

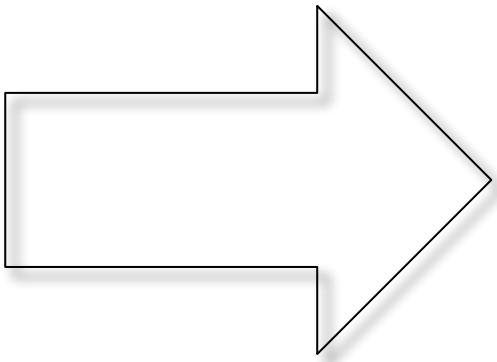
Einführung in Web Engineering

- Bachelor Informatik -
- Bachelor Wirtschaftsinformatik

Teil 1: WWW, HTTP, URI

Manfred Kaul

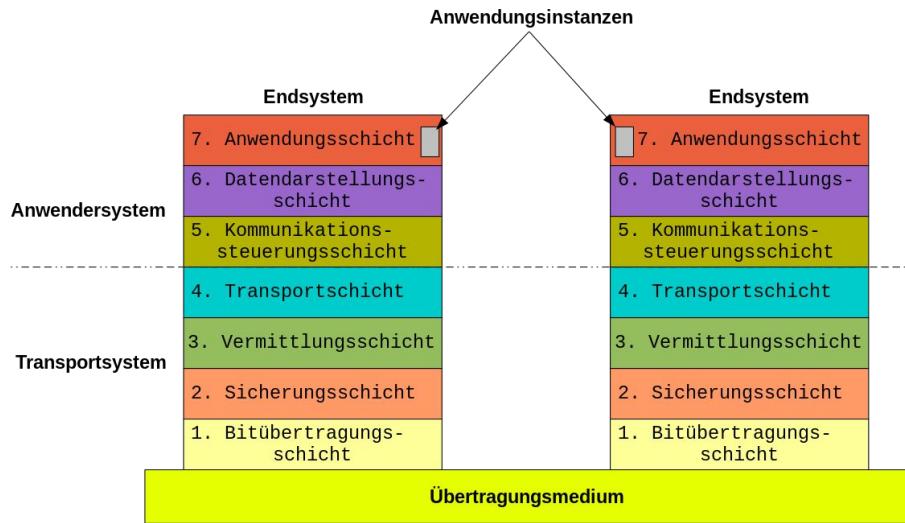
Waren Sie heute schon im Internet?



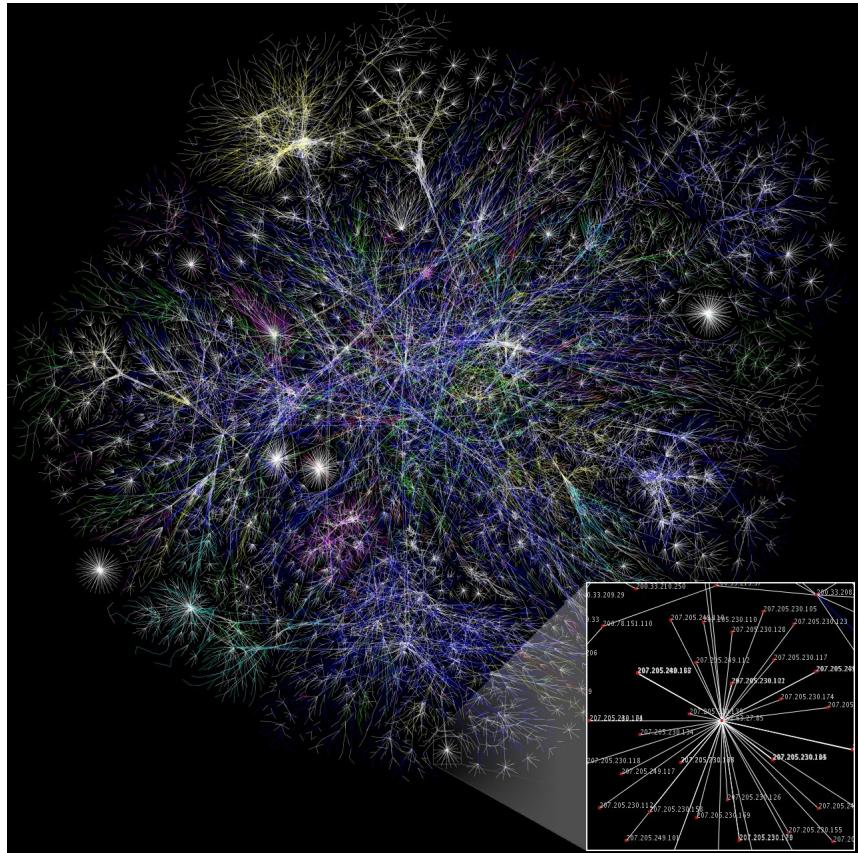
Unterscheide:
Internet ≠ WWW



Internet = Netz der Netze

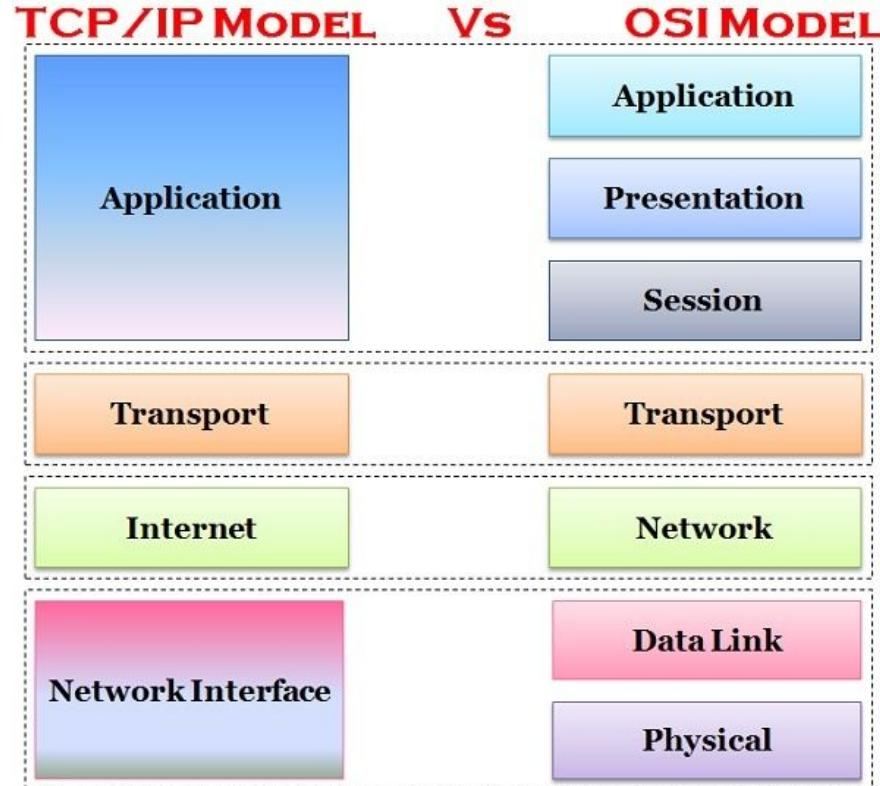


Quelle: Wikipedia,
[https://commons.wikimedia.org/wiki/File:ISO-OSI-7-Schichten-Modell\(in_Deutsch\).svg#/media/File:ISO-OSI-7-Schichten-Modell\(in_Deutsch\).svg](https://commons.wikimedia.org/wiki/File:ISO-OSI-7-Schichten-Modell(in_Deutsch).svg#/media/File:ISO-OSI-7-Schichten-Modell(in_Deutsch).svg)



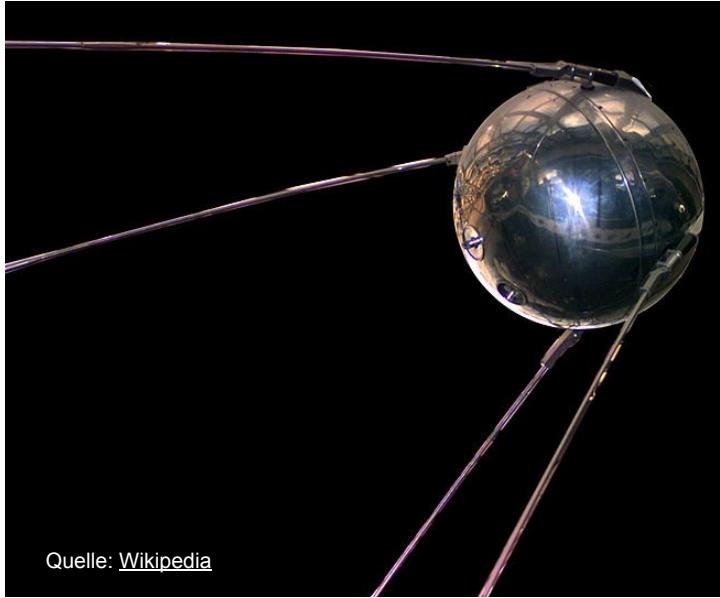
Quelle: Wikipedia, CC BY 2.5,
<https://commons.wikimedia.org/w/index.php?curid=1538544>

Internet Protocol Stack



Quelle: <https://techdifferences.net/difference-between-tcp-ip-and-osi-model/>

Anfänge des Internet

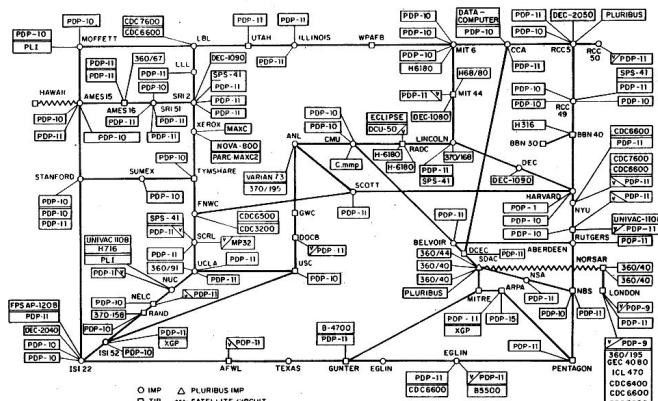


Quelle: [Wikipedia](#)

- Sputnik-Schock 1957
- US-Verteidigungsministerium gründet (Defense) Advanced Research Projects Agency (ARPA, 1958)
 - Mondlandung 1969
 - ARPANET 1969

ARPA Networking Project (ARPANET)

ARPANET LOGICAL MAP, MARCH 1977



- ARPA Networking Project (ARPANET) ab 1969
- anfangs Rechnernetz zwischen 4 Unis
 - Stanford, Utah, Los Angeles, Santa Barbara
- Paketvermittlung (“packet switching”)
- E-Mail, FTP, telnet

Quelle: Wikipedia,

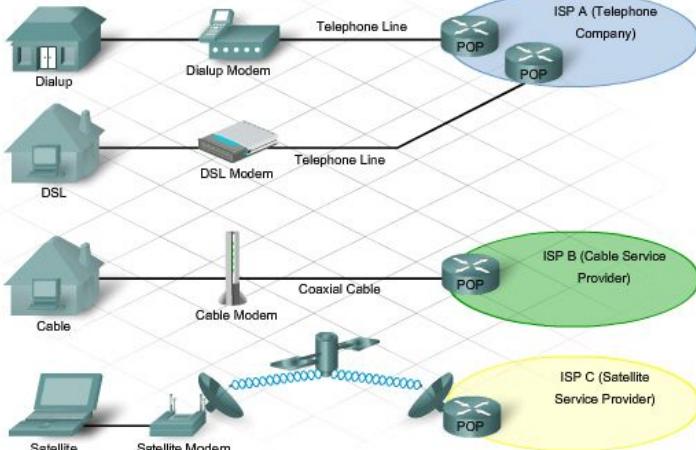
<https://commons.wikimedia.org/w/index.php?curid=9990864>

