

SelfLinux-0.13.1



Grundlagen WWW



Autor: Rolf Schmidt (rolf.frogs@t-online.de)
Formatierung: Alexander Fischer (Selflinux@tbanus.org)
Lizenz: GFDL

Dieser Artikel beschreibt kurz die Geschichte des Internets und erläutert ganz allgemein, was das Internet ist. Es werden das World Wide Web, Browser und Suchmaschinen sowie die Gefahren des Internets vorgestellt.

Inhaltsverzeichnis

1 Einleitung

- 1.1 Was ist das Internet?
- 1.2 Geschichte des Internets

2 Anschluss an das Internet

- 2.1 Voraussetzungen
 - 2.1.1 Hardware
 - 2.1.2 Software
 - 2.1.2.1 Netzwerksoftware
 - 2.1.2.2 Netzprogramme
 - 2.1.2.3 Dienstanbieter
 - 2.1.2.4 Anbindung über einen Proxyserver
- 2.2 Dienste, die im Internet angeboten werden

3 Das World Wide Web

- 3.1 Web Browser und das WWW
- 3.2 Was ist ein Browser?
- 3.3 Hypertext Markup Language
- 3.4 Welche Browser gibt es?
- 3.5 Erweiterungen der Browser
 - 3.5.1 Script-Sprachen
 - 3.5.2 PlugIns
 - 3.5.3 Verschlüsselung
 - 3.5.4 Cookies
 - 3.5.5 Sinnvoller Schutzmechanismus gegen Cookies
- 3.6 Suchen und Finden
 - 3.6.1 Informationsbeschaffung im Internet
 - 3.6.2 Kataloge
 - 3.6.3 Suchmaschinen
 - 3.6.4 Metasuchmaschinen
 - 3.6.5 Website bei Suchmaschinen anmelden
- 3.7 Sicherheit im Internet
 - 3.7.1 Gefahren des Internets
 - 3.7.2 Wie kann man sich schützen?
 - 3.7.2.1 Generell
 - 3.7.2.2 E-Mail-Verkehr
 - 3.7.2.3 Download aus dem Internet
 - 3.7.2.4 Anmelden auf Rechnern
 - 3.7.3 Firewall

1 Einleitung

1.1 Was ist das Internet?

Das Internet ist kein eigenständiges, großes Rechnernetzwerk, sondern die Verbindung verschiedener privater oder öffentlicher Rechner(-netze) mittels einer gemeinsamen **Netzwerkssprache**, dem Protokoll [TCP/IP](#).

Die lateinische Vorsilbe **inter** bedeutet **zwischen**, und somit deutet schon der Name "Internet" an, dass es sich hier um ein Netz zwischen verschiedenen Netzen handelt.

Daraus folgt, dass das Internet auch keiner einzelnen Institution, Firma oder Person gehört, sondern nur **lebt**, solange Firmen und/oder Personen verschiedene Rechner miteinander verbinden und anderen Institutionen und/oder Personen den Zugriff auf diese Rechner gestatten. Allerdings gibt es eine ordnende Instanz ( [Internet Engineering Task Force](#)), die die technischen Regeln der Netzkommunikation (und damit die Funktionsfähigkeit) verwaltet und die Regeln veröffentlicht.

Ähnlich wie im normalen Leben gibt es auch im Internet Arbeitsteilung. Diese ist (z. T.) hierarchisch, also in einer Baumstruktur organisiert, und teilt das Netz in überschaubare Segmente auf. Diese Aufteilung ist die Voraussetzung für die Handhabbarkeit eines so gigantischen Netzes, an das inzwischen Millionen von Rechnern angeschlossen sind.

Dank dieser verteilten Architektur kann es praktisch nur dann zu einem Totalausfall des Internets kommen, wenn weltweit alle Stromnetze gleichzeitig versagen. Selbstverständlich kann die verteilte Struktur auch nur durch eine verteilte Administration verwaltet werden. Dafür müssen die Administratoren kommunizieren und kooperieren. Da dies, unter anderem wegen divergierender Interessen, nicht immer perfekt klappt, hat das Internet eine anarchisch anmutende, häufig als chaotisch empfundene Struktur und ist auch Spielwiese für alle nur denkbaren Interessengruppen.

