

Web-Technologie

Benno Stein

Inhalt

- I. Einführung
- II. Rechnerkommunikation und Protokolle
- III. Dokumentsprachen
- IV. Client-Technologien
- V. Server-Technologien
- VI. Architekturen und Middleware-Technologien
- VII. Semantic Web

Ziele

- ❑ Grundbegriffe von Web-basierten Systemen kennen und einordnen können
- ❑ Zusammenhänge zu angrenzenden Gebieten herstellen können
- sich selbst weiterbilden können
- ❑ Erwerb von Kenntnissen über Bausteine, Architektur und Funktionalität von Web-basierten Systemen
- ❑ Verständnis für spezifische Eigenschaften Web-basierter Systeme
- ❑ Beherrschung der Grundlagen ausgewählter Web-basierter Sprachen
- ❑ Umsetzung eines Web-basierten Systems mit Java

Angrenzende Gebiete

1. Web-Engineering [Modelle, Methodologien]
2. Software-Engineering
3. Rechnerarchitekturen und -netze [Algorithmen]
4. (verteilte) Datenbanken
5. Information Retrieval und Information Extraction
6. Machine Learning und Data Mining
7. Logik, automatisches Beweisen, Wissensverarbeitung
8. Sicherheit und Kryptografie
9. CMS / Wissens- und Dokumentenmanagement [Anwendungen]
10. E-Business, E-Government, E-Learning
11. CSCW, Groupware

Literatur

Java:

- ❑ Ullenboom.
Java ist auch eine Insel.
10. Auflage, Galileo Computing, 2012. openbook.galileocomputing.de

Verteilte Systeme:

- ❑ Comer.
Computer Networks and Internets with Internet Applications.
5. Auflage, Pearson Prentice Hall, 2008. ISBN 9780136061274.
- ❑ Coulouris/Dollimore/Kindberg.
Verteilte Systeme: Konzepte und Design.
3. Auflage, Pearson Studium, 2005. ISBN 3827371864.
- ❑ Meinel/Sack.
Internetworking: Technische Grundlagen und Anwendungen.
Springer, 2012. ISBN 978-3540929390.
- ❑ Tanenbaum.
Computernetzwerke.
4. Auflage, Pearson Studium, 2003. ISBN 3827370469.

Literatur

Web-Engineering:

- ❑ Ceri/Fraternali/Bongio/Brambilla/Comai/Matera.
Designing Data-Intensive Web Applications.
Morgan Kaufmann Publishers, 2003.
- ❑ Dumke/Lothar/Wille/Zbrog.
Web Engineering.
Pearson Studium, 2003. ISBN 3827370809.

Web-Technologie:

- ❑ Ayala/Browne/Chopra/Sarang/Apshankar/McAllister.
Open Source Web Services.
Wrox Press, 2002. ISBN 9781861007469.
- ❑ Comer.
Computer Networks and Internets with Internet Applications.
5. Auflage, Pearson Prentice Hall, 2008. ISBN 9780136061274.
- ❑ Meinel/Sack.
Web-Technologien.
Springer, 2013. ISBN 978-3-540-92945-1.

Literatur

XML:

- ❑ Harold/Means.
XML in a Nutshell.
3. Auflage, OReilly, 2004. ISBN 9780596007645.

Semantic Web:

- ❑ Daconta/Obrst/Smith.
The Semantic Web.
Wiley, 2003. ISBN 0471432571.
- ❑ Antoniou/van Harmelen.
A Semantic Web Primer.
2. Auflage, MIT Press, 2008. ISBN 0262012421.

Weitere Literatur, auf die im World Wide Web direkt zugegriffen werden kann, ist in den Kapiteln angegeben und verlinkt.

Kapitel WT:I

I. Einführung

- ❑ Begriffsklärung
- ❑ Geschichte des Internet
- ❑ Geschichte des World Wide Web
- ❑ Organisation von Internet und World Wide Web
- ❑ Beispiele für Web-basierte Informationssysteme
- ❑ Verteilte Systeme
- ❑ Web-Technologien und Web-Engineering

II. Rechnerkommunikation und Protokolle für Web-Systeme

III. Dokumentsprachen

IV. Client-Technologien

V. Server-Technologien

VI. Architekturen und Middleware-Technologien

VII. Semantic Web

Begriffsklärung

Definition 1 (Informationssystem [\[Wikipedia\]](#))

Ein Informationssystem dient der rechnergestützten Erfassung, Speicherung, Verarbeitung, Pflege, Analyse, Benutzung, Verbreitung, Disposition, Übertragung und Anzeige von Information.

Informationssysteme sind

- ❑ offen
- ❑ dynamisch
- ❑ komplex

Begriffsklärung

Definition 1 (Informationssystem [\[Wikipedia\]](#))

Ein Informationssystem dient der rechnergestützten Erfassung, Speicherung, Verarbeitung, Pflege, Analyse, Benutzung, Verbreitung, Disposition, Übertragung und Anzeige von Information.

Informationssysteme sind

- ❑ offen
Sie interagieren mit der Umwelt und anderen Informationssystemen.
- ❑ dynamisch
Sie verändern sich im Zeitablauf, z.B. durch Lernprozesse.
- ❑ komplex
Sie verknüpfen eine große Anzahl unterschiedlicher Elemente.

Informationssysteme vereinen personelle, organisatorische und technische Komponenten. Stichwort: soziotechnisches System

Begriffsklärung

Definition 2 (Web-basiertes Informationssystem)

Ein Web-basiertes Informationssystem ist ein Informationssystem, das auf Spezifikationen des World Wide Web Consortium [W3C](#) beruht und im World Wide Web genutzt wird.

- (a) Zugriff, Eingabe und Aktualisierung von Informationen unter Verwendung des World Wide Web: Mensch-Computer-Interaktion
- (b) Synchronisation und Austausch mit anderen Informationssystemen über das World Wide Web: Computer-Computer-Interaktion

Begriffsklärung

Definition 2 (Web-basiertes Informationssystem)

Ein Web-basiertes Informationssystem ist ein Informationssystem, das auf Spezifikationen des World Wide Web Consortium [W3C](#) beruht und im World Wide Web genutzt wird.

- (a) Zugriff, Eingabe und Aktualisierung von Informationen unter Verwendung des World Wide Web: Mensch-Computer-Interaktion
- (b) Synchronisation und Austausch mit anderen Informationssystemen über das World Wide Web: Computer-Computer-Interaktion

Definition 3 (World Wide Web, WWW)

Ein Netz von Dokumenten unterschiedlicher Typen im Internet, das durch Hypertext-Verknüpfungen, die in den Dokumenten enthalten sind, verbunden wird.

Begriffsklärung

Definition 2 (Web-basiertes Informationssystem)

Ein Web-basiertes Informationssystem ist ein Informationssystem, das auf Spezifikationen des World Wide Web Consortium [W3C](#) beruht und im World Wide Web genutzt wird.

- (a) Zugriff, Eingabe und Aktualisierung von Informationen unter Verwendung des World Wide Web: Mensch-Computer-Interaktion
- (b) Synchronisation und Austausch mit anderen Informationssystemen über das World Wide Web: Computer-Computer-Interaktion

Definition 3 (World Wide Web, WWW)

Ein Netz von Dokumenten unterschiedlicher Typen im Internet, das durch Hypertext-Verknüpfungen, die in den Dokumenten enthalten sind, verbunden wird.

Definition 4 (Internetwork, Internet [Tanenbaum 2003])

Eine Gruppe miteinander verbundener Rechnernetze. Ein Rechnernetz sind mehrere, mit einer bestimmten Technologie verbundene, autonome Computer.

Bemerkungen:

- ❑ Schwerpunkt dieser Vorlesung ist die technische Realisierung eines Web-basierten Informationssystems.
- ❑ Andere Begriffe, die oft (auch hier in der Vorlesung) synonym verwendet werden:
 - Web-basiertes Informationssystem
 - Web-basiertes System
 - Web-System
 - Web-Anwendung
- ❑ Das World Wide Web ist nur einer von vielen Internet-Diensten.

Geschichte des Internet

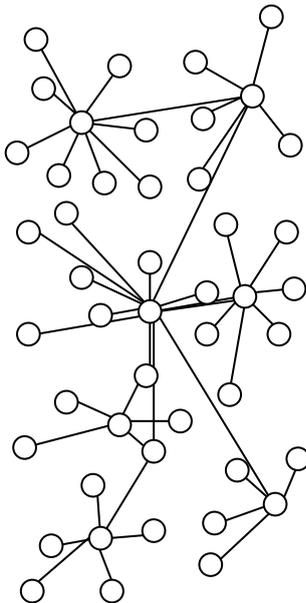
1957 Erfolgreicher Start von Sputnik – die Sowjetunion gewinnt eine Runde im Wettlauf um das All. Eine Reaktion der USA war die Gründung der Advanced Research Project Agency ARPA unter Präsident Eisenhower.

Geschichte des Internet

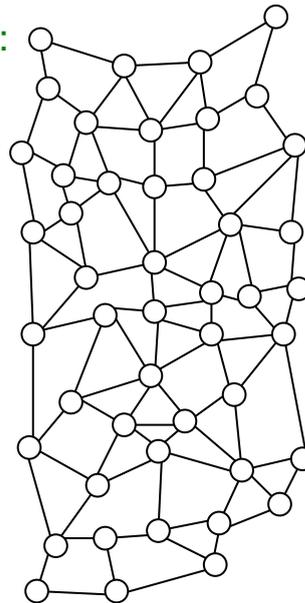
1957 Erfolgreicher Start von Sputnik – die Sowjetunion gewinnt eine Runde im Wetlauf um das All. Eine Reaktion der USA war die Gründung der Advanced Research Project Agency ARPA unter Präsident Eisenhower.

1967 Die ARPA widmet sich der Erneuerung des militärischen Kommunikationsnetzes. Paul Baran's Idee von 1960 wird aufgegriffen.

vorher:



nachher:



[Paul Baran]

1969 Erste Version des ARPANET mit 4 Rechnern (Hosts) ist lauffähig.

Bemerkungen:

- Neu beim ARPANET war vor Allem die Art der Vermittlung:
 - Es sind keine festen Verbindungen vorgegeben → Ausfallsicherheit
 - Nachrichten werden zum Übertragen vom Sender in einzelne Pakete zerlegt.
 - Jedes Paket wird vor seiner Weiterleitung vollständig empfangen.
Stichwort: Speichervermittlungsnetz
 - Empfänger setzt die Pakete wieder zusammen.

Geschichte des Internet

1969 4 Hosts (angebundene Rechner) im ARPANET.

1970 Auf Hawaii entsteht das erstes Funknetz, genannt ALOHANET.

1973 35 Hosts inklusive England und Norwegen.

1975 Die erste Satellitennetzwerkverbindung wird geschaltet.



1983 Es gibt mehr als 500 Hosts. Das Netz wird aufgeteilt in das militärische MILNET und das zivile ARPANET.

1986 Die National Science Foundation NSF baut ein Backbone-Netz, das NSFNET, um ihre 6 Superrechenzentren zu verbinden. Lokale Netze werden hieran angeschlossen. Kommunikationstechnologie ist TCP/IP.

1988 Der erste Internet-Wurm taucht auf und befällt 10% der 60.000 Hosts.

1989 150.000 Hosts. Das alte ARPANET wird abgeschaltet. Das NSFNET ist jetzt das INTERNET.

Geschichte des Internet

1989 world.std.com erster kommerzieller Internet-Anbieter.

1991 „Geburtsstunde“ des World Wide Web.