Internet

Merke: Internet \neq WWW

50 Jahre alt

- 1969 Arpanet \approx Militärisches US-Rechnernetz
- 1983 Internet \approx Vernetzte UNIX-Rechner
- 1989 WWW als Erfindung von Tim Berners-Lee**



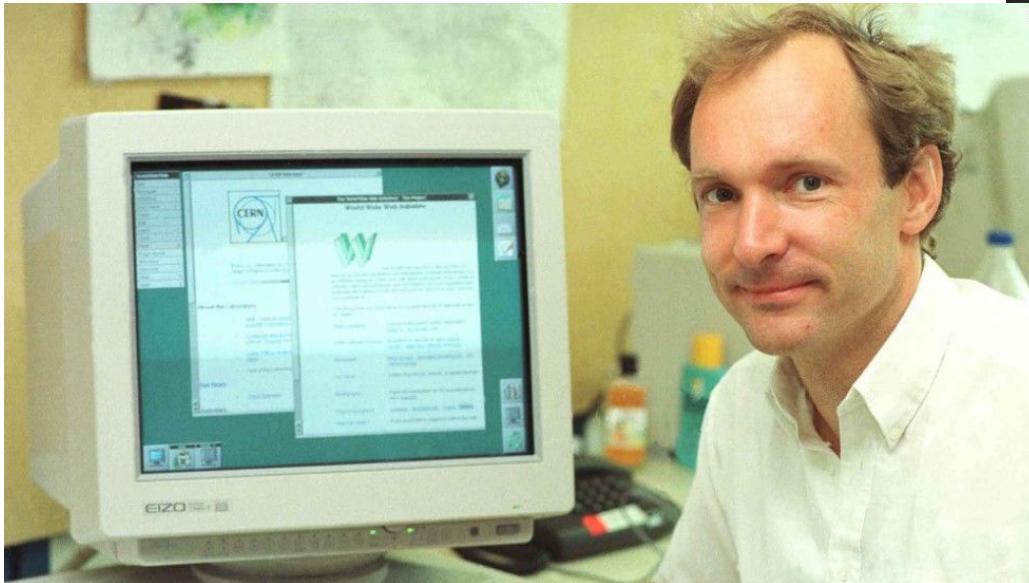
Quelle: <https://de.wikipedia.org/wiki/Internet>

30 Jahre alt



Internet und WWW werden heute in der Alltagssprache häufig als Synonyme verwendet. Das ist sachlich nicht korrekt!

1989 Tim-Berners-Lee @ CERN



1990 HTTP, HTML, URL, WWW, Browser, Server, Suchmaschine
1994 Gründung des [W3C.org](http://www.w3.org)
Tim-Berners-Lee



The first web server 1990: NeXT von Steve Jobs
<http://info.cern.ch/>

Der erste Web-Server

... steht heute im Museum



Quelle: https://de.wikipedia.org/wiki/World_Wide_Web

The World Wide Web project
<http://info.cern.ch/hypertext/WWW/TheProject.html>

World Wide Web

The WorldWideWeb (W3) is a wide-area [hypermedia](#) information retrieval initiative aiming to give universal access to a large universe of documents.

Everything there is online about W3 is linked directly or indirectly to this document, including an [executive summary](#) of the project, [Mailing lists](#) , [Policy](#) , November's [W3 news](#) , [Frequently Asked Questions](#) .

What's out there?
Pointers to the world's online information, [subjects](#) , [W3 servers](#), etc.

Help
on the browser you are using

Software Products
A list of W3 project components and their current state. (e.g. [Line Mode](#) ,[X11](#) [Viola](#) ,[NeXTStep](#) ,[Servers](#) ,[Tools](#) ,[Mail robot](#) ,[Library](#))

Technical
Details of protocols, formats, program internals etc

Bibliography
Paper documentation on W3 and references.

People
A list of some people involved in the project.

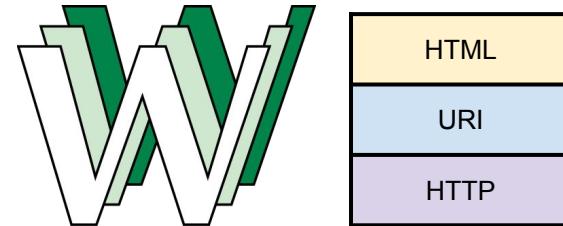
History
A summary of the history of the project.

How can I help ?
If you would like to support the web..

Getting code
Getting the code by [anonymous FTP](#) , etc.

<http://info.cern.ch/hypertext/WWW/TheProject.html>

Die anfängliche Definition von World Wide Web (WWW)



Das World Wide Web (WWW) ist eine Sammlung von Technologie-Standards, die das Lesen verteilter und vernetzter Dokumente und Medien (Hypertexte und Hypermedien) weltweit über das Internet mittels Browser als Lesegerät ermöglicht.

Das WWW basierte am Anfang auf **drei Kernstandards**:

- HTTP
- HTML
- URI und URL

URI	http://www.example.com/people/alice
URL	http://www.example.com/doc/people/alice.html

(Am Anfang also noch kein CSS und kein JavaScript)



WWW Timeline

- **Das Problem:** Physiker-Dokumente innerhalb von CERN: **Fehlende Integration** sogar innerhalb eines Betriebs.
- 1989 Tim-Berners-Lee schreibt das WWW-Konzept und legt es seinem Chef vor.
⇒ **Kein Interesse.** *Kommentar des Chefs: Interessant, aber ein bisschen vage.*
- 1990 Tim-Berners-Lee implementiert als **U-Boot-Projekt** den ersten WWW-Server, das HTTP-Protokoll, den ersten WWW-Browser für HTML-Dokumente auf NeXT
- 1991 Tim-Berners-Lee reicht sein Konzept-Papier bei der weltgrößten Hypertext-Konferenz ein ⇒ **Ablehnung**, jedoch Erlaubnis einer Poster-Session auf den Fluren der Konferenz. Wird von Marc Andreesen, NCSA, entdeckt.
- 1993 National Center for SuperComputing Applications (NCSA) baut freien WWW-Browser "Mosaic"
- 1994 Erste WWW-Konferenz & Tim-Berners-Lee gründet WWW-Consortium ([W3C](#))
- Danach: Das WWW explodiert:
 - 1994 - 3.000 Websites <https://archive.org/details/Munnecke-MarkFrischesMessageToTimBernersLee313>
 - 1995 - 25.000 Websites <https://hapgood.us/2015/04/21/that-time-berners-lee-got-knocked-down-to-a-poster-session/>
 - 1997 - 1,2 Mio Websites <http://info.cern.ch/Proposal.html>
 - Januar 2001 - 30 Mio. Websites <https://www.heise.de/ix/artikel/Gedankenspiel-507001.html>



Tim Berners-Lee hat seine
"World Wide Web"-Idee
gegen den Widerstand
der wiss. Community durchgesetzt.

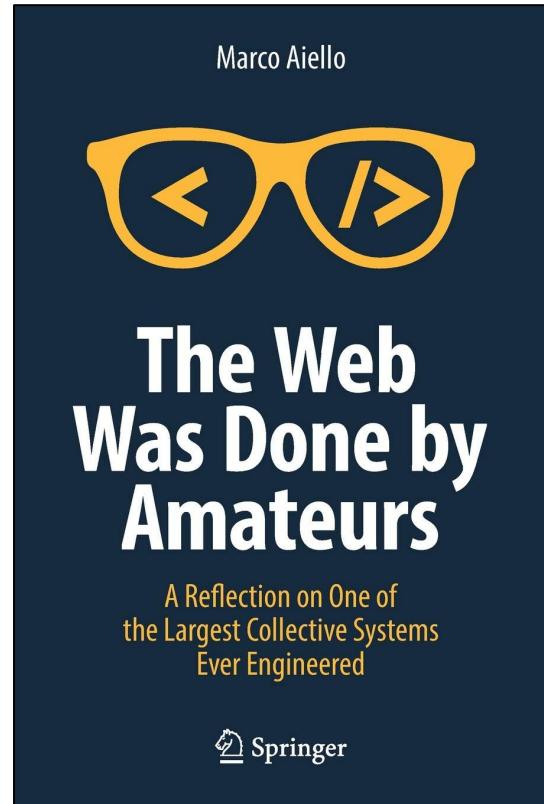
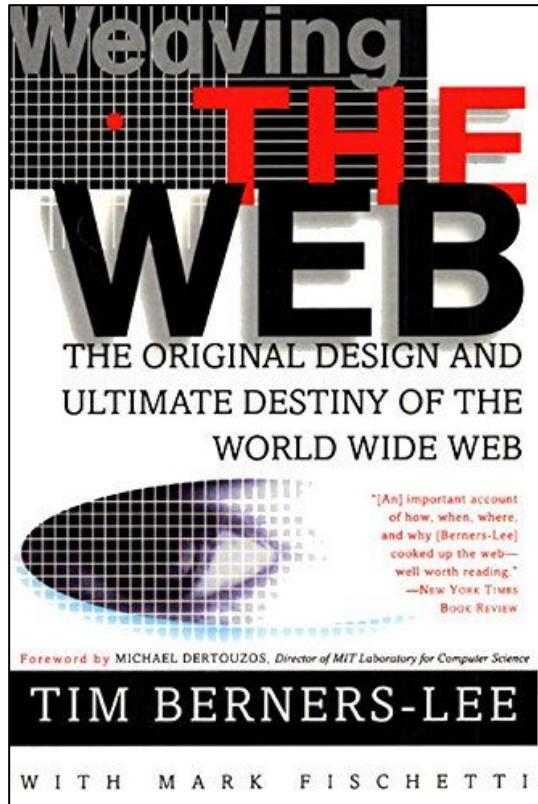
⇒ Was hatte die wiss. Community nicht verstanden?



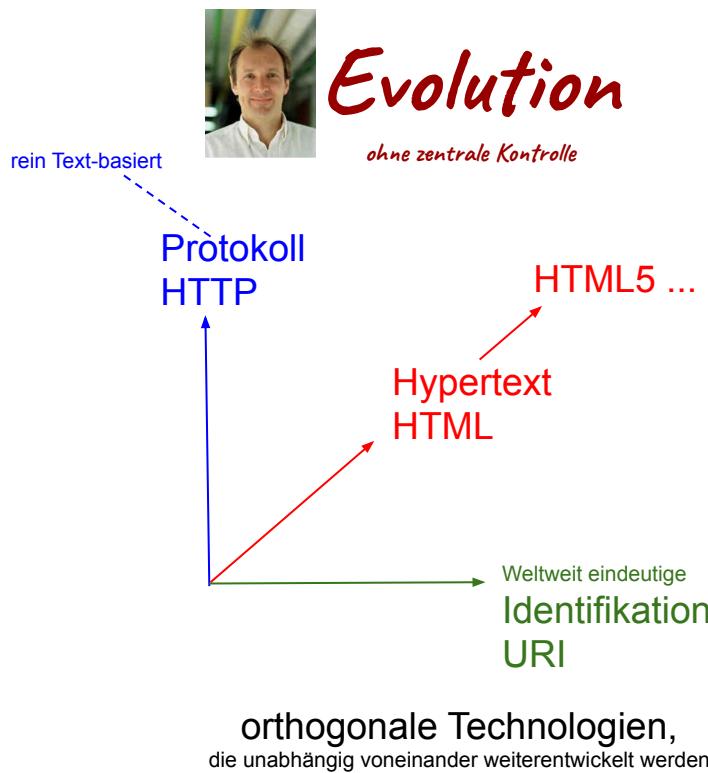
Mit welchen fachlichen Argumenten wurde das
WWW-Proposal von TBL abgelehnt?

Was sind die fachlichen Argumente, warum das
WWW dennoch ein Erfolg wurde?