1.2 Geschichte des Internets

Als Urheber des Internets wird immer wieder das amerikanische Militär genannt. Unter anderem, weil für die Entwicklung des Internets Mittel aus dem US-Militärhaushalt zur Verfügung gestellt wurden. Relativ sicher scheint allerdings, obwohl es dazu widersprüchliche Aussagen gibt, dass der Nutzen der universitären Entwicklung des Internets zuerst vom Militär erkannt wurde.

Später nahm das *Department of Defense* der USA, das amerikanische Verteidigungsministerium (DoD), sicherlich Einfluss auf die Entwicklungsrichtung und finanzierte auch das erste größere Netz.

Die Entwicklung führte zu einer Netzwerktechnik, die sich insbesondere durch die folgenden Merkmale auszeichnet:

- * Hohe Ausfallsicherheit des Gesamtnetzwerkes
- * Verteilung von Rechenleistung, Last, Daten und Datenübertragung
- * Hierarchisches Kommunikations- und Datenverteilungssystem
- * Einfache, rechnerunabhängige Kommunikationsstrukturen
- * Keine feste Bindung an ein Übertragungsmedium
- * Systemunabhängiges Benennungs- und Adressierungsschema

Die folgende Tabelle gibt einen groben Überblick über die Meilensteine der Geschichte des Internets.

1961

Die erste Veröffentlichung zur paketorientierten Datenübertragung wird ausgearbeitet.

1969

Die Advanced Research Projects Agency (ARPA) erteilt einen Forschungsauftrag zum Aufbau eines dezentralen Netzwerkes. Dieses Netz wird ARPANET genannt. Es soll eine robuste, zuverlässige und herstellerunabhängige Datenkommunikation ermöglichen.

1975

Aus dem experimentellen  [ARPANET](#) wird ein normal einsetzbares Netz. Die Verantwortung für die Administration wird der Defence Communications Agency (DCA) übertragen. In dieser Zeit wird auch das TCP/IP-Protokoll entwickelt.

1983

Das Verbindungsprotokoll wird zum militärischen Standard erklärt, sodass sich jeder Hersteller, der dem Militär Netztechnologie verkaufen will, an diesen Standard halten muss.

Im Laufe der Zeit werden immer mehr Subnetze aufgebaut und an das Internet angeschlossen. Daraus leitet sich auch der Name "Internet" ab. Diese Technologie verbindet Netze, wirkt also zwischen den Netzen (inter = zwischen).

Das Internet ist also die Verbindung und gleichzeitig die Verbindungsmethode zwischen vielen - heute zig Millionen - Netzwerken, die über die Erde verteilt sind.

1993

Erste öffentliche Web-Seite. Davor gab es ausschließlich Seiten von Universitäten bzw. wissenschaftliche Veröffentlichungen. (Hier finden Sie eine Kopie der ersten  [Site](#) von  [Tim Berners Lee](#) aus dem Jahre 1992. Er ist einer der Väter des World Wide Web.)

2 Anschluss an das Internet

Der Anschluss an das Internet stellt heute für niemanden in der westlichen Welt ein großes Problem dar. Die Diensteanbieter - *Provider* genannt - stellen in aller Regel die benötigte Software sowie Installationsunterstützung zur Verfügung.

Die Verbindung wird meist über eine Wählverbindung (Analog- oder ISDN-Technik) hergestellt. Es gibt aber auch permanente Verbindungen mit Standleitung oder über DSL.

2.1 Voraussetzungen

2.1.1 Hardware

Selbstverständlich benötigt man als erstes einen Rechner. Es muss nicht das allerneueste Modell sein (abhängig vom verwendeten Betriebssystem).

An diesen Rechner muss ein [Modem](#) (auch ISDN-Modem) oder eine [ISDN](#)- bzw. Netzwerkkarte angeschlossen sein (sofern diese nicht eingebaut sind). Beim Anschluss über [DSL](#) oder in Firmennetzen wird eine ([Ethernet](#)-) [Netzwerkkarte](#) benötigt.

2.1.2 Software

2.1.2.1 Netzwerksoftware

Als erstes muss nun die Treibersoftware für das Modem oder die ISDN- bzw. Ethernetkarte installiert werden. Dafür sollten Sie die beigelegten Treiber verwenden, falls es keine automatische Unterstützung durch das Betriebssystem gibt.

Nun wird noch die [TCP/IP-Netzwerksoftware](#) benötigt. (Falls der Rechner im Firmennetz läuft, sollte man überprüfen, ob dieses sog. Netzwerkprotokoll eingerichtet ist.) Diese Software ist heute in aller Regel Bestandteil des Betriebssystems bzw. wird mit diesem ausgeliefert.

