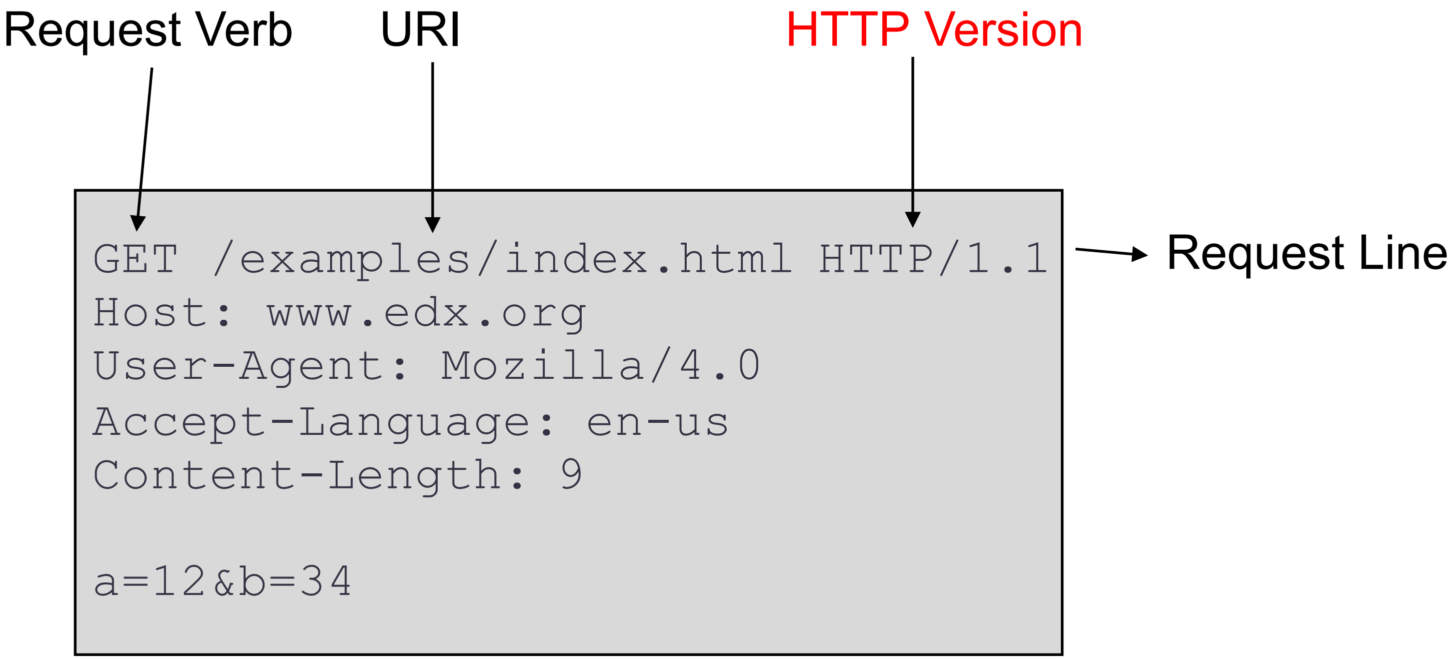
Request Verb

URI

Request Line

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

# HTTP Request Example



Request Verb

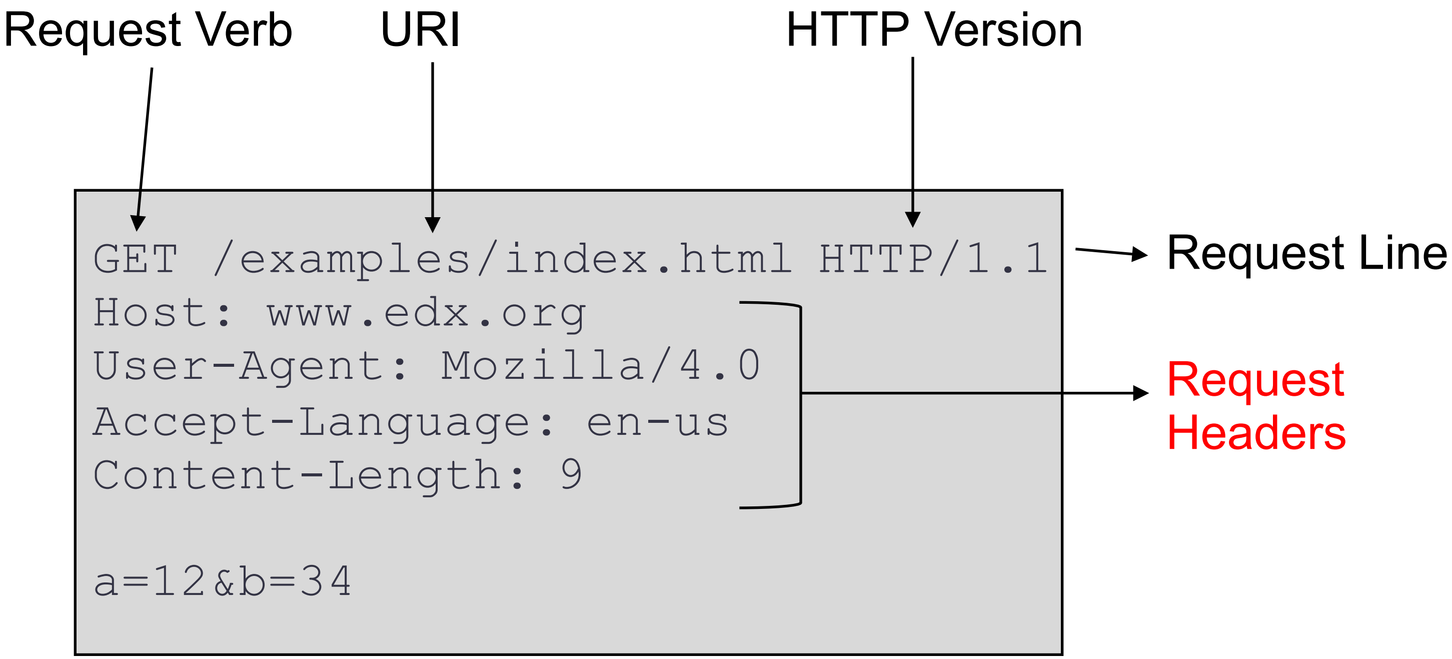
URI

HTTP Version

Request Line

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

# HTTP Request Example



Request Verb

