

TEIL C: Internet – Recherche

Inhaltsverzeichnis

Begriffserklärung	1
Informationsrecherche im Internet	1
Suchverfahren	1
Welche Arten von Suchdiensten gibt es und wann eignen sie sich besonders?	1
Woher bekommen die Suchmaschinen die Informationen?	2
Welche Grenzen gibt es bei der Suche in Suchmaschinen?	2
Warum wird eine Internetseite nicht indexiert oder nicht als Treffer angezeigt?	2
Rechercheprozess	2
Wie kann ich die Qualität einer Trefferliste beurteilen?	3
Übung 1	4
Übung 2	6
Übung 3	7
Übung 4:	7
Geschichte der Internetentwicklung	8
Zugang zum Internet	10
Übung 5	11
Internetdienste	11
World Wide Web – WWW – W3	13
E-Mail	13
FTP – File Transfer Protocol	15
URL / Domain	16
Aufbau einer URL (Domainkonzept)	16
Domain-Registrierung	18
Ablauf einer Registrierung	18
Übung 6	18
Grundsätze bei der Domainvergabe	19

Begriffserklärung

Das Wort „Internet“ besteht zum einem aus dem Teil „inter“, welches aus dem lateinischen hergeleitet die Bedeutung „zwischen“ hat und zum anderen aus dem Wortteil „net“, welches als Kurzform für „networking“ (vernetzen) steht. Im IT-Bereich bedeutet demnach also „Internet“ die **Vernetzung zwischen Computernetzen**.

Ein oft angenommener Irrglaube ist, dass das Internet und das WorldWideWeb (WWW) dasselbe seien. Der Begriff „Internet“ umfasst viele unterschiedliche Funktionen. Zum einen wäre das **WorldWideWeb** zu nennen, welches in erster Linie zur passiven Informationsbeschaffung benutzt wird. Des weiteren bietet das Internet ua. den Dienst, elektronische Post per „**E-Mail**“ zu versenden sowie das „**File Transfer Protocol**“ (FTP), welches unter anderem für die Übermittlung von großen Datenmengen genutzt wird.

Informationsrecherche im Internet

Suchen im Internet ist kein Problem – das *Finden* dagegen schon ...

Suchverfahren

Welche Arten von Suchdiensten gibt es und wann eignen sie sich besonders?

- **Suchmaschinen** – Aktuelle Informationen oder Fakten finden.
Google (www.google.de), Bing (www.bing.com), Yahoo (<http://search.yahoo.de>).
- **Spezialsuchmaschinen** – gezielte Eingrenzung auf einen Themenbereich.
BASE (<http://www.base-search.net>) für wissenschaftliche Internetseiten, Google News (<http://news.google.de>) für aktuelle Nachrichten, IMDB (<http://www.imdb.com/>) für Filme.
Tipp: Nutze Suchmaschinen, um eine spezielle Suchmaschine oder eine Linkssammlung zu einem Thema zu finden (Thema + Stichwort *suchmaschine* oder Phrase „*search engine*“ eingeben, zB *wissenschaft suchmaschine* oder *film "search engine"*).
- **Linkssammlungen** – Überblick auf ein Thema, von Menschen ausgewählte Informationen.
Diigo (www.diigo.com), Mr. Wong (www.mister-wong.de), ODP (www.dmoz.org), Fachinformation UB Bielefeld (www.ub.uni-bielefeld.de/portals/).
- **Metasuchmaschinen** – gleichzeitig in vielen verschiedenen Suchdiensten suchen.
Yippy (<http://search.yippy.com>), Search.com (www.search.com), Metager (<http://www.metager.de/>).
- **Nachschlagewerke** – Alternative zu Suchdiensten bei der Suche nach Fakten.
Wikipedia (<http://de.wikipedia.org/>), Encyclopedia Britannica (www.britannica.com/).
- **Rechnende Wissensmaschinen** sind ein auf Software basierender Internetdienst zum Auffinden und Darstellen von Informationen (zB Wolfram Alpha, <http://www.wolframalpha.com>)
- **Foren** (zB bei Programmierprobleme)



Woher bekommen die Suchmaschinen die Informationen?

Die Suchmaschinen erhalten ihre Daten mit Hilfe von **Robots** (auch Spider oder Crawler genannt), oder von den Anwendern selbst, die ihre Website anmelden. Robots sind vollautomatische Programme, die das Internet absuchen und alles Neue Wort für Wort in den Datenbanken der Suchmaschinen abspeichern.

Mit der Verwendung von **Meta-Informationen** in Webseiten enthalten spätere Recherchen deutlich weniger „Rauschen“ als wenn der Robot einen Volltextindex angelegt hätte, der zudem noch deutlich größer wäre. Metadaten oder Metainformationen sind Daten, die Informationen über Merkmale anderer Daten enthalten, aber nicht diese Daten selbst (im <head> einer Webseite anzugeben).

Welche Grenzen gibt es bei der Suche in Suchmaschinen?

Keine Suchmaschine kennt alle Internetseiten. Alles, was eine Suchmaschine nicht kennt, kann auch nicht von ihr **indexiert** werden. Alles, was nicht indexiert wurde, kann nicht gefunden werden.

Warum wird eine Internetseite nicht indexiert oder nicht als Treffer angezeigt?

- Auf die Internetseite führt kein Link bzw. die Internetseite ist zu versteckt.
- Die Internetseite ist zu aktuell (noch nicht indexiert).
- Die Internetseite hat keinen lesbaren Text (nur Bild / Video) / das Dateiformat ist unbekannt.
- Die Internetseite darf nicht indexiert werden / die Internetseite ist nicht frei zugänglich.
- Die Seite wurde aus dem Index aus gefiltert (Dublette, Rechtsverstoß, Zensur).

Recherche prozess

• Strukturierung im Vorfeld

Mit einer gezielten Strukturierung des Themas legt man den Grundstein für eine erfolgreiche Recherche. Dazu sollte man zu Beginn – vor allem bei sehr umfangreichen Themen – eine **Arbeitshypothese**¹ aufstellen. Danach sollte man auf einem Datenträger eine der folgenden **Ordnerstrukturen** anlegen, die das Thema inhaltlich strukturiert. Gefundene Informationen sollten im entsprechenden Ordner abgelegt werden.

