

WebTech (1)

Prof. Dr.-Ing. S. Gössner

University of Applied Sciences Lippe & Höxter

Inhalt

- [Inhalt](#)
- [Überblick](#)
- [Web-Technologien - Definition](#)
- [Entstehung des Web](#)
- [Funktionsweise des Web](#)
- [TCP/IP - Protokoll](#)
- [IP-Adressen](#)
- [Domain Name System](#)
- [Selbstorganisation des Internet](#)
- [Client/Server - Prinzip \(1\)](#)
- [Client/Server - Prinzip \(2\)](#)
- [HTTP - Hypertext Transport Protocol](#)
- [URI - Uniform Resource Identifier](#)
- [Webstandards - Dokumentformate](#)
- [Webstandards - W3C](#)
- [Webstandards - im Fokus](#)
- [HTML \(1\)](#)
- [HTML \(2\)](#)
- [HTML \(3\)](#)
- [CSS \(1\)](#)
- [CSS \(2\)](#)
- [Web-Programmiersprachen](#)
- [Serverseitige Programmiersprachen](#)
- [Clientseitige Programmiersprachen](#)
- [Werkzeuge](#)

Überblick

- Web-Technologien
 - Entstehung und Funktionsweise
 - Client/Server-Prinzip
 - Basiskomponenten und -dienste
 - Webstandards
 - HTML + CSS
 - XML und Co.
 - Server- und clientseitige Programmiersprachen
 - Wikis und Weblogs
 - Webanwendungen
- Multimedia
 - Einführung
 - Audio und Video
 - Grafik und SVG
 - Animation und SMIL
- Ausarbeitungen und Präsentationen

Web-Technologien - Definition

Als *Web-Technologien* bezeichnen wir die Summe aller Technologien, mittels derer das *World Wide Web* betrieben wird.

Entstehung des Web

- Tim Berners-Lee entwickelte am europäischen Kernforschungslabor [CERN](#) in Genf einen ersten Webbrowser und Webserver als experimentelles System. Heute ist er Direktor des World Wide Web Consortiums.
 - 1990 formulierte er schliesslich die drei Säulen des WWW:
 - HTTP
 - URI
 - HTML
- Marc Andreessen programmierte als Student 1993 den ersten Webbrowser *Mosaic*, der außer Text auch Grafiken anzeigen konnte, ohne jene extra laden zu müssen. Seine spätere Fa. *Netscape* erreichte dann mit dem *Navigator* einen Marktanteil von 90%.
- [W3C](#) Das World Wide Web Consortium (*W3C*) wurde 1994 von Tim Berners-Lee gegründet. Dieser [W3C](#) Zusammenschluss interessierter Firmen, Personen und Hochschulen erarbeitet freie offene Webstandards.



Tim Berners-Lee



Marc Andreessen

Funktionsweise des Web

- TCP/IP-Protokoll
- IP-Adressen
- DNS – Domain Name Service
- Selbstorganisation des Internet
- Client/Server-Technologie

TCP/IP - Protokoll

Mit diesem Übertragungsprotokoll für Daten im Netz werden

- beim Absender Teildatenpakete gebildet und mit Adresse und Sequenznummern versehen
- beim Empfänger alle Datenpakete in ihrer ursprünglichen Reihenfolge wieder zusammengesetzt.
- E-Mails, Webseiten und Dateien verschickt.

TCP/IP-Schicht	Protokolle (Auswahl)
Anwendung	HTTP, FTP, SMTP, POP3, Telnet, SSH
Transport	TCP
Netz	IP
Netzzugang	Ethernet, WLAN

Mittlerweile unterstützen alle Betriebssysteme das TCP/IP-Protokoll.

IP-Adressen

- Jeder Rechner, der – wenn auch nur momentan – Bestandteil des Internet ist, muss eine weltweit eindeutige IP-Adresse besitzen.
- IP-Adressen werden als vier — durch Punkte getrennte — Bytes in Dezimaldarstellung notiert. (Beispiel: **193.16.112.83**)
- Netzbetreiber oder ISP's können bei einer Einwahl eines Netzbenutzers diesem eine dynamische IP-Adresse zuweisen.
- Da der Adressraum allmählich eng wird, soll er von momentan 32 Bit (IPv4) auf 128 Bit (IPv6) erweitert werden.

