

# Webtechnologien

Prof. Dr.-Ing. Georg J. Schneider

Multimedia und Medieninformatik  
Fachbereich Informatik  
Hochschule Trier

# Organisation

**Art der Veranstaltung:** 2 SWS Vorlesung  
2 SWS Zentralübung

**Vorlesungsunterlagen,  
Übungen:** StudIP

**Leistungsnachweis/:  
Prüfungsleistung** Zu Beginn des Semesters wird eine Aufgabe zur Realisierung einer Webpräsenz gestellt. Die Aufgabe soll in Gruppenarbeit im Laufe des Semesters realisiert und dokumentiert werden. Nähere Infos in StudIP.

**Ende des Semesters:** Abgabe des Projektes, Prüfungsgespräch

# Literatur

**Stud.IP**

Bücher oder Links werden zu jedem  
Kapitel separat bekannt gegeben.

# Lernziele

- Systematisches Verständnis für das Phänomen WWW
- Verständnis der technischen Grundlagen des WWW
- Überblick über das WWW als Informations- und Kommunikationssystem
- Fähigkeit zum systematischen Design von WWW-Applikationen

# Aufbau

- Vorgehensmodell, Organisation der Inhalte, Barrierefreiheit
- Beschreibungssprache und Repräsentation von Webseiten ((X)HTML, DOM)
- Formatierung des Seiteninhalts (CSS)
- Programmierung dynamischer Inhalte: Client (Javascript)
- Programmierung dynamischer Inhalte: Server (PHP)
- Adressierungsschema (URI, URN, URL)
- Transport der Daten (HTTP)

# Wichtige Informationsquellen

- FAQ = Frequently Asked Question
  - Die wichtigsten Fragen und Antworten zu einem Thema
  - z.B. für Programmierung, Umgang mit Software, etc.
- RFC = Request For Comment
  - Internet Standards
  - z.B. für Protokolle, Verschlüsselung, etc.

# Wichtige Informationsquellen

## WWW

- W3C - [www.w3.org](http://www.w3.org)
- Kataloge: [suchnase.de](http://suchnase.de), [bunte-suche.de](http://bunte-suche.de)
- Suchmaschinen: google, bing, yahoo (Technologie von MS)
- Besondere Suchmaschinen: [blinde-kuh](http://blinde-kuh.de) (für Kinder), Yandex (Russland), Baidu (China)
- Metasuchmaschine [dogpile](http://dogpile.com), [metager](http://metager.com)
- Semantische Suche [wolframalpha](http://wolframalpha.com)
- Ohne Weiterverfolgung: [duckduckgo](http://duckduckgo.com), [metager](http://metager.com) (Metasuchmaschine)

# Einführung und Überblick



# Was ist das World Wide Web

## Definitionen in der Literatur

- „an internet-wide distributed hypermedia information retrieval system“ [Liu et al. 1994]
- „the World Wide Web is a global, seamless environment in which all information (text, images, audio, video, computational services) that is accessible from the Internet and can be accessed in a consistent and simple way by using a standard set of naming and access conventions“ [WebMaster Magazine 1996]
- „the World Wide Web (known as "WWW", "Web" or "W3") is the universe of network-accessible information, the embodiment of human knowledge“ [W3C 1999]\*

## Ideen und Ziele des Web I

- Lokalisierung von Information mit Hilfe einer uniformen Adressierungsmethode
- Uniformer Zugang (lesen und schreiben) über eine Standard-Benutzungsschnittstelle
- Inhalte als Hypermedia-Dokumente, visualisierbar/abspielbar auf unterschiedlichsten Rechnern

## Ideen und Ziele des Web II

- Integration externer Informationsquellen (z.B. Datenbanken)
- Unterstützung von Transaktionen als Grundlage für interaktive Anwendungen (Client/Server)
- Jeder kann Informationen in das Web einfügen, inhärente Verteilung

# Historischer Abriss

- 1945 Memex (Vannevar Bush)
- 1961 Paketvermittlung (Leonard Kleinrock)
- 1965 Begriffe Hypertext und Hypermedia (Ted Nelson)
- 1969 ARPANET (mit 4 Knoten)
- 1974 TCP (Vinton Cerf, Bob Kahn; Ablösung von NCP 1982)
- 1981 Xanadu (Ted Nelson)
- 1983 Begriff Internet
- 1989 World Wide Web (Berners-Lee, Cailliau; Release 1991)
- 1993 Mosaic Browser (Web hat 341634% jährliche Wachstumsrate)
- 1995 Web überholt FTP im Transfervolumen
- Ca. 2002 Social Networks