Unterschiedliche Sichtweisen

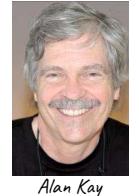


Unterschiedliche Sichtweisen



Computer Science

wissenschaftlicher Ansatz der Profis



- mathematische Rigorosität
- keine “broken links”
- Verteilte Objekte statt verteilter Texte
- Objekt- statt Text-Protokoll

“So etwas wie HTML ist genau das, was wir verhindern wollten.”

“Texte ohne klar definierte Semantik sind die Ursache allen Übels.”

“Die Geschichte des WWW ist eine lange Liste von Patches.”

Der Web-Standard heute

Liste aller 796 Browser APIs

The screenshot shows the MDN Web API Reference page, which is a comprehensive list of 796 browser APIs. The page is organized into sections by category, such as A, C, G, K, L, M, N, P, S, and D. Each section contains a list of APIs with their names and brief descriptions. The APIs are represented by icons and colored boxes indicating their status or type. The page also includes a search bar, navigation links for Technologies, References & Tutorials, and Feedback, and a footer with links to MDN's Terms of Service and Privacy Policy.

<https://developer.mozilla.org/de/docs/Web/API>

... und WWW Zukunft: Web Incubator Com. Group (WICG)

The screenshot shows the WICG website homepage. At the top, there's a banner with developer code snippets and the WICG logo. Below the banner, there are three main navigation links: GitHub, Join, and Discourse. The GitHub link has a tooltip "Join our community of developers and browser makers to define the next generation of web technologies." To the right of the GitHub link is a large, stylized "W" logo composed of colored bars (orange, red, purple, blue).

Getting started

```
npm install -g wicg  
wicg init "The Awesome API"
```

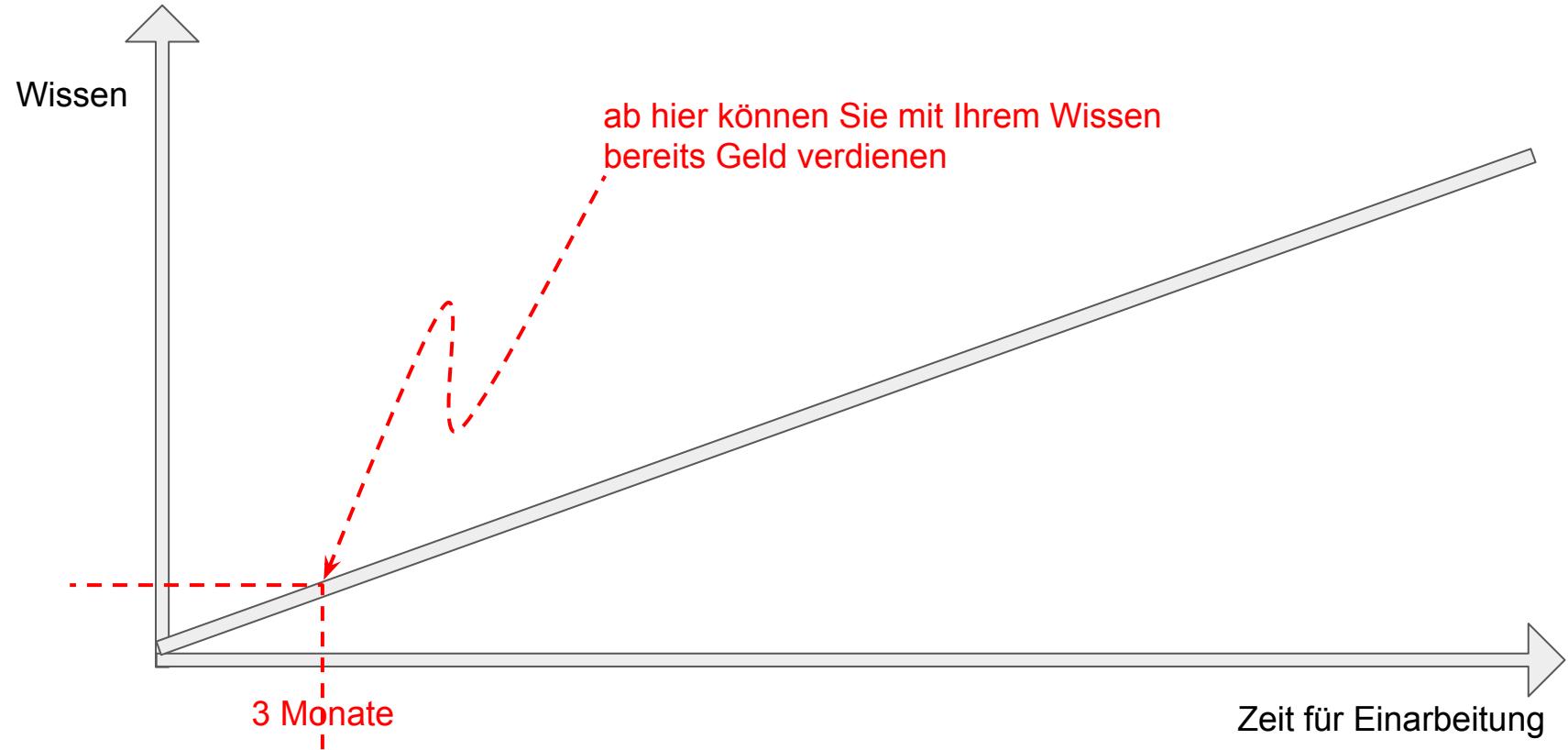
Incubations

Native File System	Focus Visible	Ink Enhancement	Element Timing	JS Self Profiling
Open UI	Sanitizer API	Change Password URL	Proposals	Input for Workers
Portals	Construct Stylesheets	Speech API	Entries API	Indexed DB Observers
UA Client Hints	Local Font Access	Is Input Pending	Interventions	Page Lifecycle
Scroll to Text Fragment	Trust Token API	Web Locks	CORS RFC 1918	Request Post Animation Frame
Display Locking	Serial	Video Editing	Keyboard Map	Intervention Reporting
Web Package	Web HID	Performance Measure Memory	WICG GitHub	CQ Use Cases
Web Monetization	Inert	Main Thread Scheduling	Window Controls Overlay	Periodic Background Sync
Cookie Store	Origin Isolation	Storage Buckets	Contact API	Lang Client Hint
Web Transport	Idle Detection	Shape Detection API	Visual Viewport	Container Queries
Web Codecs	AOM	Data Cues	Priority Hints	Raw Clipboard Access
Conversion Measurement API	Sparrow	Digital Goods	Display Override	Media Latency Hint
Media Feeds	Client Hints Infrastructure	SMS One Time Codes	Overscroll Scroll Events	Save Data
Turtle Dove	Web USB	Frame Timing	Deprecation Reporting	Permissions Request
Video RFC	PWA URL Handler	Cross-Origin Embedder Policy	KV Storage	Audio Focus
Layout Instability	Content Index	Virtual Scroller	Encrypted Media Encryption Scheme	Encrypted Media
Import Maps	Spatial Navigation	Canvas Color Space	Input Device Capabilities	DevTools Protocol
Event Timing	Largest Contentful Paint	Custom State Pseudo Class	HDCP Detection	Controls List
Origin Policy	Keyboard Lock	File Handling	SW Launch	Media Source
Web OTP	Background Fetch	Crash Reporting	Before Install Prompt	Discourse WICG Theme
WebID	Compression	Get Installed Related Apps	CSS Parser API	Image Output
Raw Sockets	Background Sync	NetInfo	Accessible Loading and Searching of Content	Permissions Revoke

Useful Resources

Charter Participants Join Mailing list

Lernkurve des WWW



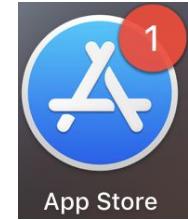
Marktanteile 2015



Safari



vs.



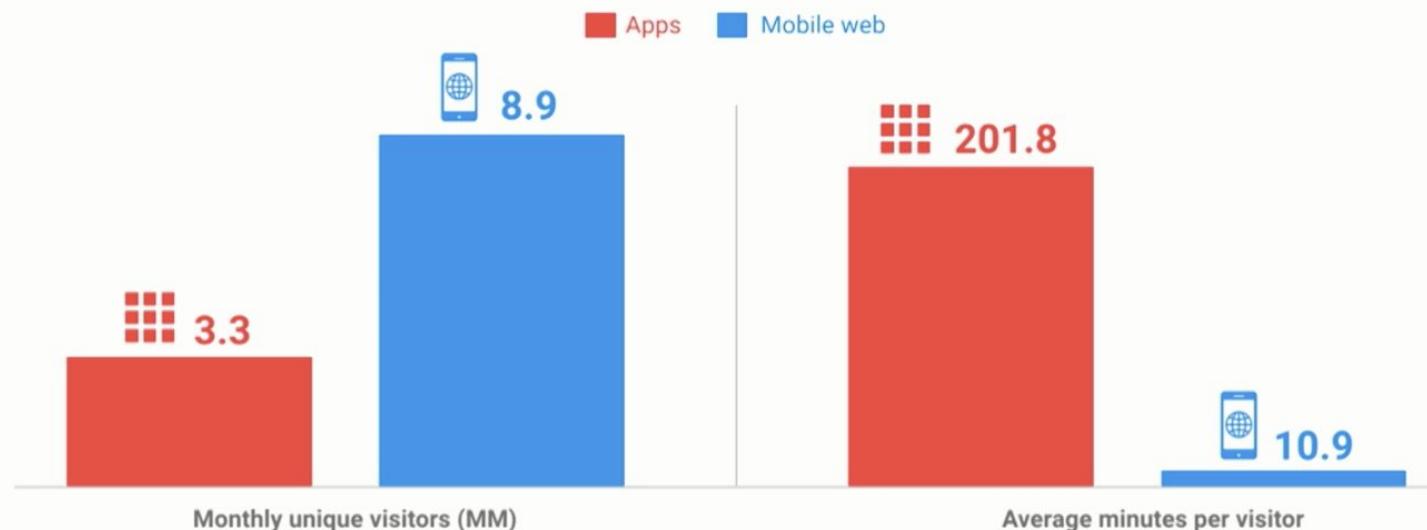
App Store

Source: comScore Mobile Metrix, U.S., Age 18+, June 2015

Mobile Web: Große Reichweite, wenig Nachhaltigkeit

Today's Mobile Web: Broad Reach, Low Engagement

Top 1000 mobile apps vs. top 1000 mobile web properties



80%

**OF TIME SPENT IS IN
USERS' TOP 3 APPS**

Source: comScore Mobile Metrix, U.S., Age 18+, June 2015

Apple Is Trying to Kill Web Technology

The company has made it extremely difficult to use web-based technology on its platforms, and it hopes developers won't bother



Owen Williams

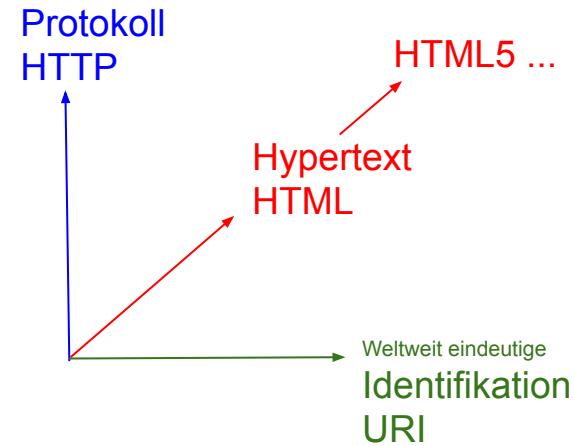
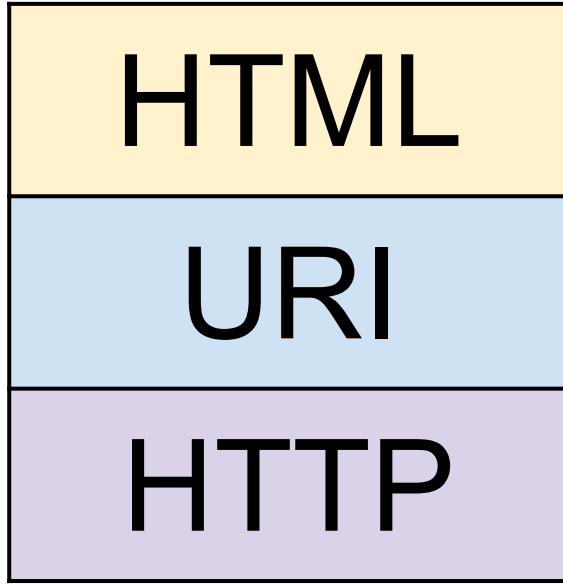
Follow

