

# 1 Einführung ins Internet

„Das Internet ist für uns alle Neuland! “  
Angela Merkel, am 19.06.2013

## 1.1 Entstehung des Internets

### 1.1.1 ARPA-Net

Die Defense Advanced Research Projects Agency (ARPA), eine seit 1958 bestehende wissenschaftliche Einrichtung, deren Forschungsergebnisse in militärische Zwecke einfließen, entschloss sich 1966 zur Vernetzung der ARPA-eigenen Großrechner. Dabei wurde die Idee des "dezentralen Netzwerks" wieder aufgegriffen. Ende 1969 waren die ersten vier, drei Jahre später bereits 40 Rechner an das ARPA-Net angeschlossen. Aus ihm sollte später das Internet erwachsen.

→ Idee eines dezentralen Netzwerkes, dass miteinander kommuniziert bzw. Daten austauscht. Falls ein Rechner ausfällt, sind die Daten auf den anderen noch vorhanden.

### 1.1.2 Wissenschaftliche Einrichtungen

→ Für die Wissenschaft war der Austausch neuer Daten interessant.

Wegen der offenen Architektur des ARPA-Net stand einer solchen Verwendung nichts im Wege. Wissenschaftler konnten von den frühen 70er Jahren an Forschungsergebnisse anderer Institute über das ARPA-Net abrufen oder anderen angeschlossenen Instituten eigene Daten zur Verfügung stellen. Da die angeschlossenen Rechner verschiedene Typen hatten, also Großrechner, Unix-Rechner oder auch PC's, musste ein unabhängiges Datenübertragungsprotokoll für das Netz entwickelt werden.

Das **TCP/IP**-Protokoll wurde dann 1973/74 entwickelt und wurde dann am 01.01.1983 vollständig im Arpanet etabliert. Erst als ein einheitliches Netzwerkstandard durch Netzwerkprotokolle (OSI-Modell) erarbeitet wurde, konnten Verbindungen zwischen Host unterschiedlicher lokaler Netze sichergestellt werden.

### 1.1.3 Das World Wide Web (WWW)

Der Informatiker Tim Berners-Lee, der in den 80er und 90er Jahren am CERN gearbeitet hat, schrieb ein **Hypertext**-Programm, mit dem man Textdateien editieren konnte. Diese waren in Knoten unterteilt und es gab zu Knoten eine zugehörige Liste mit Links zu anderen Knoten. Man konnte auf jegliche Art von Querbeziehung linken, die man kannte oder fand. Links auf Ziele innerhalb einer Datei wurden vom Enquire-Programm automatisch bidirektional dargestellt – das heißt, auch wenn der Link nur von A nach B gesetzt war, fand man bei B in der Liste den Rückverweis auf A.

→ Hypertext bedeutet in etwa so viel wie Übertext. Gemeint ist jedoch das Hypermedium als das den konkreten Text und seine Medien übergreifende Medium. Hypertext bleibt dabei ein abstraktes Konzept, da es keine konkrete Summe aller involvierten Texte

in Hypertext geben kann.

Im Herbst des Jahres 1990 schrieb Berners-Lee eigenhändig die ersten Versionen der drei Säulen seines Konzepts:

- 1. die Spezifikation für die Kommunikation zwischen Web-Clients und Web-Servern – das so genannte HTTP-Protokoll (HTTP = Hypertext Transfer Protocol)
- 2. die Spezifikation für die Adressierung beliebiger Dateien und Datenquellen im Web und im übrigen Internet – das Schema der sogenannten URIs (URI = Universal Resource Identifier, universeller Quellenbezeichner).
- 3. die Spezifikation einer Auszeichnungssprache für Web-Dokumente, der Berners-Lee den Namen HTML gab (Hypertext Markup Language, Hypertext Auszeichnungssprache).

Berners-Lee schrieb auch die erste Web-Server-Software die unter `info.cern.ch` erreichbar war.