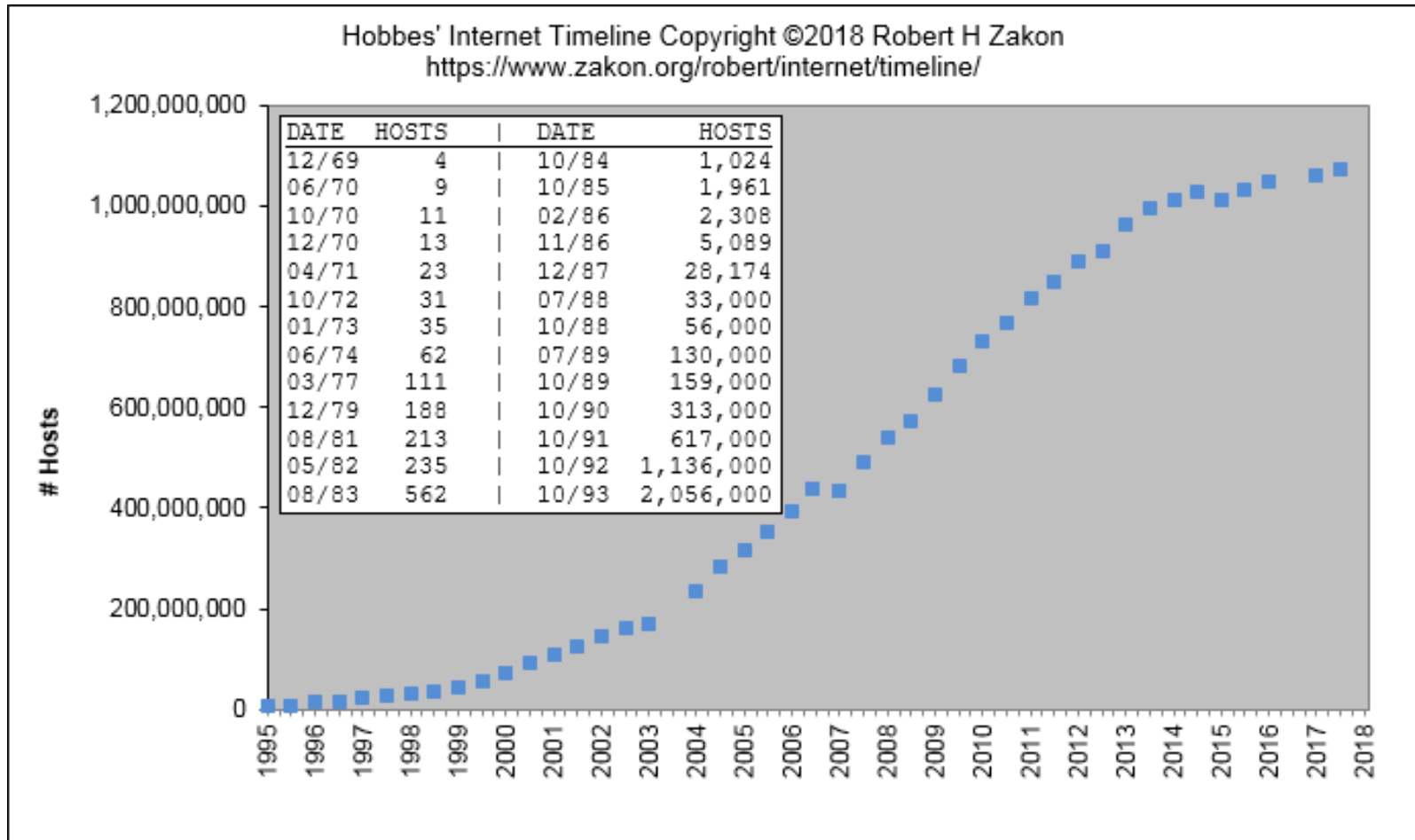
# Statistik und Demoskopie

Alle Zahlen sind Schätzungen!

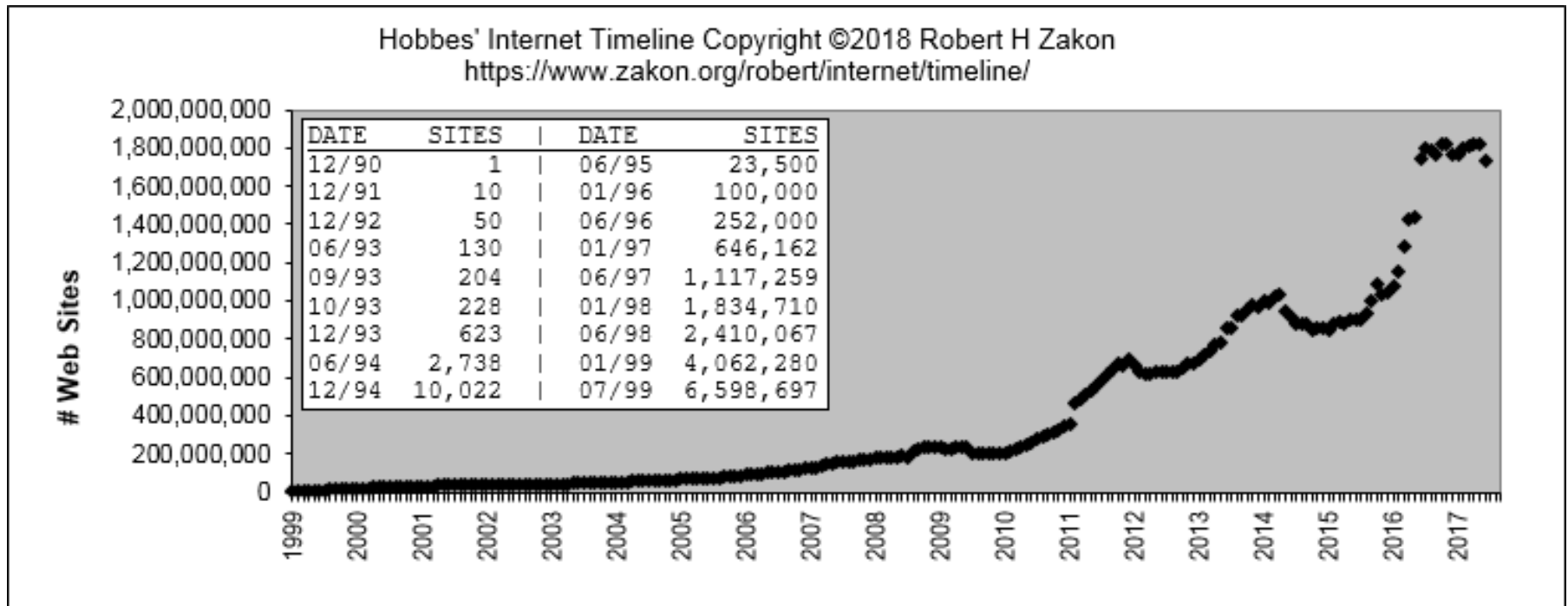
*„The Internet is distributed by nature. This is its strongest feature, since no single entity is in control ...“*

[Marc Abrams (Editor). World Wide Web Beyond the Basics. Prentice Hall 1998. Seite 40]