1992 Die Internet Society wird gegründet; über 1 Million angebundene Rechner.

1993 Das Weiße Haus geht online.

1994 E-Commerce hält Einzug ins WWW.



1995 Der [Vatikan](http://www.vatican.va) geht online.

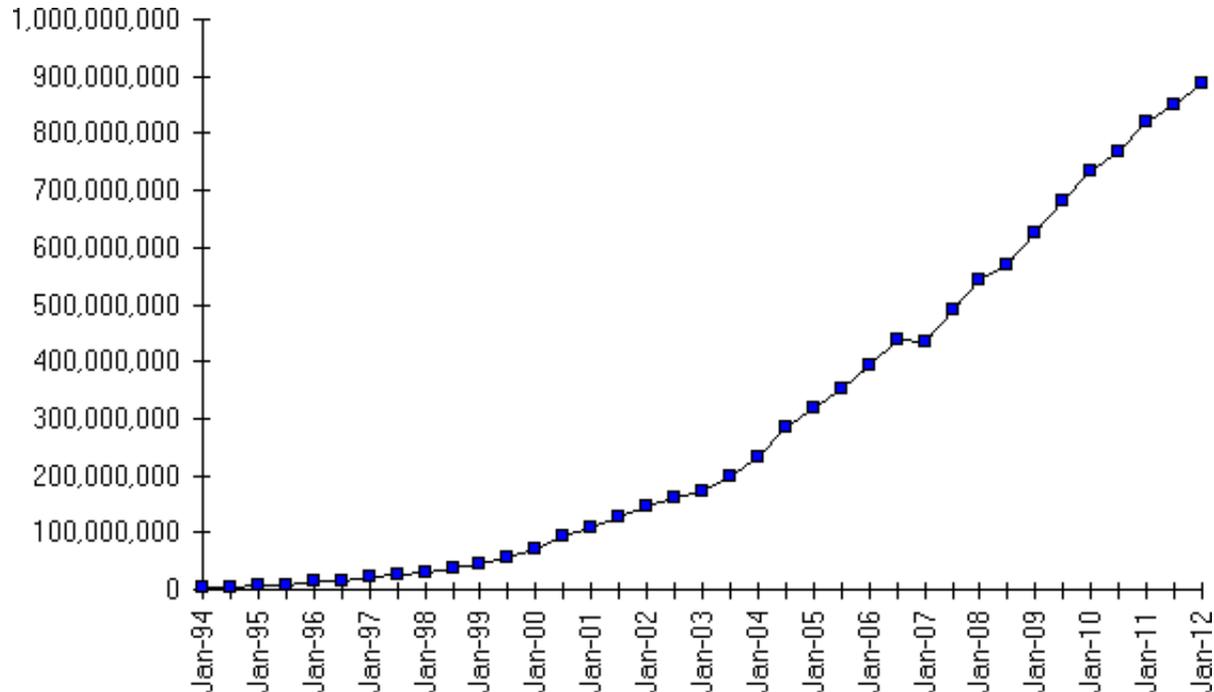
1996 Der Domainname tv.com wird für 15.000 USD verkauft.

1999 In der Auseinandersetzung zwischen Serbien und Kosovo kommt eine neue Art der Kriegsführung zum Einsatz. Stichwort: Cyberwar

2010 732.740.444 Hosts im Januar 2010.

Geschichte des Internet

2012 888.239.420 Hosts (angebundene Rechner) im Januar 2012.



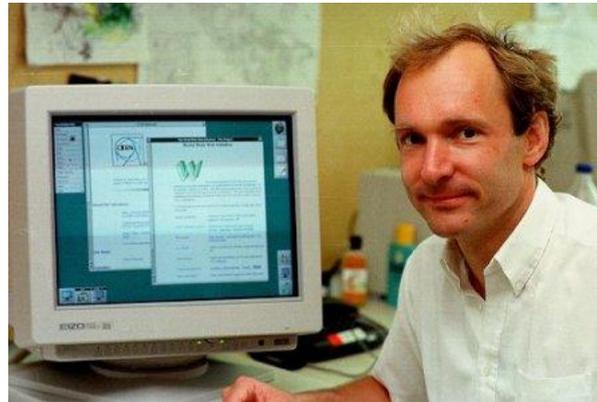
“The Domain Survey attempts to discover every host on the Internet by doing a complete search of the allocated address space and following links to domain names.”

[[Internet Software Consortium, www.isc.org](http://www.isc.org)]

Geschichte des World Wide Web

1945 Vennevar Bush schlägt das Hypertext-System [Memex](#) vor. [\[YouTube\]](#)

1990 Berners-Lee entwickelt einen Web-Client, konzipiert HTML und programmiert einen Web-Server. Er nennt das System „WorldWideWeb“. Das Telefonverzeichnis des [CERN](#) ist die erste Anwendung.



1993 Die Studenten Andreessen und Bina entwickeln den Browser [Mosaic](#).

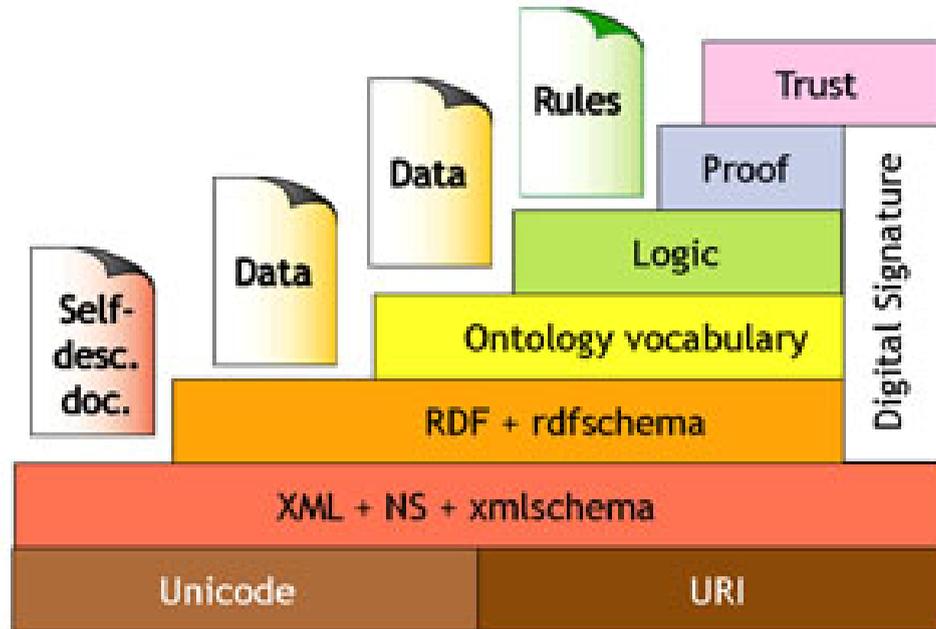
1994 Andreessen und Clark gründen die Firma [Netscape](#).

1994 Gründung des World Wide Web Consortiums [W3C](#).

Geschichte des World Wide Web

1997 Tim Bray stellt die XML-Spezifikation vor.

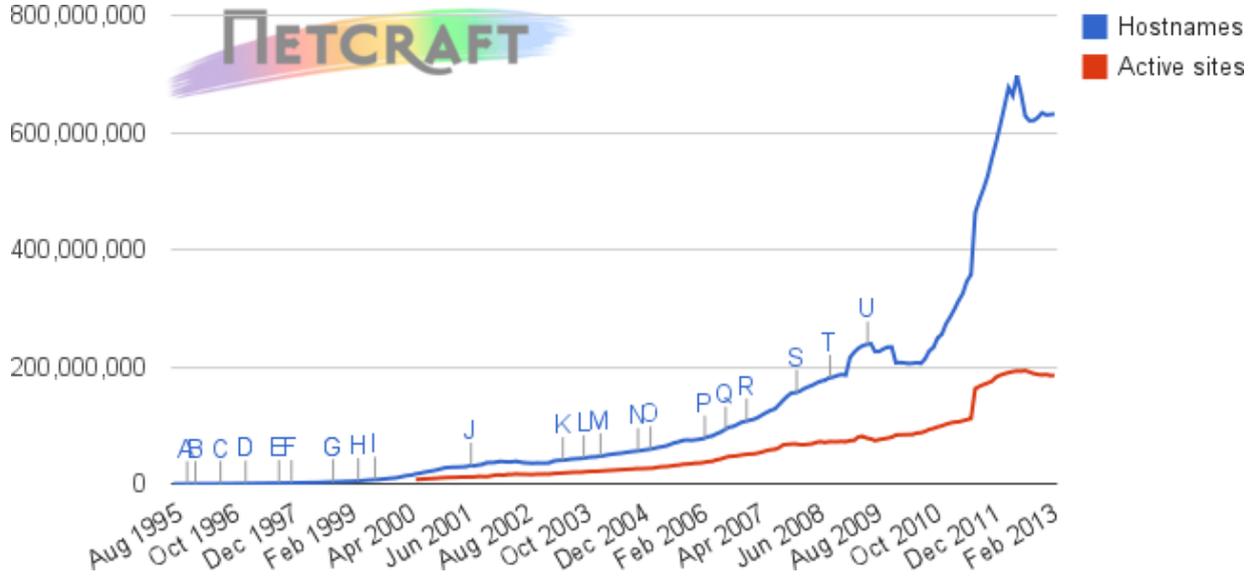
1998 Berners-Lee stellt seine Vision vom Semantischen Web vor: „Allgemeine Plattform für die Zusammenarbeit beliebiger Teilnehmer mit beliebigen Intentionen.“



2001 Standards zur Implementierung eines Semantischen Webs werden entwickelt, z.B. die DARPA Agent Markup Language DAML für Web-basierte Ontologien.

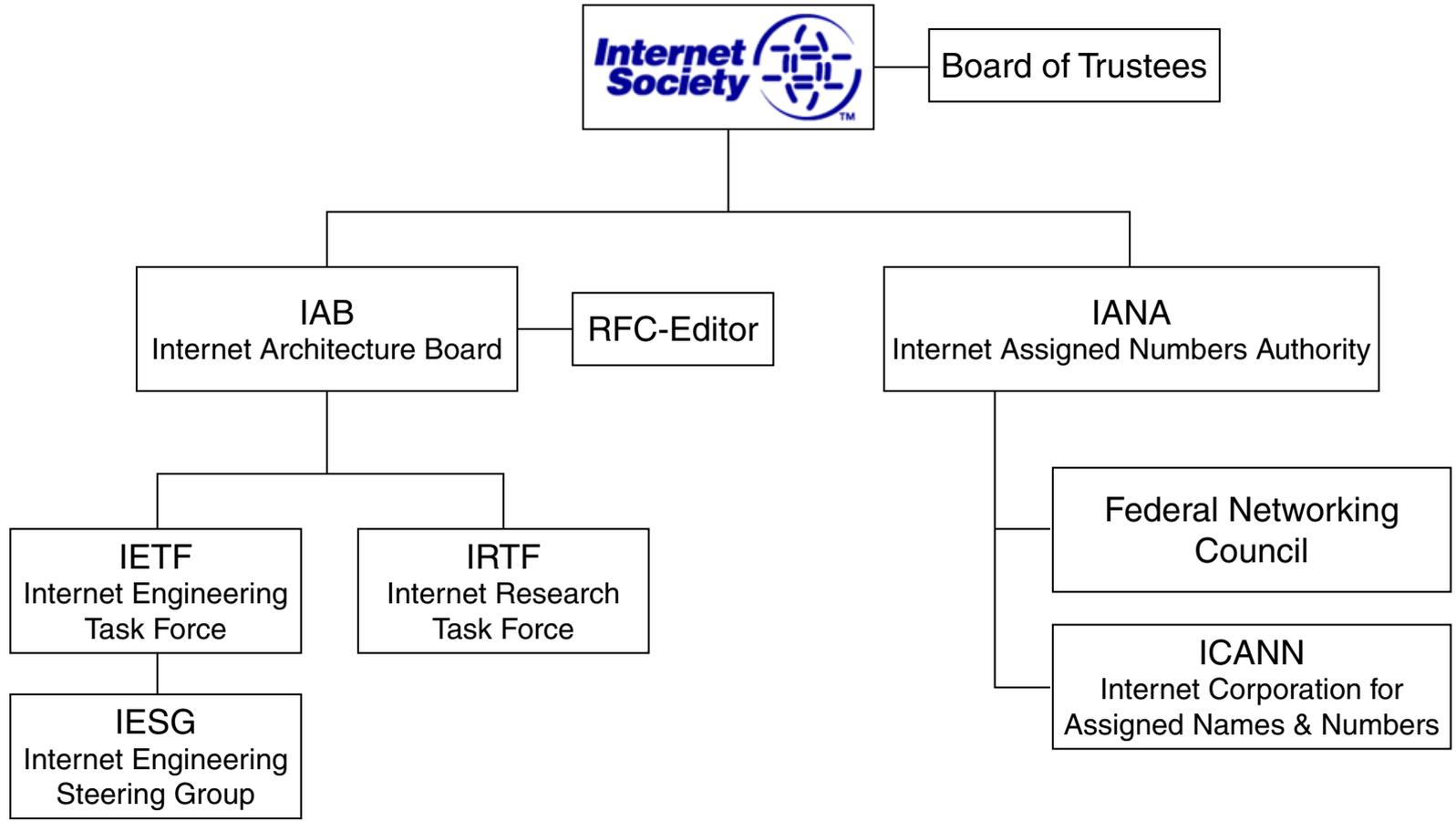
Geschichte des World Wide Web

2013 631.521.198 Websites im Internet / World Wide Web im März 2013.