#### **1.1.4 Webbrowser**

Wirklich benutzbar wurde das Internet erst durch die Entwicklung von Web-Browsern, mit denen man Webseiten ohne komplizierte Befehle laden und anschauen konnte. Da aber unterschiedliche Unternehmen, wie Microsoft oder Netscape immer wieder neue Features einführten, die die Konkurrenz nicht unterstützte, erarbeiteten das World Wide Web Consortium (W3C) gemeinsame Standards. Da diese aber zu langsam waren, gründeten Apple, Google und Mozilla 2004 die Web Hypertext Application Technology Working Group (WHATWG), die neue Standards unter dem Namen **HTML 5** entwickeln sollten. Derzeit beherrscht Google mit seinem Chrom-Browser den Markt. (vgl. Abbildung 1)

### **1.2 Aufbau des Internets und Adressierung**

Das Internet ist ein weltumspannendes, dezentrales Daten-Netz aus Millionen verbundenen Computern und Computernetzwerken, die miteinander kommunizieren können. Prinzipiell kann jeder Computer, ob Smartphone, Tablet oder PC, ans Internet angeschlossen werden. Um einen Computer mit dem Internet zu verbinden, muss man eine Funk, Daten- oder Telefonverbindung zu einem Internetdienstanbieter (engl. Internet Service Provider) aufbauen. Dieser Provider ist über Datenleitungen mit anderen Providern und Internet-Knoten verbunden, über die man auf Webseiten aus der ganzen Welt zugreifen kann.

Auf z.B. <https://global-internet-map-2021.telegeography.com/> findet sich eine sog. Internetkarte mit den unterschiedlichen Knotenpunkten.

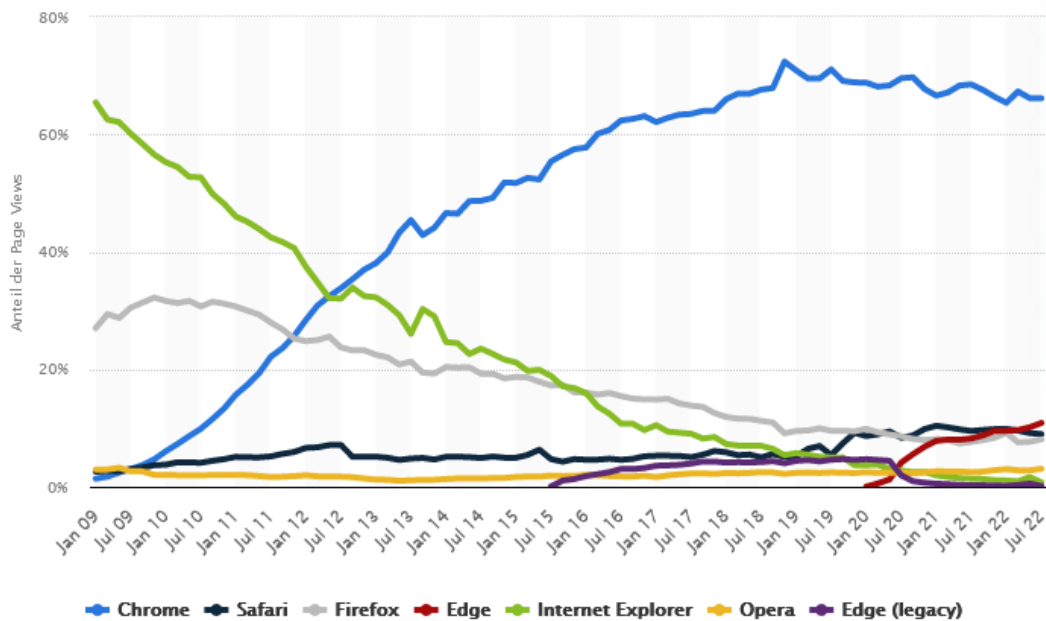


Abbildung 1: Marktanteile der führenden Webbrowser von 2009 bis 2018.

### 1.2.1 Internet-Provider

Der Internet Service Provider, den man im Alltag als Internet Provider bezeichnet, ist der Anbieter, der dem Nutzer das Internet bereitstellt. Ein solcher Provider stellt die Verbindungstechnik bereit, zu der in den Augen der Normalnutzer vor allem der Router gehört. Diese Provider kümmern sich aber auch um die Leitungen bzw. um die Funktechnik. Man unterscheidet zwischen Tier-1-, Tier-2- und Tier-3-ISP, wobei die Klassifizierung eine Aussage über die Netzwerkgröße zulässt. Tier-1-Provider sind die lokalen Anbieter, während die Tier-3-ISP auf der globalen Ebene agieren.