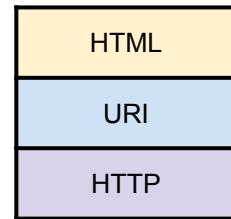
Nov 7 · 4 min read



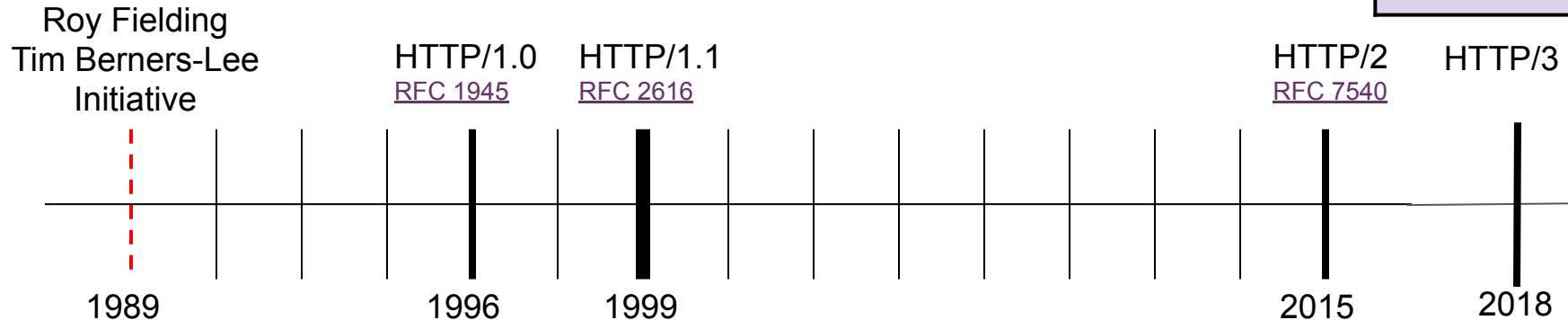
<https://onezero.medium.com/apple-is-trying-to-kill-web-technology-a274237c174d>

Credit: SOPA Images/Getty Images





HTTP Timeline



HTTP/1.0 - eine TCP-Verbindung pro HTTP-Request

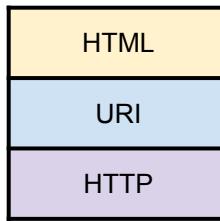
HTTP/1.1 - Keep-Alive im Header ⇒ mehrere Requests pro Verbindung

HTTP/2 - Multiplexing & Datenkompression & Server Push

HTTP/3 - User Space Congestion Control over UDP

WWW ist eine Client-Server-Architektur

HTTP ist ein zustandsloses Protokoll.

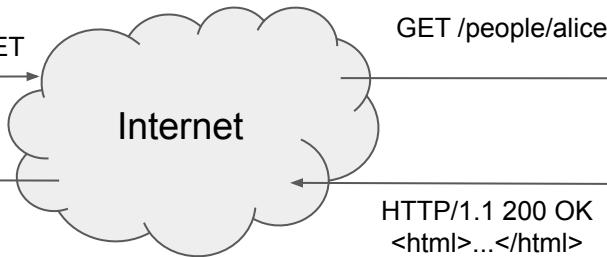


Browser als HTTP-Client

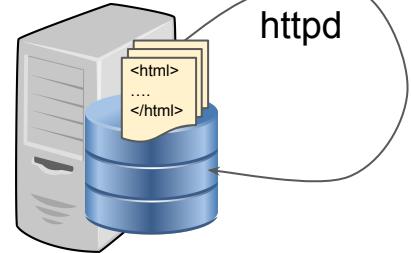


HTTP-Request: GET

HTTP-Response:
HTML



httpd als HTTP-Server



OSI Layer	TCP/IP Protocol Stack
Application	HTTP
Transport	TCP
Network	IP
Physical	Ethernet, WLAN

Klartext

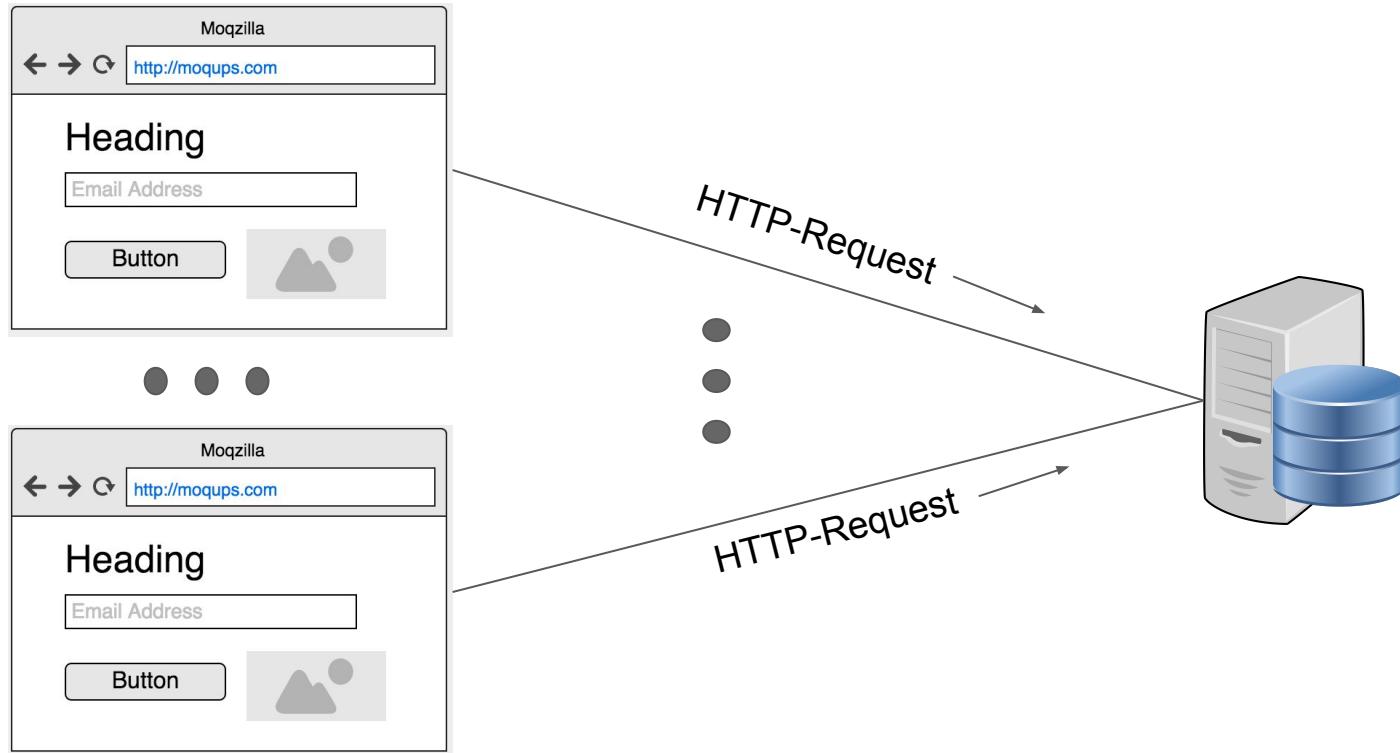
GET /people/alice HTTP/1.1
Host: www.example.com
Accept: text/html, application/xhtml+xml
Accept-Language: en, de

The server could answer:

HTTP/1.1 200 OK
Content-Type: text/html
Content-Language: en
Content-Location: http://www.example.com/people.en.html

HTTP-Request
HTTP-Response
HTTP-Response

Viele Clients können den gleichen Server nutzen



Eigenschaften des HTTP-Protokolls



1. zustandslos
2. verbindungslos: Nach jedem Request-/Response-Zyklus ist die Verbindung beendet.

(stimmt bei HTTP/1.0. Seit HTTP/1.1 kann die gleiche TCP-Verbindung wiederverwendet werden.)

3. medientyp-unabhängig, d.h. für alle Medien geeignet, für Texte, Bilder, Video, VR, AR, ...

```
$ telnet www.perdu.com 80
Trying 208.97.177.124...
Connected to www.perdu.com.
Escape character is '^['.

GET / http/1.1
Host: www.perdu.com

HTTP/1.1 200 OK
Date: Sat, 17 Aug 2013 12:14:56 GMT
Server: Apache
Accept-Ranges: bytes
X-Mod-Pagespeed: 1.1.23.1-2169
Vary: Accept-Encoding
Cache-Control: max-age=0, no-cache
Content-Length: 204
Content-Type: text/html

<html><head><title>Vous Etes Perdu ?</title></head><body><h1>Perdu sur l'Internet ?</h1><h2>Pas de panique, on va vous aider</h2><strong><pre> * <----- vous &ecirc;tes ici</pre></strong></body></html>
Connection closed by foreign host.
$
```

Verbindsaufbau zum Server

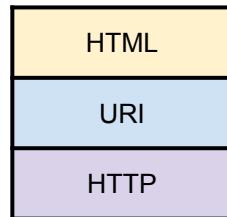
HTTP-Anfrage

Serverantwort: Header

Serverantwort: Nachrichtenrumpf

Verbindungsende

Quelle: https://de.wikipedia.org/wiki/Hypertext_Transfer_Protocol



Wichtigste HTTP-Methoden

: mehr als **CRUD**
Server darf beliebige Teilmenge implementieren, meistens GET und POST.