URI

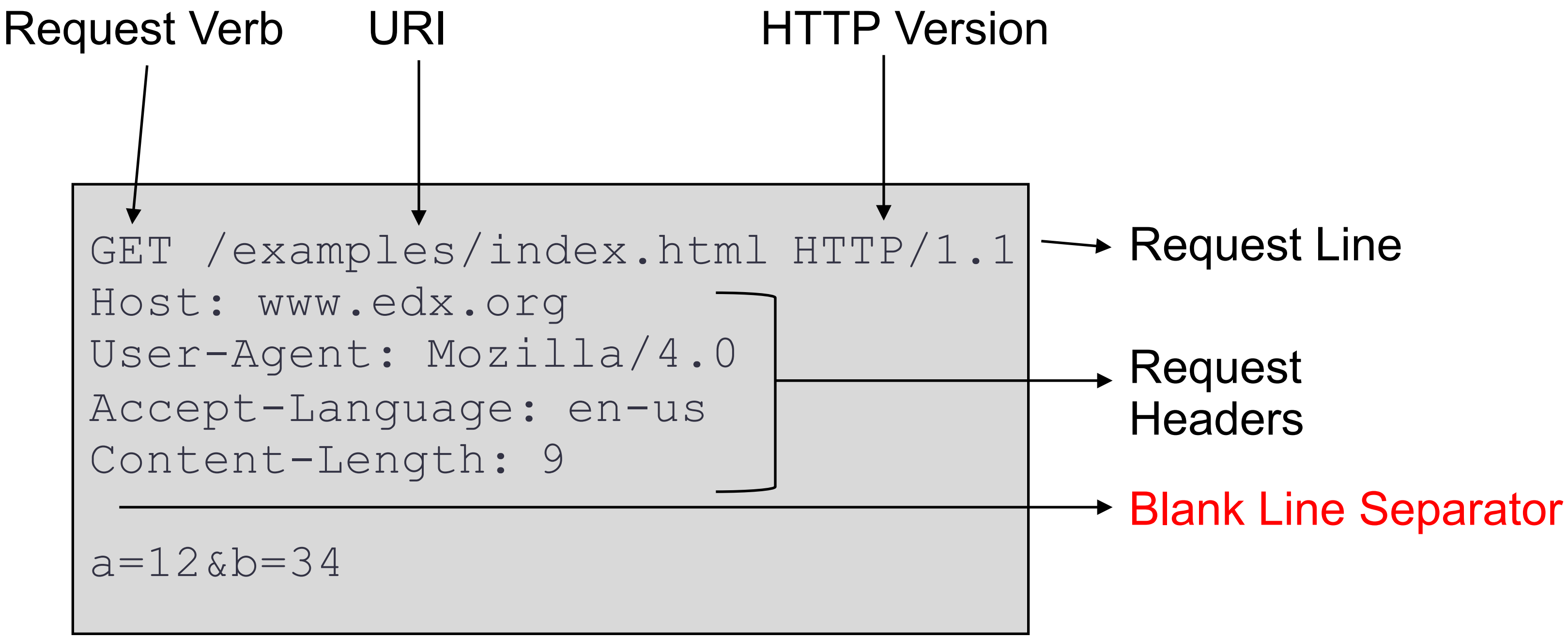
HTTP Version

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

Request Line

Request Headers

# HTTP Request Example



Request Verb

URI

HTTP Version

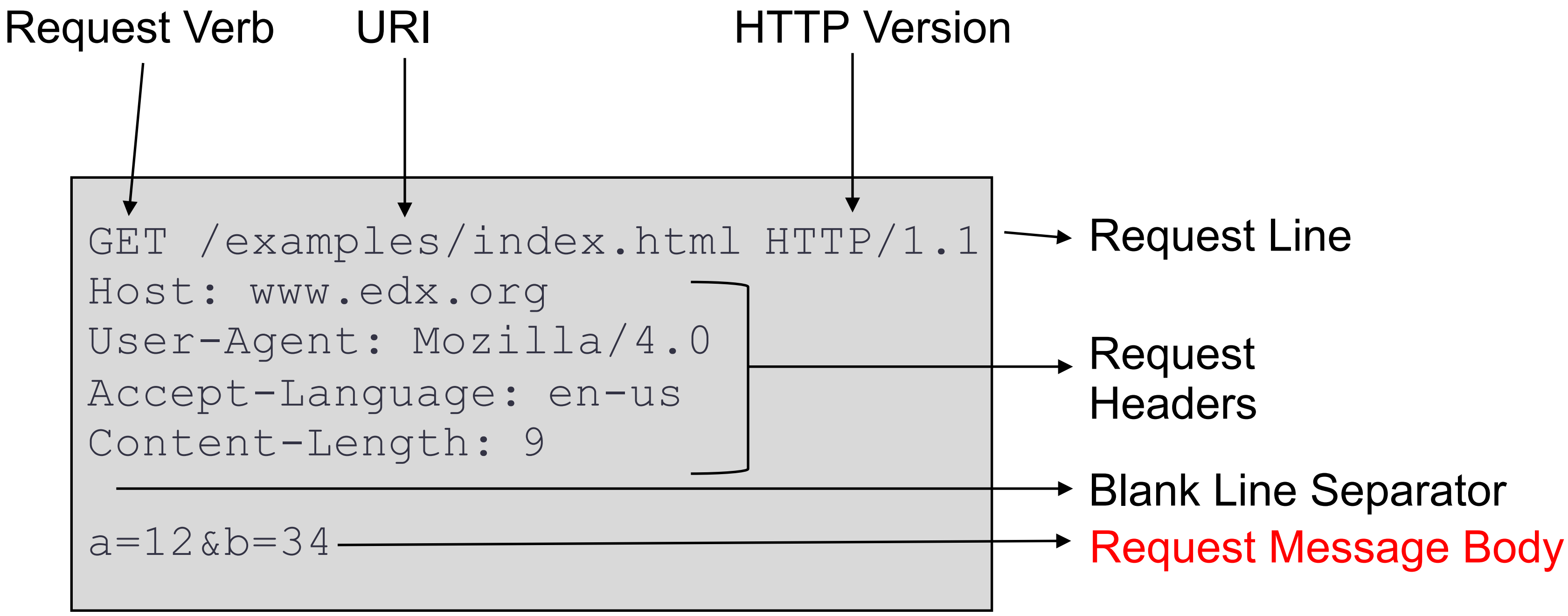
```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

Request Line