### 1.2.2 Client-Server-Technologie

Für die einzelnen Internetdienste wie WWW, E-Mail, FTP usw. muss auf einem Host-Rechner, der anderen Rechnern diese Dienste anbieten will, eine entsprechende Server-Software laufen. Ein Host-Rechner kann einen Internet-Dienst nur anbieten, wenn eine entsprechende Server-Software auf dem Rechner aktiv ist, wenn der Rechner online ist und wenn keine schützende Software (Firewall) den Zugriff von außen verhindert bzw. einschränkt.

Server sind Programme, die permanent darauf warten, dass eine Anfrage eintrifft, die ihren Dienst betreffen. So wartet etwa ein Web-Server darauf, dass Anfragen eintreffen, die Web-Seiten auf dem Server-Rechner abrufen wollen. Clients sind dagegen Software-Programme, die typischerweise Daten von Servern anfordern. Ihr Browser ist

beispielsweise ein Client. Wenn Sie etwa auf einen Verweis klicken, der zu einer HTTP-Adresse führt, startet der Browser, also der Client, eine Anfrage an den entsprechenden Server auf dem entfernten Host-Rechner. Der Server wertet die Anfrage aus und sendet die gewünschten Daten. Um die Kommunikation zwischen Clients und Servern zu regeln, gibt es entsprechende Protokolle.

→ Internetdienste müssen von Host-Rechnern 'angeboten' werden. Dies geschieht durch Server-Software, die darauf warten bis eine Anfrage durch einen Client gestellt wird. Die Kommunikation zwischen Client und Servern regeln entsprechende Protokolle:

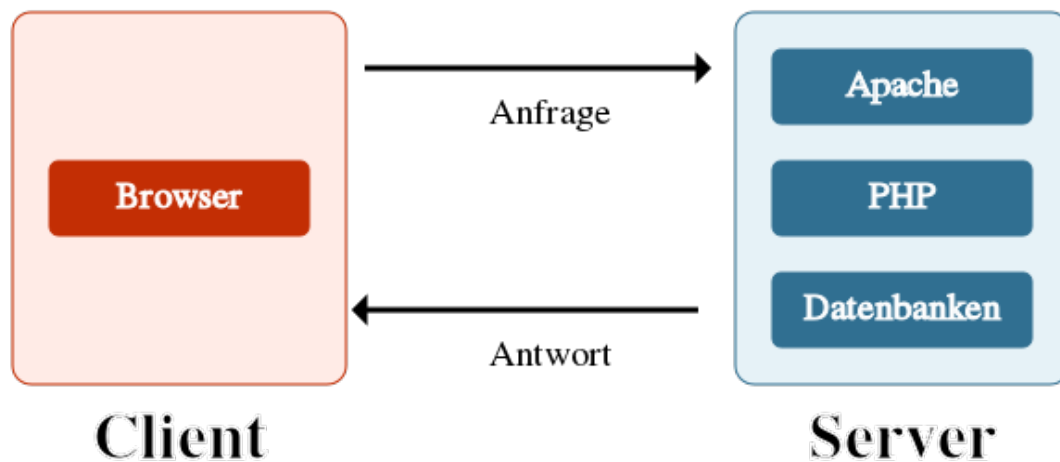


Abbildung 2: Client-Server-Modell

### 1.3 Protokolle

Protokolle wurden entwickelt, damit es einheitliche Verfahrens- und Ablaufvorschriften gibt. Dies gewährleistet die Kommunikation zwischen unterschiedlichen Computern und Netzwerksystemen.

#### 1.3.1 Internet Protocol (IP)

Jeder Rechner im Internet besitzt eine IP-Adresse und ist dadurch eindeutig identifizierbar. Eine solche Adresse besteht bei IPV4 aus 32 Bit, die in vier durch Punkte getrennte Zahlen zwischen 0 und 255 dargestellt werden (gepunktete Dezimalform), oder bei IPV6 durch acht Vierergruppen von hexadezimalen Digits, also  $8 \times 16 = 128$  Bit. Die einzelnen Gruppen werden hier durch Doppelpunkte getrennt.

### 1.3.2 Domain Name System (DNS)

Ein Domainname besteht aus mehreren, durch Punkt getrennten Labels. Er wird immer von rechts nach links gelesen und ausgewertet. Am wichtigsten ist die Top Level Domain, ein Buchstabenkürzel, das das Land (oder eine weltweite Organisation) kennzeichnet.