www.netcraft.com

Organisation von Internet und World Wide Web

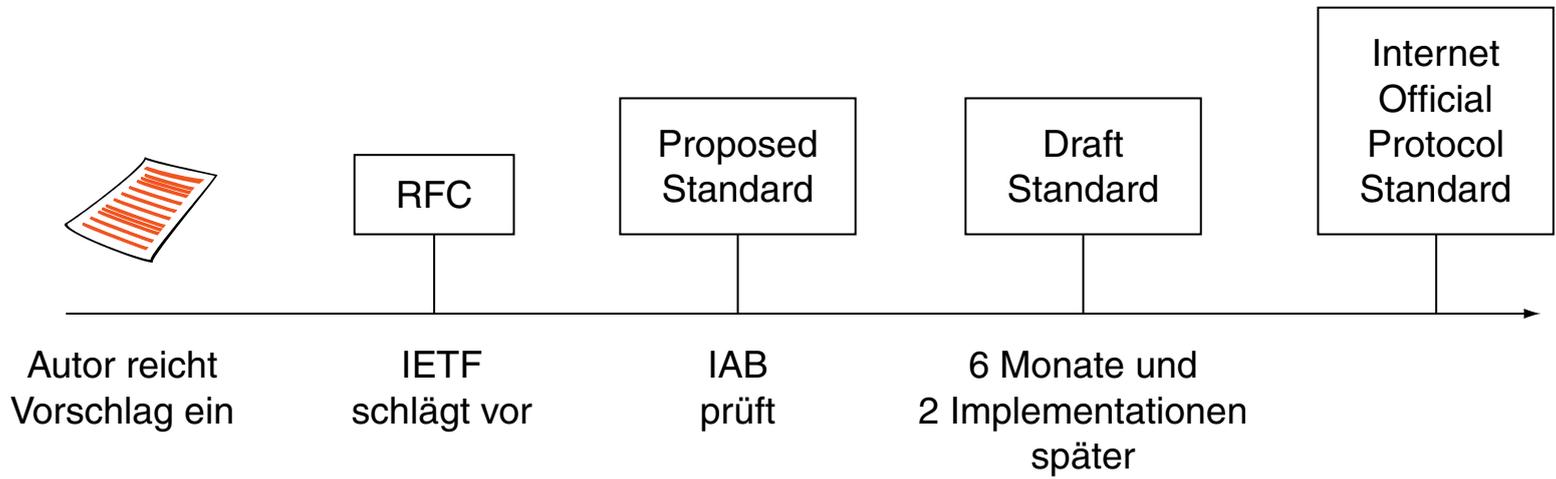


[IANA, IETF]

[Wikipedia: [Internet Society](#), [IAB](#), [IANA](#)]

Organisation von Internet und World Wide Web

Request for Comment RFC



[W3C, Sack 2004]

- ❑ RFCs sind Dokumente des RFC-Editors zum Internet
- ❑ RFCs durchlaufen ein öffentliches Diskussions- und Bewertungsverfahren
- ❑ die RFC-Reihe wurde 1969 begonnen und fortlaufend durchnummeriert
- ❑ alle RFCs sind im Web unter www.ietf.org/rfc.html frei verfügbar

Beispiele für Web-basierte Informationssysteme

Beispiele für Web-basierte Informationssysteme

Nachrichtenportale

- sich häufig verändernde, multimediale Inhalte
- teilweise automatisches Layout
- Kapazität für viele Benutzer gleichzeitig
- große multimediale Datenbank im Hintergrund
- Funktionen wie Suchen in der aktuellen Ausgabe, im Archiv, in Foren
- ...

SPIEGEL ONLINE

► NEWSLETTER | RSS ► ABO-SERVICE
► PDA | i-mode | WAP ► IMPRESSUM SCHLAGZEILEN ►

Montag, 11. April 2005

► HOME
POLITIK
WIRTSCHAFT
PANORAMA
SPORT
KULTUR
NETZWELT
WISSENSCHAFT
UNISPIEGEL
REISE
AUTO
DER SPIEGEL
ENGLISH SITE

SCHLAGZEILEN
WETTER
FORUM
ARCHIV
DOSSIERS
LÄNDERLEXIKON
NEWSLETTER
SHOP
ABO

SPIEGEL TV
XXP
KulturSPIEGEL
manager magazin
SPIEGEL-Gruppe
SCHULE@SPIEGEL

DEUTSCH-RUSSISCHE WIRTSCHAFTSABKOMMEN

Schröder und Putin bekräftigen Zusammenarbeit

Auf der Hannover Messe haben Bundeskanzler Schröder und Russlands Präsident Putin angekündigt, die Wirtschaftsbeziehungen zwischen beiden Ländern auszubauen. BASF und Gazprom vereinbarten, bei der Erschließung von Gasfeldern zu kooperieren. Am Rande des Treffens protestierten Menschenrechtler gegen Putin. [mehr...](#)

- **Arbeitsmarkt:** Parteien schmieden Pakt gegen Billiglöhne
- **Warnstreik:** Ver.di kämpft für die 38-Stunden-Woche

VERKEHRSABGABE

Grüne wollen Maut-Pflicht für Kleinlaster

Hundert Tage sind seit dem Start der Maut für Lastkraftwagen vergangen, nun sprechen sich die Grünen für eine Ausweitung der Abgabepflichten aus. Auch kleinere Fahrzeuge ab 3,5 Tonnen sollten bald für die Benutzung von Autobahnen zahlen müssen. [mehr...](#)

- **Stolpes neue Mautpläne:** "Wir müssen Lkw-Fahrer vergrämen"

SPEIHEL ONLINE SPEZIAL

Einer dieser Männer wird der nächste Papst

Wenn sich die Kardinäle in der Sixtinischen Kapelle versammeln, könnten sie theoretisch jeden katholischen Mann zum Papst wählen. Doch seit 627 Jahren kam der neue Stellvertreter Christi stets aus dem Kardinalskollegium. SPIEGEL ONLINE stellt

Beispiele für Web-basierte Informationssysteme

E-Commerce Plattformen

Ziel: Digitale Abwicklung der Geschäftsprozesse zwischen Unternehmen und zu deren Kunden auf Basis des World Wide Web.

Rollen:

- ❑ B ~ Business
- ❑ C ~ Consumer, auch Citizen
- ❑ A, G ~ Administration, Government

Beispiele für Geschäftsbeziehungen:

- ❑ Business to Business (B2B). Bestellung eines Unternehmens bei Zulieferer
- ❑ Business to Consumer (B2C). Kauf im Online-Shop

Formen des E-Commerce:

- ❑ Web-Shops, elektronische Marktplätze, Supply-Chain-Management, Reservierungs- und Buchungssysteme, Wertpapiergeschäfte

Beispiele für Web-basierte Informationssysteme

Web-Shops

- ❑ Produktpräsentation und Werbung
- ❑ Produktsuche
- ❑ Produktbestellung und Status-Tracking
- ❑ sichere Abwicklung der Bezahlung
- ❑ Produktbewertung, Recommender-Systeme
- ❑ Kundenverwaltung

The screenshot shows the Amazon.de homepage. At the top, there's a navigation bar with the Amazon logo and a shopping cart icon. Below it, a banner for a DVD recorder is displayed with a price of 149 €. A green navigation bar contains categories like HOME, MEIN SHOP, BÜCHER, ENGLISH BOOKS, ELEKTRONIK & FOTO, MUSIK, DVD, VHS, and SOFTWARE. Below this, a green bar features 'ERWEITERTE SUCHE', 'STÖBERN', 'BESTSELLER', 'NEUHEITEN', 'HÖR-BÜCHER', and 'TASCHEN BÜCHER'. A large green banner for 'Grillsaison eröffnet!' is visible. The main content area is titled 'Bücher' and includes a search bar with 'Deutsche Bücher' selected. Below the search bar is a 'STÖBERN' section with a list of book categories. To the right, there are 'Preis-Hits' and 'Aktuelles' sections with promotional text and book covers.

DVD-Rekorder amazon.de

ab 149 €
Hier klicken!

HOME MEIN SHOP **BÜCHER** ENGLISH BOOKS ELEKTRONIK & FOTO MUSIK DVD VHS SOFTWARE

ERWEITERTE SUCHE | STÖBERN | BESTSELLER | NEUHEITEN | HÖR-BÜCHER | TASCHEN BÜCHER

Grillsaison eröffnet! Grills jetzt bis zu 40% gi

SCHNELLSUCHE

Deutsche Bücher

LOS

STÖBERN

- [Antiquarische Bücher](#)
- [Beletristik](#)
- [Biografien & Erinnerungen](#)
- [Börse & Geld](#)
- [Business & Karriere](#)
- [Computer & Internet](#)
- [Erotik](#)
- [Fachbücher](#)
- [Film, Kultur & Comics](#)
- [Geschenkbücher](#)
- [Kinder- & Jugendbücher](#)
- [Kochen & Lifestyle](#)
- [Krimis & Thriller](#)
- [Lernen & Nachschlagen](#)
- [Musiknoten](#)
- [Naturwissenschaften & Technik](#)
- [Politik & Geschichte](#)
- [Ratgeber](#)
- [Reise & Sport](#)
- [Religion & Esoterik](#)
- [Science Fiction, Fantasy &](#)

Bücher

Alle Bücher **kostenlos liefern lassen**

Hallo. Sind Sie Neukunde? [Hier geht's los.](#)

Preis-Hits

- [Business-Bestseller bis zu 50% reduziert!](#)
- [20 Kinderbuch-Klassiker von GEOlino für nur je 4,95 EUR!](#)
- [Jetzt sichern: Bild-Bestseller-Bibliothek für 4,99 EUR pro Band!](#)

Aktuelles

- [Es grünt so grün für Garten & B:](#)
- [Ist Deutschland Aktuelle Büche](#)
- [100 Jahre Rel zum Einstein-](#)

Aktuell in den Medien

- [Bücher & DVDs zu Papst Johannes Paul II.](#)
- [Elke Heidenreich: Lesen!](#)
- [60 Jahre Kriegsende](#)
- [Alles relativ Einstein-Jahr](#)

Romane & Krimis

- [Paulo Coelho über sein neues Buch](#)
- [Neue Romane & Schmöcker](#)
- [Neue Krimis & Thriller](#)
- [Dan Brown](#)

Top

- [Hör](#)
- [30%](#)
- [Gar](#)
- [70%](#)

Beispiele für Web-basierte Informationssysteme

Auktionen

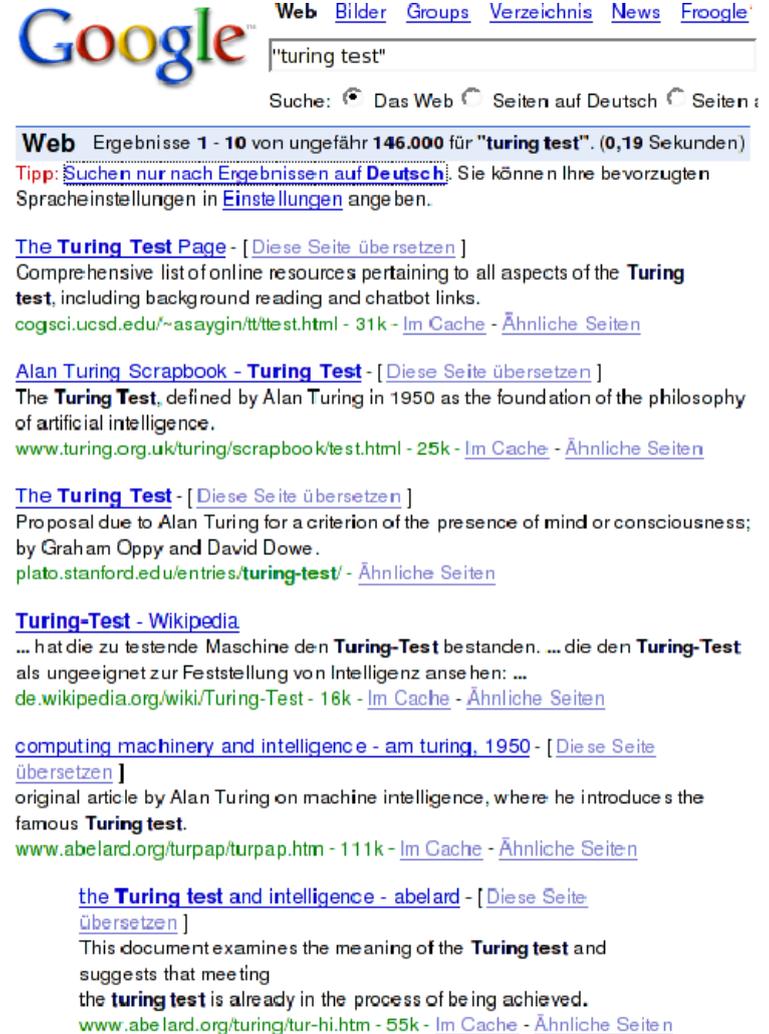
- ❑ sehr große Anzahl von Benutzern, Produkten und Transaktionen
- ❑ im Gegensatz zu Nachrichtenportalen und Web-Shops existieren viele mögliche Kommunikationsflüsse zwischen Benutzern
- ❑ Funktionalität zum Bewerten der Teilnehmer
- ❑ Überprüfung von juristischen Grenzen
- ❑ ständige Erweiterung der Funktionalität

The screenshot displays the eBay website interface. At the top, the eBay logo is prominent on the left, and a navigation menu includes links for 'Startseite', 'Artikel bezahlen', 'Anmelden', 'Service', and 'Übersicht'. Below this, a secondary menu features 'Kaufen', 'Verkaufen', 'Mein eBay', and 'Gemeinschaft'. A user greeting 'Hallo! Einloggen/Ausloggen.' is visible. The main header area contains the text 'Der weltweite Online-Markt' and a search bar with 'Alle Kategorien' as a placeholder. A vertical sidebar on the left lists various categories such as 'Antiquitäten & Kunst', 'Audio & Hi-Fi', 'Auto & Motorrad', 'Baby', 'Beauty & Gesundheit', 'Briefmarken', 'Bücher', 'Büro & Schreibwaren', 'Business & Industrie', 'Computer', 'Feinschmecker', 'Filme & DVDs', 'Foto & Camcorder', 'Handy & Organizer', 'Haushaltsgeräte', 'Heimwerker & Garten', 'Immobilien', 'Kleidung & Accessoires', 'Modellbau', 'Möbel & Wohnen', 'Münzen', 'Musik', 'Musikinstrumente', 'PC- & Videospiele', 'Reise', 'Sammeln & Seltenes', 'Software', 'Spielzeug', and 'Sport'. The main content area features several promotional banners: 'Jetzt Auto verkaufen! Ab sofort bis zu 55% weniger Gebühren*' with an image of a car, 'Traumhafte Dessous' with a list of brands like Donna Karan/DKNY and Victoria's Secret, and 'Trendige Strumphosen' with brands like Vive Maria & Pussy Deluxe and Calvin Klein. At the bottom, there are small images of a Playboy bunny and a computer monitor.