Ordnerstruktur Variante 1	Ordnerstruktur Variante 2
1. Problemstellung	1. Das Phänomen
2. These	2. Seine Ursachen
3. Pro	3. Meinung A
4. Contra	4. Meinung B
5. Schlussfolgerung	5. Fazit

¹ Eine Arbeitshypothese ist im wissenschaftlichen Bereich eine noch zu präzisierende Annahme, die meist vorläufigen Charakter hat.



- Vor der Suche Gedanken machen über die Verwendung „richtiger“ **Suchbegriffe**.
- **Mehrere Recherchen** zu einem Thema mit jeweils 2-3 Suchbegriffen durchführen.
- Eingabe der Suchbegriffe in Kleinbuchstaben und am besten im **Singular**.
- Auf die richtige Suchsyntax bei der **Verknüpfung von Suchbegriffen** achten oder die „erweiterten Suche“ (Suchfelder) verwenden.
- **Die ersten 10 Treffer** näher anschauen (Kriterien zur Bewertung der Qualität einer Trefferliste und einer Internetseite finden Sie auf der zweiten Seite).
- Interessante Internetseiten entdeckt? Dann **mit neuen Suchbegriffen**, die du aus den Treffern ermittelt hast, den Recherche-Prozess noch einmal durchführen, um weitere (bessere) Treffer unter den Top 10 zu finden.

Wie kann ich die Qualität einer Trefferliste beurteilen?

Probleme bei der Bewertung / Grundsätzliche Hinweise zur Bewertung:

- In Suchmaschinen gibt es keine verlässliche Suche oder Sortierung nach **Erscheinungsdatum**.
- Versteckte **Werbung** in der Trefferliste, Verfälschung der Relevanzbewertung durch Betrüger.
- Unterschiedliche Treffermengen und -reihenfolge je nach **Standort**
- Daher: Setze **verschiedene Suchdienste** ein und vergleiche die Ergebnisse (Blindtest zum Vergleich von Google und Bing: <http://blindsearch.fejus.com/>)

Zu viele (unpassende) Treffer gefunden:

- **Fehlerhafte** (automatische) **Wortkorrektur**? Mit Anführungszeichen um das Wort abschalten.
- Weitere Suchbegriffe hinzufügen, **Synonyme** oder weniger allgemeinere Begriffe verwenden.
- **Unpassende Begriffe** / Websites mit „-“ vor dem Wort **ausschließen**.
- Statt einer allgemeinen Suchmaschine eine Spezialsuchmaschine für das Thema verwenden.

Zu wenige passende Treffer gefunden:

- **Schreibfehler**? Gibt es verschiedene Schreibformen? Begriffe kontrollieren (Wörterbuch), Begriffe mit OR (Synonymensuche) kombinieren.
- Einzelne Suchbegriffe weglassen oder allgemeinere Suchbegriffe verwenden.
- Thema zu aktuell? Spezialsuchmaschine für **aktuelle Nachrichten** verwenden.

Wie kann ich die Qualität einer Internetseite beurteilen?

- Unterscheidet sich der Inhalt von dem, was in Trefferliste stand / was du erwartet hast?
- Wer betreibt das Internetangebot? Angaben im **Impressum**, Internetadresse (firma-xy.de).
- Wer ist der **Verfasser**? Angaben unter „Wir über uns“ oder in der Wikipedia.



- Wie aktuell sind die Informationen? Letztes **Änderungsdatum** (unsicher) / Inhalt prüfen.
- Sind **Quellen** angegeben / wie ist die Qualität der Quellen einzuschätzen?
- Gibt es direkte Kommentare von anderen Nutzern? Lob, Kritik, Hinweise auf Fehler beachten.
- Die „**Glaubwürdigkeit**“ einer Website messen zB Dienste wie „Web of Trust“ (<http://mywot.com>). Grundsätzlich gilt: Verlasse dich nicht nur auf eine Quelle; schaue mehrere Webseiten zu einem Thema an. Qualität hängt nicht davon ab, ob es etwas kostenlos oder kostenpflichtig ist.

Interessante Internetseiten weiter verarbeiten / zitieren:

- Speichere interessante Seiten ab / setze **Lesezeichen** (Bookmark / Favorit).
- Besuche von interessanten Seiten aus die **verlinkte Internetseiten** zum Thema.
- Entdecke **neue Begriffe**, suchen mit diesen noch einmal in einer Suchmaschine.
- Erstelle eine **Zusammenfassung**, die die Fülle der gefundenen Informationen in verdichteter Form enthält.
- **Zitierregeln** für Internetquellen sind unterschiedlich – Abrufdatum immer mit angeben, wenn vorhanden unveränderliche Internetadresse benutzen, zB id (Wikipedia), DOI (Zeitschriftenartikel)

Übung 1

Nutze für die Recherche verschiedene Suchmaschinen, Wikipedia etc.

Aufgabe 1

Finde mindestens sechs Homepages von Tageszeitungen und notiere die dazugehörigen Adressen.

http://
http://
http://

http://
http://
http://

Aufgabe 2

Finde die Namen aller österreichischen Bundeskanzler der 2. Republik heraus.

Aufgabe 3

Finde eine Liste aller amerikanischen Präsidenten und notiere 8 davon.

Aufgabe 4

Finde ein Kinoprogramm für ein Kino in Wien-Favoriten und notiere vier Filme, die heute ab 20:00 Uhr laufen.

Aufgabe 5

Wann und wo wurde die Friedensnobelpreisträgerin Mutter Teresa geboren?

Aufgabe 6

Wie, wo und wann kam der Musiker Ritchie Valens um Leben?

Aufgabe 7

Finde zwei Seiten mit Informationen zum Thema Hautkrebs. Notiere die dazugehörigen Internetadressen.

http://

http://

Aufgabe 8

Finde einen Stadtplan mit einer Karte zu Bielefeld. Notiere die dazugehörige Internetadresse.

http://

Aufgabe 9

Wie hoch sind die Einwohnerzahlen folgender Länder: Niederlande, Australien, Südafrika, Indien und Kolumbien?

