

Kapitel I.4: Das World Wide Web und seine Browser¹

Dr. Frank Krüger

Fachbereich 23, Johannes Gutenberg-Universität Mainz

e-mail: krueger@acm.org

Inhalt

- 1 Das [World Wide Web](#)
- 1.1 [Allgemeines](#)
- 1.2 [Generalisierte Zugriffsmethode \(URL-Adressen\)](#)
 - [Absolute URL](#), [relative URL](#) und [Dateitypen](#)
- 1.3 [Browser und Hilfsprogramme](#)

In Teil II wird ergänzend auf folgende Themen eingegangen:

- [Kapitel II.3](#): Andere Informationssysteme bzw. -dienste (WAIS, Gopher, [Usenet News](#) und [Chat](#) sowie Suchhilfen
- [Kapitel II.4](#): Erstellung eigener WWW-Angebote mit HTML
- [Kapitel II.5](#): Hilfsmittel und Tips für das Design eigener Angebote

Im Anhang Bx: Funktionsbeschreibung der [Browser](#) am Fachbereich: [Netscape 3.0 \(Mac/Windows\)](#), Microsoft Internet Explorer und [Lynx](#)

1. Bildschirmaufbau ([Mac](#), [Windows](#)) bzw. Befehlsübersicht ([Lynx](#))
2. Die wichtigsten Funktionen: Die Symbolleiste ([Mac](#), [Windows](#))
3. Menüs und Tastaturkürzel ([Mac](#), [Windows](#)) bzw. Optionen ([Lynx](#))
4. Lesezeichen und History: Verwalten von Information ([Mac](#), [Windows](#) und [Lynx](#))

sowie Editorfunktionen ([Mac/Win](#)) und die Benutzung anderer Informationsdienste mit [Netscape 1.1](#) (WAIS, Gopher), [2.0](#) (auch [Usenet News](#); ebenso mit [Lynx](#)) und [3.0](#).

Anhang G: [Einige wichtige Adressen](#) im Internet (insbesondere WWW)

1. Das World Wide Web (WWW, W3)

A. Allgemeines

Das **World Wide Web** ist heute das populärste, aber keineswegs das einzige verteilte Informationssystem im Internet. So zeichnen sich die anderen Informationssysteme, wie GOPHER, USENET NEWS und WAIS (siehe [Kapitel 2.3](#)), durch ähnliche technische Rahmenbedingungen aus.

Die wesentlichen Gründe für die Popularität des WWW sind:

- die Überwindung der rein zeichen-orientierten Benutzungsoberfläche mit umfangreichen Befehlssprachen durch eine moderne, intuitiv nutzbare **graphische Benutzerschnittstelle**,
- der **multimediale Inhalt**, d.h. die Information wird nicht nur in Texten dargeboten, sondern auch in Graphiken und Bildern, als ausführbare Programme sowie Ton- und Videodateien mit zahlreichen Varianten (z.B. variierbare 3-D-Bilder nach Art der virtuellen Realität) und die
- **Integration** und **Standardisierung** aller verschiedenen Dienste und Dateitypen auf den unterschiedlichsten Betriebssystemplattformen (UNIX, MAC, WINDOWS).

a) Charakterisierung

(vgl. [andere verteilte Informationssysteme](#), Kapitel 2.3.B)

<i>Grundprinzip</i>	Hypertextstruktur, d.h. im fortlaufenden Text oder anderen Dokumenten (Graphik, Video, Audio) sind Verweise zu anderen Dokumenten bzw. -arten eingebettet.
<i>Ursprung</i>	Von Tim Berners-Lee (CERN , Schweiz) als Hilfsmittel zur Unterstützung von räumlich verteilten Forschungsgruppen über das Internet entwickelt; heute koordiniert (aber keineswegs zentral verwaltet) von der eigens gegründeten W3-Organisation (http://www.w3.org), die als gemeinnütziges Konsortium von verschiedenen Firmen und staatlichen Stellen/Hochschulen getragen wird.
<i>Inhalt</i>	Integratives, multimediales Angebot aller Arten von Informationen der unterschiedlichsten Quellen - von speziellen Forschungsberichten über persönliche Angaben bis zu kommerzieller Werbung. Durch spezielle, programmierbare Schnittstellen (sog. cgi - common gateway interface) und die modulare Programmiersprache JAVA wird auch der dynamische Aufbau von Dokumenten aus Datenbankinhalten ermöglicht (z.B. Bibliothekskatalog (OPAC) der UB Mainz)