Host/Dienst	Second-Level-Domain	Top-Level-Domain
www	youtube	de

→ Es gibt in Inhaltsverzeichnis um das Ziel von hinten aufzuschlüsseln. Dieses wird auf Root-Servern gespeichert, von denen es 13 auf der Welt gibt.

**Iterative-Suche:** Bei einer Anfrage wird nun das Inhaltsverzeichnis nach dem zuständigen Nameserver für die TLD gefunden. Dieses ermittelt nun den zuständigen SLD-NS.

**Rekursive Suche:** Bei einer Anfrage wird der nächste erreichbare NS angefragt, des Internet-Anbieters. Falls dieser keinen Eintrag hat, fragt er den NS, der für ihn zuständig ist, das geht dann bis zum Root-Server und dann startet die iterative.

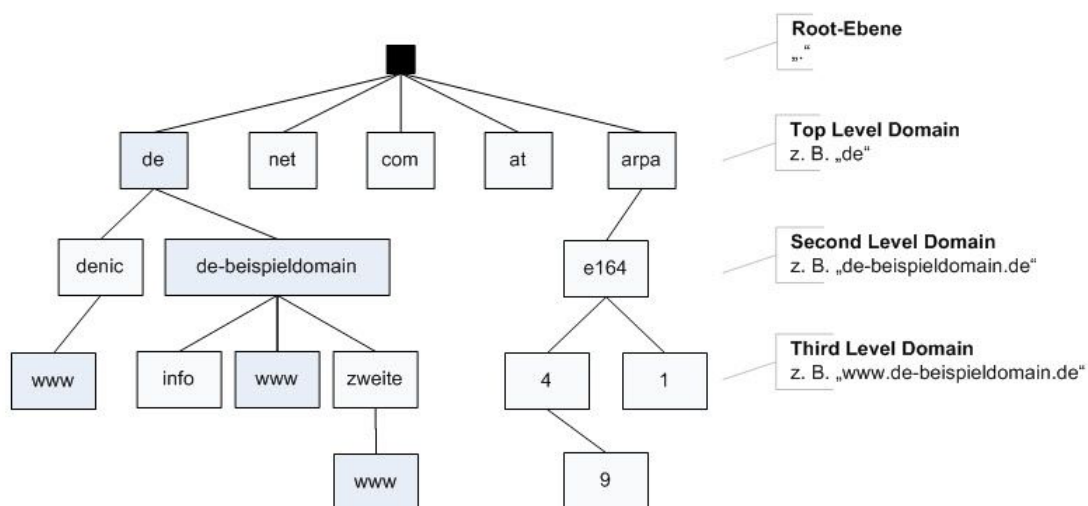


Abbildung 3: Schematischer Aufbau einer Namensauflösung.

### 1.3.3 Uniform Resource Locator (URL)

Auf Deutsch: einheitlicher Ressourcenzeiger. Im allgemeinen Sprachgebrauch auch als Internetadressen oder Webadressen bezeichnet.

→ identifiziert und lokalisiert eine Ressource, z.B. eine Webseite oder Datei.  
`https://www.youtube.com/watch?v=dQw4w9WgXcQ`

- Übertragungsprotokoll: https:
- Adresse: www.youtube.com
- evtl. Pfad: watch?v=dQw4w9WgXcQ

### 1.3.4 TCP/IP

Da Texte, Bilder und Videos oft aus zu großen Dateien bestehen, werden sie für den Transport mit dem TCP/IP-Protokoll in Datenpakete zerlegt und dann verschickt. Dabei nehmen die Pakete nicht den direkten Weg, sondern werden je nach Verfügbarkeit zwischen den einzelnen Internet-Knoten hin und her geroutet (deshalb der engl. Begriff: Router). So läuft ein Großteil „innerdeutscher“ E-Mails und Seiten-, bzw. Suchanfragen über Router, die im Ausland stehen.

Damit die Datenpakete wieder auffindbar sind, bestehen sie immer aus zwei Teilen, dem Header und der Nutzlast (payload). Die Nutzlast enthält die zu übertragenden Daten, die wiederum Protokollinformationen der Anwendungsschicht wie HTTP oder FTP entsprechen können. Der Header enthält für die Kommunikation erforderliche Daten sowie die Dateiformat-beschreibende Information.