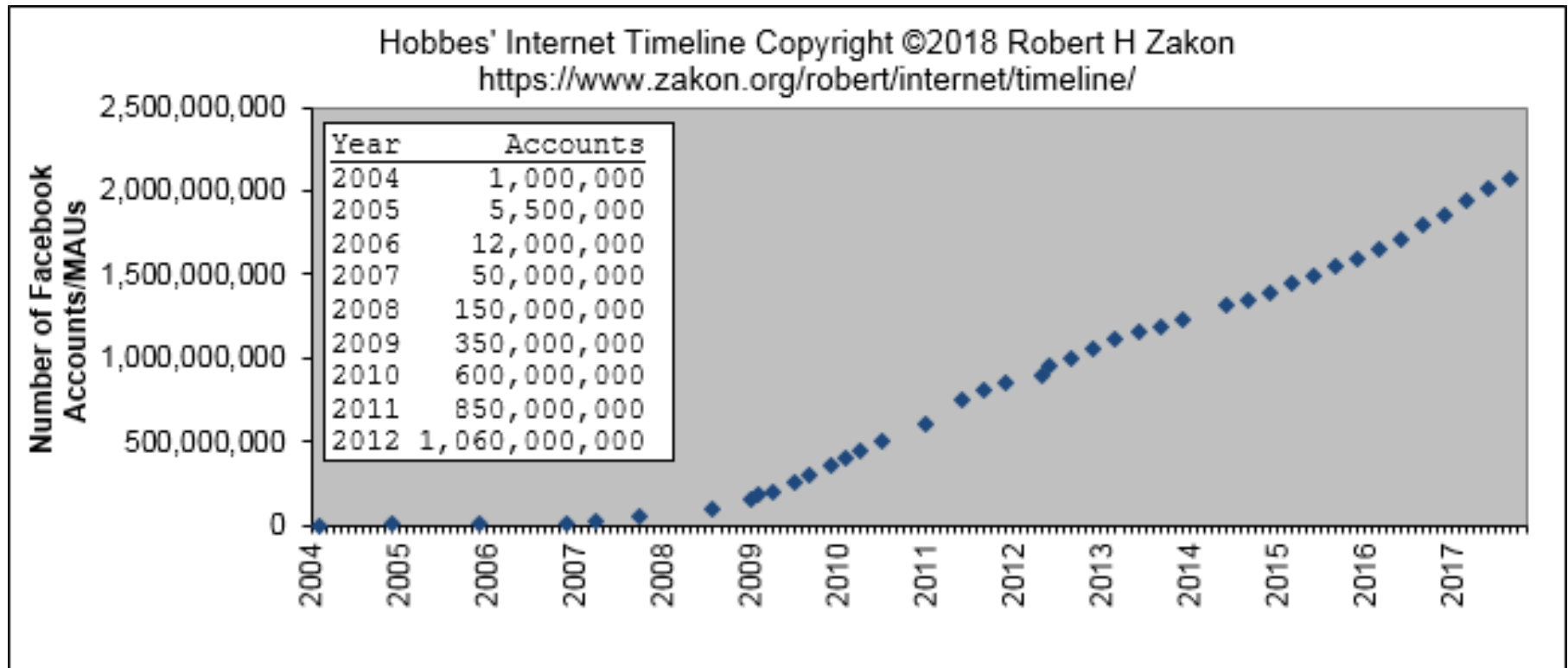
# Wachstum des Web I



## Wachstum des Web II

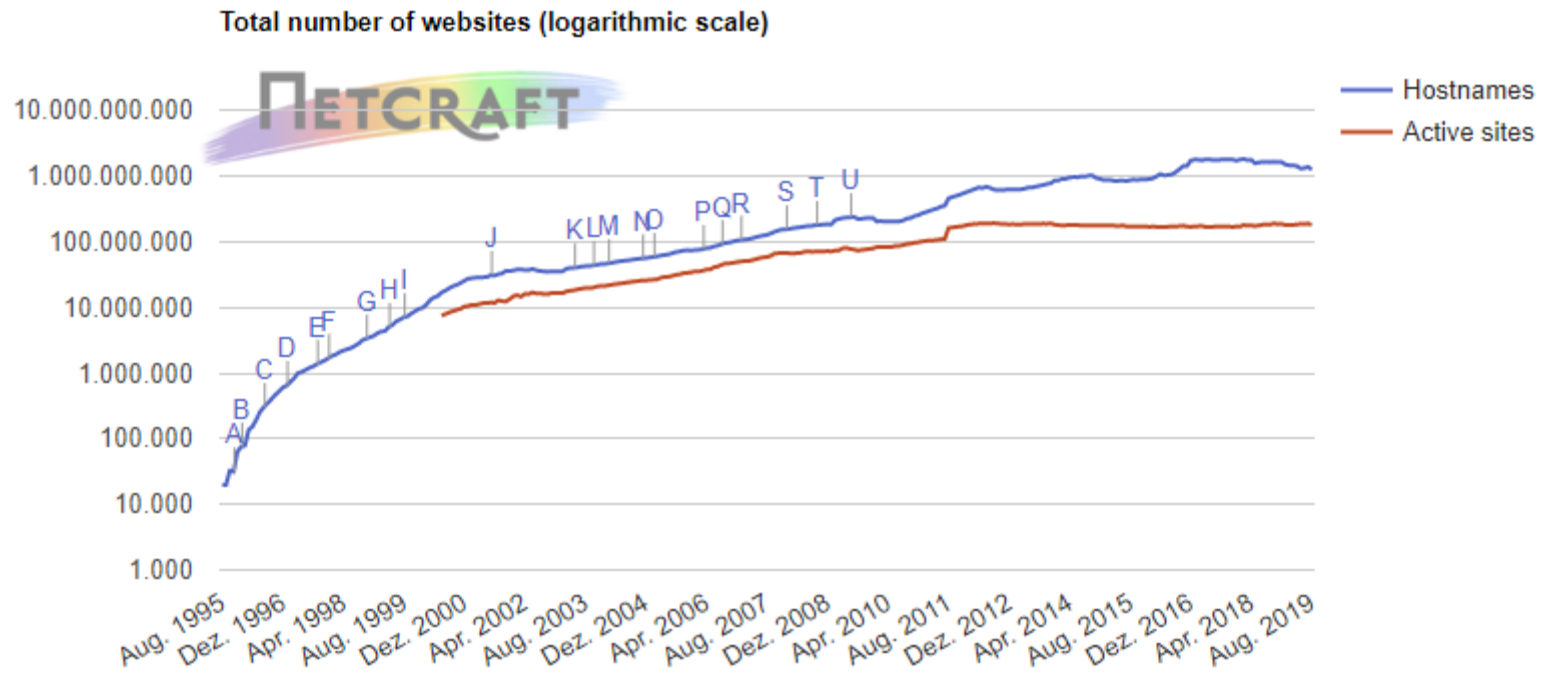


## Wachstum des Web III



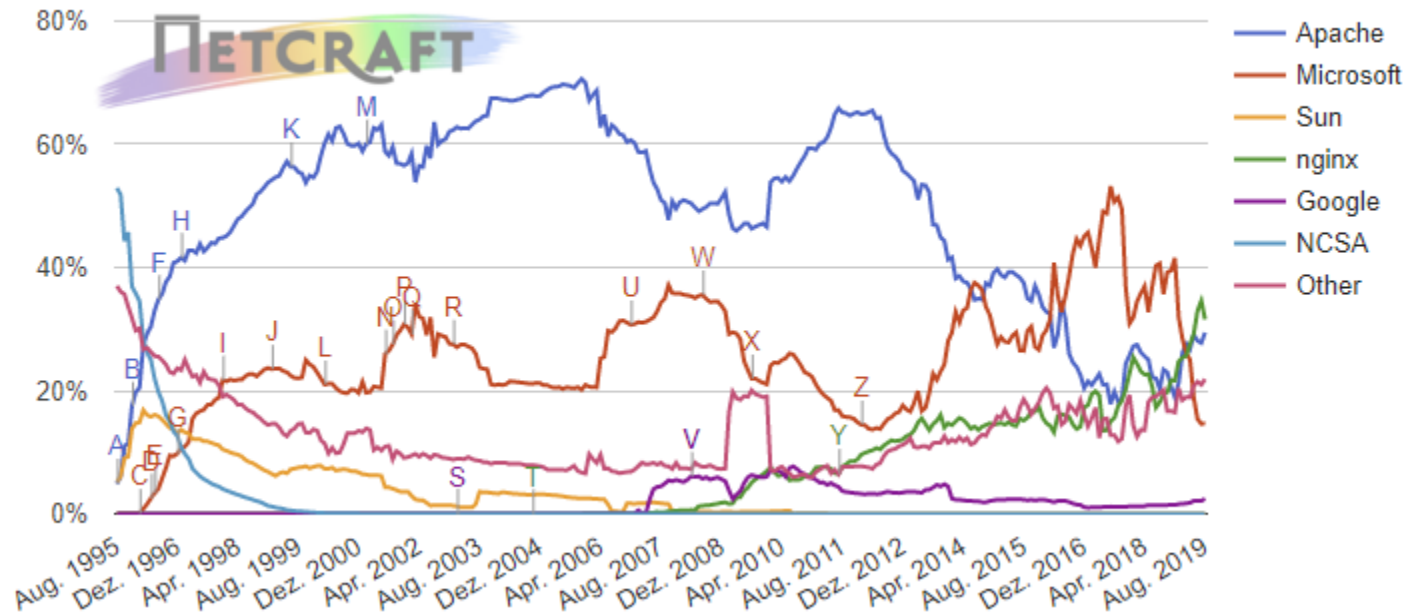