¹ : Eine elektronische Version dieses Dokumentes befindet sich auf dem WWW-Server des Fachbereiches unter der Adresse <http://www.fask.uni-mainz.de/cafl/kurse/komm/komm-14.html> (Kleinschreibung beachten!).

Verzeichnisse systematisch: Yahoo oder Virtual Library,
 Geographisch: Virtual Tourist, Master-Index
 Suchhilfen: Lycos, AltaVista und viele andere
 => nähere Information zu Suchhilfen siehe [Kapitel 2.3, A](#)

Verbreitung exponentielles Wachstum weltweit in allen Bereichen, Januar 1996
 225 Tsd. Server, und bis zu 30 Millionen Benutzer (geschätzt)

2. Generalisierte Zugriffsmethode (URL)

Ein wesentliches Charakteristikum des WWW ist die Integration und Standardisierung der Zugriffsmechanismen und technischen Protokolle.

Dazu wurden die sog. **Unified Resource Locators (URL)** entwickelt, die eine einheitliche Spezifizierung von Informationseinheiten unterschiedlichster Dienste ermöglicht. Daneben ermöglicht der sog. **Multimedia Information Mail Exchange (MIME)** Standard, ursprünglich zum Austausch multimedialer E-mail geschaffen, eine einheitliche Spezifizierung verschiedener Dateiformate.

a) Absolute, vollständige Adresse

{METHODE}://{HOST}{/PFAD}{/DOK.EXT}{#FRAGMENT}[?{AUSDRUCK}]

Dabei bezeichnet

- **{Methode}** das verwendete Protokoll (z.B. HTTP, FTP, news),
- **{Host}** den Namen oder die IP-Nummer des angesprochenen Rechners
- **{Pfad}** die Unterverzeichnisse, auf die zugegriffen wird
- **{Dokumentext}** den Dokumentnamen mit einer Erweiterung, die den Dokumentnamen angibt (etwa .HTML, .TXT, .GIF oder .MPEG); bei **{#Fragment}** handelt es sich um den Bestandteil einer Datei **{?Ausdruck}** eine Suchformulierung, wobei nur ein eingeschränkter Zeichensatz direkt dargestellt wird (Leerzeichen durch +, alle andere Zeichen hexadezimal codiert)

Zugriffsmethoden mit der oben dargestellten Struktur sind z.B.:

- **http** - Hypertext Transfer Protocol (in Verbindung mit HTML, s. [Kap. II.4](#)), sowie
- **ftp** - Dateitransfer, **gopher** - Gopher-Dienst und **wais** - Wide Area Information Service

Von der Standardstruktur weichen - protokollbedingt - etwas ab:

- **news** - Usenet Newsgroups mit der Form NEWS:NEWSGROUP-STRUKTUR, z.B. *news:soc.culture.italian* und *news.soc.culture.** für den Hierarchiebereich aller kulturspezifischen Gruppen.

- **mailto** - **Elektronische Post**, mit der Form einer E-mail Adresse

Hinweis: In den letzten beiden Fällen sind unbedingt gültige **Einstellungen** für die Mail- (SMTP bzw. POP3-Protokolle) bzw. News-Server (NNTP) notwendig, auf die zugegriffen werden soll. Außerdem verlangen viele News-Server, daß der zugreifende Client-Rechner mit seinem Namen in einem Name Domain Server registriert ist.