Request Headers

Blank Line Separator

# HTTP Request Example



Request Verb

URI

HTTP Version

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

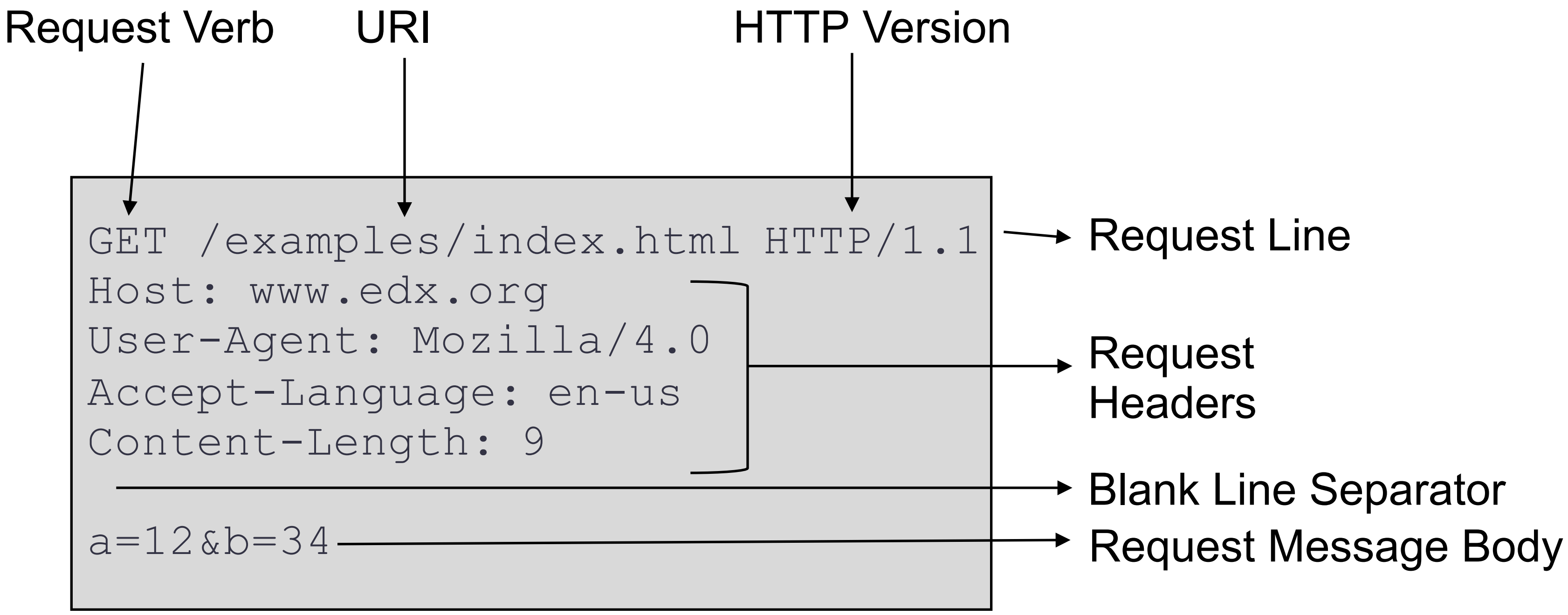
Request Line

Request Headers

Blank Line Separator

Request Message Body

# HTTP Request Example



Request Verb

URI

HTTP Version