Aufgabe 10

Welches ist die größte Katzenart und wo lebt das Tier?

Aufgabe 11

Wer ist amtierender Finanzminister des Landes Nordrhein-Westfalen?

Aufgabe 12

Wie heißt die Hauptstadt von Kanada und wie viele Menschen leben dort?

Übung 2

Aufgabe 1

Recherchiere die besten Informationsquellen zu deinen persönlichen Interessensgebieten:

Thema „Hobby“:

Thema „Beruf“:

Thema „Österreich“

Aufgabe 2

Recherchiere aktuelle Seiten zum EU-Datenschutz und achte dabei besonders auf die Aktualität:



Aufgabe 3

Suche Informationen über Schneelagen in Österreich. Gehe dazu auf <https://www.bergfex.at/> und nutze die Linkssammlungen zu diesem Thema!

Übung 3

Recherchiere die **historischen Wurzeln des Internet** (1 Seite)

Achte auf:

- die Nutzung einer zuverlässigen Quelle
- Benutze mehr als eine Suchmaschine
- Verdichtung der Daten
- Quellenangaben
- Aufbereitung der Information

Übung 4:

Du wirst gebeten, zum Thema „**Telearbeit/Teleworking**“ Stellung zu nehmen. Da du zu dieser Thematik kaum Informationen hast, greifst du bei der Informationsbeschaffung auf das Internet zurück. Nutze bei deinen Recherchen die in diesem Kapitel erworbenen Fähigkeiten bezüglich Suchoptimierung, Suchstrategien, Archivierung und Analyse von Webinhalten.

Erstelle eine kurze Powerpoint-Präsentation, in der du die Vor- und Nachteile der Telearbeit erörterst.

Zusätzliche Anforderungen:

- Folienmaster verwenden
- Fußzeile: Datum, Name, Foliennummer
- Mindestens ein Diagramm
- Mindestens 3 Bilder zum Thema einbinden



Geschichte der Internetentwicklung

1957 begannen die ersten Entwicklungen, als die USA – aufgerüttelt durch den Start des ersten Sputnik-Satelliten – mit zahlreichen Forschungsprojekten nach Überlegenheit in der Militärtechnologie strebte. Die kalifornische Firma RAND (**Paul Baran**) wurde damals beauftragt, ein **Konzept für ein militärisches Netzwerk** zu entwickeln, das auch dann noch **funktionsfähig** bliebe, wenn etwa durch einen atomaren Erstschlag ein Teil seiner Infrastruktur zerstört würde. Auf diesem Konzept aufbauend, entwickelte die Advanced Research Projects Agency (**ARPA**) die **paketorientierte Datenübertragung**. 1969 nahm ARPA das erste **paketorientiert** arbeitende Netz in Betrieb (**ARPANET**), dem zwei Jahre später immerhin 30 Teilnehmer angehörten. Ende **1969** wird auch das **Telnet**-Protokoll implementiert kurze Zeit später **FTP**.

Auch in zivilen Institutionen (**Universitäten** und andere Forschungseinrichtungen) entsteht ein **Interesse an der Entwicklung eines solchen Netzes**, denn die damals in Relation sehr teuren Rechner-Kapazitäten ließen sich vernetzt effektiver nutzen. Die wertvollen Ressourcen ließen sich teilen, und die **Ergebnisse konnten ausgetauscht (E-MAIL; 1971; Ray Tomlinson)** werden.

Auch eine Konzentration bzw. Spezialisierung der Rechner für bestimmte Aufgaben war möglich. So wurden **vier Uni-Rechner miteinander verbunden**: Sie standen in der University of California in Los Angeles, am Stanford Research Institute, an der Universität in Santa Barbara und an der University of Utah. Dieses „Netz“ mit vier Knoten wird von vielen als die **Keimzelle des Internet** angesehen. Allerdings verwendete ARPANET **noch kein TCP/IP**; TCP entstand erst **1975** im Zusammenhang mit dem Versuch, verschiedene paketorientierte Netze zu verbinden und wurde **1980** um IP erweitert.

1973 entstand ein Backbone für Europa (Großbritannien und Norwegen wurden mit dem Netzwerk verbunden).

1980 wurde das ARPANET mit einem **Satelliten-** und einem **Funknetz** sowie dem von XEROX-PARC entwickelten **Ethernet** über **TCP/IP** verbunden - **das Internet war geboren!**

1983 wurde das **ARPANET selbst auf TCP/IP umgestellt**. In diesen Jahren stießen immer mehr Netze zum Internet, etwa USENET, BITNET und der 1982 in Europa gegründete Ableger EUnet. **Das "Netz der Netze" nahm Gestalt an**. Es besteht aus vielen Teilnetzen, die von verschiedenen Firmen und staatlichen Einrichtungen betrieben werden.

1986 erfolgte eine Trennung in ein rein militärisches, nicht öffentliches (MILNET) und dem zivilen, forschungsorientierten ARPANET. Das **zivile Netz bekam den Namen Internet** (von „interconnected set of networks“) und verbreitete sich schnell auch in anderen Ländern.

1989 wurde das **ARPANET** vom US-Verteidigungsministerium offiziell **aufgelöst**.

1990 überlegten Wissenschaftler am Genfer Hochenergieforschungszentrum CERN (**Tim BERNERS-LEE**), wie man für die zivile Nutzung **Dokumente weltweit abrufen und Grafiken einbinden** kann. Vor allem aber sollte eine **Hypertextfunktionalität** eingebaut werden, die es ermöglicht, dass ein Dokument auf andere verweisen kann, auch wenn sie auf anderen Internet-Rechnern gespeichert sind. Heraus kam ein neues Internet-Protokoll: **HTTP (HyperText Transfer Protocol)**. Das Projekt wurde **World Wide Web (WWW)** (weltweites Netz) getauft.