Beispiele für Web-basierte Informationssysteme

Suchmaschinen

- ❑ automatisches Sammeln von Inhalten (indizieren)
- ❑ automatische und manuelle Aufbereitung der Inhalte
- ❑ neben textbasierter Indizierung auch Suche nach Bildern, Musik und Videos
- ❑ extrem hohe Last durch Millionen von Anfragen täglich
- ❑ hohe Anforderungen bzgl. Ausfallsicherheit



Google Web Bilder Groups Verzeichnis News Froogle

"turing test"

Suche: Das Web Seiten auf Deutsch Seiten

Web Ergebnisse 1 - 10 von ungefähr 146.000 für "turing test". (0,19 Sekunden)

Tipp: Suchen nur nach Ergebnissen auf Deutsch: Sie können Ihre bevorzugten Spracheinstellungen in [Einstellungen](#) angeben.

[The Turing Test Page](#) - [[Diese Seite übersetzen](#)]
Comprehensive list of online resources pertaining to all aspects of the **Turing test**, including background reading and chatbot links.
cogsci.ucsd.edu/~asaygin/tt/test.html - 31k - [Im Cache](#) - [Ähnliche Seiten](#)

[Alan Turing Scrapbook - Turing Test](#) - [[Diese Seite übersetzen](#)]
The **Turing Test**, defined by Alan Turing in 1950 as the foundation of the philosophy of artificial intelligence.
www.turing.org.uk/turing/scrapbook/test.html - 25k - [Im Cache](#) - [Ähnliche Seiten](#)

[The Turing Test](#) - [[Diese Seite übersetzen](#)]
Proposal due to Alan Turing for a criterion of the presence of mind or consciousness; by Graham Oppy and David Dowe.
plato.stanford.edu/entries/turing-test/ - [Ähnliche Seiten](#)

[Turing-Test](#) - Wikipedia
... hat die zu testende Maschine den **Turing-Test** bestanden. ... die den **Turing-Test** als ungeeignet zur Feststellung von Intelligenz ansehen: ...
de.wikipedia.org/wiki/Turing-Test - 16k - [Im Cache](#) - [Ähnliche Seiten](#)

[computing machinery and intelligence - am turing, 1950](#) - [[Diese Seite übersetzen](#)]
original article by Alan Turing on machine intelligence, where he introduces the famous **Turing test**.
www.abelard.org/turpap/turpap.htm - 111k - [Im Cache](#) - [Ähnliche Seiten](#)

[the Turing test and intelligence - abelard](#) - [[Diese Seite übersetzen](#)]
This document examines the meaning of the **Turing test** and suggests that meeting the **turing test** is already in the process of being achieved.
www.abelard.org/turing/tur-hi.htm - 55k - [Im Cache](#) - [Ähnliche Seiten](#)

Bemerkungen:

- ❑ Technologien und Forschungsaspekte für die Suche in großen Dokumentkollektionen:
 - focused Crawl: Finden relevanter Seiten durch intelligente Suchbegriffexpansion und heuristische Link-Auswahl
 - effizientes Parsen: Trennung von Text und Auszeichnungssprache
 - Natural Language Processing: Stammwortbildung, Stopworteliminierung, Part-of-Speech-Analyse
 - Ähnlichkeitsanalyse: Klassifikation relevanter und nicht-relevanter Dokumente
 - Personalisierung: Erkennen und Verwalten von Benutzervorlieben
 - Ranking-Verfahren
 - Informationsvisualisierung: Darstellung der Suchergebnisse
 - Social Media Mining
 - Adversarial Information Retrieval AIR

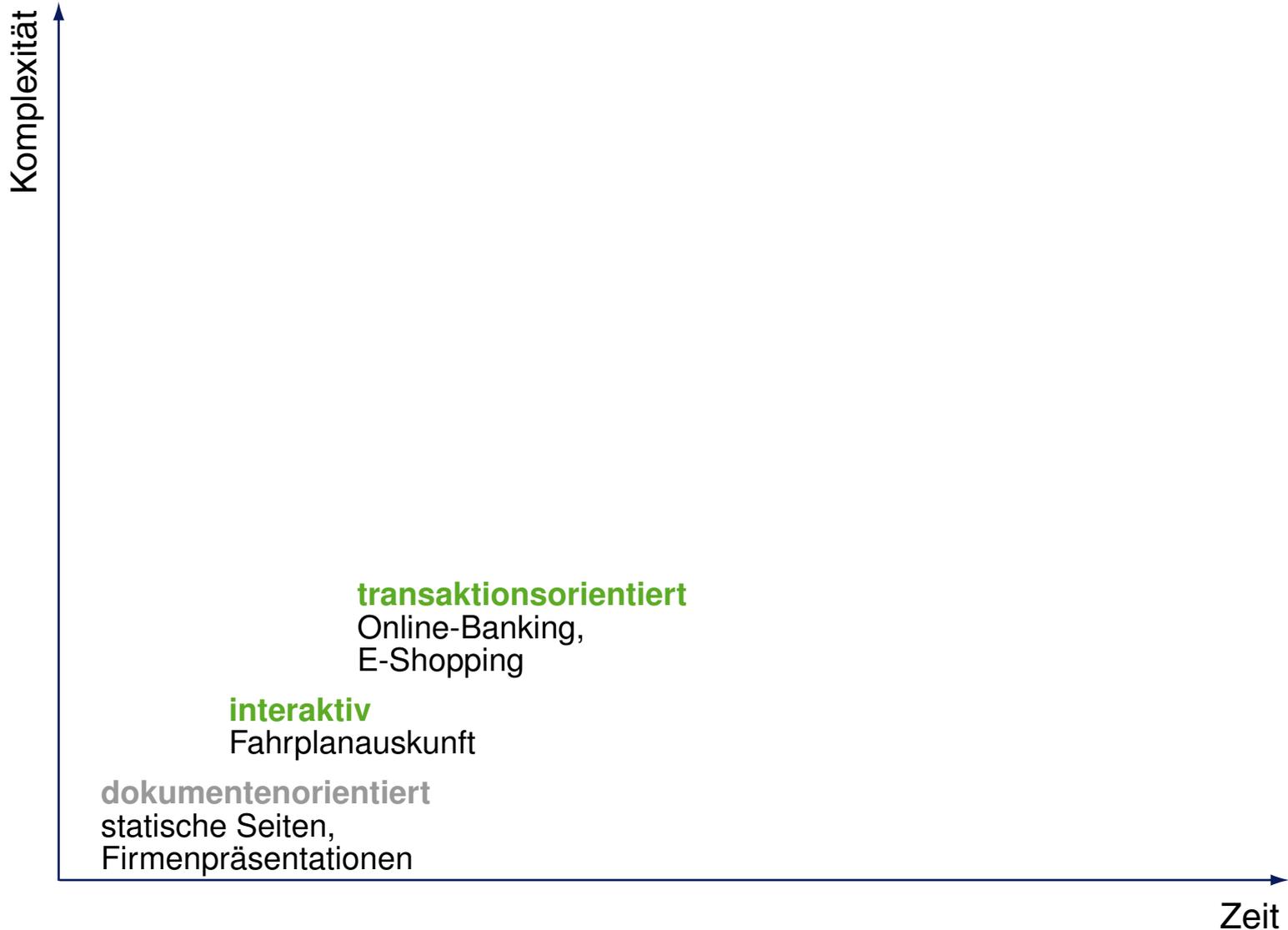
- ❑ Unsere Arbeitsgruppe [webis](#) betreibt Forschung in diesem Bereich.

Evolution Web-basierter Informationssysteme



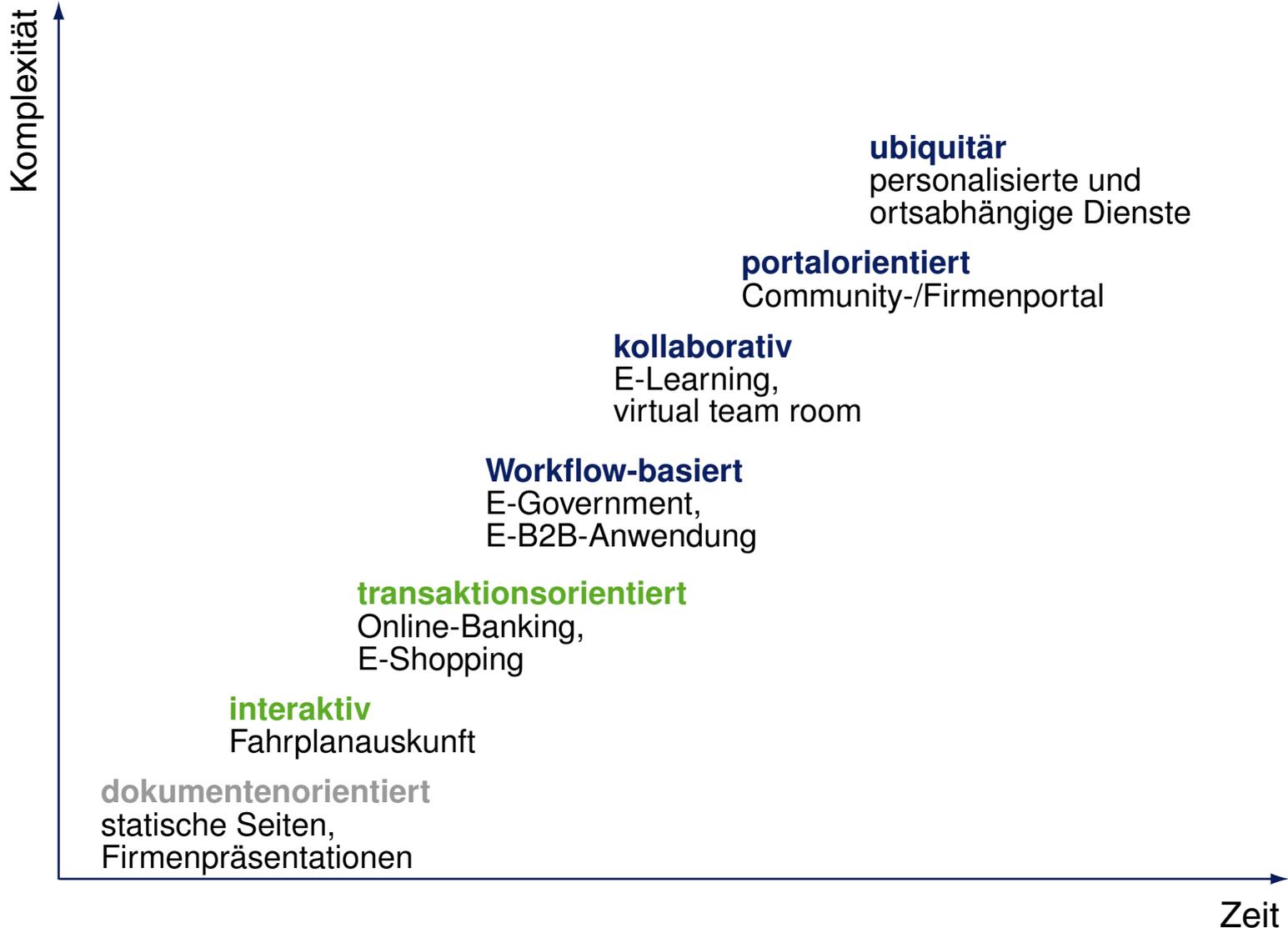
[vgl. Koch 2004]

Evolution Web-basierter Informationssysteme



[vgl. Koch 2004]

Evolution Web-basierter Informationssysteme



[vgl. Koch 2004]