```
GET /examples/index.html HTTP/1.1
Host: www.edx.org
User-Agent: Mozilla/4.0
Accept-Language: en-us
Content-Length: 9

a=12&b=34
```

Request Line

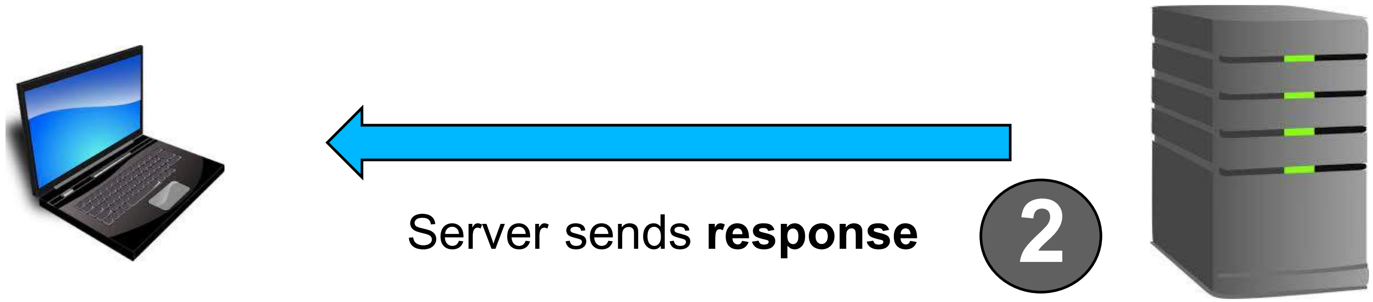
Request Headers

Blank Line Separator

Request Message Body

# Anatomy of an HTTP Response

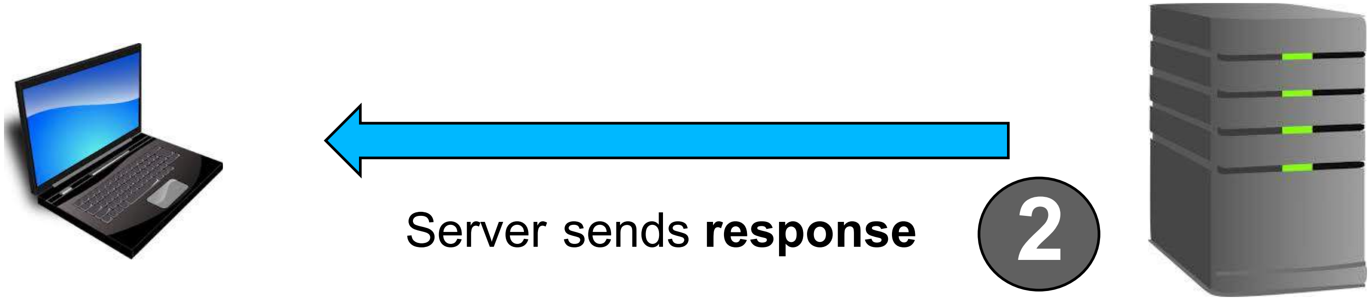
- First line is always protocol and **status code**
  - 1XX – information only
  - 2XX – success
  - 3XX – client redirect
  - 4XX – client error
  - 5XX – server error



Server sends **response**

# Most Common Status Codes

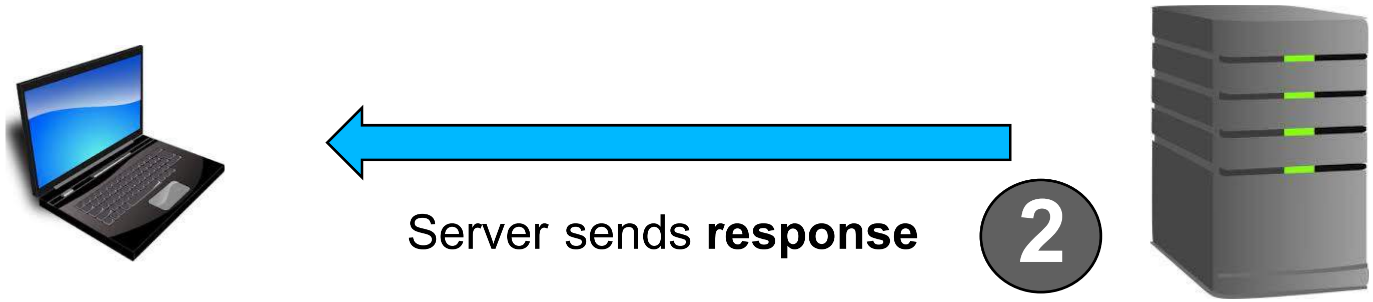