2.1.2.2 Netzprogramme

Um die angebotenen Internetdienste tatsächlich nutzen zu können, benötigen Sie noch Software in Form von sogenannten Client-Programmen für die entsprechenden Dienste.

Bei der Verbreitung von Client-Programmen steht heute der sog. Browser für den WWW-Dienst an erster Stelle. Außerdem kann ein separater FTP-Client nützlich sein, obwohl in die meisten Browser bereits ein FTP-Client integriert ist.

Wahrscheinlich wollen Sie auch E-Mails senden und empfangen, sodass Sie auch dafür eine entsprechende Software benötigen. Diese kann ebenfalls im Browser eingebaut oder eine separate Software sein.

2.1.2.3 Dienstanbieter

Wenn Sie nicht mit Call-by-Call auf das Internet zugreifen möchten, benötigen Sie noch einen Vertrag mit einem Dienstanbieter (Provider) wie z. B.  [T-Online](#),  [Arcor](#),  [1&1](#). Dieser Dienstanbieter stellt Ihnen gegen

Bezahlung eine Telefonnummer, einen Benutzernamen (Account), ein Passwort und meist auch eine E-Mail-Adresse zur Verfügung. Über die angegebene Telefonnummer können Sie sich dann in das Netzwerk des Providers einwählen und mit dem Benutzernamen und dem Passwort bei diesem Netzwerk anmelden. Nun erst können Sie auf das Internet zugreifen.

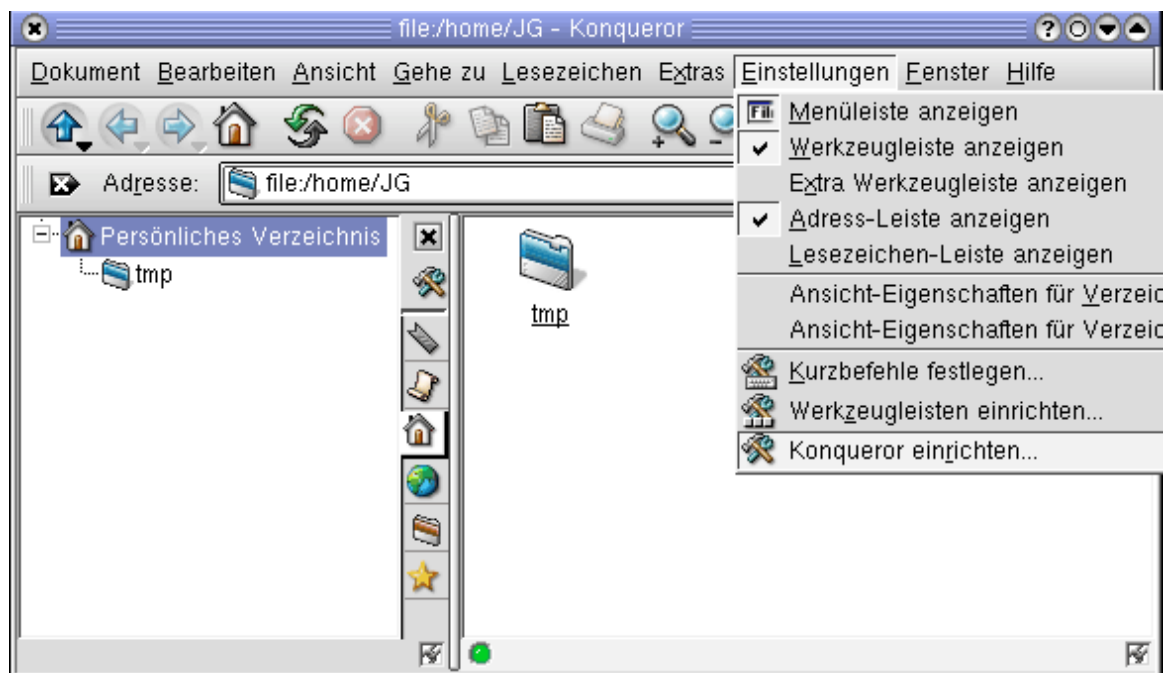
Beim Abschluss eines entsprechenden Vertrages sollte man darauf achten, wie die Kosten berechnet werden. Entweder wird die Verbindungsdauer in Rechnung gestellt oder die übertragene Datenmenge, wobei beide Übertragungsrichtungen angerechnet werden. Hinzu kommen normalerweise noch die Telefongebühren.