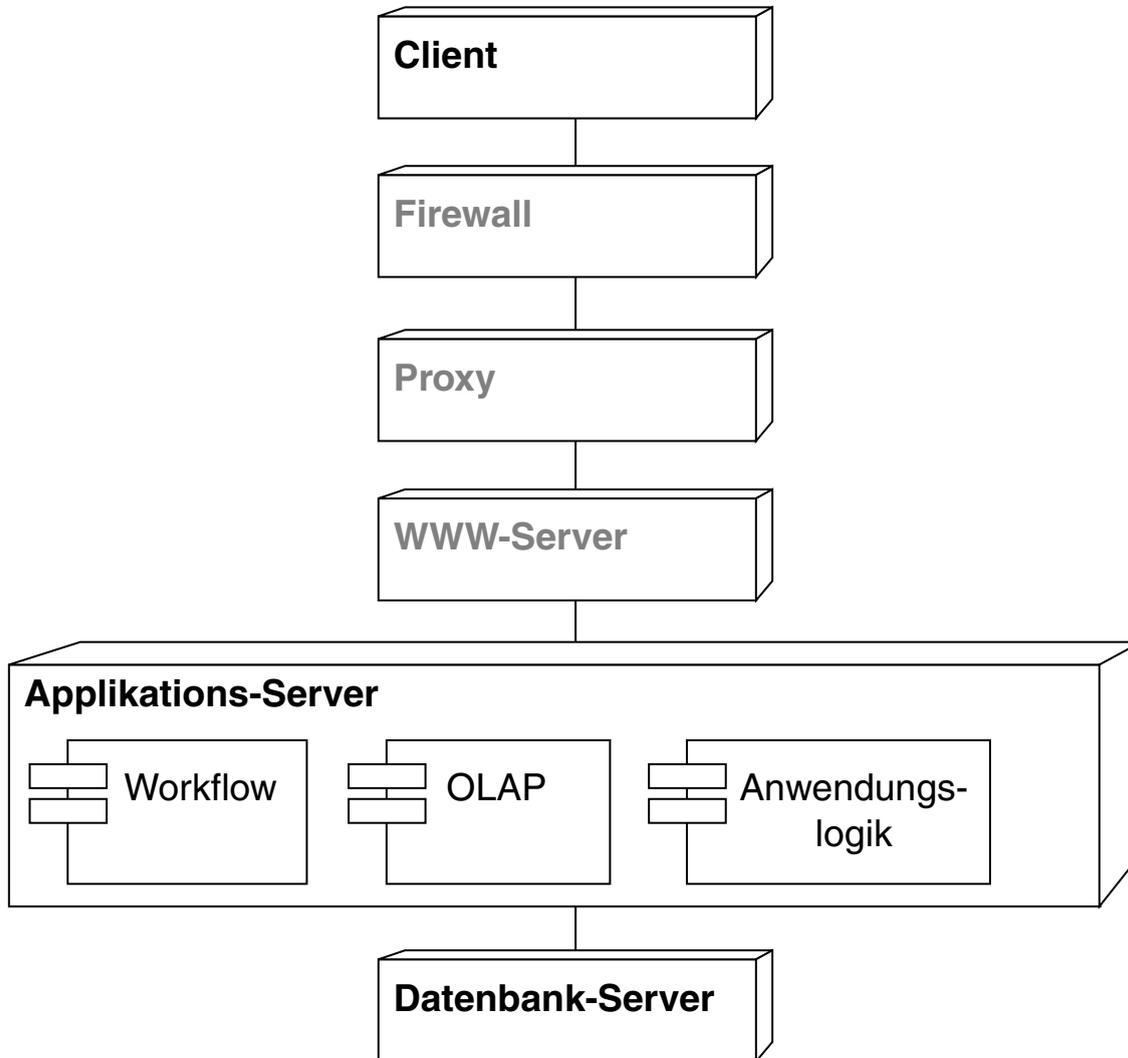
Evolution Web-basierter Informationssysteme



[vgl. Koch 2004]

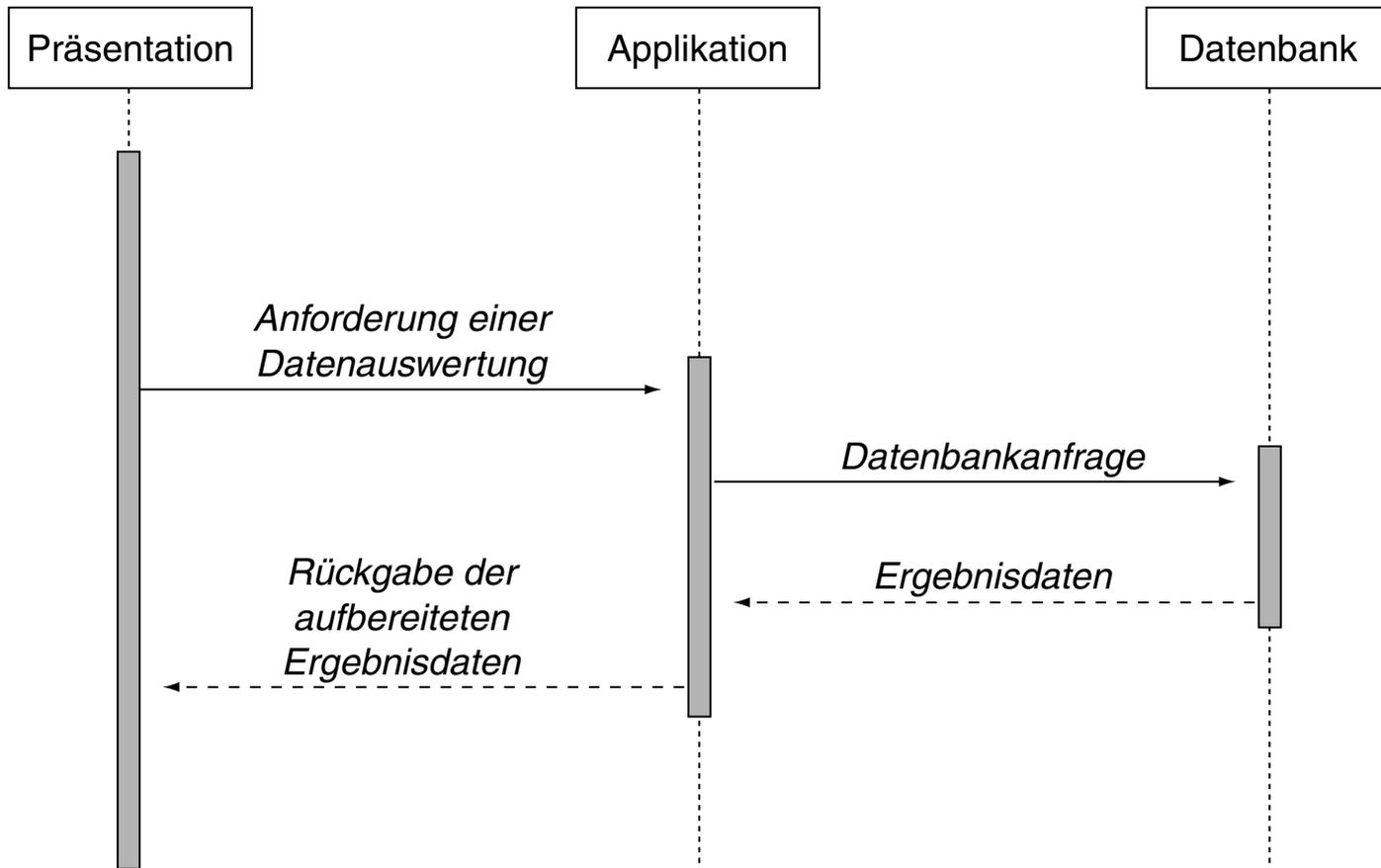
Verteilte Systeme

Web-Systeme sind verteilte Systeme



Verteilte Systeme

Web-Systeme sind verteilte Systeme



Aufbau von Web-Systemen oft als 3-Tier- (allgemein: n-Tier-) Architektur.

Verteilte Systeme

Definition 5 (Verteiltes System [Coulouris 2001])

Ein System, bei dem sich die Hardware- und Softwarekomponenten auf vernetzten Rechnern befinden und nur über den Austausch von Nachrichten kommunizieren und ihre Aktionen koordinieren. Dabei nimmt der Benutzer nur *eine* Ressource wahr.

Aber auch:

Ein verteiltes System ist ein System, mit dem man nicht arbeiten kann, weil irgendein Rechner abgestürzt ist, von dem man nicht einmal weiß, dass es ihn überhaupt gibt... ; -)

[Lamport]

Verteilte Systeme

Definition 5 (Verteiltes System [Coulouris 2001])

Ein System, bei dem sich die Hardware- und Softwarekomponenten auf vernetzten Rechnern befinden und nur über den Austausch von Nachrichten kommunizieren und ihre Aktionen koordinieren. Dabei nimmt der Benutzer nur *eine* Ressource wahr.

Aber auch:

Ein verteiltes System ist ein System, mit dem man nicht arbeiten kann, weil irgendein Rechner abgestürzt ist, von dem man nicht einmal weiß, dass es ihn überhaupt gibt... ; -)

[Lampport]

Nutzen verteilter Systeme:

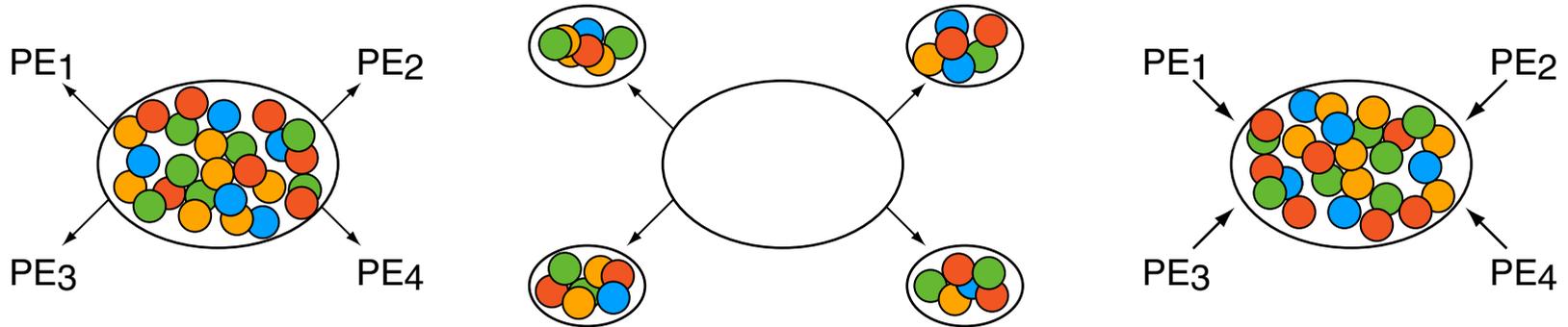
- ❑ Teilung von Ressourcen (Drucker, Festplatten, Datenbanken, Web-Dienste)
- ❑ ortsunabhängiger Zugang zu Ressourcen
- ❑ Ausfallsicherheit durch Redundanz
- ❑ Beschleunigung der Verarbeitung

Verteilte Systeme

Beschleunigung der Verarbeitung

Prinzip (vgl. [Map-Reduce](#), [Hadoop](#)):

1. Aufteilung der zu verarbeitenden Daten in (disjunkte) Teilmengen
2. Verteilung der Teilmengen über mehrere Rechner PE
3. Simultane Verarbeitung und Zusammenfassung der Teilergebnisse

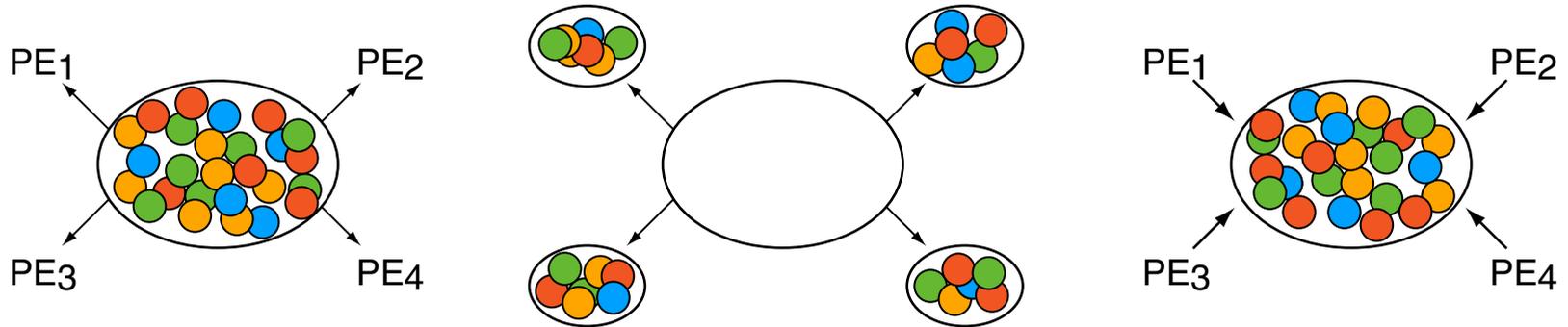


Verteilte Systeme

Beschleunigung der Verarbeitung

Prinzip (vgl. [Map-Reduce](#), [Hadoop](#)):