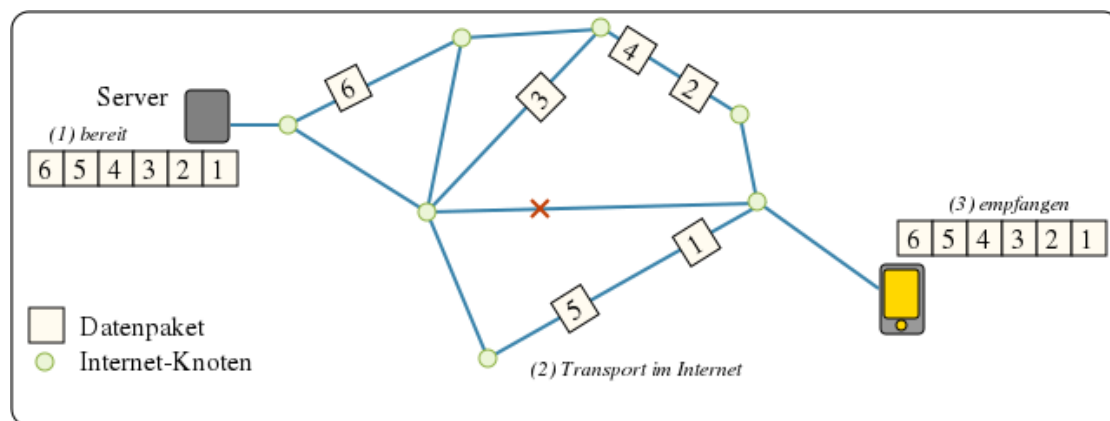


Abbildung 4: Beispiel-schema für ein Datentransport durch das Internet, mit Hilfe des TCP/IP Protokolls.

→ Daten werden in kleinere Pakete getrennt und werden bis zum Zielort hin und her geroutet. Die Datenpakete bestehen aus dem **Header**, der die Informationen für die Kommunikation beinhaltet und der **Nutzlast**, die wiederum die übertragenden Daten hat.

## **1.4 Organisationen des Internets**

### **1.4.1 W3C (World Wide Web Consortium):**

Diese Organisation überwacht die Entwicklung von Standards im World Wide Web und dokumentiert diese. Das W3C arbeitet mit der ISOC zusammen, ist ihr aber nicht untergeordnet.

### **1.4.2 IGF (Internet Governance Forum):**

Dieses Beratungsgremium wurde von der UN-Organisation für wirtschaftliche und soziale Angelegenheiten (UN-DESA) 2004 ins Leben gerufen. Das IGF berät die Gremien der ISOC und versucht, den Einfluss der Vereinten Nationen auf das Internet sicherzustellen.

### **1.4.3 ICANN (Internet Corporation for Assigned Names and Numbers):**

Diese Non-Profit-Organisation koordiniert die Vergabe von einmaligen Namen und Adressen im Internet. Dazu gehört die Koordination des Domain Name Systems und die Zuteilung von IP-Adressen. Die ICANN hat ihren Hauptsitz in Los Angeles und wurde 1998 gegründet.

### **1.4.4 IANA (Internet Assigned Numbers Administration):**

Die Vergabe von Internet-Adressen unterstand bis 1998 einem einzelnen Mann: Jon Postel. Nach dem Tode des Informatikers wurde die Kontrolle über das Domain Name System an die ICANN übergeben.

### **1.4.5 IRTF (Internet Research Task Force):**

Diese Gruppe widmet sich der Forschung an neuen Internet-Technologien. Sie wurde 1986 gegründet, um die Forschung und Entwicklung im Bereich der Netzwerke und deren Techniken zu fördern. Ihre Arbeit wird koordiniert von der IRSG (Internet Research Steering Group).

### **1.4.6 IETF (Internet Engineering Task Force):**

Diese Organisation ist eine Vereinigung von Informatikern und Fachleuten, die Standards wie das Internet Protocol und Übertragungsstandards wie TCP oder UDP entwickeln und dokumentieren. Seine Arbeit wird gesteuert von der IESG (Internet Engineering Steering Group).

### **1.4.7 IAB (Internet Architecture Board):**

Das IAB ist ein Komitee, das den Überblick über die Arbeit der einzelnen ISOC-Unterorganisationen hält und mit dem RFC Editor die Dokumentation von Internet-Standards organisiert. → RFC: Request for Comments - Dokumentation.