# Wachstum des Web IV



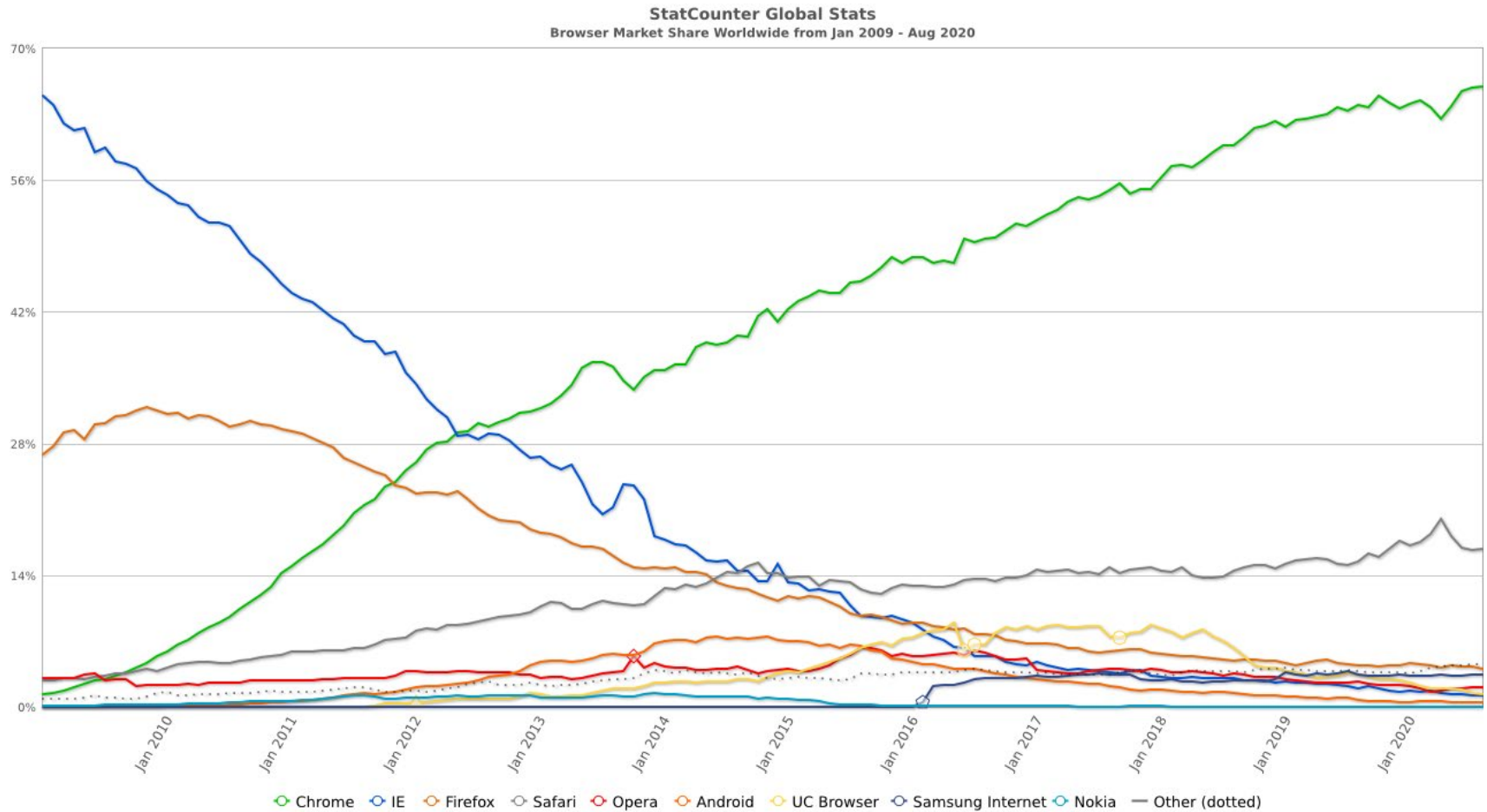
# Webserver

Web server developers: Market share of all sites



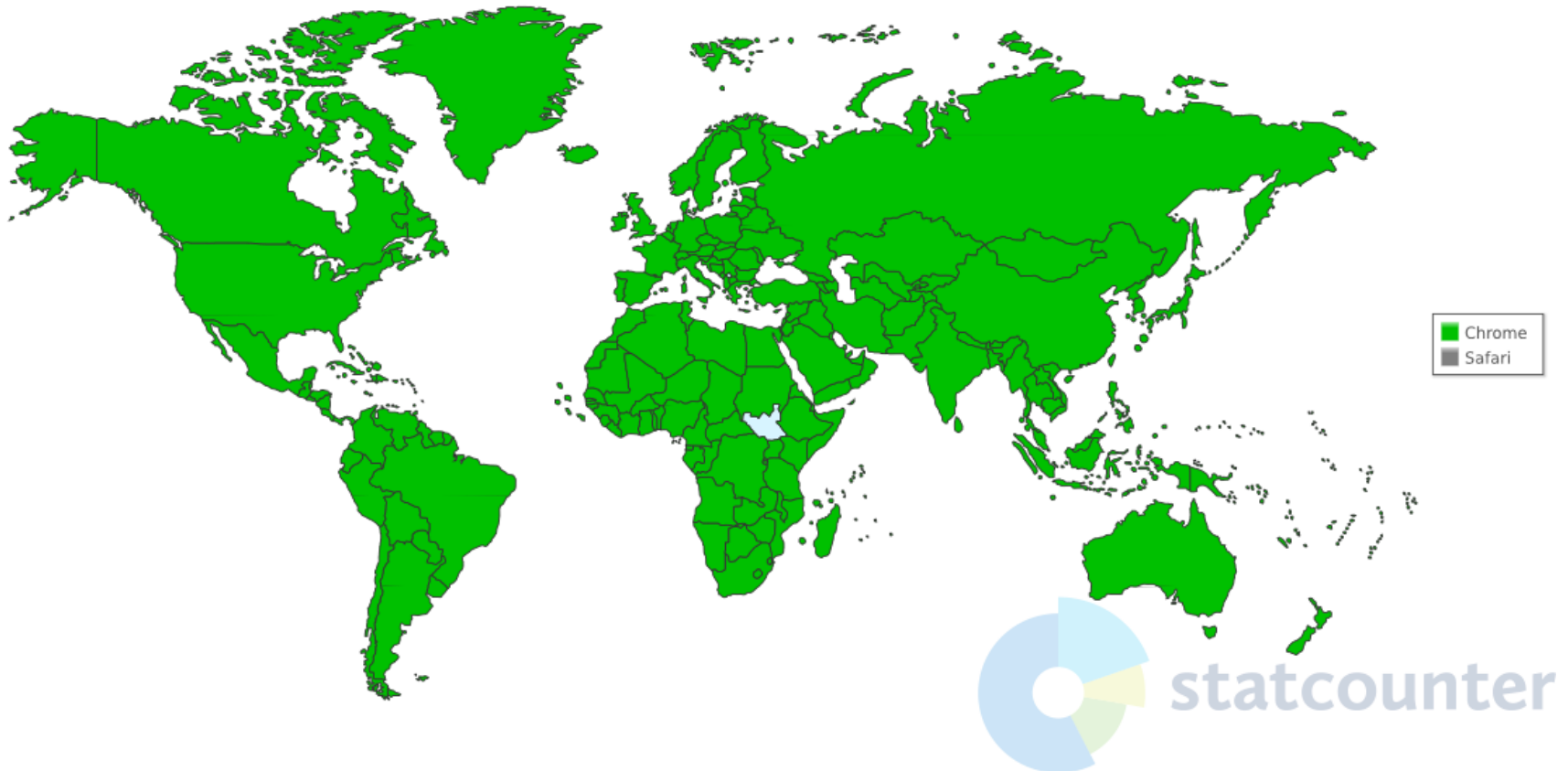
Developer	July 2019	Percent	August 2019	Percent	Change
nginx	482,877,275	34.59%	401,454,029	31.56%	-3.03
Apache	387,366,826	27.75%	374,277,243	29.43%	1.68
Microsoft	203,673,344	14.59%	187,109,423	14.71%	0.12
Google	29,385,065	2.11%	30,969,259	2.43%	0.33