- **200 OK** – request succeeded, resulting resource (as stated in request) will be included in message body
- **404 Not Found** – requested resource does not exist
- **500 Server Error** – Error on the server side in processing request



Server sends **response**

# Anatomy of an HTTP Response

- Following protocol and status code will be other **header information** regarding the response and/or the server
- Then a blank line
- Then the response body, i.e. the resource that was requested



# HTTP Response Example

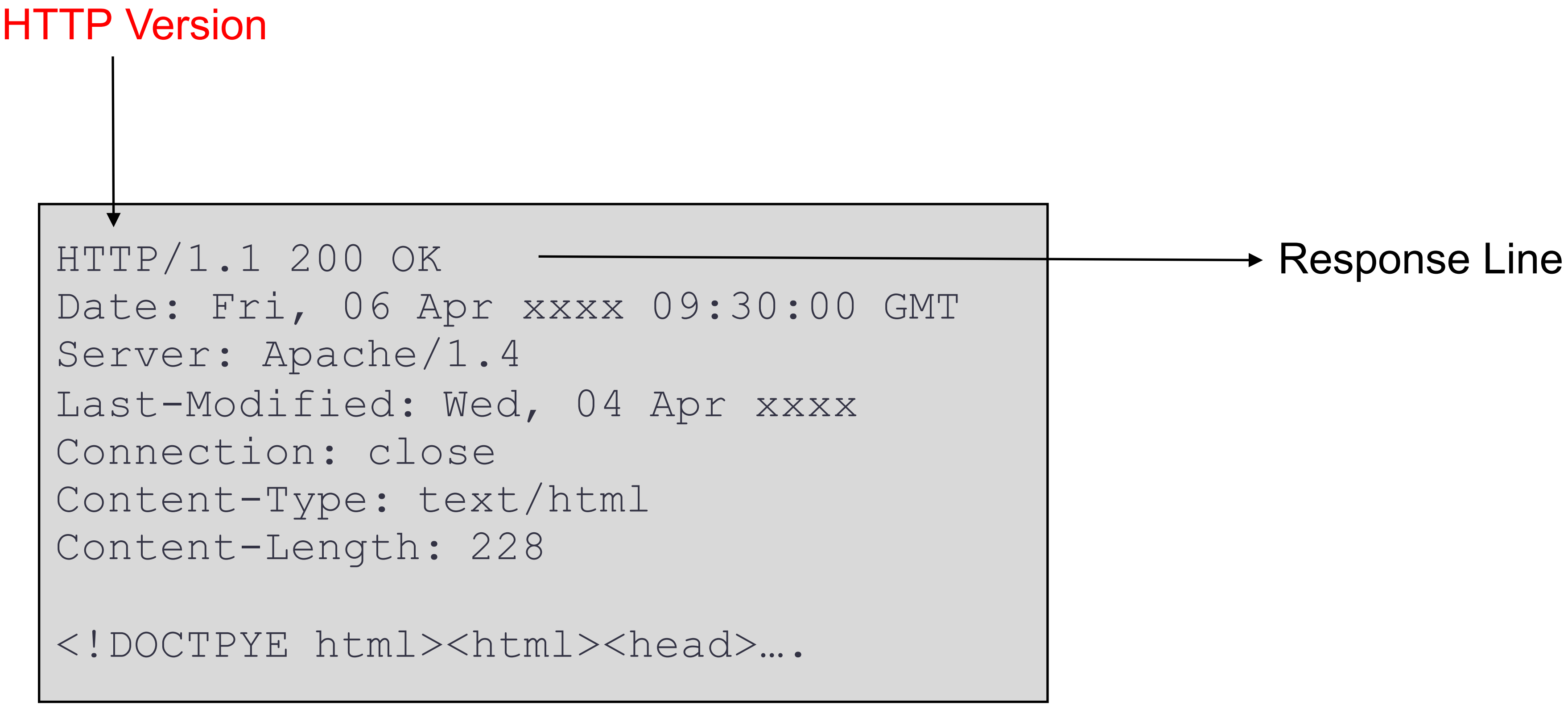
```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