1993: Die Popularität des WWW stieg vor allem durch die Einführung des **WWW-Browsers** (NSCA Mosaic; Netscape; Die Firma Netscape wurde vom Mosaic-Entwickler Marc Andreessen gegründet). Diese einfach zu bedienenden Browser machten das WWW für die breite Masse zugänglich. Nachdem man einen Trend verschlafen hatte, entwickelte in den Folgejahren Microsoft unter Hochdruck seinen eigenen Browser: den **Microsoft Internet-Explorer**. Bald hat dieser den Netscape-Communicator verdrängt.

Mitte der 90er Jahre begann das **Internet immer schneller zu wachsen** (schnellere, billigere Zugänge, schnellere, breitere Backbones) – und war spätestens zu diesem Zeitpunkt auch schon immer größeren Teilen der (nicht-akademischen) Bevölkerung ein Begriff.



Zugang zum Internet

Normalerweise haben **Privatpersonen keinen direkten Zugang** zum Internet. Hierfür ist ein relativ großer technischer Aufwand nötig und vor allem sehr viel Fachkenntnis. Der übliche Weg ins Internet geht über einen so genannten **ISP, Internet Service Provider**. Das ist eine Firma, die DEN Internetzugang vermittelt. Die Verbindung von deinem Computer zu dem Internet Service Provider wird häufig über das **Telefonnetz** hergestellt.

Internetzugangstechnologien:

- Modem / ISDN → veraltet
- **DSL, ADSL und SDSL** (*über 60 % der Anschlüsse; Stand 2018*)
 - **DSL** steht für Digital Subscriber Line und ist der Oberbegriff für eine Breitband-Übertragungstechnik, die schnelles Surfen und großen Datentransfer im Internet über das Telefonnetz ermöglicht.
 - **ADSL** (Asymmetric Digital Subscriber Line) ist eine Übertragungsart bei der die Datenübertragung zum Nutzer (der Download) wesentlich höher ist als andersherum (der Upload). Bei einem Internet XL Anschluss von A1 (*Stand 2020*) beträgt die Download-Rate in der Regel 150 MBit/s, die Upload-Rate hingegen nur 20 MBit/s.
 - Die **SDSL-Technik** (Symmetric Digital Subscriber Line) wird häufig von Firmen verwendet, da sie viel teurer ist und eine schnellere Datenübertragung möglich ist. Hier ist die Übertragung symmetrisch. Das bedeutet, dass **Upload und Download gleich schnell** sind. Für Videokonferenzen und den Upload großer Datenmengen auf Webserver ist diese Übertragungstechnik wichtig.
- **Kabel** (*über 30 % der Anschlüsse; Stand 2018*)
Die Daten werden mit Kabelmodems auf die analogen Signale des Kabelfernsehnetzes aufmoduliert und so über diese Koaxialkabel übertragen.
- **Glasfaser**
Den Endkunden direkt per Glasfaser anzubinden ermöglicht hohe Datenraten (größer 1 GBit/s) über große Entfernung, verlangt aber hohe Investitionskosten.
- **Satellit**
Reine Satellitenverbindungen (2-Wege-Satellitenverbindung) sind unabhängig von landschaftlichen Gegebenheiten oder anderer Infrastruktur praktisch überall auf der Erdoberfläche verfügbar und eignen sich damit besonders für entlegene Gebiete und Schiffe.
- **Elektrizitätsnetz**
Mittels Trägerfrequenzanlagen können Internetzugänge über das Stromnetz realisiert werden, auch unter dem englischsprachigen Begriff *Powerline Communication* (PLC) bekannt. Hat sich in der Praxis nicht durchgesetzt.
- **Terrestrische Funktechnologien (Mobilfunk)**
Mobiles Internet bzw. die Nutzung der Mobilfunknetze als Zugang zum Internet liegt nahe, weil Mobilfunk fast überall verfügbar ist. Doch weil die **Bandbreite begrenzt** und Geschwindigkeit der Verbindungen **von der Anzahl der Nutzer innerhalb einer Funkzelle** und von der Entfernung zur nächstgelegenen Basisstation ist, schwankt die Übertragungsgeschwindigkeit zwischen der von Analog-Modems und DSL. Die verwendeten Mobilfunkstandards sind **5G, LTE, UMTS (HSDPA)** oder **EDGE**. Der A1 5GigaCube L bietet zB. im Optimalfall 500 Mbit/s im Upload und 70 Mbit/s im Download (*Stand 2020*).

Übung 5

Recherchiere **Zugangsmöglichkeiten (Technologien) zum Internet** in Österreich.

Erstelle eine Tabelle die eine kurze Beschreibung der Technologie enthält und in der Up-/Download-Geschwindigkeiten (von – bis), benötigte Geräte und sonstige Vor- bzw. Nachteile gegenübergestellt werden.

Ermittle weiters **2 Provider**, die für deinen Wohnort in Frage kommen und vergleiche die beiden Angebote hinsichtlich obiger Kriterien (und zusätzlich bezüglich Kosten, E-Mail-Adressen, Webspeicherplatz etc).

Maximal 2 Seiten

Internetdienste

Das **Internet** selbst stellt lediglich die **Infrastruktur** zur Verfügung. Ein Nutzen für die Anwender entsteht erst dadurch, dass, basierend auf der Struktur des Internets, dem Anwender verschiedene **Dienste** zur Verfügung stehen. Um die unterschiedlichen Dienste nutzen zu können, braucht man die entsprechende **Client-Software** (ein Mail-Programm, ein Programm für FTP usw.).

Wer einen sog. **Browser** (*Firefox, Safari, Internet Explorer*) auf dem eigenen Rechner, dem **Client**, installiert hat, kann mit diesem Universalwerkzeug die meisten Dienste nutzen. Die Rechner, die einen Dienst im Internet anbieten, nennt man **Server**. Auf dem Server ist die dem Dienst entsprechende **Server-Software** installiert.