- **file** - Lokale Dateien, wobei der Rechnername entfallen kann und Laufwerksbezeichnungen mit zum Pfad gezählt werden. Sonderzeichen werden durch % plus einen speziellen Code umschrieben:
 z.B. *file:///Macintosh%20HD/Dokumente/Test.HTML*, wobei %20= Leerzeichen

b) Relative, partielle Adresse

Nicht immer muß die vollständige Adresse angegeben werden. Alle Angaben, die im Vergleich zum aktuellen Dokument gleichbleiben, werden automatisch ergänzt.

Ist z.B. die aktuelle Adresse *schema://host/x/y/z* so wird

die Adresse ergänzt zu

- **#marke** schema://host/x/y/z#marke
- **datei** schema://host/x/y/datei aber (Achtung!):
 - **/datei** schema://host/datei (Hauptverzeichnis)
 - **../w/datei** schema://host/x/w/datei (Schwesterverz.)

c) Dateitypen








Die meisten **Browser** und Server im WWW - z.B. NETSCAPE - sind in der Lage ganz verschiedene Dateitypen (die technische Bezeichnung für Dokumentarten) zu verwalten. Die entsprechenden Dokumenttypen müssen sowohl vom Server als auch dem Client bzw. Browser erkannt werden können, um korrekt übertragen zu werden. Auf der Serverseite spielt hierbei HTTP (Hypertext Transfer Protocol) eine große Rolle, da bei der Übertragung zunächst der sog. MIME-Typ (siehe unten) übertragen wird.

Die "Entschlüsselung" oder Verarbeitung des MIME-Typs kann entweder im Browser selbst erfolgen - teils auch durch zusätzlich installierte sog. **Plugins** oder - v.a. bei weniger gebräuchlichen und plattformabhängigen - Datenformaten - mit Hilfe sog. **Helper Applications**.



Der Unterschied zwischen beiden Formen der Erweiterung des Browsers ist, daß die Plugins direkt im Browser-Fenster dargestellt werden (und damit auch auf einen bestimmten Browser zugeschnitten sind), während die *Helper Applications* lediglich als weitere Anwendung im Rahmen des Programm-Multitaskings auf dem Rechner geladen werden und die Dokumente unabhängig vom Browser dargestellt werden. Ist für einen MIME-Typ keine Helper Application definiert, so kann das Dokument auch zunächst abgespeichert werden.



Einige der wichtigsten Helper Applications und Plug-ins werden im [Anhang D.1](#) für [DOS](#), [Macintosh](#) bzw. [Windows](#) vorgestellt und kurz beschrieben.

Folgende Dateitypen werden in [NETSCAPE](#) und den meisten anderen [Browsern](#) direkt erkannt. [LYNX](#) zeigt jedoch im Gegensatz zu den graphischen Browsern (etwa [NETSCAPE](#) und [INTERNET EXPLORER](#)) nur das Vorhandensein dieser Dateien an und kann diese zwar abspeichern, aber nicht darstellen.

Symbol	Art	Beschreibung
	ASCII-Text	Normaler Text der MIME-Formate <i>text/plain</i> (Endung .TXT) oder HTML-Seiten (MIME: <i>text/html</i> , Endung: .HTML, .HTM)
	Graphik	bzw. speziell die MIME -Formate GIF (für Zeichnungen und Vektorgraphiken allgemein, Endung.GIF) und JPEG (für Fotos und Bitmaps allgemein, .JPEG)
	Verzeichnis	Typisches Menüsymbol für Gopher und FTP, dagegen seltener im eigentlichen Web verwendet.
	News	weltweite Diskussionsgruppen im sog. Usenet
	Gopher-Suche	VERONICA-Suchsystem, mit Suchmöglichkeit nur in Verzeichniseinträgen oder in allen Titeln (z.B. an der Uni Köln)
	Dialogverbindung	Mit dem Programm Telnet aktivierbare Terminalverbindung z.B. für Bibliothekskataloge (muß in den meisten Browsern als "Application" eingetragen sein, vgl. Netscape Einstellungen).
	unbekannt	alle anderen Dateitypen, die sonst nicht definiert sind