# HTTP Response Example

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

Response Line

# HTTP Response Example

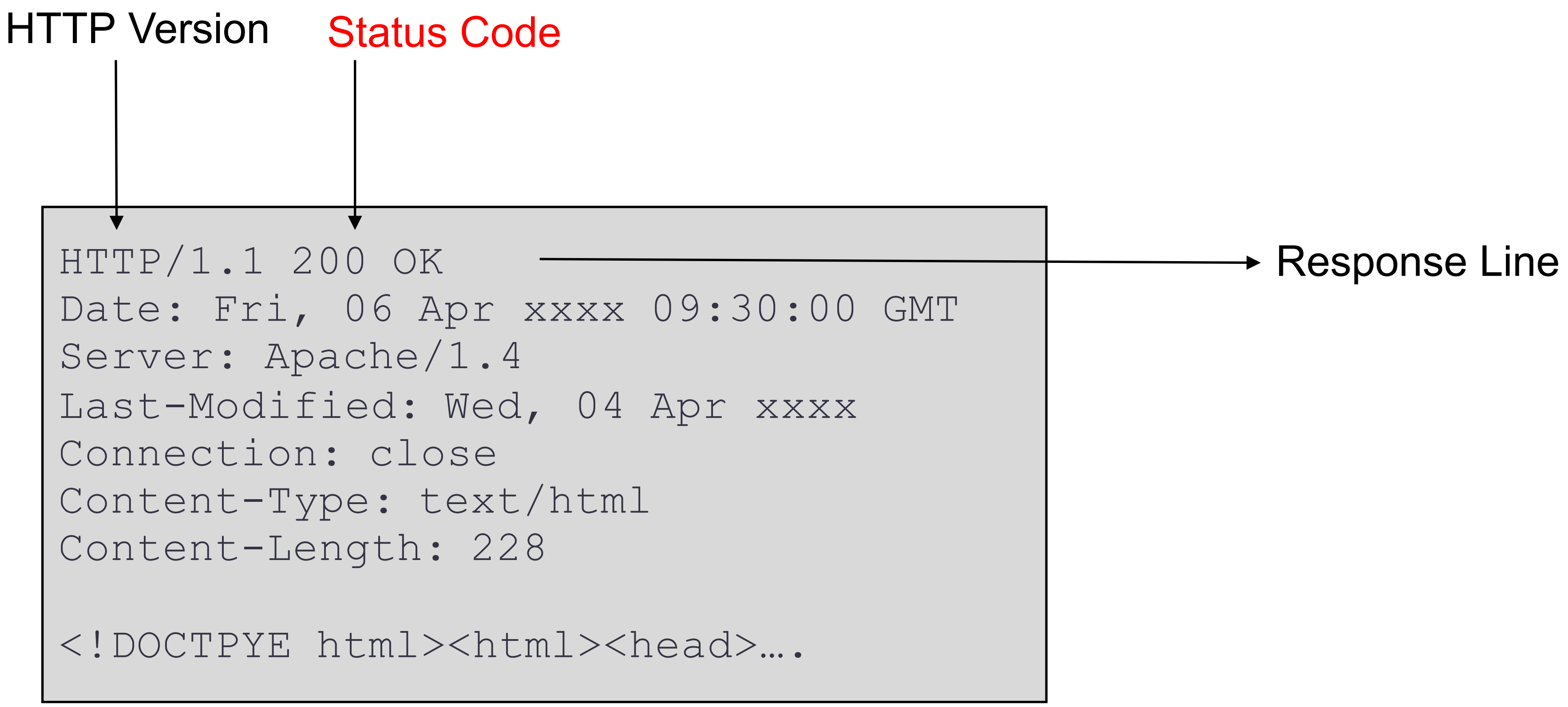


HTTP Version

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

Response Line

# HTTP Response Example



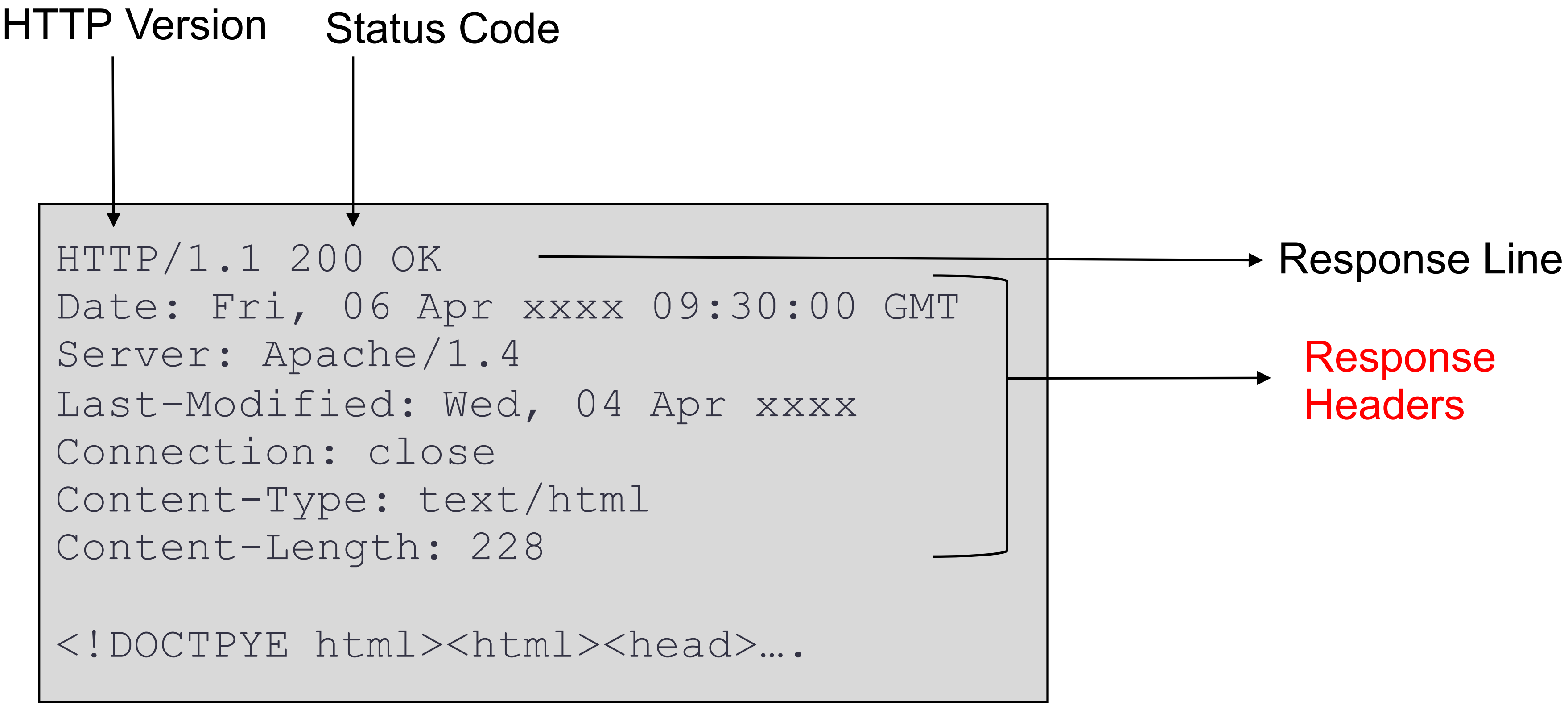
HTTP Version

Status Code

Response Line

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

# HTTP Response Example



HTTP Version

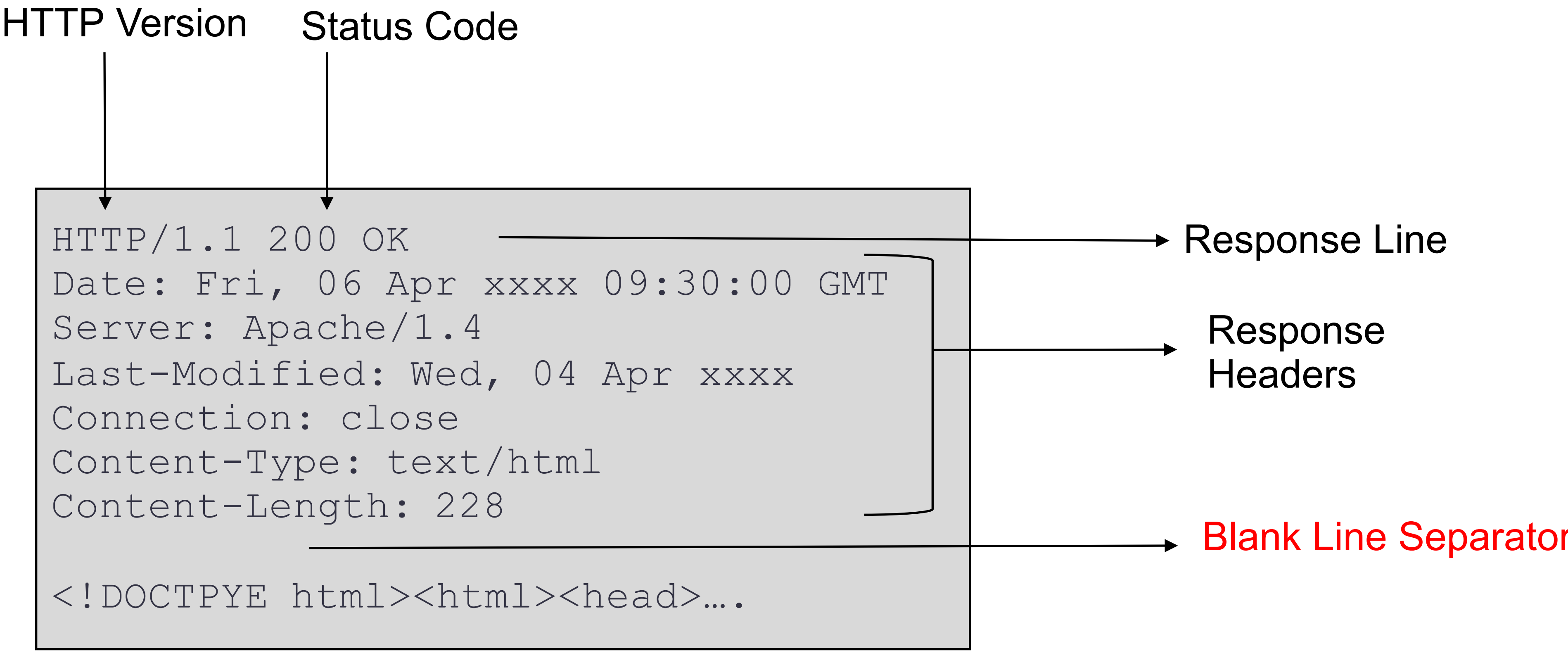
Status Code

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