2.1.2.4 Anbindung über einen Proxyserver

Ein Proxyserver ist eine Software die als Stellvertreter, an Stelle des Browsers, die Web-Seiten im Internet abholt und an den Browser weiterleitet. Anders ausgedrückt: Man versteckt das eigene Netz hinter dem Proxy, da keine Datenpakete aus diesem Netz im Internet auftauchen. Außerdem können Proxyserver häufig aufgerufene Seiten zwischenspeichern, sodass für diese Seiten nicht jedes Mal eine Internetverbindung aufgebaut werden muss und der Zugriff beschleunigt wird.

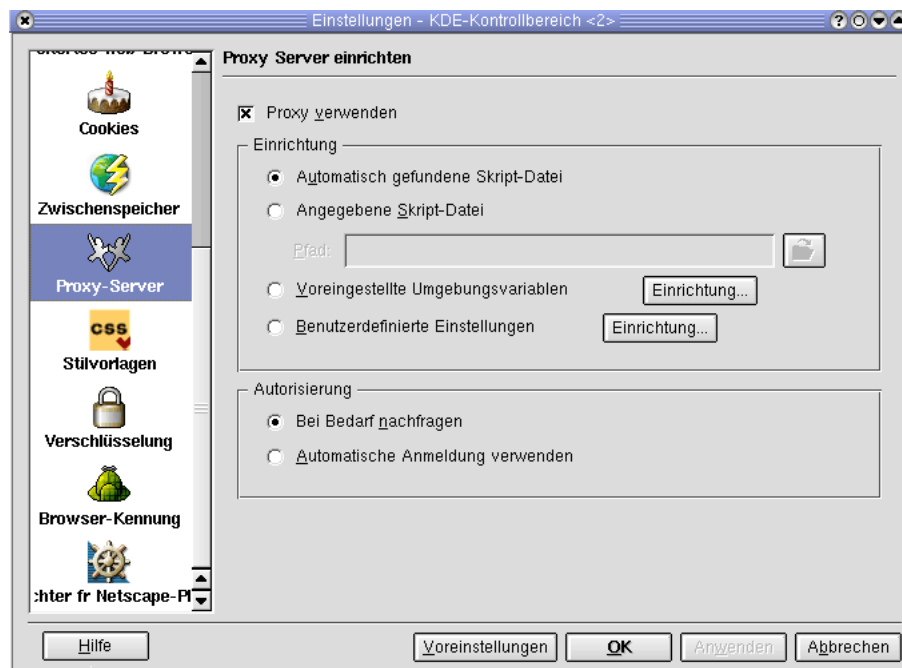
In vielen Firmen ist der Zugriff auf das Internet nur über einen Proxy möglich. Falls Sie ihren Browser selbst konfigurieren wollen, benötigen Sie die IP-Adresse und Portnummer des Proxyservers Ihrer Firma oder Ihres Providers.

Hier wird die Einrichtung des Proxys am Beispiel des Browsers [Konqueror](#) erklärt. Wählen Sie im Konqueror-Hauptmenü **Einstellungen** und dann **Einrichten**.

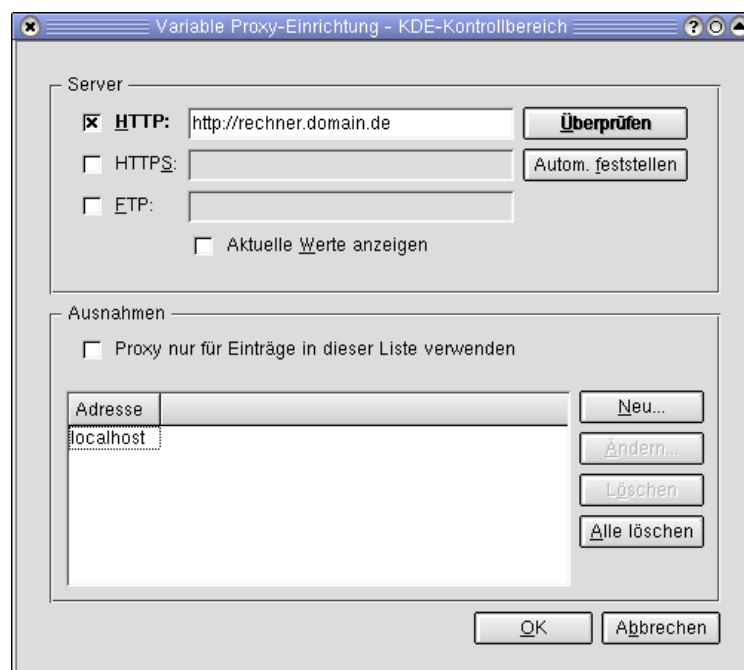


Proxy einrichten beim Konqueror (Hauptmenü)