Methode	Funktion	Safe	Idem-potent	Cacheable
GET	READ: WWW-Ressource anfordern	✓	✓	✓
POST	CREATE: Ressourcen erzeugen (neue ID)	✗	✗	✗
HEAD	nur Header anfordern	✓	✓	✓
OPTIONS	Server-Optionen abfragen	✓	✓	✗
PUT	UPDATE: Ressource ändern oder ergänzen	✗	✓	✗
DELETE	DELETE: Ressource löschen	✗	✓	✗

Safe = "*Retrieval Only*", keine Seiteneffekte, keine Änderungen an den WWW-Ressourcen

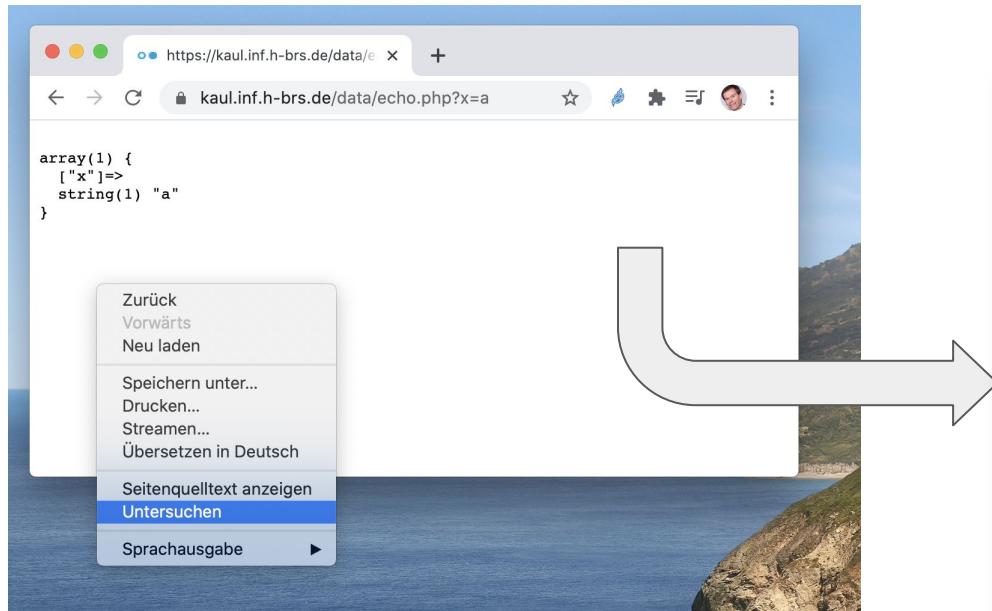
Idempotent = jedesmal das gleiche Ergebnis, daher gut für Caching geeignet

HTTP-Methoden

https://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol

HTTP method	RFC	Request has Body	Response has Body	Safe	Idempotent	Cacheable
GET	RFC 7231 ↗	Optional	Yes	Yes	Yes	Yes
HEAD	RFC 7231 ↗	Optional	No	Yes	Yes	Yes
POST	RFC 7231 ↗	Yes	Yes	No	No	Yes
PUT	RFC 7231 ↗	Yes	Yes	No	Yes	No
DELETE	RFC 7231 ↗	Optional	Yes	No	Yes	No
CONNECT	RFC 7231 ↗	Optional	Yes	No	No	No
OPTIONS	RFC 7231 ↗	Optional	Yes	Yes	Yes	No
TRACE	RFC 7231 ↗	No	Yes	Yes	Yes	No
PATCH	RFC 5789 ↗	Yes	Yes	No	No	No

HTTP-Kommunikation in "Network" in DevTools



Ein Rechts-Klick
öffnet das **Kontextmenü**

A screenshot of the Chrome DevTools Network tab. The URL is `https://kaul.inf.h-brs.de/data/echo.php?x=a`. The Network table shows two requests:

Name	Initiator	Size	T...	Waterfall
echo.php?...	Other	141 B	2...	
favicon.ico	Other	942 B	1...	

Below the table, the status bar indicates: 2 requests, 1.1 kB transferred, 984 B resources, Finish: 61 ms, DOMContentLoaded: 36 ms, Load: 0 ms.

HTTP Headers in "Network" in DevTools

array(1) { ["x"] => string(1) "a" }

Request URL: https://kaul.inf.h-brs.de/data/echo.php?x=a

Status Code: 200

Content-Type: text/html; charset=UTF-8

Date: Sun, 04 Oct 2020 18:53:58 GMT

Server: Apache/2.4.29 (Ubuntu)

X-Powered-By: PHP/7.2.24-0ubuntu0.18.04.6

Authorization: kaul.inf.h-brs.de

Method: GET

Path: /data/echo.php?x=a

Scheme: https

Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange;v=b3;q=0.9

Accept-Encoding: gzip, deflate, br

Accept-Language: de-DE,de;q=0.9,en-US;q=0.8,en;q=0.7

Cache-Control: max-age=0

Cookie: SSESS3be9484fe9510e2d7b0f181666f65=kT0GzDZh2kexC31_A0myH8-n05FcCuIy3isX20gA1A

DNT: 1

See-On-User: "(Not:A)(&and:v=v99", "Google Chrome";v=85", "Chromium";v=85"

See-On-User-Mobile: 70

See-Fetch-Dest: document

See-Fetch-Mode: navigate

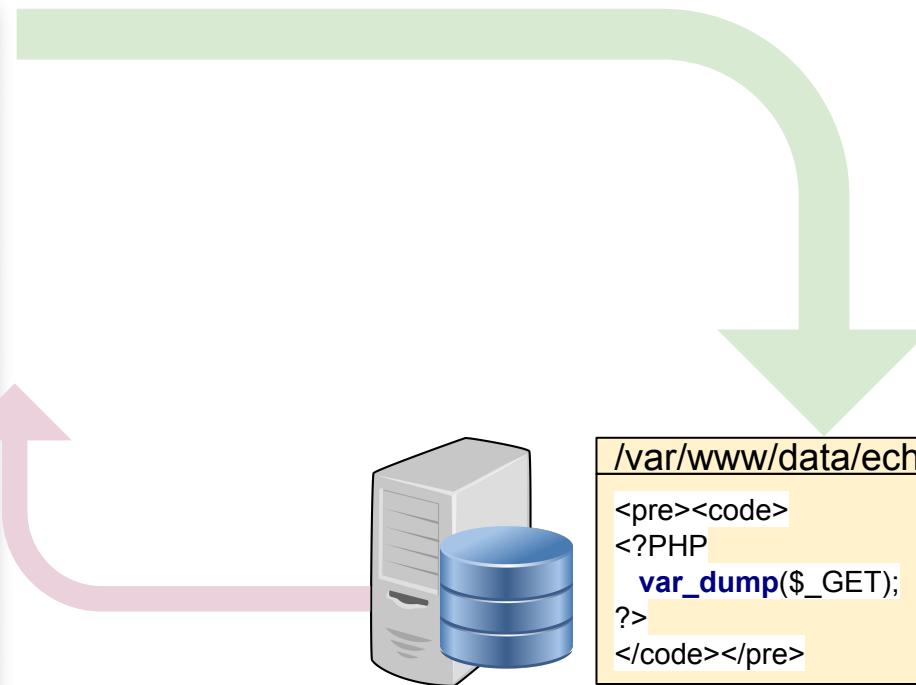
See-Fetch-Site: none

See-Fetch-User: 71

Upgrade-Insecure-Requests: 1

User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_15_6) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/85.0.4183.121 Safari/537.36

X: a



HTTP-Request POST mit Parametern

The screenshot shows a browser window with the URL <https://kaul.inf.h-brs.de/data/formular.html>. The page contains a form with two text inputs: `x1=payload1` and `x2=payload2`, and a submit button labeled `Eingaben absenden`.

```
<!doctype html>
<form method="POST" action="post.php">
  x1=<input type="text" name="x1">
  x2=<input type="text" name="x2">
  <button type="submit">Eingaben absenden</button>
</form>
```

HTTP/1.1 200 OK
Date: Thu, 12 Apr 2018 07:33:13 GMT
Server: Apache/2.4.7 (Ubuntu)
X-Powered-By: PHP/5.5.9-1ubuntu4.24
Content-Length: 93
Content-Type: text/html
<pre><code>
array(2) {
 ["x1"]=>
 string(8) "payload1"
 ["x2"]=>
 string(8) "payload2"
}</code></pre>

POST /data/post.php HTTP/1.1
Host: kaul.inf.h-brs.de
Content-Length: 23
Content-Type: application/x-www-form-urlencoded
User-Agent: Mozilla/5.0 (Macintosh; Intel Mac OS X 10_13_4)
AppleWebKit/537.36 (KHTML, like Gecko) Chrome/65.0.3325.181
Safari/537.36
Accept:
text/html,application/xhtml+xml,application/xml;q=0.9,image/web
p,image/apng,*/*;q=0.8
Referer: <https://kaul.inf.h-brs.de/data/formular.html>

▼ Form Data

x1=payload1&x2=payload2

view parsed

codiert im
HTTP Body



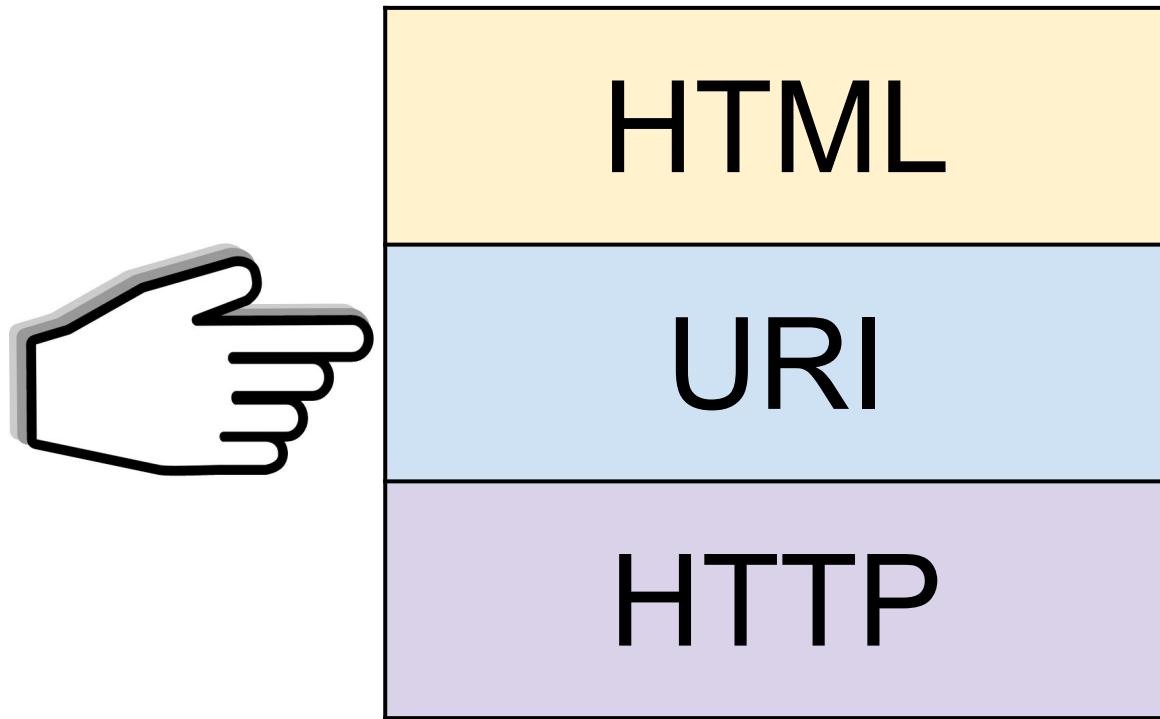
/var/www/data/post.php
<pre><code>
<?PHP
 var_dump(\$_POST);
?>
</code></pre>

HTTP Response: HTTP Status Codes

HTTP-Prinzip: Fehler melden!