Domain Name System

- Wir Menschen können mit IP-Adressen aus Dezimalzahlen nicht besonders gut umgehen.
- Also wird jeder IP-Adresse eine Namesadresse (Domain Name) zugeordnet. Beispiel: **www.fh-luh.de**.
- Zu einer Namens- oder IP-Adresse kann über einen **Dienst** Information abgerufen werden.
- DNS ist eine verteilte Datenbank, die den Namensraum im Internet verwaltet.
- Die hierarchische Struktur der Namensadressen wird über einzelne durch Punkte getrennte Namensteile abgebildet:
 - **de** — Top-Level-Domain
 - **fh-luh.de** — Domain
 - **www.fh-luh.de** — Sub-Level-Domain
- Die Domainvergabe wird zentral über NIC bzw. über organisations- oder länderspezifische Verwaltungsorgane gesteuert (**NIC**).

Selbstorganisation des Internet

- Das Internet — und das Web — ist anarchisch (*herrschaftsfrei, führerlos*).
- Ein Funktionieren basiert auf der grosszügigen Bereitschaft aller Teilnehmer, Informationen kostenlos bereitzustellen und weiterzuleiten. Anfallende Kosten tragen die Netzbetreiber selbst.
- Die wenigen notwendigen Steuerungsfunktionen übernehmen internationale oder nationale Organisationen.

ICANN

[Oberste Organisation](#) für die technische Regulierung des Internets (z.B. neue Top-Level-Domains)

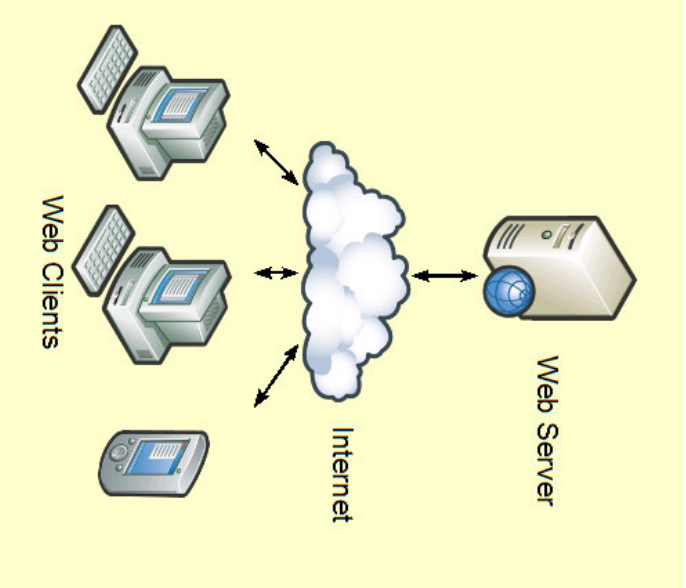
IETF

[Organisation](#) zur Erarbeitung technischer Standards

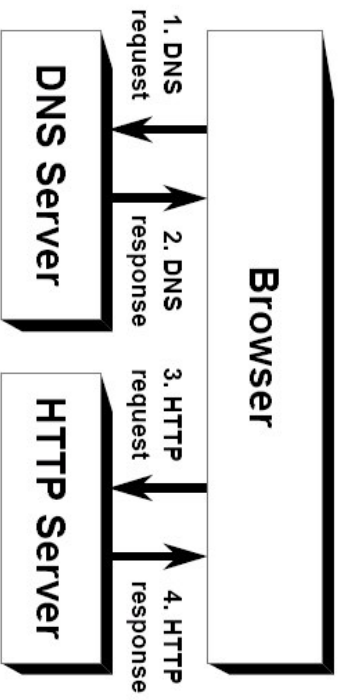
W3C

[Organisation](#) zur Erarbeitung von Webstandards.

Client/Server - Prinzip (1)



Client/Server - Prinzip (2)



HTTP - Hypertext Transport Protocol

- 1989 von [Tim Berners-Lee](#) am [Cern](#) definiert und erstmalig implementiert.
- bildet zusammen mit [URL](#) und [HTML](#) die 3 Säulen des *World Wide Web*.
- HTTP erlaubt eine bidirektionale Kommunikation in der Form:
 1. Anfrage (*Request*) seitens des Web-Clients an den Web-Server
 2. Antwort (*Response*) vom Web-Server zurück an den Web-Client.
- HTTP stellt verschiedene Request-Methoden bereit:
 - **GET** — Inhalte von Server anfordern (Webseite).
 - **POST** — Daten an den Server senden (Formulare).
 - **PUT** — Daten/Datei auf den Server "hochladen".
 - **DELETE** — Daten/Datei vom Server löschen.
 - **HEAD** — wie GET, jedoch nur *Header*-Daten ohne Inhalte anfordern.
- HTTP ist ein zustandsloses Protokoll.
- [HTTP in action](#)

URI - Uniform Resource Identifier

- Ein URI ist eine Zeichenfolge zur Identifizierung einer Resource im Internet.
- Damit werden Webseiten, Dateien, E-Mail Adressen und Webservices angesprochen.
- Eine Untermenge der URI's bilden die URL's.
- URI's als Verweise innerhalb von Webseiten auf jeweils andere bezeichnen wir auch als *Hyperlinks*.