1. Aufteilung der zu verarbeitenden Daten in (disjunkte) Teilmengen
2. Verteilung der Teilmengen über mehrere Rechner PE
3. Simultane Verarbeitung und Zusammenfassung der Teilergebnisse



SETI@home [\[Wikipedia\]](#) :



Verteilte Systeme

Herausforderung Heterogenität

- ❑ gemeinsame, standardisierte und offene Netzprotokolle
- ❑ austauschbare, Hardware-unabhängige Formate für Daten
- ❑ Standards zum Austausch von Nachrichten, Datenbankabfragen, etc.
- ❑ **Middleware**
Softwareschicht, die eine Programmierabstraktion bereitstellt und die Heterogenität darunter liegender Komponenten verbirgt. Beispiele: CORBA, Web-Services, MOM
- ❑ **virtuelle Maschinen** [[Turingmaschine](#)]
Compiler erzeugt Code für eine „Software-Maschine“, nicht für die Zielhardware. Beispiel: [Java Virtual Machine](#), [Native Client](#)

Verteilte Systeme

Herausforderung Skalierbarkeit

- effizientes und effektives Arbeiten auch bei steigender Anzahl von Benutzern und Komponenten

	Hosts	Web-Sites / Server	Anteil
1979	188	0	0%
1989	130.000	0	0%
1999	56.318.000	6.598.697	12%
2005	317.646.084	62.286.451	19%
2009	625.226.456	185.497.213	30%
2011	818.374.269	273.301.445	33%
2012	888.239.420	555.482.744	63%

- automatische Anpassung an erhöhte Last
- bestellbare Rechen- und Speicherleistung für bestimmte Aufgaben
Stichworte: Grid-Computing, Cloud-Computing
- Erweiterung muss zu vernünftigen Kosten möglich sein.
- Ressourcen sollen zukünftige Erweiterungen berücksichtigen.
Aktuell: Umstellung von 32-Bit Internet-Adressen auf 128-Bit

Verteilte Systeme

Herausforderung Sicherheit

Datenübertragung über öffentliche Netze sowie ein möglicher Zugang von außen stellen hohe Anforderungen an die Sicherheit.

- ❑ Vertraulichkeit
- ❑ Integrität
- ❑ Verfügbarkeit
- ❑ Sicherheit mobilen Codes
- ❑ Denial-of-Service-Angriffe

Verteilte Systeme

Herausforderung Sicherheit

Datenübertragung über öffentliche Netze sowie ein möglicher Zugang von außen stellen hohe Anforderungen an die Sicherheit.

- ❑ **Vertraulichkeit**

Schutz der Ressourcen gegenüber nicht-berechtigten Personen

- ❑ **Integrität**

Schutz gegen Manipulation oder Beschädigung

- ❑ **Verfügbarkeit**

Reaktion auf Störungen und Überlastung durch Sonderereignisse.

Beispiel: Überlastung von Bankrechnern in turbulenten Börsenzeiten

- ❑ **Sicherheit mobilen Codes**

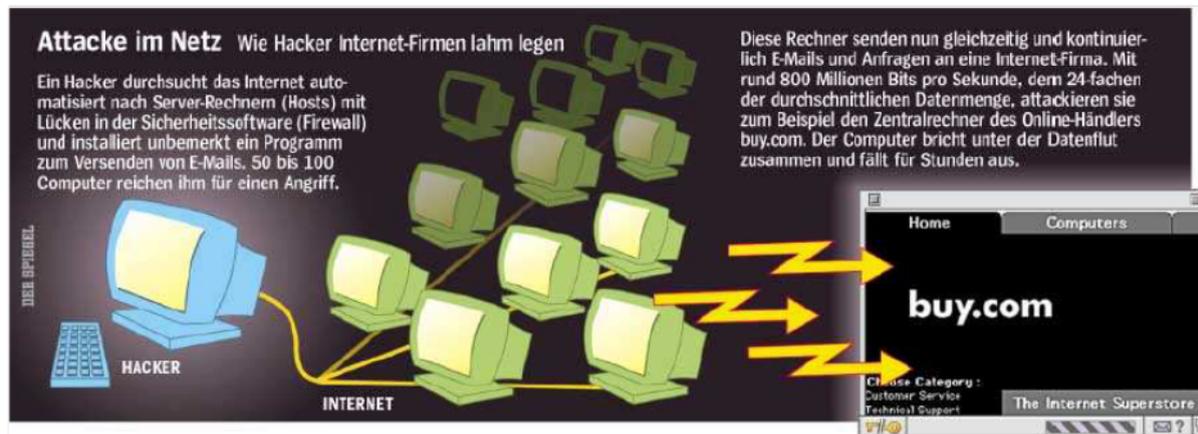
Wie erkennt man, ob ein mitgeliefertes Skript einen Virus enthält?

- ❑ **Denial-of-Service-Angriffe**

Ein Server wird mit sinnlosen Anfragen überflutet und ist für ernsthafte Anfragen nicht verfügbar. Beispiel: Root-(Name)Server-Angriff

Verteilte Systeme

Herausforderung Sicherheit (Fortsetzung)



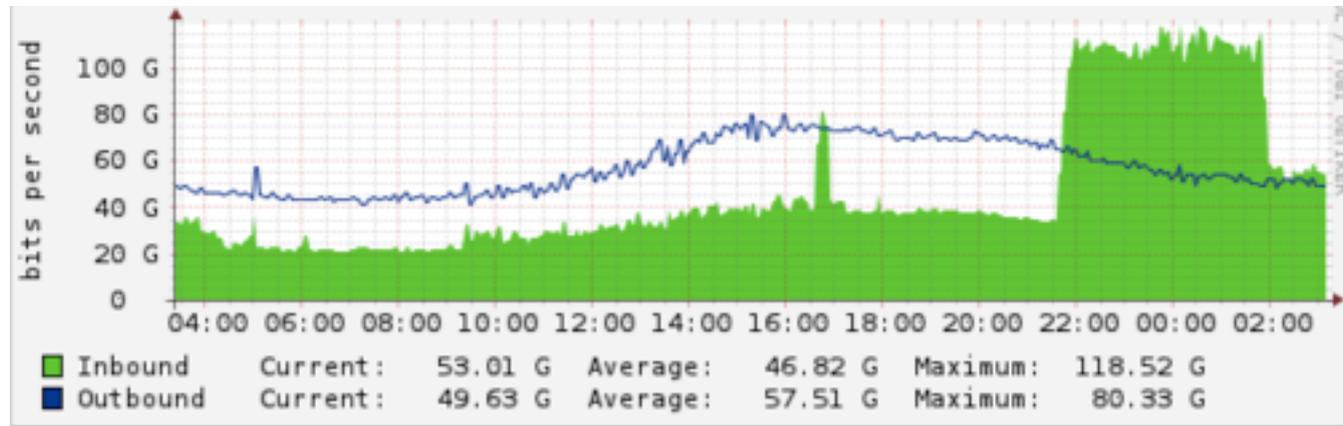
„Die größte und höchst entwickelte Attacke, die das Internet je erlebte [bis 19.03.2013], begann am Montag [21.10.2002] um 10 Uhr. Sie fegte neun der 13 zentralen Server zeitweilig aus dem Netz – und kein Mensch bemerkte es.“

[Spiegel-Online 23.10.2002]

- ❑ Situation: 13 Root-(Name-)Server (10 in USA) verteilen DNS-Tabellen.
- ❑ Dank Caching war der Angriff von 1 Stunde mit 2 Millionen Anfragen/s nahezu wirkungslos.

Verteilte Systeme

Herausforderung Sicherheit (Fortsetzung)



„Nahezu unbemerkt von der Öffentlichkeit ereignete sich in der vergangenen Woche [19.03.2013] die bislang heftigste Distributed-DoS-Attacke in der Geschichte des Internet. [...] Ziel des Angriffs war die Antispam-Organisation Spamhaus.“

[www.heise.de]

- ❑ Situation: Spamhaus pflegt „Blacklists“ von Spam-Versendern.
- ❑ Der Angriff erzeugt eine Datenflut mit Spitzen von 300 GBit/s. Kurz nach Beginn der Attacke bat Spamhaus das Security-Unternehmen Cloudflare um Abwehrunterstützung.

Verteilte Systeme

Weitere Herausforderungen

- ❑ Koordination und Synchronisation von Komponenten
- ❑ Transparenz
 - Verbergen der räumlichen Trennung der einzelnen Komponenten im verteilten System vor Benutzern/Anwendungen, das System wird als eine Einheit wahrgenommen.
 - Zugriffstransparenz
 - Positionstransparenz
 - Mobilitätstransparenz
- ❑ Diagnose
 - Fehler und Ausfälle in Rechnerknoten, Verbindungsstrukturen oder der Kommunikation sind wahrscheinlich.

Verteilte Systeme

Weitere Herausforderungen

- ❑ Koordination und Synchronisation von Komponenten

- ❑ Transparenz

Verbergen der räumlichen Trennung der einzelnen Komponenten im verteilten System vor Benutzern/Anwendungen, das System wird als eine Einheit wahrgenommen.