Aus dem sich öffnenden Menü wählen Sie nun



den Punkt **Proxies**. Nun öffnet sich ein weiteres Fenster,



in dem Sie individuelle Einstellungen vornehmen können.

Auch der Navigator der Firma  [Netscape](#) ermöglicht den Zugriff über einen Proxy. Die entsprechenden Einstellungen werden im Menü **Bearbeiten** unter **Einstellungen** vorgenommen. In dem sich dann öffnenden Fenster findet man in der Kategorie **Erweitert** den Eintrag für die Konfiguration des Proxys. Zu detaillierteren Einstellungen gelangt man über die Schaltfläche **Anschaun**. Hier kann man zum Beispiel die Rechner angeben, die ohne Proxy erreicht werden können.

2.2 Dienste, die im Internet angeboten werden

Praktisch alles, was auf Computern machbar ist, kann auch im Internet gemacht werden. Allerdings ist nicht alles sinnvoll, und so haben erst wenige Dienste eine weite Verbreitung gefunden.

- * Dazu gehört an erster Stelle der älteste Dienst, nämlich die elektronische Post, kurz *E-Mail*. Diese unterscheidet sich lediglich durch das Übertragungsmedium von der herkömmlichen physischen Post (**snailmail**).
Über E-Mail ist der Zugang zu Mailinglisten möglich, einer Möglichkeit für Gruppendiskussionen im Internet.
- * Den zweithäufigsten Dienst stellt die Übertragung von Daten in Form von Dateien dar. Dazu wird das File Transfer Protocol (FTP) verwendet.
Für die Suche nach Dateien auf **FTP-Servern** sind die Dienste **Gopher**, **Archie** und **Veronica** entwickelt worden.
- * Erst an dritter Stelle folgt die Übertragung bzw. Darstellung komplexer Dokumente mit Text, Bild, Ton usw. Dies ist das so genannte World Wide Web, das zwar nicht der am meisten genutzte, wohl aber der populärste Dienst ist und dem Internet zu seinem aktuellen Boom verholfen hat.
- * Weitere Dienste sind der **News-Dienst** für Diskussionsforen, das Sound- und Video-Streaming, Telnet für das Anmelden und Arbeiten auf entfernten Rechner, Internet Relay Chat (IRC) für die Online-Kommunikation, diverse online Multiplayer-Spiele und viele andere.
- * Mit dem **Whois-Dienst** kann man feststellen, wem eine bestimmte Internetadresse gehört oder ob eine gesuchte Adresse noch frei ist.

3 Das World Wide Web

3.1 Web Browser und das WWW

Das World Wide Web wird heute häufig mit dem Internet gleichgesetzt. Das zentrale Werkzeug für diesen Dienst ist der Browser. Browser werden von unterschiedlichen Firmen und Organisationen angeboten - z. B.  [Microsoft](#),  [Netscape](#),  [Opera](#),  [Mozilla](#),  [Firefox](#),  [WWW-Consortium](#) - und stehen in verschiedenen Varianten zur Verfügung, z. B. als reiner Text-Browser, der keine Bilder darstellt und deshalb wesentlich schneller ist (ein Beispiel ist  [Lynx](#)). Der Browser öffnet das Internet gewissermaßen für die massenhafte Nutzung.

Neben der ansprechenden Darstellung der Dokumente wird mit dem Browser der Zugriff auf andere Rechner und die dort liegenden Dokumente über Links erheblich vereinfacht. In diesem Zusammenhang wurden verschiedene Möglichkeiten entwickelt, Dienste, Rechner und Elemente auf einzelnen Rechnern zu adressieren. Diese Dienste haben die Handhabung vernetzter Rechner erheblich erleichtert.

Eine solche Adressangabe wird als URL (Uniform Resource Locator) bezeichnet. Dabei werden sowohl der Dienst als auch der Rechner angegeben, mit dem man eine Verbindung herstellen möchte. Optional kann noch eine Portnummer angegeben werden, unter der ein Dienst angesprochen werden kann (zum Beispiel Port 80 für den WWW-Dienst mit `http://rechnername:80`). Die folgende Tabelle listet die möglichen URL- Dienste bzw. Verlinkungsmöglichkeiten auf.