Die bekanntesten Internet-Dienste sind:

Dienst	Verwendetes Protokoll	Übliche Ports	Beschreibung	Anwendungen
World Wide Web	Hypertext Transfer Protocol (HTTP), Hypertext Transfer Protocol Secure (HTTPS)	80, 443	Zur Übertragung von Webseiten	Webbrowser
E-Mail	Simple Mail Transfer Protocol (SMTP), Post Office Protocol Version 3 (POP3), Internet Message Access Protocol (IMAP)	25, 110, 143	Zum Versand elektronischer Briefe (E-Mails)	E-Mail-Programm
Dateiübertragung (File Transfer)	File Transfer Protocol (FTP)	20, 21	Zur Übertragung von Dateien	FTP-Server und -Clients wie FileZilla
Namensauflösung	Domain Name System (DNS)	53	Mit diesem Dienst werden Namen zB de.wikipedia.org in IP-Adressen übersetzt	Meistens im Betriebssystem integriert
SSH	SSH Protocol	22	Zur verschlüsselten Benutzung entfernter Rechner	ssh, unter Windows zB PuTTY oder WinSCP
Instant Messaging	Verschiedene, proprietäre Protokolle	je nach Protokoll	Kurznachrichten von Person zu Person, Clients unterstützen oft auch IP-Telefonie	Verschiedene Programme, z. B. WhatsApp, Skype, Telegram, Signal

Peer-to-Peer-Systeme	eDonkey, Gnutella, FastTrack	6881 bis 6889, 4661, 4662, 4665, 4672 (eDonkey). 6346 (Gnutella), 1214 (FastTrack)	zB Tauschbörsen zum Austausch von Dateien	eMule, FrostWire, Kazaa Lite K++, Vuze, µTorrent
Internet-Telefonie (VoIP)	H.323, Session Initiation Protocol (SIP)	5060	Telefonieren	Beispiele siehe Kategorie VoIP
Video-Chat	H.264, QuickTime-Streaming		Video-Telefonie	Beispiele siehe Kategorie VoIP
Virtual Private Network VPN	GRE, IPsec, PPTP		Kopplung von LANs durch das Internet, optional mit Verschlüsselung und Authentifizierung	OpenVPN
Internetradio	Hypertext Transfer Protocol (HTTP)		Radio hören/senden	Diverse Medienspieler
Netzwerk-administration	Simple Network Management Protocol (SNMP)	161	Dient der Fernkonfiguration, -wartung und -überwachung von Netzwerkkomponenten wie zB Routern	
Zeit-synchronisation	Network Time Protocol (NTP)	123	Dient dem Zeitabgleich von Computern und Netzwerkkomponenten.	ntpdate bzw. ntp-client (GNU/Linux)
Usenet	Network News Transfer Protocol (NNTP)	119	Diskussionsforen zu allen erdenklichen Themen	News Client, zB Microsoft Outlook Express oder Mozilla Thunderbird
Telnet	Telnet Protocol	23	Zur Benutzung entfernter Rechner	unter den meisten Betriebssystemen standardmäßig verfügbar: telnet
WAIS	Z39.50		Ein früherer Internet-Suchdienst	
Gopher	Internet Gopher Protocol	70	Hypertext-ähnlicher Informationsdienst	

World Wide Web – WWW – W3

Das *World Wide Web* (Abkürzungen *Web* oder *WWW* oder *W3*) überträgt Webseiten.

Zur Anzeige der Seiten wird ein sogenanntes **Browser-Programm** (zB Internet Explorer, Firefox, Opera oder Google Chrome) verwendet.

Seit 1990 hat sich das World Wide Web ständig weiterentwickelt und bietet neben **Bebilderungen, Ton, Animationen** und **Videos** auch **interaktive Inhalte** aller Art.

Eine typische **WWW-Adresse** (der Einfachheit halber auch „Internetadresse“ genannt) ist beispielsweise <http://www.sz-ybbs.ac.at/>.

Das WWW basiert auf drei **Kernstandards**:

- **HTTP** als Protokoll, mit dem der Browser Informationen vom Webserver anfordern kann.
- **HTML** als Dokumentenbeschreibungssprache, die festlegt, wie die Information gegliedert ist und wie die Dokumente verknüpft sind (Hyperlinks).
- **URLs** als eindeutige Bezeichnung einer Ressource, die in Hyperlinks verwendet wird. Durch Anklicken eines **Links** (Verweises) wählt man eine neue **Adresse (URL)** im Internet an.

Folgende **Standards** kamen später dazu:

- **Cascading Style Sheets (CSS)** legen das Aussehen der Elemente einer Webseite fest, wobei Darstellung und Inhalt getrennt werden.
- **Hypertext Transfer Protocol Secure (HTTPS)** ist eine Weiterentwicklung von HTTP, bei dem das Protokoll SSL zwischen TCP und HTTP geschoben wird und in der Folge der Datentransfer komplett verschlüsselt wird.
- **Document Object Model (DOM)** als Programmierschnittstelle für externe Programme oder Skriptsprachen von Webbrowsern.

Nicht vom W3-Konsortium standardisiert ist die am weitesten verbreitete Skript- bzw. Makrosprache von Webbrowsern **JavaScript**.

E-Mail

Die elektronische Post [(e)lectronic (**Mail**)] ist einer der wichtigsten und beliebtesten Dienste. Damit überträgt man auf elektronischem Weg in Computernetzwerken eine briefähnliche Nachricht.

Funktionsprinzip

Wie bei der klassischen Briefpost ist bei der elektronischen Kommunikation eine Empfänger- und Absender-Adresse notwendig. Die **E-Mail-Adresse** besteht aus **drei Teilen**,

- dem Namen des Empfängers
- dem Trennungszeichen @ [at] und
- dem Domain-Namen und der Top Level Domain (siehe auch Domänenkonzept)

unter der der Internet-Provider die Mails seiner Benutzer verwaltet.