- Zugriffstransparenz

Beispiel: identische Zugriffsoperationen für lokalen oder Netzwerkdrucker

- Positionstransparenz

Beispiel: keine exakte Kenntnis der Druckerposition

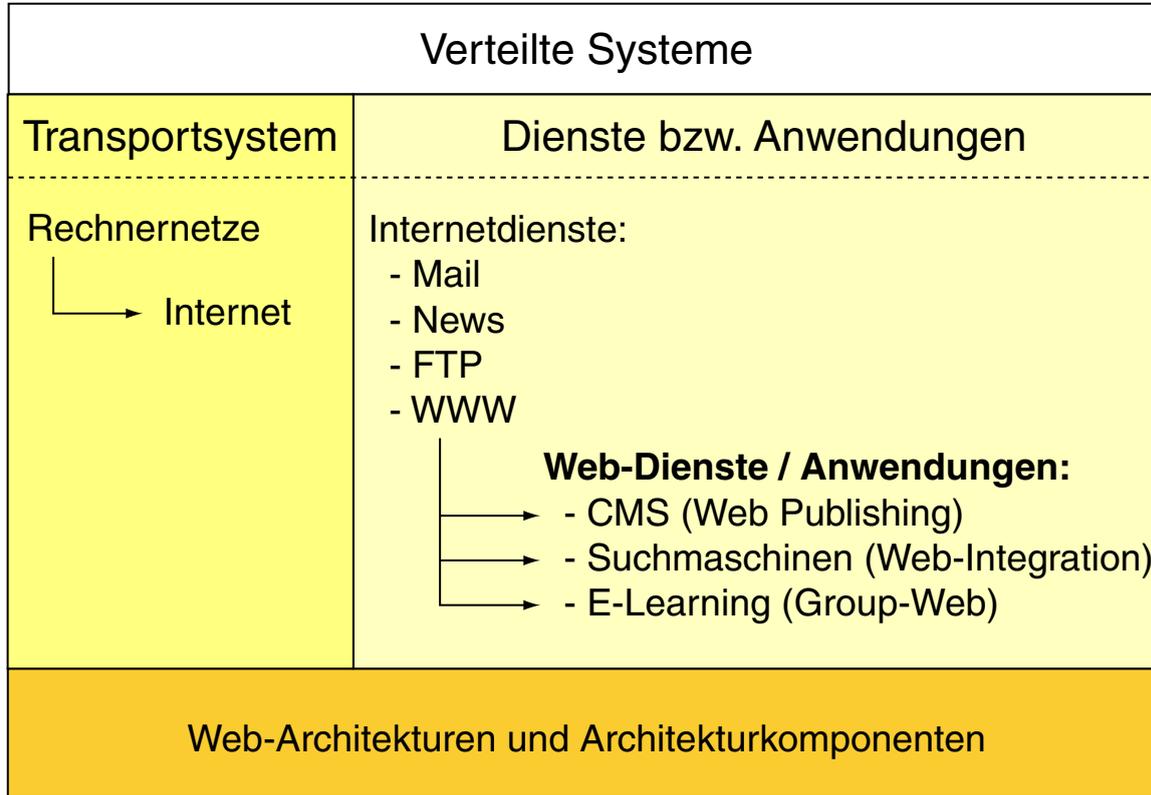
- Mobilitätstransparenz

Beispiel: Verschiebung eines Handygesprächs zwischen Zonen

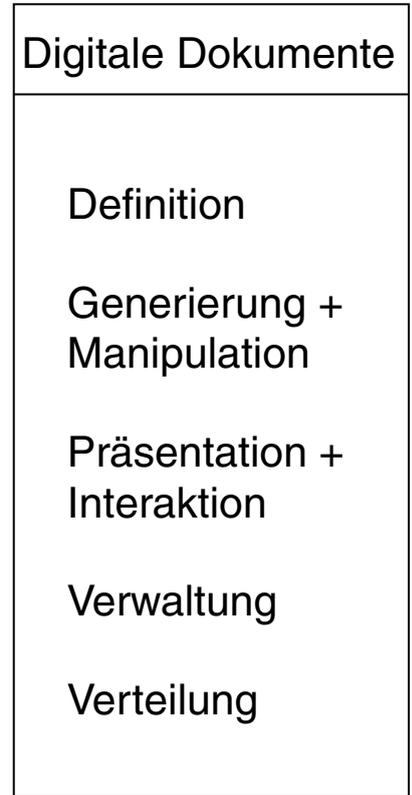
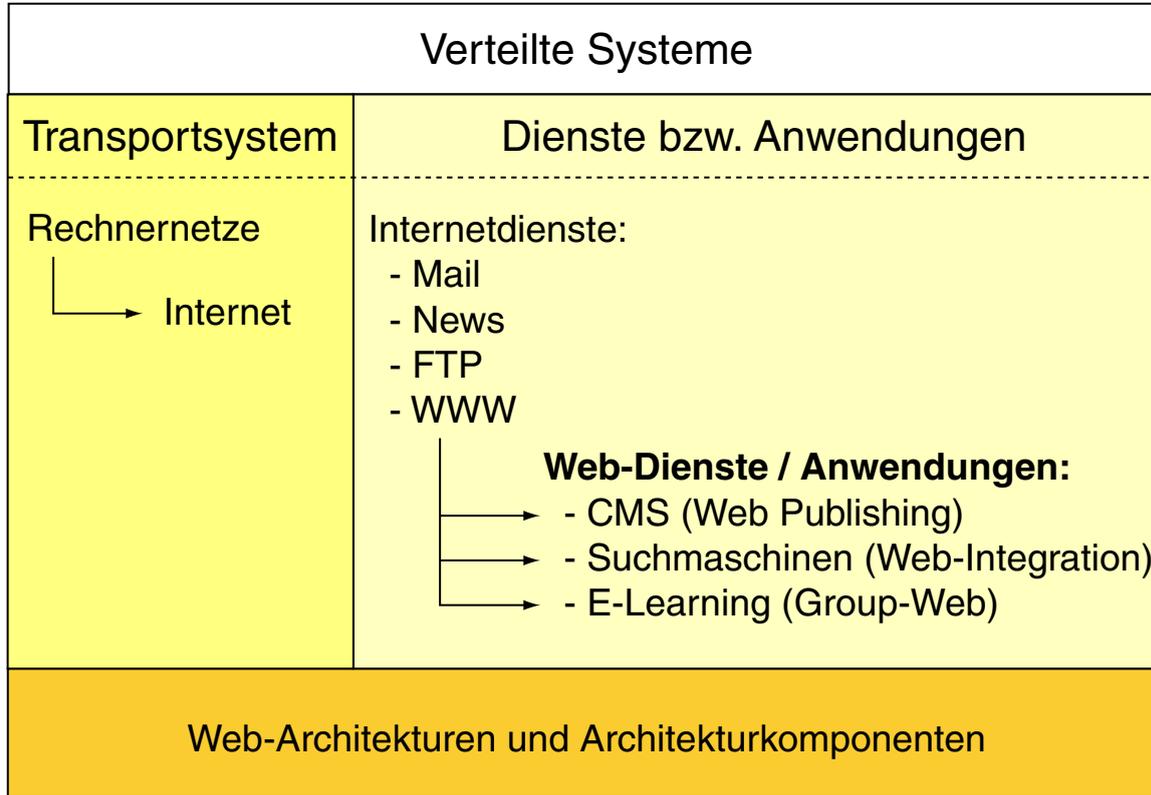
- ❑ Diagnose

Fehler und Ausfälle in Rechnerknoten, Verbindungsstrukturen oder der Kommunikation sind wahrscheinlich.

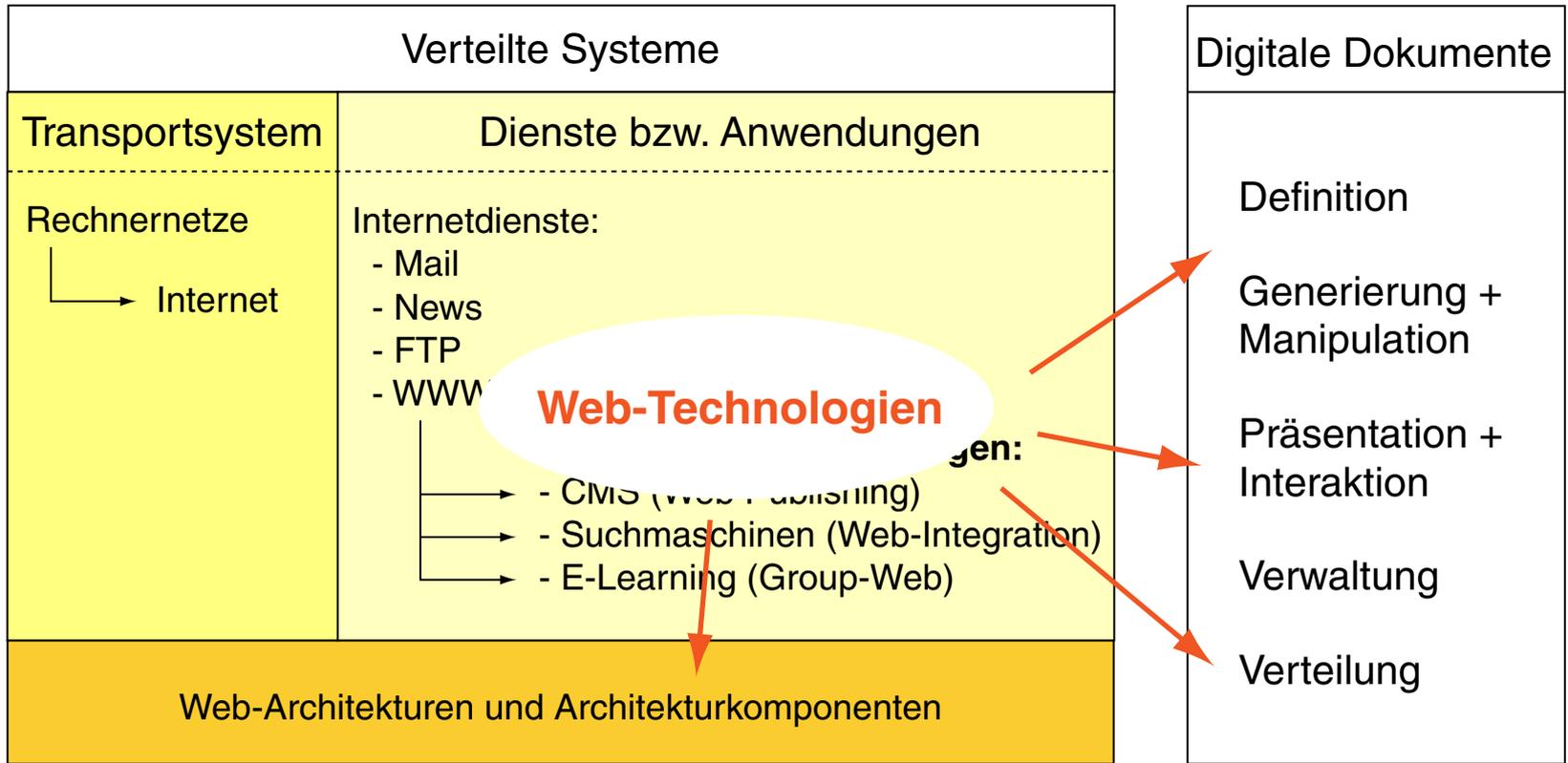
Web-Technologien und Web-Engineering



Web-Technologien und Web-Engineering



Web-Technologien und Web-Engineering



Definition 6 (Web-Technologien) [Dumke 2003]

Web-Technologien sind implementierte Methoden und Verfahren, die für die Entwicklung und Anwendung von Systemen, die im World Wide Web genutzt werden, die Grundlage bilden.

Web-Technologien und Web-Engineering

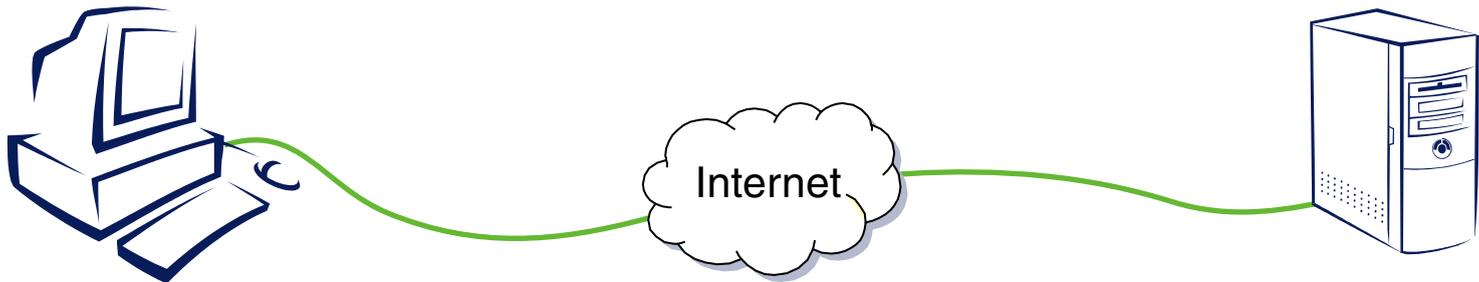
Aufteilung von Web-Technologien

II. Rechnerkommunikation und Protokolle für Web-Systeme

- Rechnernetze
- Netzsoftware und Kommunikationsprotokolle
- Client-Server-Interaktionsmodell
- Hypertext-Transfer-Protokoll HTTP

Sicherheitstechnologien

- symmetrische und asymmetrische Verschlüsselung
- Public Key Infrastruktur
- digitale Signaturen
- SSL und TLS

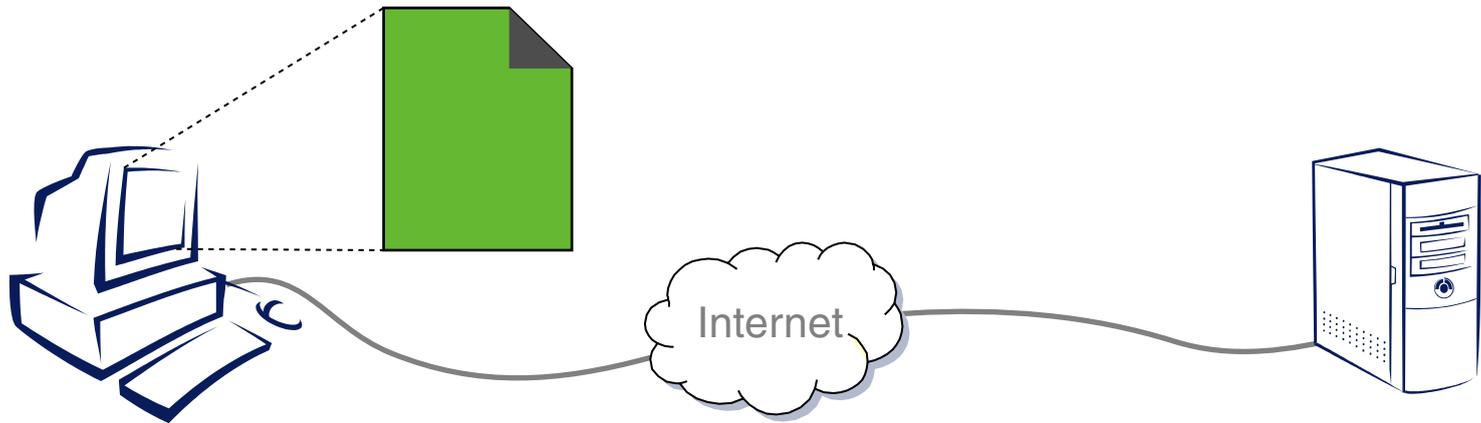


Web-Technologien und Web-Engineering

Aufteilung von Web-Technologien (Fortsetzung)

III. Dokumentsprachen

- HTML, Cascading Stylesheets CSS
- XML-Grundlagen: Syntax, DTDs, Namensräume
- XML-Schema
- Die XSL-Familie: XPath, XSLT
- XML-Erweiterungen: XLink, XPointer, XQuery
- Document Object Model DOM, die Parser DOM und SAX