# Web Browser I



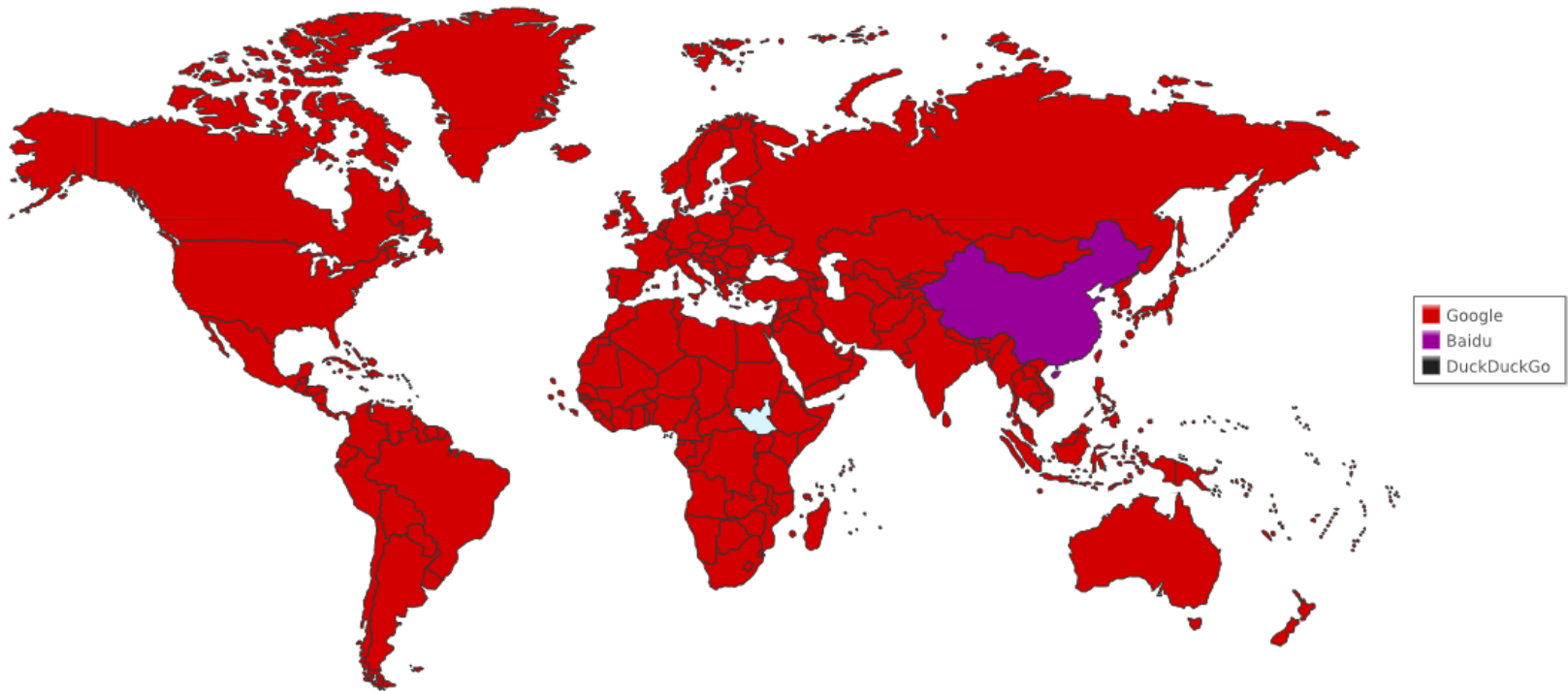
# Web Browser II

StatCounter Global Stats  
Browser Market Share Worldwide, Aug 2020



# Search Engine

StatCounter Global Stats  
Search Engine Market Share Worldwide, Aug 2020



# Typen von Suchmaschinen

- Crawler oder Spider basierte
- „Human-Powered Directories“
- Hybride Suchmaschinen oder Mixed Results

# Ranking bei Suchmaschinen

- Ort und Häufigkeit von Schlüsselwörtern auf einer Seite
- Link Popularität
- Meta Tags (später)

# Link Popularität bei Google (Page Rank)

$$PR(A) = (1-d) + d (PR(T1)/C(T1) + \dots + PR(Tn)/C(Tn))$$

Hierbei ist:

PR(A) der PageRank einer Seite A,

PR(Ti) der PageRank der Seiten Ti, von denen ein Link auf die  
Seite A zeigt,

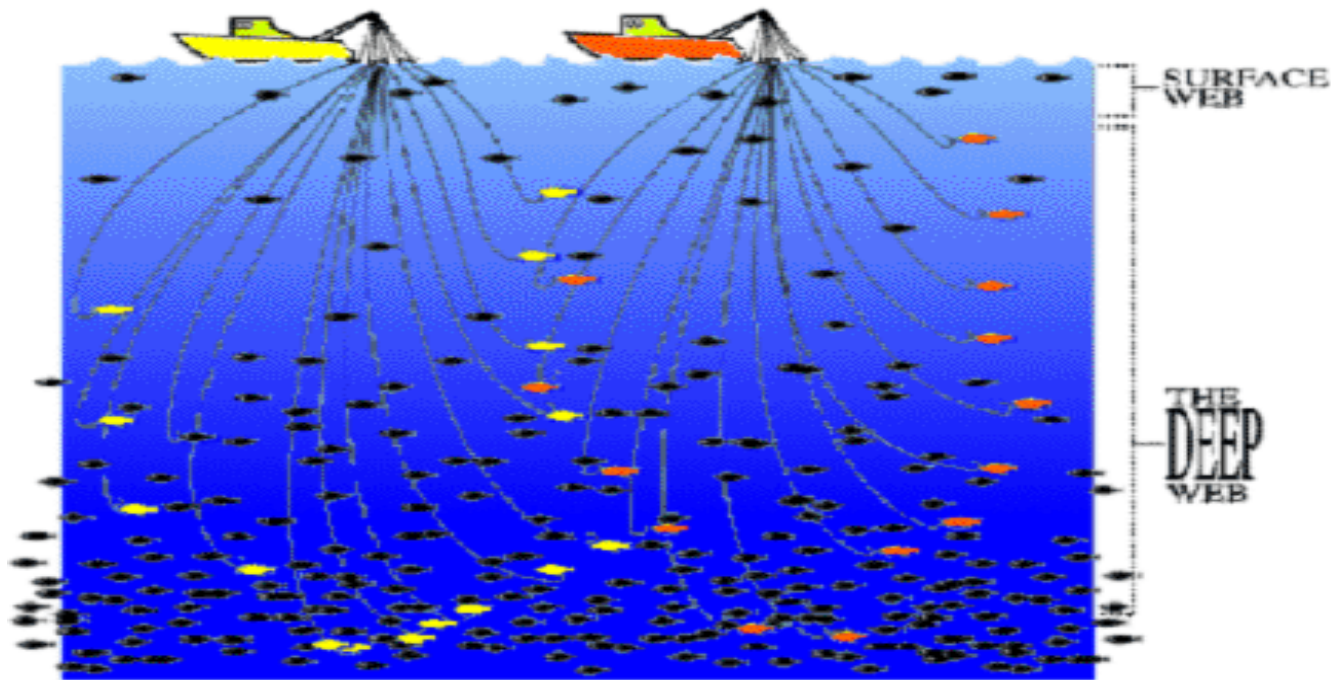
C(Ti) die Gesamtanzahl der ausgehenden Links von Seite Ti und

d ein Dämpfungsfaktor (Damping Factor),

wobei  $0 \leq d \leq 1$  ist.