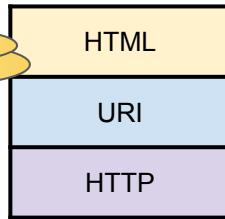
- nicht verschweigen
- kein Error Recovery, kein Error Correction

S.N.	Code and Description
1	1xx: Informational 100 Request received and continuing process.
2	2xx: Success 200 OK, 204 OK, aber ohne Inhalt.
3	3xx: Redirection 302 Moved Temporarily ⇒ Erneuter GET auf einer anderen Adresse
4	4xx: Client Error 400 Fehlerhafte Anfrage, 403 Fehlende Zugriffsberechtigung, 404 Nicht gefunden.
5	5xx: Server Error 500 Internal Server Error ⇒ Programmierfehler auf dem Server





Cool URIs don't change



URI (Uniform Resource Identifier)

Ein URI ist ein Identifikator und besteht aus einer Zeichenfolge, die zur Identifizierung einer abstrakten oder physischen Ressource dient. URIs werden zur Bezeichnung von Ressourcen (wie Webseiten, sonstigen Dateien, Aufruf von Webservices, aber auch z. B. E-Mail-Empfängern) im Internet und dort vor allem im WWW eingesetzt. Der aktuelle Stand 2016 ist als RFC 3986 publiziert, in BNF:

URI ::= scheme ":" authority ["/" path] ["?" query] ["#" fragment]

http://kaul.inf.h-brs.de:80/wordpress/review4711?q=www#subpart



Schema
Protokoll

Autorität
Server

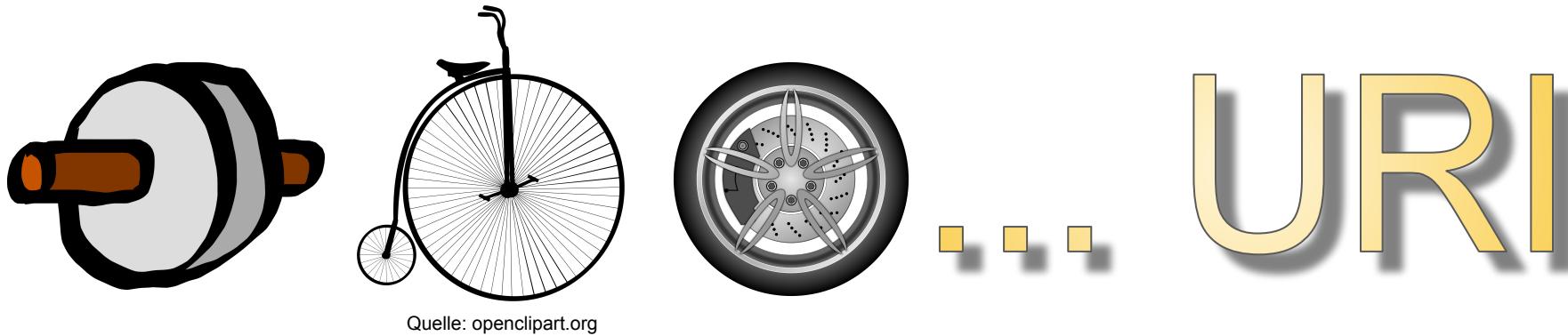
Pfad

Query Fragment

https://de.wikipedia.org/wiki/Uniform_Resource_Identifier

Ressourcen-Identifikation:

Das Rad der Namensschemata wurde leider immer wieder neu erfunden ...

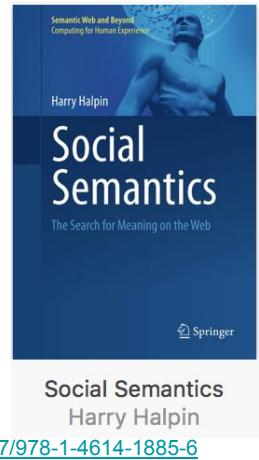


besonders in der Informatik:

Die Aufgabe, ein einheitliches Namensschema für alle Objekte der Betrachtung zu entwickeln, stellt sich immer wieder ...

URI

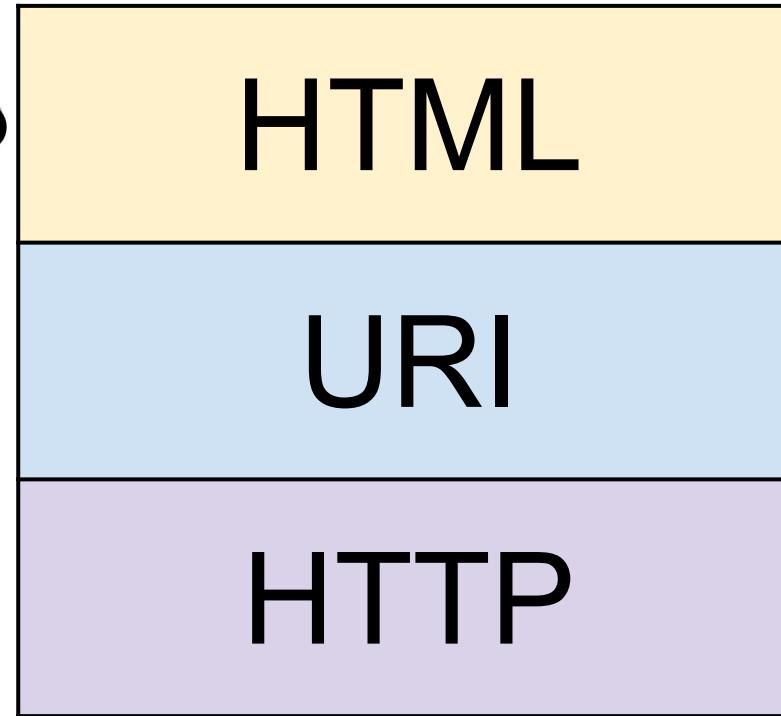
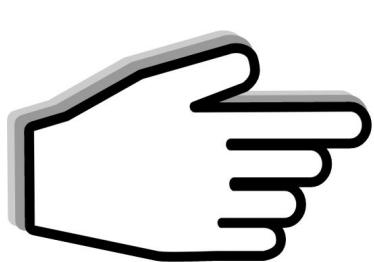
- Erfolgreicher ISO-Standard zur Identifizierung von Ressourcen aller Art in einem weltweiten "universellen Informationsraum" (*universal information space*)
- Universelles Format
 - Protokoll-übergreifend
 - Server-übergreifend, Port-übergreifend
 - Verzeichnis-übergreifend, d.h. für beliebige Verzeichnistiefe
 - Format-übergreifend, d.h. für alle Datei-Formate
 - auch geeignet für Auszeichnung von Fragmenten innerhalb einer Ressource
 - Hersteller-übergreifend (*vendor neutral*)
- Dezentrale Namensvergabe
 - ohne zentrale Kontrolle
 - ohne zentrale Datenbank
 - ⇒ keine Gewähr von **referentieller Integrität**
 - d.h. Gefahr von "*broken links*"
- Grundlage für die Integration von Informationssystemen zu einem globalen Netz, dem World Wide Web



Unterschied URI vs. URL

- URI = Uniform Resource Identifier
 - Identifikation, nicht notwendig auch Ort
 - z.B.

```
<html xmlns="http://www.w3.org/1999/xhtml">
    ...
</html>
```
- URL = Uniform Resource Locator
 - Ort der Ressource
 - dient dem Abruf der Ressource
 - Die Menge aller URLs sind eine Teilmenge der Menge der URIs



HTML

Falls die Doctype-Angabe fehlt, so wird der [Quirks-Modus](#) verwendet.

<https://quirks.spec.whatwg.org/#history>

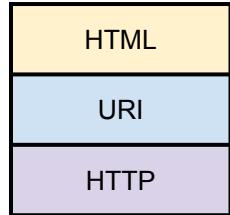
```
<!DOCTYPE html>
<html>
<head>
  <title>Titel wird in Kopfzeile angezeigt</title>
  <meta charset="utf-8">
</head>
<body>
  <h1>Überschrift</h1>
  
  <p>Inhalt eines Paragraphen mit
    <a href="seite2.html">Hyperlink</a>
  </p>
  <ul>
    <li>erster Listeneintrag</li>
    <li>zweiter Listeneintrag</li>
  </ul>
</body>
</html>
```

<https://www.w3.org/TR/html52/>
<https://html.spec.whatwg.org>

HTML Tags

- <h1> ... <h6> header level 1 ... 6
- hyperlink anchor
- image
- <p> paragraph
-
 line break
- <hr> horizontal rule
- <pre> preformatted
- unordered list
- ordered list
- list item
- <table> table
- <th> table header
- <tr> table row
- <td> table cell
-
-
- <div>
-
- ...

<https://wiki.selfhtml.org/wiki/HTML/Tutorials/HTML5-Grundger%C3%BCst>



Grundsätzliche Unterscheidung:

Block versus Inline Element



Gruppierung von Content mit
Absatz (*block element*)

z.B. <h1>, <p>, <div>

Block-Elemente werden
untereinander angeordnet und
belegen die gesamte Breite.

Gruppierung von Content ohne
Absatz (*inline element*)

z.B. <a>, , ,

Inline-Elemente werden
nebeneinander angeordnet im
Fluss des Textes mit
Zeilenumbruch wie jeweils
erforderlich.

HTML-Elemente

Eine Auswahl

Kommentare

<!-- Kommentar -->

Abschnitte (*sections*)

- <h1> Größte Überschrift </h1>
- <h2> bis <h6> Unterüberschriften
- neu in HTML5: Layout-Elemente:
<header>, <footer>, <nav>, <address>,
<article>, <section>, <aside>, <hgroup>, ...

Gruppierung von Content

- <p> Paragraph
- Ordered List
- Unordered List
- List Item

Block-Elemente

Block-Elemente

Gruppierung von Content (Fortsetzung)

- <pre> a block of preformatted text
- <dl> description list
- <dt> description term (inside dl)
- <dd> description, definition, or value (inside dl)
- <figure> mit <figcaption>
- <div> Gruppierung von Content mit Absatz (*block content*), neutrales Tag, wenn kein anderes Tag passen würde, ohne Semantik.

Inline-Elemente

Text-level semantics

- <a> Anker-Element, meist für Hyperlinks
- emphasis
- Betonung, "*strong importance*"
- <mark> markiert (*marked or highlighted*)
- <small> small print
- <sub> Tieferstellen (*subscript*)
- <sup> Höherstellen (*superscript*)
- <i> Kursiv (*italics*)
- Fett (*bold*)
- <u> Unterstrichen (*underline*)
-
 Zeilenumbruch (*break*)
- Gruppierung von Content ohne Absatz (*inline content*), neutrales Tag, wenn kein anderes Tag passen würde, ohne Semantik.

<https://www.w3.org/TR/html52/>

<https://html.spec.whatwg.org>

HTML ist eine Auszeichnungssprache (*Markup-Sprache*)

1986: **SGML** = *Standard Generalized Markup Language* ist eine Metasprache, mit der Auszeichnungssprachen (engl. markup languages) für Dokumente definiert werden können, ISO-Standard: ISO 8879:1986. In SGML beschreibt eine Dokumenttypdefinition (DTD), was ein valides Dokument ist und was nicht.

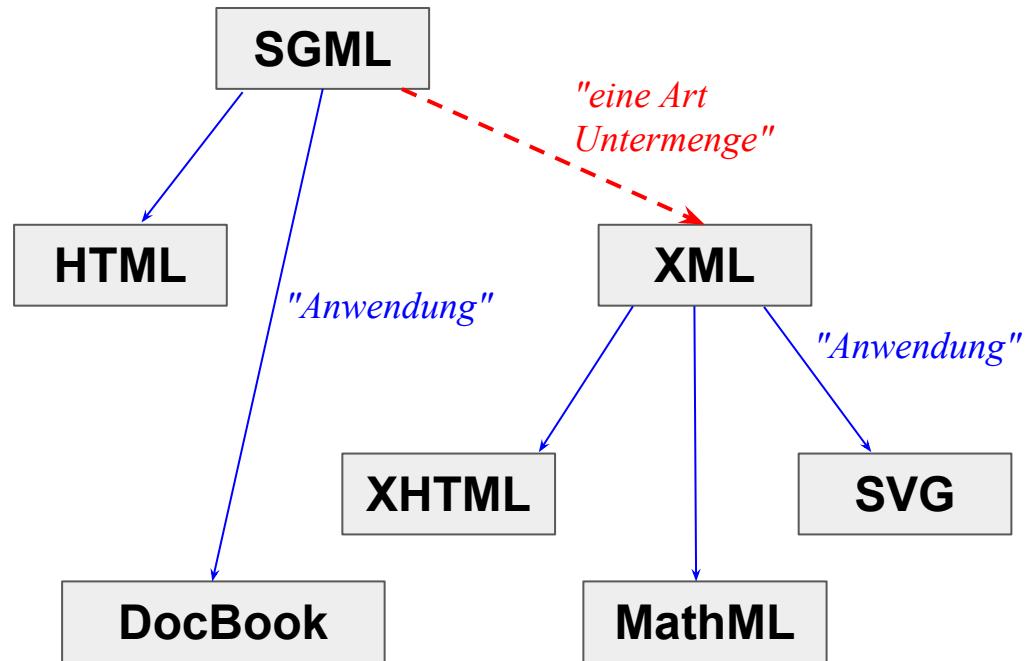
1998: **XML** = *Extensible Markup Language* ist eine Metasprache, auf deren Basis durch strukturelle und inhaltliche Einschränkungen anwendungsspezifische Sprachen (Domain Specific Languages, DSL) definiert werden.