Syntax:

```
Schema://Benutzer:Passwort@Server:Port/Pfad?Anfrage#Fragment
```

Beispiel:

<http://www.fh-luh.de:80/fb7#fuss>

Webstandards - Dokumentformate

Von einem standardisierten Dokumentformat sprechen wir, wenn ...

- die Dokumentstruktur festgelegten Regeln unterliegt.
- eine gewisse Verbreitung vorliegt.

ein Standard ist offen, wenn ...

- die Dokumentstruktur vollständig offengelegt ist.
- die Festlegung der Dokumentstruktur durch mehrere Personen und/oder Firmen erfolgt.
- Eine Erweiterung der Dokumentstruktur möglich/erlaubt ist.
- die Verarbeitung eines Dokuments nicht an ein einzelnes Programm gebunden ist.

... ansonsten nennen wir den Standard *proprietär*.

Webstandards - W3C

Das W3C ist ...

- das **World Wide Web Consortium** ([W3C](#)).
- 1994 gegründet und bis heute geleitet von *Tim Berners-Lee*.
- ein Gremium zur Standardisierung von Web-Technologien — insbesondere Datenformaten
- ein Konsortium aus unterschiedlichen Mitgliedern (Industrie, Hochschulen, Forschungseinrichtungen, unabhängige Experten).
- selbstverpflichtet, nur Standards frei von Lizenzgebühren zu entwickeln.
- keine zwischenstaatlich anerkannte Organisation wie bspw. die ISO.
- daher nicht berechtigt, *Standards* festzulegen und bezeichnet ihre Arbeiten als *Recommendations* (*Empfehlungen*).
- dem Entwicklungsprozess verbunden: *Working Draft* -> *Last Call* -> *Candidate Recommendation* -> *Proposed Recommendation* -> *Recommendation*
- gegenüberlich zunehmend in der Kritik ...
 - ... weil der Standardisierungsprozess zu lange dauert.
 - ... wegen angeblicher Überbewertung von Mitgliederinteressen.
 - ... wegen Nichtberücksichtigung aktueller *De-Facto-Standards*.
- gegenüberlich konfrontiert mit einem Interessensverbund von Browserherstellern — [WHATI](#)

Webstandards - im Fokus

- Hypertext Markup Language (HTML)
- Extensible Markup Language (XML)
- Extensible Hypertext Markup Language (XHTML)
- Cascading Style Sheets (CSS)
- Portable Network Graphics (PNG)
- Scalable Vector Graphics (SVG)
- Synchronized Multimedia Integration Language (SMIL)
- Mathematical Markup Language (MathML)
- Resource Description Framework (RDF)
- Document Object Model (DOM)
- Web Accessibility Initiative (WAI)
- XML Path Language (XPath)

HTML (1)

HTML ist ...

- die **Hypertext Markup Language**
- seit 1990 das Dokumentformat der Webseiten.
- entworfen worden, um Dokumente untereinander zu verlinken (*Hypertext*).
- eine Auszeichnungssprache (*Markup*).
Beispiel: `Fett` und `<i>schräg</i>`
- reines Textformat und damit hochgradig portabel.
- bewusst einfach gehalten.
- ein Abkömmling der **SGML**.
- ein Format zur Beschreibung der logischen Dokumentstruktur.
- keine Seitenbeschreibungssprache wie Postscript oder PDF.
- **multimedial**.
- in seiner gegenwärtigen Version 4.01 nicht mehr zeitgemäss.

HTML (2)

HTML genügt nach Meinung des W3C nicht mehr den künftigen Anforderungen.

- neue Geräte (Digital TV, mobile Geräte, Automobi, Kühlschränk, Kleidung) erfordern ...
 - einen erweiterten Sprachumfang mit moderner, leistungsfähiger Funktionalität, oder ...
 - einen reduzierten Sprachumfang für einfache, kleine Geräte.
- Internationalisierung wird nicht ausreichend unterstützt (asiatische Schriften).
- die heutigen interaktiven Fähigkeiten reichen nicht aus.
- die geforderten Erweiterungen in Hinblick auf Grafik, Audio, Video, math. Formeln, etc. würden die Sprache überfrachten.
- das auf XML basierende XHTML ist erklärter Nachfolger von HTML.