# Deep Web



**Deep Web betrachtet auch dynamisch generierte Datenbankabfragen.  
Schätzung: ca. 500 x größer als Surface Web**

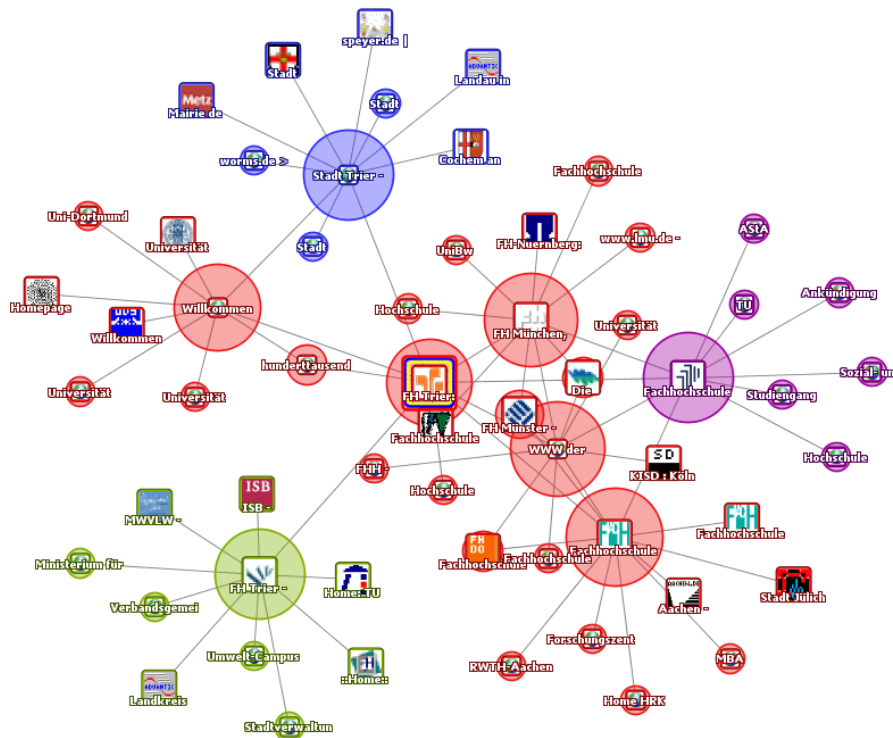
## Deep Web contd.

Wenn in der internen Linkstruktur Ihrer Site nicht alle Seiten verknüpft sind, finden unsere Suchprogramme möglicherweise nicht alle Seiten auf Ihrer Site. Google folgt den Verknüpfungen von einer Seite zur nächsten. Seiten, auf die nicht von anderen Seiten aus verwiesen wird, können also unbemerkt bleiben<sup>1</sup>

Siehe auch: <https://support.google.com/webmasters/answer/35769> (Aufruf 9/2020)

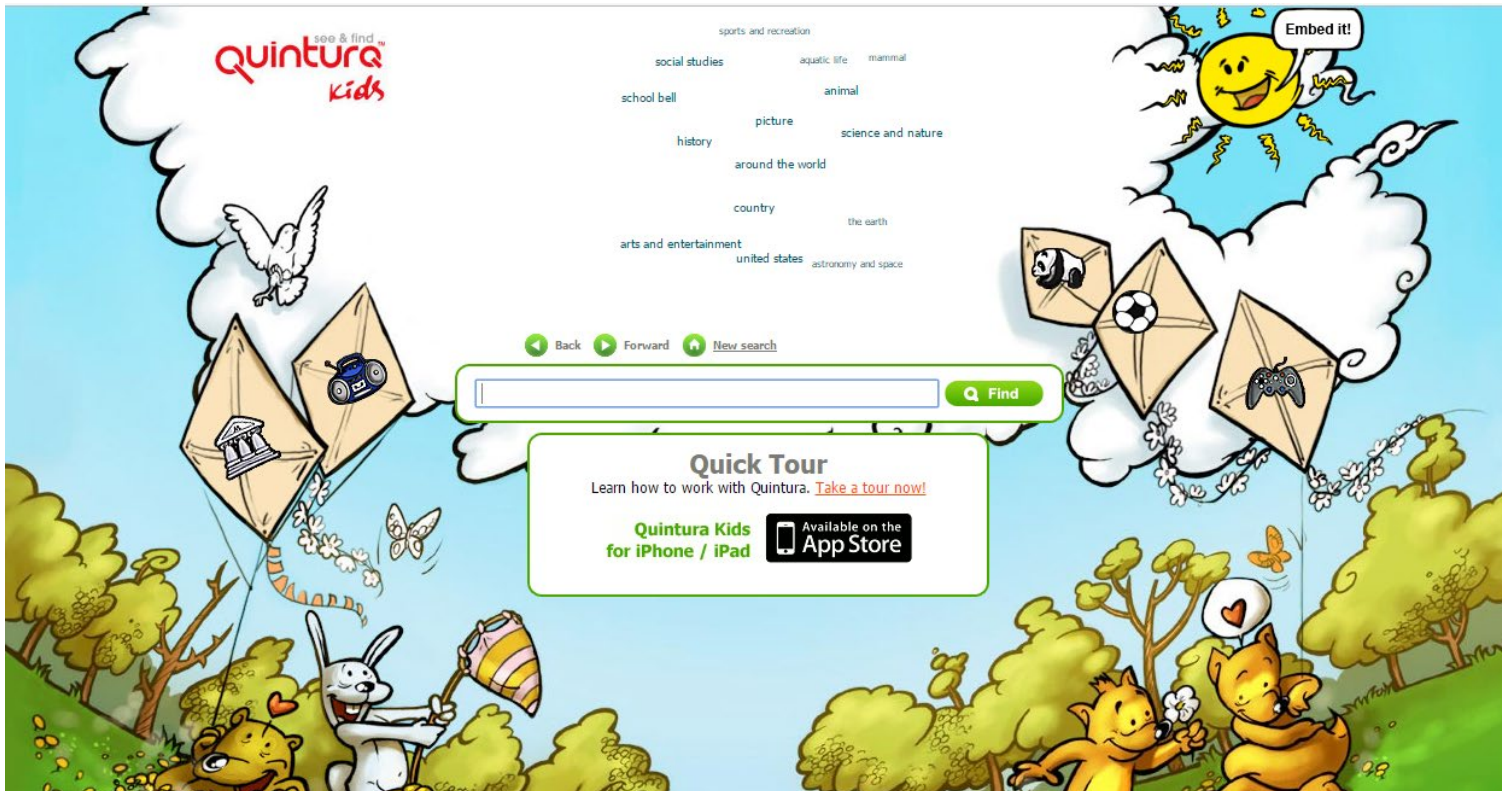
<sup>1</sup><http://www.google.de/intl/de/webmasters/2.html>, 10/2013