Verlinken von

Webseiten
verschlüsselte Webseiten
EMail
Newsgruppe
FTP
Gopher
Telnet

Aufruf in einer URL

`http://rechnername/Seite`
`https://rechnername/Seite`
`mailto:emailadresse`
`news:newsgruppe`
`ftp://rechnername/Datei`
`gopher://rechnername/`
`telnet://rechnername/`

3.2 Was ist ein Browser?

Ursprünglich war das Word Wide Web eine  [Präsentationsplattform](#) für wissenschaftliche Forschungsergebnisse. Dabei lag der Schwerpunkt zunächst auf den Texten, die auf jedem Rechner lesbar und recherchierbar sein sollten. Hierfür wurde eine Markierungsmöglichkeit entwickelt, die die wichtigen inhaltlichen Elemente der Dokumente mit bestimmten (Text-) Auszeichnungen versah. Wie diese Elemente auf den verschiedenen Computern dargestellt werden, sollte diesen überlassen bleiben, da die Umsetzung von den Fähigkeiten des jeweiligen Ausgabegeräts abhängig ist.

Später kamen Möglichkeiten zur Einbindung von Grafiken, Bildern und Formularen hinzu.

Als Darstellungsprogramm (Client) für diesen Dienst wurde der Browser entwickelt. Er soll das Material auf dem Computerbildschirm optimal darstellen. Deshalb muss sich der Browser um den Zeilenumbruch in dem darzustellenden Texten kümmern, der von der Anzahl der Zeichen abhängt, die in einer Zeile dargestellt werden können. Somit sind zwar die Browser naturgemäß rechnerabhängig, nicht aber die Dokumente. (Die einzige Ausnahme von dieser Regel ist, dass auf reinen Textbildschirmen, z.B. bei Großrechnern, keine Bilder dargestellt werden können.)

3.3 Hypertext Markup Language

Für die **rechnerunabhängige** Darstellung von Texten wurde die *Hypertext Markup Language* (HTML) entwickelt. Mit dieser Auszeichnungssprache werden die Überschriften, der Autor, die verschiedenen Absätze, Tabellenreihen, Tabellenfelder, Bildelemente usw. im Dokument gekennzeichnet. Anders ausgedrückt: Mit HTML wird der Inhalt ausgezeichnet, nicht die äußere Form.

Das Wort **"Hypertext"** bedeutet in dieser Auszeichnungssprache, dass einzelne Elemente verknüpft sind und ein direkter Wechsel zu diesen Elementen möglich ist. Man nennt diese Elemente in einem HTML-Dokument **Links** oder **Hyperlinks**.

HTML ist eine Sprache mit sehr engen Grenzen, denn ihr Wortvorrat und die Benutzungsregeln (**Syntax**) sind stark begrenzt und können natürlich nur auf internationaler Ebene (bzw. durch Marktmacht) geändert werden. Mit Hilfe der HTML-Auszeichnung erstellt man einen leeren Dokumentrahmen, der erst durch das Hinzufügen von Daten zu einem Dokument wird.

Der Standard der Sprache **HTML** wird vom **World Wide Web Konsortium (W3C)** gepflegt und erweitert, wie auch andere standardisierte Sprachen, die sich mit der Zeit rund um das Netz entwickelt haben (z.B. **XML** oder **Cascading Stylesheets**).

Wer selbst eine Internet-Site realisieren möchte, dem sei als Informationsquelle zur Programmierung von Web-Seiten das **SelfHTML-Tutorial** von *Stefan Münz* empfohlen. Als Anleitung für den Betrieb eines eigenen Webserver empfiehlt sich das Selflinux Kapitel zum **Apache Webserver**.

3.4 Welche Browser gibt es?

Eigentlich sind der Aufbau und die Elemente, die in einer Web-Seite verwendet werden, vom **WWW Konsortium (W3C)** festgelegt. Die meisten Hersteller haben jedoch eigene Erweiterungen hinzugefügt. An erster Stelle sind hier **Microsoft** und der Internet Explorer zu nennen. (Im Extremfall können HTML-Dokumentationen nicht mit anderen Browsern dargestellt werden!) Aber auch der Netscape-Browser hat bis zur Version 4.XX eigene nicht standardisierte Erweiterungen eingeführt. Neuere Netscape-Versionen (ab Version 6) sind nun standardkonform, da die Rendering-Engine von Netscape auf dem Mozilla Projekt basiert.