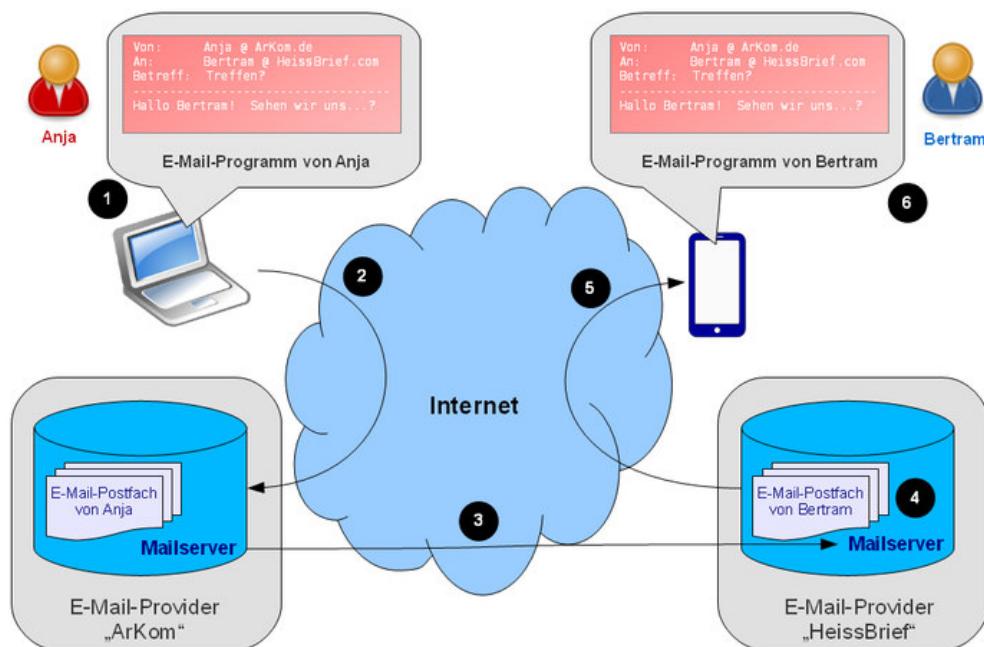


Die E-Mail-Adresse darf außer den Zeichen . - _ **keine anderweitigen Sonderzeichen und auch keine Umlaute** beinhalten.

Bsp: hansi.maier@sz-ybbs.ac.at

Vereinfacht dargestellt werden **E-Mails nach folgendem Verfahren transportiert:**

- Der Absender verfasst die Nachricht, versieht diese mit einer Empfängeradresse und übergibt diese einem speziellen Rechner, dem **Mail-Server**.
- Der Mail-Server **ermittelt** aus der Empfängeradresse den **Zielrechner** des Empfängers und nimmt mit diesem Kontakt auf.
- Die Nachricht wird danach vom Rechner des Absenders zum Rechner des Empfängers **übertragen**.
- Der Zielrechner nimmt die Nachricht entgegen und **speichert** sie für den Empfänger.
- Der Empfänger findet beim nächsten Überprüfen seines Postfaches die Nachricht vor und kann diese **Lesen** oder auf seinen lokalen Rechner herunterladen.



Vorteile

- E-Mails benötigen nur **wenigen Sekunden** vom Absender zum Empfänger, gleich wo sich der Empfänger auf der Welt befindet.
- Der **Preis** für den Versand einer E-Mail beträgt nur einen Bruchteil im Vergleich zu einem herkömmlichen Brief.
- Durch die E-Mail-Programme (E-Mail-Clients) können empfangene und gesendete Nachrichten **zwischengespeichert** werden und jederzeit recherchiert oder die Inhalte der E-Mails in anderen Dokumenten **weiterverarbeitet** werden.
- Das Versenden beliebiger elektronisch erstellter Dokumente (Texte, Bilder, Sprache und Video) über den MIME-Standard als **Anlage** zu einer E-Mail ist unproblematisch.
- E-Mail zwingt **weder Sender noch Empfänger gleichzeitig online** zu sein.

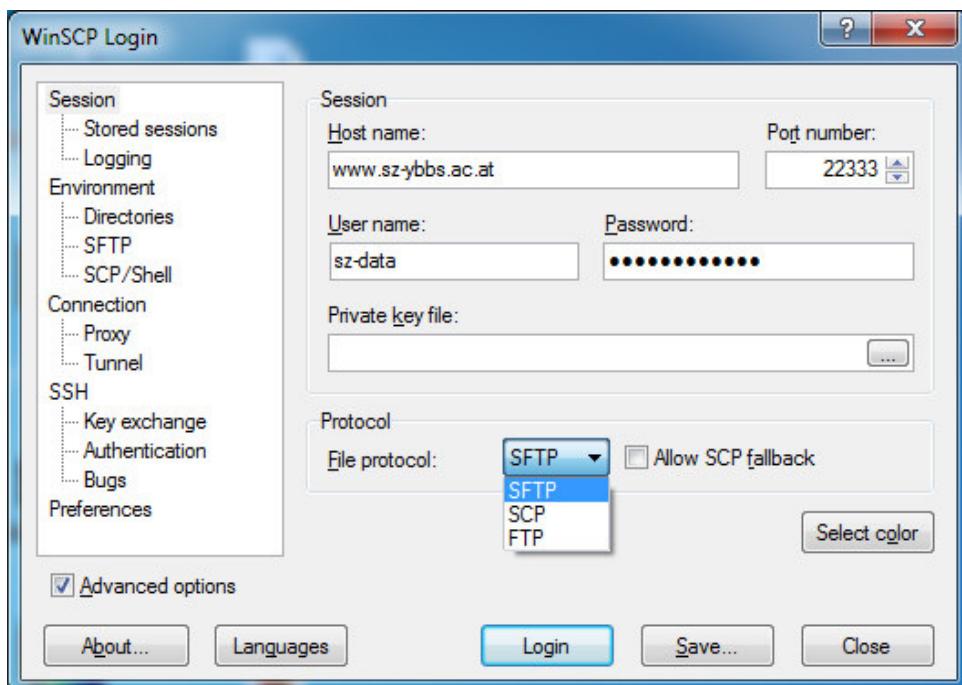
Eingesetzte Protokolle

- **SMTP** ist ein Protokoll zum **Mailversand und -transport**. Zum Versenden über ein E-Mail-Programm benötigt man den Namen eines SMTP-Servers. Dieses entspricht beim Versand eines Postbriefes dem Postbriefkasten im Stadtteil.
- **POP3** dient zum **Abruf von Mails** aus dem E-Mail-Postfach eines Mailservers. Für die Briefpost entspricht es dem Gang zum Briefschlitz an der Haustür.
- **IMAP** dient ebenfalls dazu, auf Postfächer zuzugreifen, die auf Mailservern liegen. Im Gegensatz zu POP3 ist IMAP darauf ausgelegt, die **Mails am Server zu belassen** und dort in Ordner zu verwalten.