Weitere Dateitypen sind nach dem sog. **MIME-Schema** definiert, mit dem alle Dateiformate und Medienarten einheitlich bezeichnet werden können; und zwar in der Form:

Symbol	Type/Subtype	Extension	Beschreibung
	image /tiff /pict /x-... (versch.)	.TIF, .TIFF .PICT	Tagged Image File Format Macintosh X-Windows/Unix
	audio /basic /aiff /midi	au, .SND .aiff .MID	Audio-Dateien (Sprache, Musik, Töne) Audio Image File Format (Mac) MIDI (Windows-Format)

	video /mpeg /quicktime /x-msvideo	.MPG, .MPEG .MoV .AVI	Animationen o. "Real Video" Standard für Bewegtbilder Quicktime (Mac/Windows) MS-Video (Windows)
	application /x-macbinary /x-stuffit /x-tar, /x-zip /postscript /pdf	.BIN .sit .tar, .zip .ps, .EPS .PDF	ausführbare Programme, entweder binär kodiert oder komprimiert (Mac); Unix- bzw. Dos-Kompressionsformate Postscript bzw. Encapsulated Postscript Portable Document Format (Adobe)

In Netscape findet sich die (erweiterbare) Liste der Zuweisungen dieser MIME-Formate zu den sog. **HelperApplications** unter OPTIONS/(GENERAL) PREFERENCES/HELPER APPLICATIONS.

3. Web-Funktionen der Browser

Für den Zugriff auf das World Wide Web ist auf der Seite der Benutzer ein sog. **Browser** notwendig, um die Information darzustellen. Dabei handelt es sich um den sog. **Client** im Client-Server-Modell, also eine aktive Software, die mit dem Server (hier ebenfalls lediglich ein bestimmtes Software-Programm und keineswegs der "physikalische" Rechner) Daten und Informationen aufgrund bestimmter Protokolle (HTTP, Gopher, FTP, NNTP ...) austauscht.

Zu den wichtigsten Browser gehören:

- **LYNX**: Ausschließlich zeichenorientiert, was aber immerhin zu einem wesentlich schnelleren Seitenaufbau führt - die meist unnützen Bilder werden nur auf ausdrücklichen Befehl geladen, da sie nicht direkt angezeigt werden können. Durch die geringen Anforderungen an den PC ist Lynx auch mit älteren Rechnern und über langsamere Modemverbindungen sehr gut verwendbar.
- **NETSCAPE 3.0** für [Mac](#), [Windows](#) und Unix. Gegenüber den Vorgängerversionen 1.x und 2.0 ist die Funktionalität nochmals verbessert (z.B. Druckfunktionen und Navigation in den sog. Frames) und erweitert worden. So lassen sich nicht nur andere Informationsdienste wie FTP, Usenet News und Gopher damit benutzen, sondern auch ein (einfacher) [Mail-Client](#) und [Editor](#) ist in das Programm integriert.
- **NETSCAPE COMMUNICATOR (Version 4.0.x)** wird zusätzlich einen Online-Kalender und Videokonferenzfunktionen (Collabra) sowie Ansätze von Netzwerkbetriebs-systemfunktionen (Dateiverwaltung etc.) enthalten.
- **INTERNET EXPLORER** (aktuelle Version 4.0) ist das mit großem Aufwand von Microsoft vorangetriebene Konkurrenzprodukt zu Netscape, das zwar recht gut mit den anderen Standardprogrammen von Microsoft (Office und Backoffice-Pakete) zusammenarbeitet, aber in Punkto Sicherheit v.a. bei programmierbaren Dokumenten mit Java oder Active-X-Technologie eher kritisch zu beurteilen ist.