Alle anderen Browser versuchen den einen oder anderen (privaten) Standard zu unterstützen. Im Einzelnen stehen die folgenden Browser zur Verfügung:

- * **Opera** versucht sich als dritte Kraft unter den Browsern zu etablieren. Dies ist ein kleiner und flinker Browser, der zudem kostenlos ist, im Gegensatz zu früheren Versionen seit 8.5 sogar ohne Werbefbanner.
- * **Mozilla**, ursprünglich eine Weiterentwicklung aus dem Quellcode des Netscape-Browsers.
- * **Firefox**, ein schlanger Ableger des Mozilla-Browsers, der mit vielen **XUL-Anwendungen** erweitert werden kann.
- * **Lynx** und **Links** sind Browser, die nur Text darstellen.
- * Das W3C stellt den Referenz-Browser **Amaya** zur Verfügung.
- * **Konqueror** ist der Browser des **KDE Projekts**
- * **Galeon** ist der Browser zur **GNOME** Oberfläche

3.5 Erweiterungen der Browser

Neben den Erweiterungen einzelner Hersteller zum HTML-Sprachstandard in Form zusätzlicher Elemente, stellen Script-Sprachen eine wesentliche Browser-Erweiterung dar. Diese Scripte werden auf dem Rechner des Benutzers ausgeführt. Sie haben also prinzipiell Zugriff auf den Rechner des Internetbenutzers und können

Informationen gewinnen und weiterleiten.

PlugIns sind hingegen Softwareprogramme, die den Browser um die Fähigkeit erweitern, bestimmte Datenformate darzustellen. Wie der Name suggeriert, werden PlugIns über eine vordefinierte Schnittstelle des Browsers **eingeklinkt**.

3.5.1 Script-Sprachen

HTML Seiten sind eigentlich statisch und zur Laufzeit unveränderbar. Scriptsprachen erweitern die Möglichkeiten von Webseiten um interaktive Inhalte. Script-Sprachen erfüllen verschiedene Aufgaben: Sie sorgen dafür, dass eine Seite abhängig von den Einstellungen des Anwenders optimal dargestellt wird. Es können aber auch Rückmeldungen an den Server generiert werden, z. B. um den Browser des Anwenders zu steuern. Außerdem kann man mit diesen Sprachen den Inhalt von Formularen auf Plausibilität hin überprüfen, bevor diese an den Server übertragen werden.

Die beiden wichtigsten Sprachen die in Browsern implementiert sind:

- * JavaScript
- * Java

Obwohl beide Script-Sprachen den Namensbestandteil *Java* führen, haben sie nichts gemeinsam, außer dass beide vom Browser ausgeführt werden. Auch wenn beide den Namensbestandteil *Java* haben, haben sie nichts gemeinsam, außer dass sie vom Browser ausgeführt werden.

 [Java](#) ist eine von der Firma  [SUN](#) entwickelte plattformunabhängige Programmiersprache. Bei der Entwicklung wurde großen Wert darauf gelegt, dass Java in einer geschützten Umgebung läuft und, wenn möglich, keinen Zugriff auf den Gastrechner hat, insbesondere keinen Schreibzugriff. Dieser als **Sandbox** bezeichnete, geschützte Raum stellt eine wichtige Sicherheitsfunktion dar.

3.5.2 PlugIns

PlugIns stellen Erweiterungen für neu entwickelte Datenformate dar. Damit können z. B. animierte Grafiken oder kontinuierliche Video- oder Soundstreams dargestellt werden. Hier einige Beispiele häufig verwendeter PlugIns:

Flash

wird häufig für animierte Grafiken verwendet und stammt von der Firma  [Macromedia](#).

Sound und Video (Real Audio)

Für das kontinuierliche Abspielen so genannter Streams werden häufig  [Real Audio PlugIns](#) verwendet. Im Gegensatz zu anderen Grafiken und Sounds, die vor dem Abspielen vollständig auf den Rechner übertragen werden müssen, können Streams sofort, also direkt während der Übertragung, abgespielt werden. Alternativ lässt sich für das Streaming von Sound-Daten auch das MP3 Format einsetzen.