FTP – File Transfer Protocol

Mit FTP werden sowohl der Dienst als auch die Sammlung von Befehlen, die zur Übertragung von Dateien dient, bezeichnet. Mittels **FTP** können **Dateien eines FTP-Servers** (in der Regel **authentifiziert** mit Benutzername und Passwort) auf die eigene Festplatte (**unverschlüsselt** über Port 21) **heruntergeladen** werden und umgekehrt. Diese Dateien können Programme, Daten oder andere Informationen enthalten.

Viele FTP-Server, zB von Universitäten und Fachhochschulen, bieten sogenanntes **Anonymous FTP** an. Hier ist zum Einloggen neben den realen Benutzerkonten ein **spezielles Benutzerkonto**, typischerweise „anonymous“ und/oder „ftp“, vorgesehen, für das kein (oder ein beliebiges) Passwort angegeben werden muss.



Um **Verschlüsselung** und Authentifizierung zu nutzen, kann Transport Layer Security eingesetzt werden (FTP über TLS, kurz **FTPS**). Außerdem existiert mit dem **SSH File Transfer Protocol (SFTP)** eine auf SSH (Secure Shell) aufbauende Alternative zu FTP, die ebenfalls Verschlüsselung ermöglicht.

URL / Domain

Internetinhalte benötigen eine Adresse (**URL ... Uniform Resource Locator**) mit der sie gefunden werden können. Im allgemeinen Sprachgebrauch werden URLs auch als *Internetadresse* oder *Webadresse* bezeichnet.

Bsp.: Webseite mit Pfad:

`http://www.sz-ybbs.ac.at/~sz-ybbs/sz-plus/jugendcoaching/`

Unter einer **Domain** versteht man den Klartextnamen eines Internetangebotes hinter dem sich typischerweise eine **IP-Adresse** verbirgt.

Bsp.: `sz-ybbs.ac.at`

Aufbau einer URL (Domainkonzept)

```
scheme-specific-part
|
http://hans:geheim@example.org:80/demo/example.cgi?land=de&stadt=aa#geschichte
|   |   |   |
|   |   |   host      |   | url-path      |   query      |   fragment
|   |   | password    |   port
|   user
scheme (Netzwerkprotokoll)
```

scheme

Legt fest, mit welcher technischen Methode die Ressource angesprochen werden soll. Ist meistens, aber nicht zwingend gleichlautend mit dem verwendeten **Netzwerkprotokoll**, über das die Ressource lokalisiert werden kann.

Bsp.: HTTP, HTTPS oder FTP, aber auch
mailto (zum Schreiben einer E-Mail) oder
file (zum Zugriff auf lokale Dateien).

scheme-specific-part

Je nach `scheme` sind unterschiedliche spezifische Angaben erforderlich und möglich. In den meisten Fällen beginnt er mit der Zeichenkette `://`, jedoch ist bei manchen Varianten auch lediglich der Doppelpunkt definiert (zB `mailto:hans@example.org`).

Die folgenden Beispiele beziehen sich auf das HTTP-Protokoll.

user/password

Falls benötigt, kann ein Login aus Benutzername (`user`) und Passwort (`password`) übermittelt werden. Diese werden, voneinander durch Doppelpunkt getrennt, dem Host mit einem trennenden At-Zeichen (@) vorangestellt.

Auch wenn für dieses Beispiel das Protokoll HTTP gewählt wurde, ist die Angabe von **Benutzername und Passwort als Teil der URL nicht Teil der HTTP-Spezifikation!** Bei einigen anderen Protokollen, etwa **FTP**, ist die Angabe der Benutzerdaten in der gezeigten Form dagegen völlig **korrekt** und durch die Standards abgedeckt.

host

Die Host-Komponente wird in Form einer **IPv4-Adresse** in dezimaler Schreibweise durch Punkte getrennt (zB `http://207.142.131.235`), in Form einer **IPv6-Adresse** in hexadezimaler Schreibweise durch Doppelpunkte getrennt und in eckige Klammern gesetzt (zB `http://[2001:0db8:85a3:08d3:1319:8a2e:0370:7344]/`) oder in Form eines **FQDN (Fully Qualified Domain Name)** notiert.

Bsp.: `www.sz-ybbs.ac.at`

www	Name des Servers; dieser kann auch entfallen (wenn auf dem entsprechenden Webserver die Domain mit 'www' und ohne 'www' eingetragen ist). Meistens wird der Server der Bezeichnung des Internetdienstes entsprechend mit "www" benannt (kann aber auch anders heißen).
sz-ybbs	<p>Domain-Name; Abgesehen von einigen wenigen Einschränkungen bestehen bei der Wahl des Domain-Namens wenig formale Zwänge.</p> <p><u>Sonderzeichen:</u> Abgesehen von Binde- und Unterstrich sind keine Sonderzeichen zulässig. Umlaute sind in Österreich seit 2004 erlaubt.</p> <p><u>Groß- und Kleinschreibung:</u> Durch Groß- und Kleinschreibung innerhalb des Domain-Namens ist eine Kollision gleich lautender Domains nicht zu verhindern.</p> <p>Subdomain: Unterkategorie der Domain Es ist technisch möglich, zu jeder Domain Unterkategorien (sogenannte "Subdomains") einzurichten. Formal wird der Name der Subdomain dem Domain-Namen vorangestellt und dabei wiederum durch einen Punkt abgegrenzt. Beispiel: <code>1bhit.sz-ybbs.ac.at</code>. In diesem Fall stellt „<code>1bhit</code>“ eine Subdomain dar.</p> <p>Vorteil: Gegenüber der Beantragung mehrerer Domain-Namen hat dies den Vorteil, dass Subdomains nicht bei den offiziellen Vergabestellen registriert, sondern lediglich selbst (das heißt in der Regel durch den Provider) eingerichtet werden müssen. Die mit der Registrierung der Top-Level-Domain verbundenen Kosten entfallen.</p>
ac	Second-Level-Domain; präzisiert die TLD, zB <code>ac</code> ... für academic, <code>or</code> ... organisation.
at	<p>Toplevel-Domain; entweder <u>thematisch</u> (zB .com, .edu, .gov, .mil, .net, .org, .aero (für den Bereich Luftverkehr), .biz (für Business), .coop (für genossenschaftlich organisierte und betriebene Unternehmen und Organisationen), .info, .museum, .name (für Web-Sites von Privatpersonen), .pro (für bestimmte Berufsgruppen))</p> <p>oder <u>geografisch</u> nach Ländern (zB .de (Deutschland), .uk (Großbritannien), .nl (Niederlande), .be (Belgien), .it (Italien), .fr (Frankreich), .at (Österreich), .ch (Schweiz)) nach einem Zwei-Buchstabencode</p>