HTML (3)

Die Grundstruktur von HTML ist einfach.

- eine Dokumenttypdeklaration (DTD) am Anfang der Dokumentdatei definiert den erlaubten Sprachumfang (Syntax).
- der *head*-Bereich enthält nicht sichtbare Information, Silangaben, Programmcode und Meta-Angaben.
- der *body*-Bereich beinhaltet die strukturelle Information des sichtbaren Dokumentinhalts.

```
<!DOCTYPE HTML PUBLIC "-//W3C//DTD HTML 4.01//EN"
  "http://www.w3.org/TR/html4/strict.dtd">
<html>
  <head>
    <title>Titel der Webseite</title>
    <!--Evtl. weitere Kopfinformationen-->
  </head>
  <body>
    Inhalt der Webseite
  </body>
</html>
```

CSS (1)

CSS

- bedeutet **Cascading Style Sheets** ([CSS](#)).
- ist eine deklarative Formatbeschreibungssprache des W3C für HTML und XML-Dokumente.
- hat eine ähnliche Funktion, wie Formatvorlagen eines *Word*-Dokuments.
- unterstützt die angestrebte Trennung von *Struktur* und *Aussehen* eines Dokuments.
- kann die logische Struktur eines Dokuments für unterschiedliche Ausgabemedien aufbereiten (Projektion, Druck, Sprachausgabe, ...).
- aktuelle Browser haben gegenwärtig CSS2 implementiert.
- die neue Version CSS3 ist momentan in der Entwicklung des W3C.

CSS (2)

Die CSS-Syntax unterscheidet sich erheblich von HTML.

```
selector { attribute1: value1; /* comment */  
          attribute2: value2;  
          ...  
          attributen: valuen;  
}
```

Beispiel:

```
div.note {  
  font-style: italic;  
  color: red;  
  background-color: white;  
}
```

CSS-Angaben werden üblicherweise in einer zentralen CSS-Datei abgelegt und können so von mehreren HTML-Dokumenten verwendet werden.

Web-Programmiersprachen

- Imperative Programmiersprachen können zur Erstellung einer Webseite auf Anforderung eingesetzt werden.
- Programmiersprachen können andererseits zur Steuerung einer Programmlogik *innerhalb* einer Webseite verwendet werden.
- Wir unterscheiden dementsprechend zwischen *serverseitigen* und *clientseitigen* Programmiersprachen.

Serverseitige Programmiersprachen

- Auf der Seite des Webservers kann grundsätzlich jede Programmiersprache eingesetzt werden, wenn sie nur das zugrundeliegende Betriebssystem unterstützt.
- In der Praxis haben sich allerdings einige durchgesetzt:
 - Perl
 - PHP
 - Java
 - Python
 - Basic
 - C/C++
- Es überwiegen hier die Interpretersprachen
- Aufgabe der hier erstellten Programme ist
 - der Zusammenbau der angeforderten Webseite kurz vor der Auslieferung an den Browser.
 - Zugriff auf gewünschte Datenbankinträge (SQL).
 - Daten von anderen Webservern anfordern.

Clientseitige Programmiersprachen

- Diese Programmiersprachen müssen innerhalb des Browsers verfügbar sein.
- *JavaScript* ist standardisiert (*ECMA 262*) und dominierend.
- *VBScript* ist proprietär und praktisch bedeutungslos.
- *Java Applets* sind ebenfalls proprietär und heute bedeutungslos.
- Programme innerhalb des Browsers dürfen auf keinen Fall auf das Datei- und Betriebssystem des Benutzers zugreifen (Ausnahme: *Cookies*).
- Vor einigen Jahren galt: *JavaScript + HTML = DHTML*
- *JavaScript* hatte lange Zeit ein negatives Image:
 - Sprache für Spielereien.
 - stellt Sicherheitslücke dar; diese Generalisierung basiert jedoch auf Fehlern (*Bugs*) in der Implementierung von Browserkomponenten oder der Sprache selbst.
 - Fehlende professionelle Literatur und vorbildliche Beispielanwendungen.
- Gegenwärtig gewinnt *JavaScript* zunehmend an Ansehen und ist für moderne Webanwendungen unverzichtbar (*Ajax*).

Werkzeuge

- Standards-konformer Webbrowser (IE6, Mozilla, Opera, ...)
- Texteditor ([EditPlus](#))
 - Username: stefan goessner
 - Regcode: 74D1F-84A70-B0017-B835A-F527B
- alternativer Texteditor [Notepad++](#)