Response Line

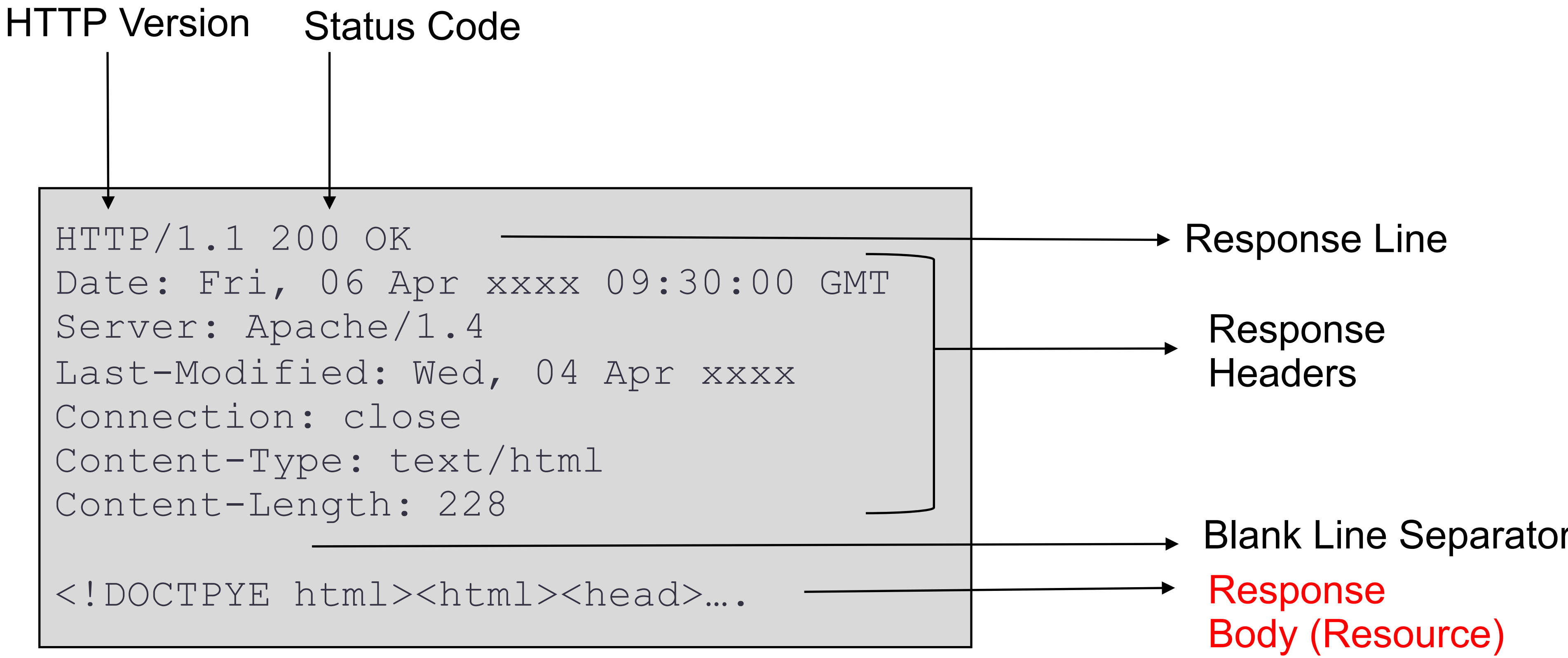
Response Headers

# HTTP Response Example



```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

# HTTP Response Example



HTTP Version

Status Code

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

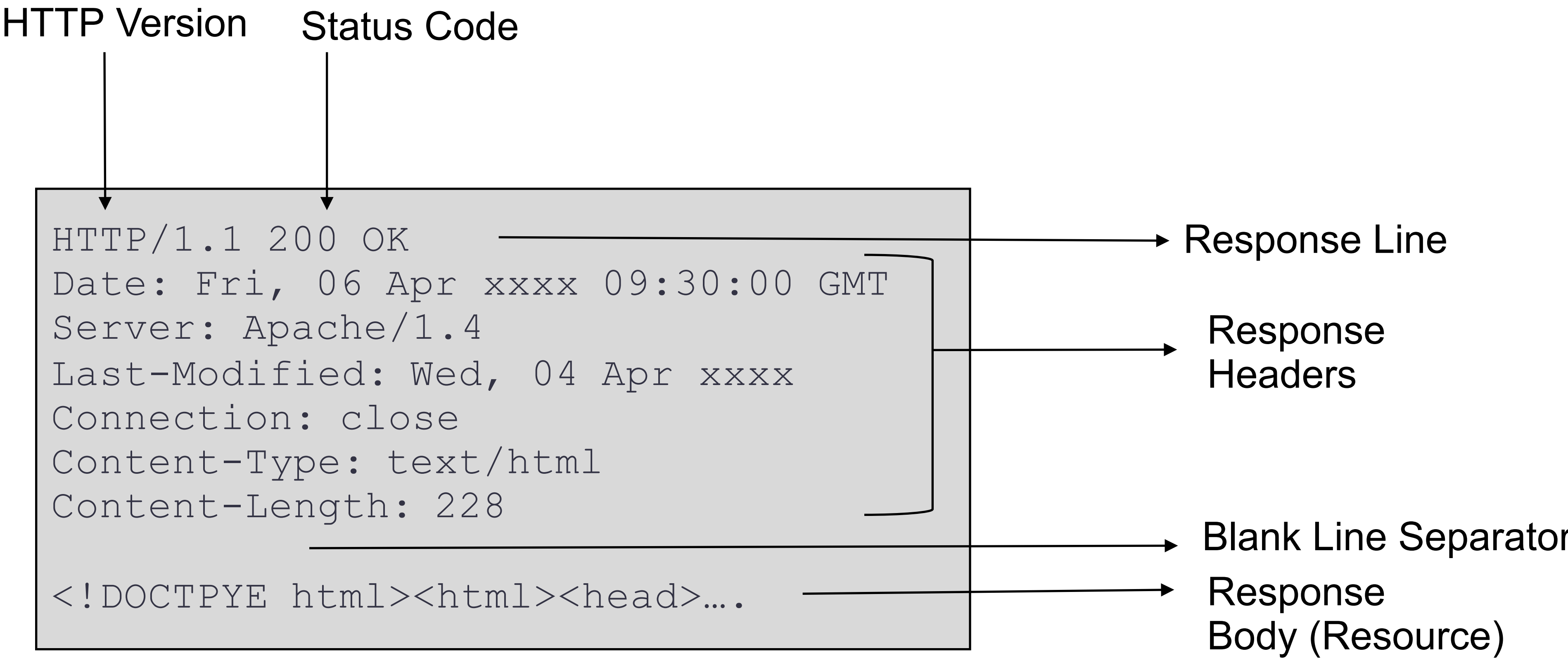
Response Line

Response Headers

Blank Line Separator

Response Body (Resource)

# HTTP Response Example



HTTP Version

Status Code

```
HTTP/1.1 200 OK
Date: Fri, 06 Apr xxxx 09:30:00 GMT
Server: Apache/1.4
Last-Modified: Wed, 04 Apr xxxx
Connection: close
Content-Type: text/html
Content-Length: 228

<!DOCTPYE html><html><head>….
```

Response Line

Response Headers

Blank Line Separator

Response Body (Resource)

# Summary

- Web browsers are used to access data on the Web
- Browsers communicate with web servers using HTTP
- HTTP is based on a client-server model:
  - Client sends **request** for resource, possibly including information about the client
  - Server sends **response**, including header (status information) and requested resource