Web-Technologien und Web-Engineering

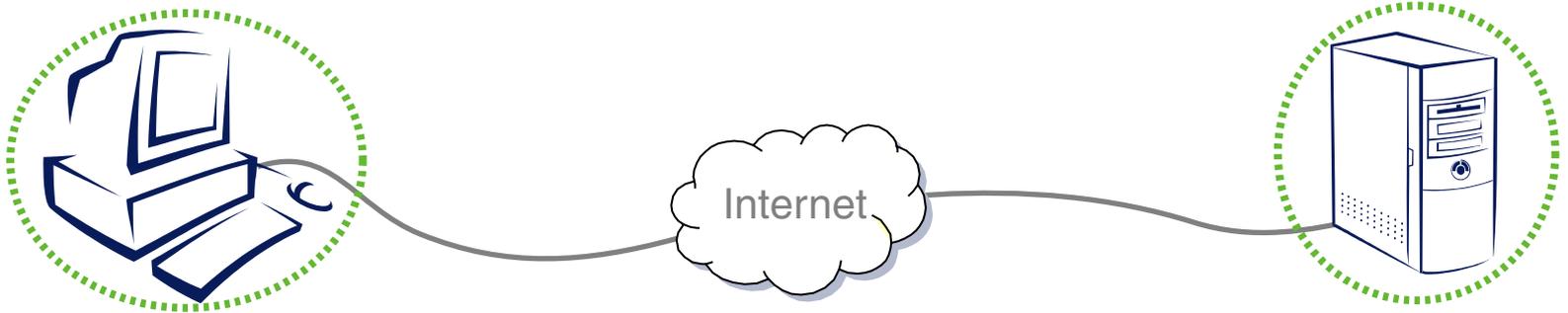
Aufteilung von Web-Technologien (Fortsetzung)

IV. Client-Technologien

- Skriptsprachen: JavaScript, VBScript
- Java Applets

V. Server-Technologien

- Common Gateway Interface CGI
- PHP Hypertext Processor PHP
- Perl, Python, Ruby
- Java-Servlets, Java-Server-Pages JSP

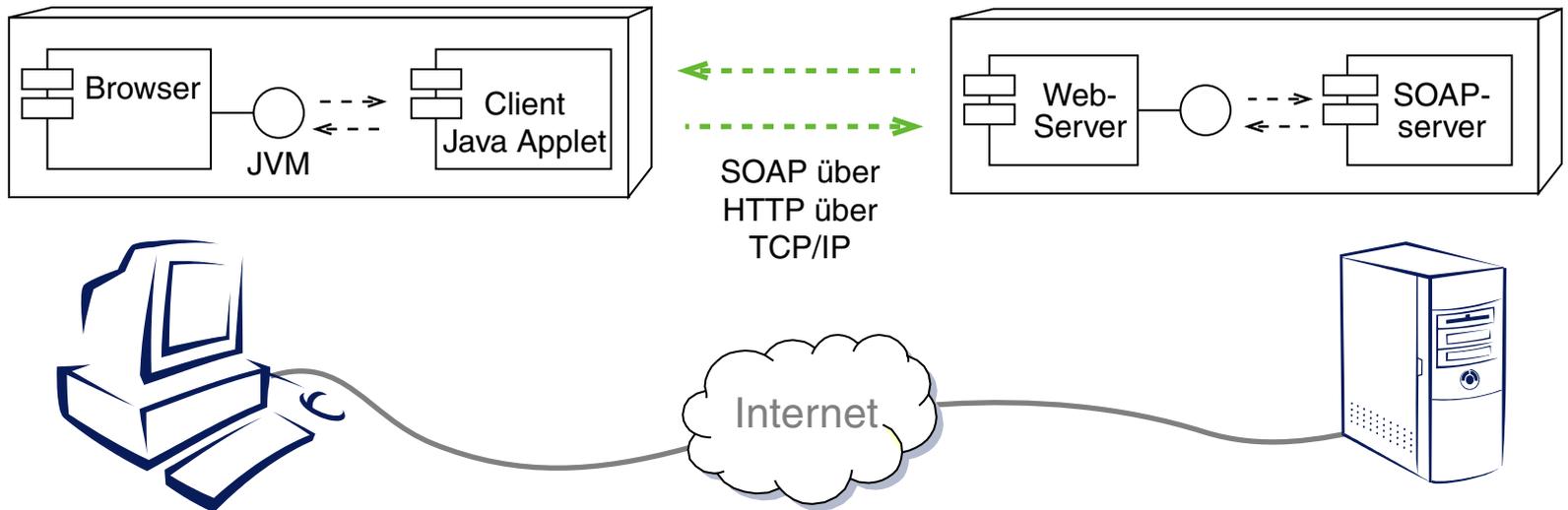


Web-Technologien und Web-Engineering

Aufteilung von Web-Technologien (Fortsetzung)

VI. Architekturen und Middleware-Technologien

- (Muster für) Client-Server-Architekturen
- RPC und RMI
- CORBA
- Message-oriented Middleware MOM
- **Web-Services: SOAP, WSDL, UDDI, WSFL**
- Enterprise Application Integration EAI

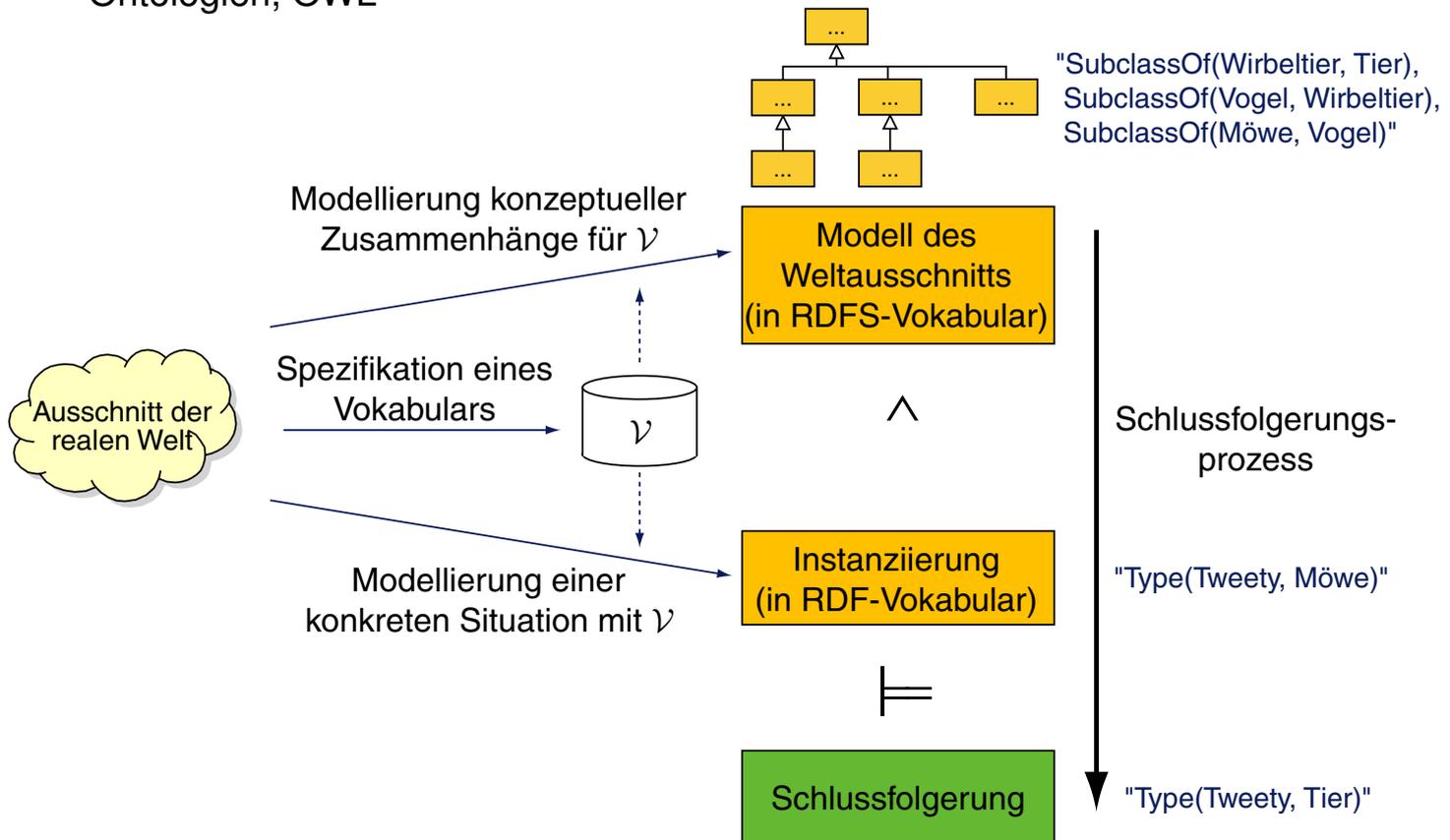


Web-Technologien und Web-Engineering

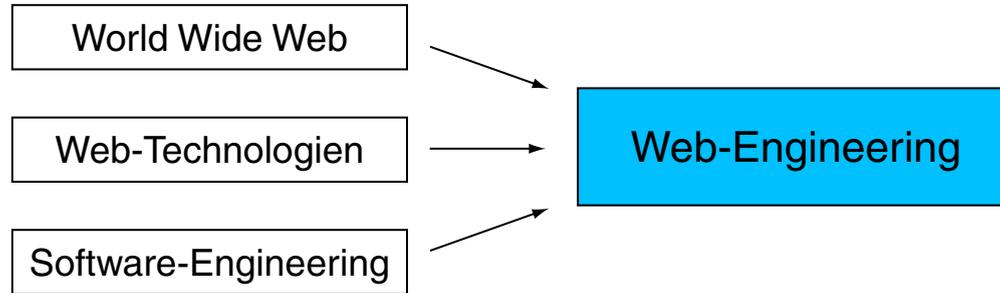
Aufteilung von Web-Technologien (Fortsetzung)

VII. Technologien für das Semantic Web

- Hintergrund und Motivation
- RDF, RDF-Schema, DAML+OIL
- Ontologien, OWL



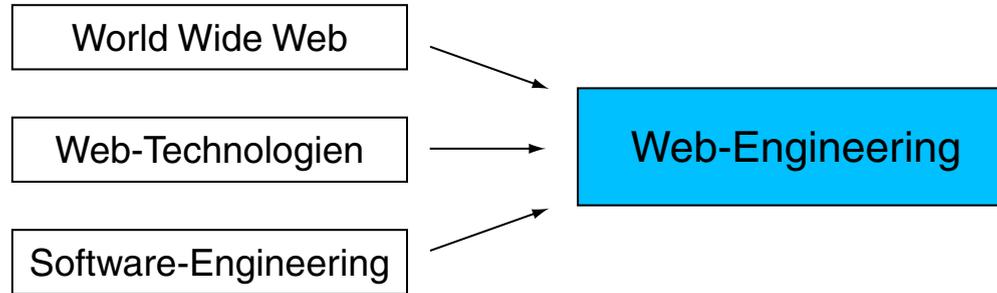
Web-Technologien und Web-Engineering



Definition 7 (Web-Engineering [Dumke 2003])

Web-Engineering ist die methodenbasierte, werkzeugunterstützte, quantifizierte, standardgerechte, erfahrungsausnutzende und Community-bezogene Entwicklung und Wartung von Web-Systemen.

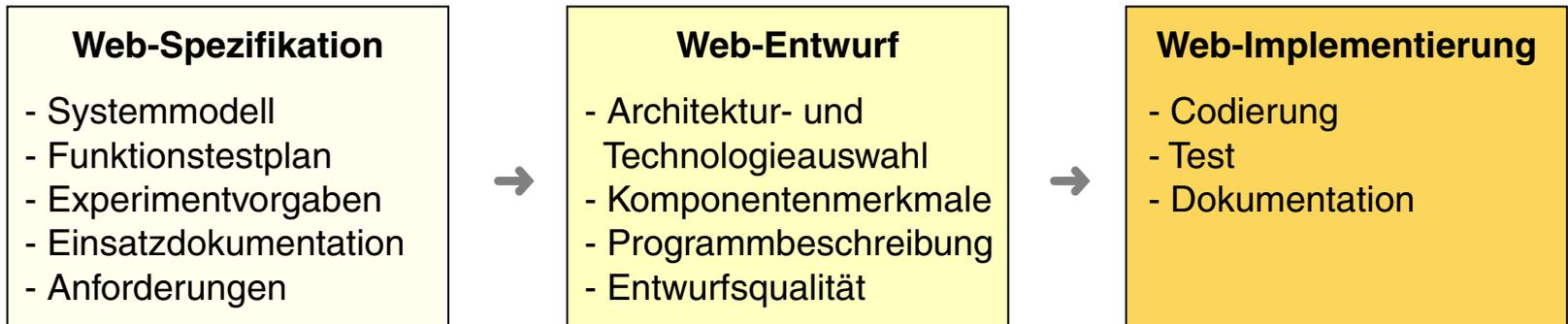
Web-Technologien und Web-Engineering



Definition 7 (Web-Engineering [Dumke 2003])

Web-Engineering ist die methodenbasierte, werkzeugunterstützte, quantifizierte, standardgerechte, erfahrungsausnutzende und Community-bezogene Entwicklung und Wartung von Web-Systemen.

Entwicklungsverlauf von Web-Systemen [Dumke 2003]:



Bemerkungen:

- Inhalt dieser Vorlesung sind Web-Technologien – und weniger die Modelle und Methoden des Web-Engineering. Verständnis für und Beherrschung von Web-Technologien ist Voraussetzung für die Anwendung von Methoden des Web-Engineering.