# Visualisierung von Suchergebnissen I

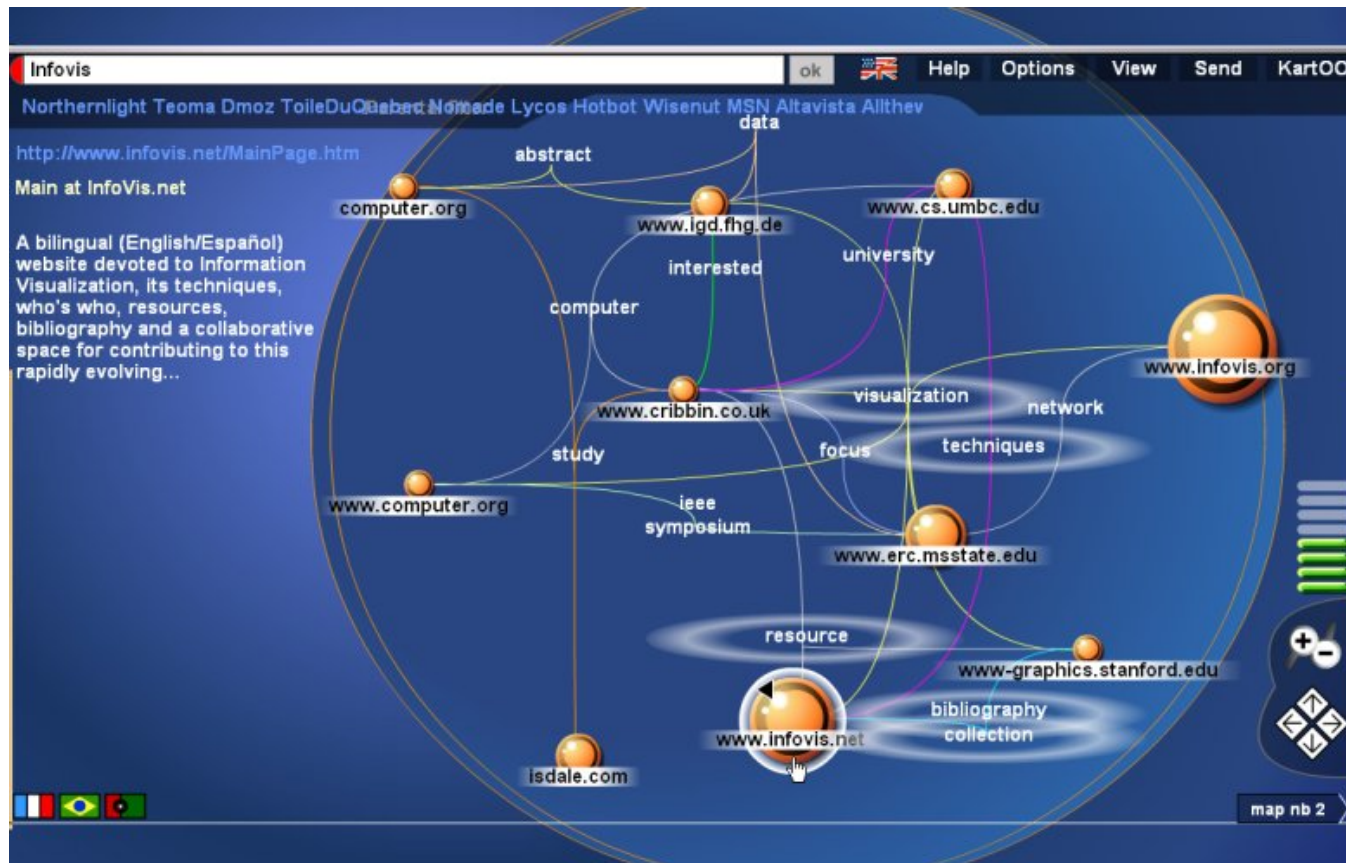


<http://www.touchgraph.com/>

# Visualisierung von Suchergebnissen II



# Visualisierung von Suchergebnissen III



Kartoo nicht mehr verfügbar

## Internet versus WWW (1)

The Internet is the entirety of all computers which are networked (using various networking technologies) and employ the Internet protocol suite on top of their networking systems.

=> **Infrastruktur**

The World Wide Web is a distributed Hypermedia system, which is built on top of some of the services provided by the Internet.

=> **Anwendungsarchitektur auf Internet**

## Internet versus WWW (2)

Internet ist eine Verknüpfung heterogener Netzwerke

- Netzwerkprotokoll IP schafft die Transparenz
  - IP ist paketvermittelnd
- verschiedene Transportprotokolle
  - TCP, UDP, RTP, ...
- unterstützt eine Vielfalt von Anwendungsprotokollen
  - Telnet, FTP, NNTP, SMTP, HTTP, ...

Web ist einer von vielen Internetdiensten



# Web versus Multimedia I

Multimedia ist das gleichzeitige Auftreten von diskreten und kontinuierlichen Medien

Multimediadokumente:

- HTML unterstützt Einbettung von Bewegtbild-, Video- und Tondokumenten
- in Erweiterungen auch besondere Formate
  - Flac, Opus, MPEG-H, ...



# Web versus Multimedia II

## Multimediakommunikation

- HTTP unterstützt Transfer jeglicher Ressourcen
- Erweiterungen für Streaming von Audio und Videodokumenten

Web kann Multimediaelemente  
enthalten und interpretieren

# Web Begriffe I

User:	Menschlicher Benutzer des Web
Client:	Programm, welches HTTP Anfragen über das Web versendet
Browser:	Client, der HTML anzeigen kann
Server:	Software, die auf HTTP Anfragen wartet und mit HTTP Antworten reagiert
Site:	Eine Sammlung von Webseiten, die organisatorisch zusammen gehören

## Web Begriffe II

### Page

- eine einzelne HTML Seite
  - kann andere Dokumenttypen enthalten

### Homepage

- Einstiegsseite zu einem inhaltlichen Zusammenhang, z.B.
  - Homepage einer Firma oder Organisation
  - thematische Homepage
  - persönliche Homepage

# Web Begriffe III

## Portal

- speziell entworfene Betriebsoberfläche für das Web, z.B.
  - Portale von Netzprovidern (z.B. aol, t-online, ...)
  - Portale von Medienunternehmen (z.B. web.de, stern, focus, ...)
  - Portale von Internetdienstleistern (z.B. yahoo, lycos, ...)