port

Die Angabe des Ports erlaubt die Ansteuerung eines TCP/IP-Ports. Wird kein Port angegeben, so wird der **Standard-Port** des jeweiligen Protokolls verwendet – zum Beispiel bei HTTP 80, bei HTTPS 443 und bei FTP 21.

url-path

Der Pfad **beschreibt eine bestimmte Ressource** (diese kann sich beispielsweise mit der Verzeichnisstruktur des Zielsystems decken, also etwa eine Datei oder ein Verzeichnis)

auf dem Server. Der Pfad kann auch leer sein. Ein leerer Pfad kann optional durch einen Slash ersetzt werden und ist zu diesem gleichbedeutend.

query string

Im Fall des HTTP kann nach dem eigentlichen Ressourcenanzeiger – getrennt durch ein **Fragezeichen** – ein Query String folgen. Damit können **zusätzliche Informationen** übertragen werden, die server- oder clientseitig weiterverarbeitet werden können.

fragment

Nach einem **Doppelkreuz** kann ein Teil der Ressource referenziert werden (typischerweise ein **Anker in einer HTML-Seite**, zu dem dann automatisch heruntergescrollt wird).

Domain-Registrierung

Die Domain-Registrierung ist der Vorgang, den ein Registrant bei einem Registrar durchführt, um eine **neue Domain unterhalb einer Top-Level-Domain** zu registrieren.

Ablauf einer Registrierung

Auswahl eines Namens

Als erstes müssen der **Name der Domain** und die übergeordnete Domain (meist eine **Top-Level-Domain** wie zB .com oder .at) festgelegt werden. In jedem Namen können Buchstaben, Ziffern oder Bindestriche enthalten sein, wobei nicht zwischen Groß- und Kleinbuchstaben unterschieden wird.

Über den **Whois-Dienst** kann festgestellt werden, ob der gewünschte Domainname überhaupt verfügbar ist. Viele Provider stellen dafür auf ihrer Internetseite mehr oder weniger umfangreiche Abfrage-Werkzeuge zur Verfügung.

Übung 6

Überprüfe ob deine Wunschdomain noch frei ist.

Vergabestellen für Domainnamen

Ebenso wie die IP-Adressen dürfen auch Domains aus technischen Gründen weltweit nur einmal vergeben werden. Aus diesem Grunde ist eine zentralisierte Vergabe der Domainnamen erforderlich. Die Registrierung eines Domain-Namens erfolgt **bei der für die jeweilige Top-Level-Domain zuständigen Vergabestelle**.

Für die Top-Level-Domain „at“ ist dies beispielsweise das **NIC (Network Information Center)** www.nic.at, für die Top-Level-Domain .gov, .org, .aero, .biz, .coop, .info, .museum, .name, and .pro die Organisation InterNIC (www.internic.com).

Registrierung über einen Provider

Eine direkte Registrierung einer Domain beim NIC geschieht eher selten (aber doch). Die meisten Interessenten registrieren die **Domain über Internet-Service-Provider** wie zum Beispiel **World4You.com**. Diese übernehmen für Ihre Kunden nicht nur die Registrierung und Freischaltung der Domain beim NIC, sondern stellen zudem **Speicherkapazitäten** zur Verfügung, auf denen Inhalte für die Domain zum Abruf bereitgestellt werden können.

Grundsätze bei der Domainvergabe

Registrar prüft nur formale Anforderungen

Wird ein Registrierungsantrag für eine Domain innerhalb der Top-Level-Domain „.at“ zum Beispiel beim NIC beantragt, prüft dieses grundsätzlich nur, ob die beantragte Domain den formalen Anforderungen entspricht und nicht bereits vergeben ist. Das dem zugrunde liegende Prinzip wird auch als „**first come, first served**“ (wer zuerst kommt, mahlt zuerst) bezeichnet, da eine Domain demjenigen zugesprochen wird, der sie zuerst beantragt hat. **Der Registrierung entgegenstehende Rechte Dritter werden nicht überprüft.** Vielmehr versichert der Antragsteller, dass er selber eine solche Kontrolle vorgenommen hat und keine Indizien für die Verletzung Rechte Dritter oder sonstiger Rechtsnormen gewinnen konnte.

Häufige Ursache für Domainstreitigkeiten

Diese Selbstbeschränkung der Vergabestellen ist vor dem Hintergrund der potenziellen Komplexität möglicher Prüfungsverfahren verständlich und hat das Registrierungsverfahren erheblich beschleunigt. Zugleich ist die Präferierung der zeitlichen Komponente gegenüber der rechtlichen Position jedoch mit ursächlich für eine nicht unerhebliche Zahl von **Domainstreitigkeiten**. Selbst in den Fällen, in denen ohne weiteres erkennbar ist, dass durch die Registrierung die Rechtsposition Dritter verletzt wird, wird der Antragsteller zunächst als Domaininhaber eingetragen. **Der Dritte muss seine Rechte dann in mitunter langwierigen gerichtlichen oder außergerichtlichen Verfahren geltend machen.**