1999: **XMLSchemata** = Alternative zur DTD, die selbst in XML geschrieben wird.

2001: **SVG** = Scalable Vector Graphics

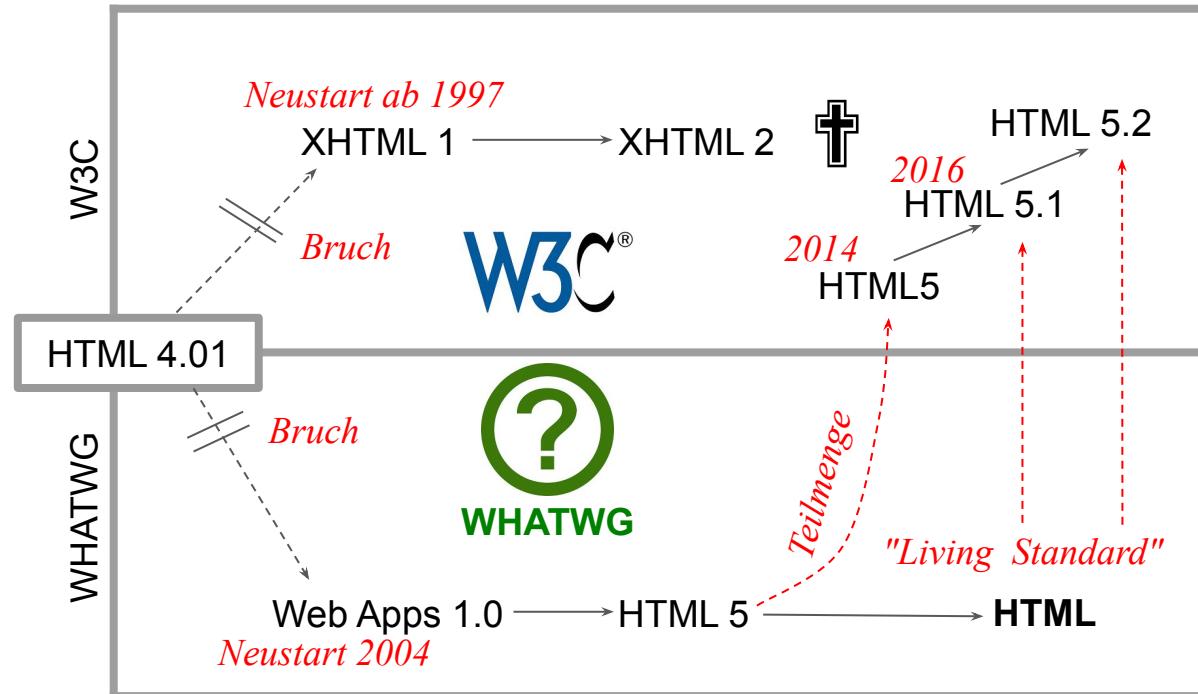
wohlgeformt: korrekt als XML-Dokument

valide: konform zur Grammatik (in DTD oder XMLSchemata)



HTML Timeline: "Standards War"

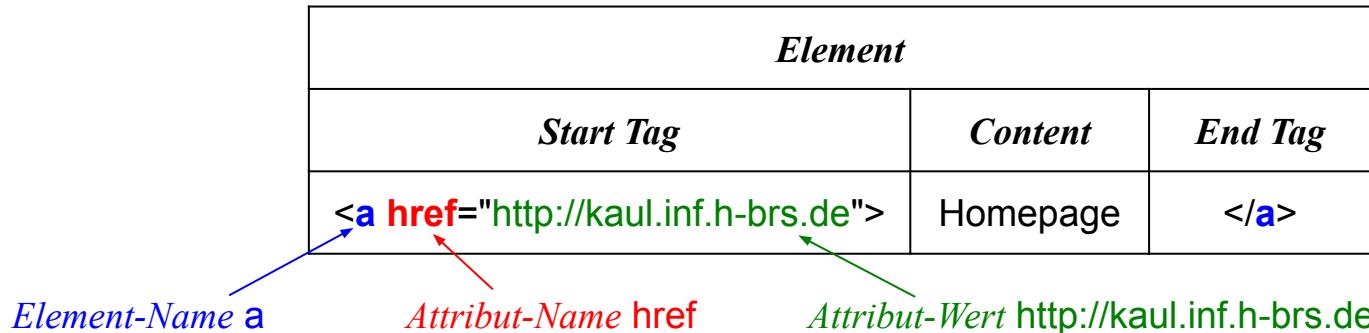
- 1989 TBL erfindet HTML
- 1990er Jahre: Browser implementieren HTML direkt und **ignorieren SGML**
- ab 1997: XHTML
- 2009: W3C erfindet XHTML 2
- ab 2004: Browserhersteller **ignorieren XML** und XHTML und gründen eigene Arbeitsgruppe, die WHATWG und entwerfen HTML 5 als "Living Standard HTML" (auch ohne Versionsnummer)
- 2012: W3C übernimmt HTML-Anteile aus HTML 5



<https://en.wikipedia.org/wiki/HTML>

Hauptunterschiede zwischen HTML und XML

HTML	XML
Schließende Klammern können als optional definiert sein: <code> <p>Lorem...<p>Ipsum...</code>	Schließende Klammern sind immer verpflichtend: <code> <p>Lorem...</p><p>Ipsum...</p></code>
Bei Element- und Attribut-Namen wird zwischen Groß- und Kleinschreibung NICHT unterschieden (<i>not case-sensitive</i>).	Bei Element- und Attribut-Namen wird zwischen Groß- und Kleinschreibung unterschieden (<i>case-sensitive</i>).
Attributwerte müssen nur dann mit Apostrophen begrenzt werden, wenn sie auch andere als alphanumerische Zeichen enthalten (z.B. Leerzeichen). " oder ' sind erlaubt.	Attributwerte müssen immer mit Apostrophen begrenzt werden. Es ist durch das doppelte Anführungszeichen " erlaubt.



⇒ Hauptunterschiede zwischen HTML und XHTML

HTML	XHTML
HTML kann als Anwendung von SGML verstanden werden. HTML wird durch eine DTD definiert.	XHTML ist eine Anwendung von XML, Dokumente können sowohl mit DTD als auch mit XMLSchema definiert werden
Schließende Klammern können als optional definiert sein: <pre> <p>Lorem...<p>Ipsum...</pre>	Schließende Klammern sind immer verpflichtend: <pre> <p>Lorem...</p><p>Ipsum...</p></pre>
Bei Element- und Attribut-Namen wird zwischen Groß- und Kleinschreibung NICHT unterschieden (<i>not case-sensitive</i>).	Bei Element- und Attribut-Namen wird zwischen Groß- und Kleinschreibung unterschieden (<i>case-sensitive</i>).
Attributwerte müssen nur dann mit Apostrophen begrenzt werden, wenn sie auch andere als alphanumerische Zeichen enthalten (z.B. Leerzeichen).	Attributwerte müssen immer mit Apostrophen begrenzt werden.

XHTML 1.0 ist nicht tot

Heute noch verbliebene Anwendungen von XHTML

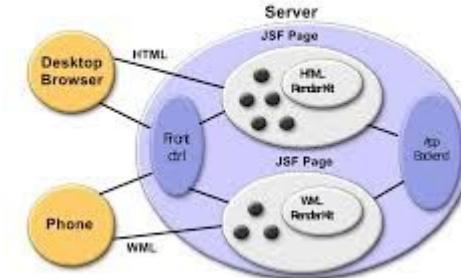
```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0
Transitional//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd">
<html xmlns="http://www.w3.org/1999/xhtml"
      xmlns:h="http://xmlns.jcp.org/jsf/html">
<head>
    <title>Hello JSF World</title>
</head>
<body>
    <h3><h:outputText value="Hello JSF World" /></h3>
    <p> <h:outputText value="Host: #{header['Host']}' /><br/>
        <h:outputText value="User-Agent: #{header['User-Agent']}' /></p>
</body>
</html>
```

<http://www.torsten-horn.de/techdocs/jsf.htm>

Manche Technologien passen
sehr gut zu dem XML-Ansatz
von HTML: XHTML.



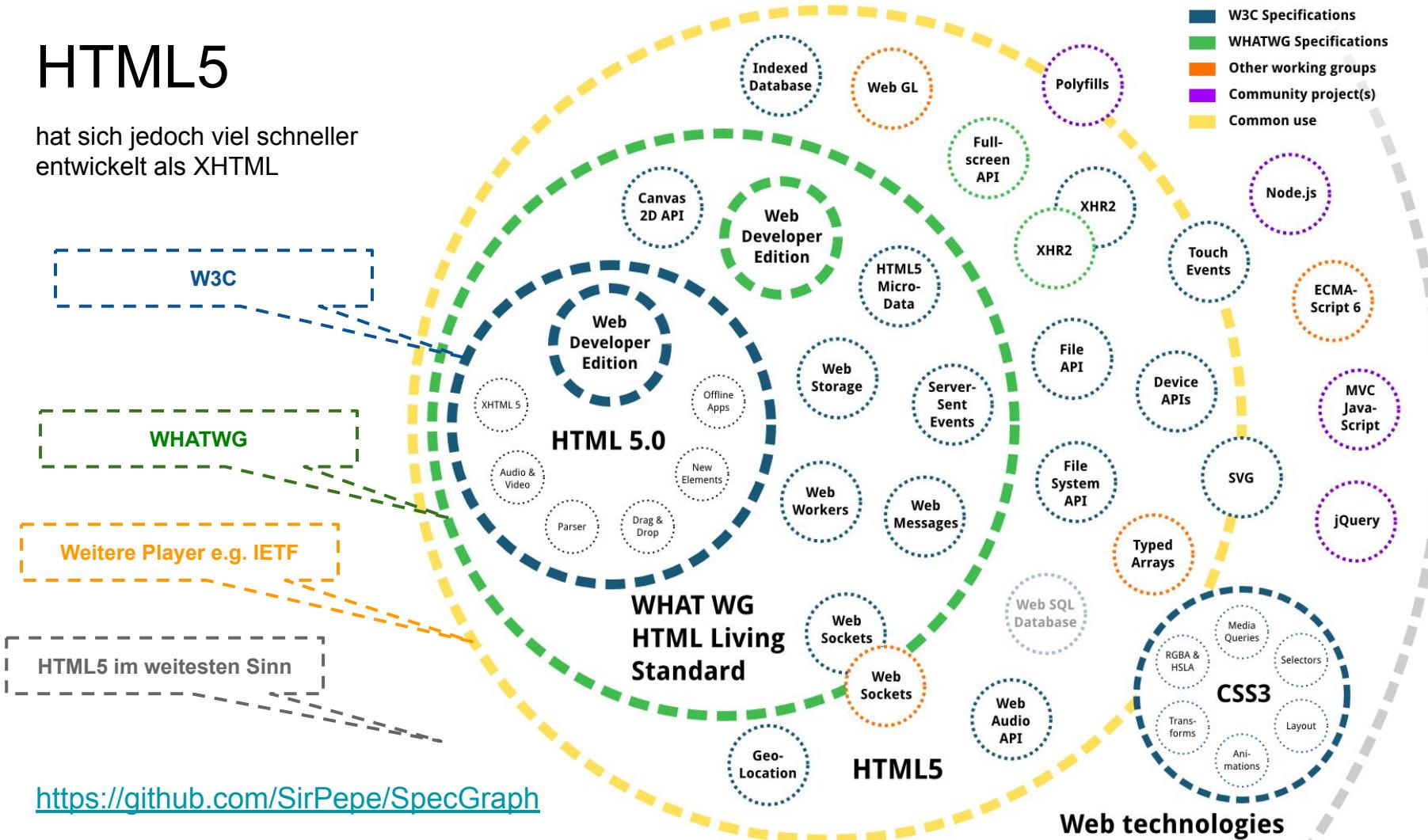
JSF Architecture



<http://www.oracle.com/technetwork/java/javaee/javaserverfaces-139869.html>

HTML5

hat sich jedoch viel schneller entwickelt als XHTML

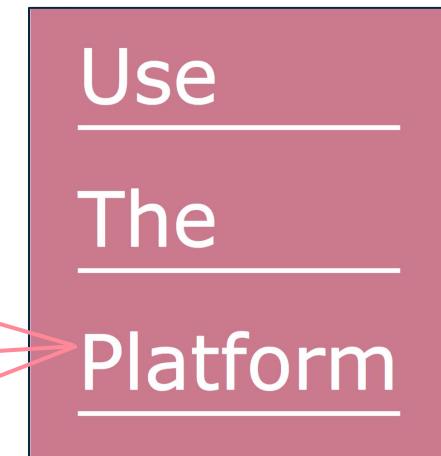


Das WWW ist die erfolgreichste Integrationsplattform

- für vernetzte Informationen
- für vernetzte Informationsarchitekturen
- für Technologien (HTML, CSS, JavaScript, WASM, XML, JSON, SVG, WebGL, ...)

Gemeint ist das WWW.

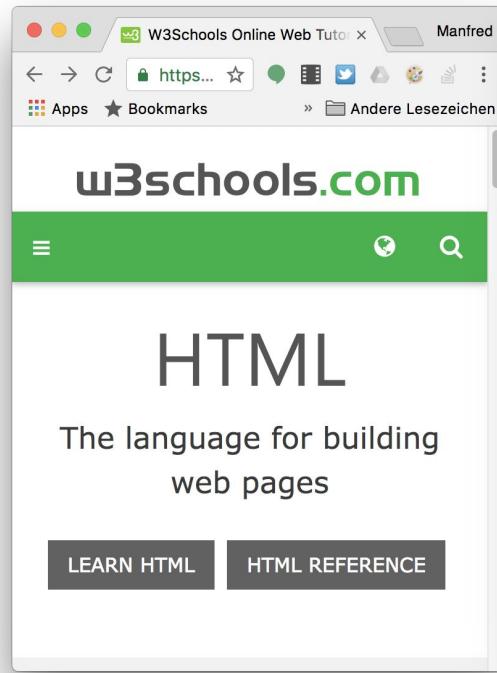
Gemeint ist der Browser-Standard.



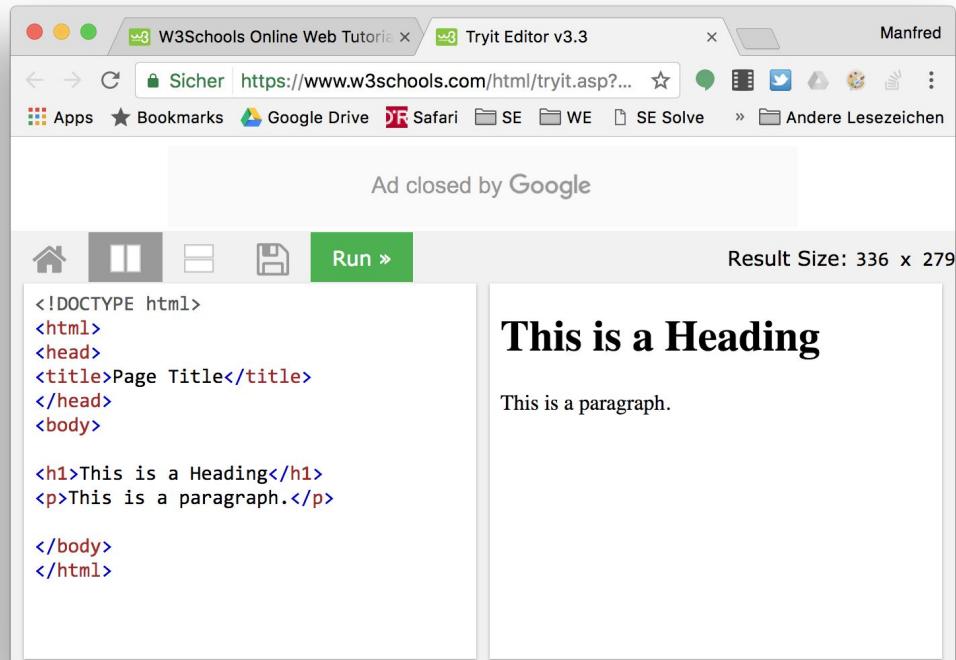
Wie lernt man HTML?



<https://wiki.selfhtml.org/wiki/HTML>



A screenshot of a web browser showing the selfhtml wiki page for HTML. The title 'HTML' is prominently displayed in large letters. Below it, the subtitle 'The language for building web pages' is shown. At the bottom, there are two buttons: 'LEARN HTML' and 'HTML REFERENCE'. The URL in the address bar is <https://www.w3schools.com>.



A screenshot of a web browser showing the W3Schools Tryit Editor. The editor interface includes a toolbar with icons for home, new file, save, and run, and a status bar indicating 'Result Size: 336 x 279'. On the left, the HTML code is displayed:

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>
<h1>This is a Heading</h1>
<p>This is a paragraph.</p>
</body>
</html>
```

On the right, the rendered result is shown: 'This is a Heading' and 'This is a paragraph.'

<https://www.w3schools.com>

Erste Woche: HTML lernen

The screenshot shows the homepage of w3schools.com. On the left, there's a sidebar with categories like 'HTML and CSS', 'JavaScript', 'Server Side', 'Programming', 'Web Building', 'XML Tutorials', and 'Learn XML'. The main content area has three large boxes: 'HTML' (The language for building web pages), 'CSS' (The language for styling web pages), and 'JavaScript' (The language for programming web). Each box contains a code example and a 'Try It Yourself' button.

HTML Example:

```
<!DOCTYPE html>
<html>
<title>HTML Tutorial</title>
<body>
<h1>This is a heading</h1>
<p>This is a paragraph.</p>
</body>
</html>
```

CSS Example:

```
body {
background-color: lightblue;
}

h1 {
color: white;
text-align: center;
}

p {
font-family: verdana;
font-size: 20px;
}
```

JavaScript Example:

```
<button onclick="myFunction()">Click Me!</button>
```

<https://www.w3schools.com>

This screenshot shows the 'Tryit Editor v3.3' interface from w3schools.com. At the top, it says 'Sicher https://www.w3schools.com/html/tryit.asp?...'. The editor window displays the HTML code from the previous screenshot. Below the code, the output is shown in a large font: 'This is a Heading' and 'This is a paragraph.'.

Result Size: 336 x 279

```
<!DOCTYPE html>
<html>
<head>
<title>Page Title</title>
</head>
<body>
<h1>This is a Heading</h1>
<p>This is a paragraph.</p>
</body>
</html>
```

This is a Heading

This is a paragraph.

Was sind die primären Nachschlagequellen?

W3C

HTML 5.2
W3C Working Draft, 6 April 2017

This version:
<https://www.w3.org/TR/2017/WD-html52-20170406/>

Latest published version:
<https://www.w3.org/TR/html52/>

Latest version of HTML:
<https://www.w3.org/TR/html/>

Editor's Draft:
<https://w3c.github.io/html/>

Previous Versions:
<https://www.w3.org/TR/2017/WD-html52-20170228/>

Editors:
Steve Faulkner (The Paciello Group)
Arron Eicholz (Microsoft)
Travis Leithead (Microsoft)
Alex Danilo (Google)
Sangwhan Moon (Invited Expert)

Former Editors:
Erika Doyle Navara (Microsoft)

WHATWG

HTML
Living Standard – Last Updated 19 April 2017

?

← 3 Semantics, structure, and APIs of HTML documents — Table of Contents — 4.8
Embedded content →

4 The elements of HTML
4.1 The document element
4.1.1 The `html` element

4.2 Document metadata
4.2.1 The `head` element
4.2.2 The `title` element
4.2.3 The `base` element
4.2.4 The `link` element
4.2.4.1 Processing the `media` attribute
4.2.4.2 Processing the `type` attribute
4.2.4.3 Obtaining a resource from a `link` element
4.2.4.4 Processing `Link` headers
4.2.4.5 Providing users with a means to follow hyperlinks created using the `link` element

4.2.5 The `meta` element
4.2.5.1 Standard metadata names

File an issue about the selected text

MDN

MDN-Web-Dokumentation
Sicher | https://developer.mozilla.org

MDN web docs Technologien Referenzen & Leitfäden Rückmeldung Anmelden

Informationen für Entwickler von Entwicklern.

Die Dokumentation durchsuchen

Web-Technologien Web-Entwicklung lernen Entwicklerwerkzeuge

Testen Sie Firefox und über 800 Kombinationen von Browser und Betriebssystem mit Mozilla und Sauce Labs Auf Sauce Labs testen

Lernen Sie das Beste aus dem Bereich Web-Entwicklung
Erhalten Sie das Neueste und Wichtigste von MDN direkt in Ihren Posteingang.
Sie@example.com Melden Sie sich jetzt an

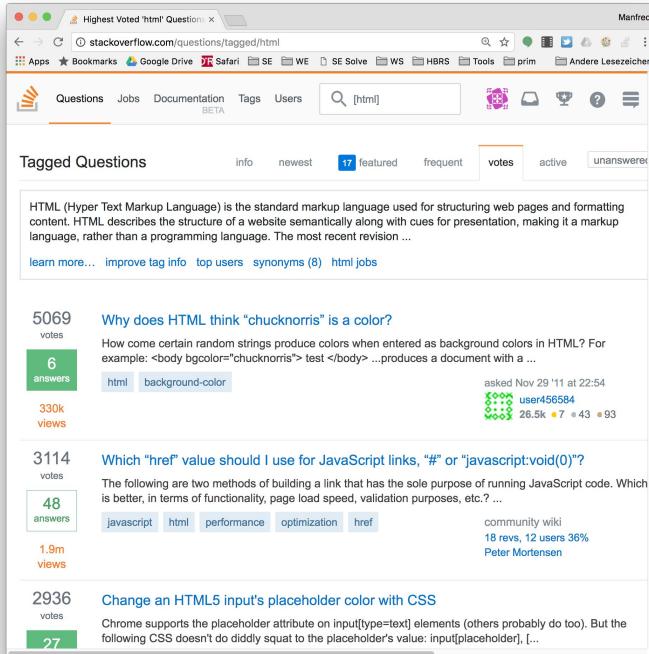
<https://www.w3.org/TR/html52>

<https://html.spec.whatwg.org>

<https://developer.mozilla.org>

<https://medium.com/swlh/w3c-whatwg-standards-437b375cb793>

Wo bekommt man Antworten?



A screenshot of a web browser window titled "Highest Voted 'html' Questions". The URL in the address bar is <http://stackoverflow.com/questions/tagged/html>. The browser's toolbar includes icons for Apps, Bookmarks, Google Drive, Safari, SE, WE, SE Solve, WS, HBRs, Tools, prim, and Andere Lesezeichen. Below the toolbar, the Stack Overflow navigation bar shows "Questions" (selected), "Jobs", "Documentation", "Tags", and "Users". A search bar contains the query "[html]". The main content area displays a list of tagged questions under the heading "Tagged Questions". The first question is "Why does HTML think "chucknorris" is a color?", with 5069 votes, 6 answers, and 330k views. The second question is "Which "href" value should I use for JavaScript links, "#" or "javascript:void(0)"?", with 3114 votes, 48 answers, and 1.9m views. The third question is "Change an HTML5 input's placeholder color with CSS", with 2936 votes, 27 answers, and 1.9m views.



<http://stackoverflow.com/questions/tagged/html>

Praxis und Übungen in der ersten Woche