Word/Excel/PDF Viewer

Mit entsprechenden Hilfsprogrammen für die Browser können diese Daten direkt im Browser-Fenster dargestellt werden. Sie müssen dazu nicht in HTML-Dokumente umgewandelt werden.

ActiveX-Controls und VBA-Makros (nur im Internet Explorer)

Hierbei handelt es sich um Elemente von Microsofts Programmiersprache Visual Basic. Diese können dynamisch eingebunden werden und haben im Prinzip alle Möglichkeiten, die auch die zugrunde liegende Programmiersprache hat. Darüber hinaus ermöglichen sie den problemlosen Zugriff auf andere Programme des

Office-Paketes.

Im Rahmen der Viren- und E-Mail-Attacken der letzten Jahre haben diese Elemente eine (traurige) Berühmtheit erlangt.

3.5.3 Verschlüsselung

Die meisten Web Browser erlauben eine sichere Datenübertragung. Dabei werden die ausgetauschten Daten nicht mehr als Klartext, sondern in verschlüsselter Form übertragen.

Hierfür wird ein so genanntes [Public-Key-Verfahren](#) benutzt. Dabei hat jeder Rechner einen privaten und einen öffentlichen Schlüssel. Die öffentlichen Schlüssel werden zwischen Server und Browser ausgetauscht. Um zu gewährleisten, dass der Schlüssel des Servers tatsächlich von diesem Server stammt, muss er von einer anerkannten Zertifizierungsfirma signiert sein. Dann gilt die alte Regel, dass "meines Freundes Freund auch mein Freund ist": Vertraut der Anwender der Zertifizierungsfirma, so vertraut er auch dem Server. Bei der aktuellen Übertragung werden die Daten dann mit einem Schlüssel verschlüsselt, der mit dem privaten und dem öffentlichen Schlüssel des anderen Rechners verschlüsselt wurde.

Die verschlüsselte Übertragung ist natürlich bei Geschäften über das Internet (*E-Commerce*) unumgänglich. Nur damit können die für das zu tätige Geschäft notwendigen Daten, z. B. Adresse und Kontonummer, vor den Augen Dritter verborgen übertragen werden.

3.5.4 Cookies

Normalerweise ist die Beziehung zwischen Web Browser und Client (relativ) anonym. Dies bedeutet, dass sich jeder Anwender ausgeben kann, wofür er möchte. Beim normalen Surfen ist dagegen auch nichts einzuwenden.

Allerdings lassen sich auf diese Art keine Geschäftsbeziehungen aufbauen, bei denen der Empfänger nur genau das erhalten will, was er bestellt, und der Lieferant einen identifizierbaren Kunden braucht, an den er die Ware schickt und der diese auch bezahlt.

Um eine eindeutige Beziehung zwischen Web Server und Client zu gewährleisten, werden Cookies (engl. für "Plätzchen") verwendet. Ein Cookie besteht aus einer Kennung sowie einer Verfallszeit, die der Server generiert und an den Browser sendet. Dieser wiederum speichert die Cookie-Daten und den Namen des sendenden Servers und überträgt den Inhalt des Cookies bei jedem erneuten Kontakt wieder an den Server. Nun kann der Server in einer Datenbank nachsehen, welche Daten er über diese Beziehung bereits gespeichert hat und sich auf diese Daten beziehen.

Leider werden Cookies nicht nur im Rahmen von E-Commerce eingesetzt. Es versuchen alle möglichen Server, den Anwendern einen solchen "Keks" unterzuschieben. Über diese Vorgänge sollte man sich als Internetnutzer im Klaren sein, denn wenn jemand gezielt und langfristig Daten über einen (zwar noch anonymen) Surfer sammelt, kann er damit doch einiges über dessen Gewohnheiten erfahren.

3.5.5 Sinnvoller Schutzmechanismus gegen Cookies

TIPP:

Möchten Sie die Annahme von Cookies generell verweigern, ohne dies in jedem Einzelfall im Browser bestätigen zu müssen, aktivieren Sie den Schreibschutz für die entsprechende Datei (z.B. [~/netscape/cookies](#)). Damit entfernen Sie alle Schreibrechte und können nun in Ihrem Browser einstellen, dass alle Cookies angenommen werden dürfen. Der Browser nimmt die Cookies dann an, kann sie aber durch den Schreibschutz nicht in die entsprechende Datei schreiben.

3.6 Suchen und Finden

3.6.1 Informationsbeschaffung im Internet

Das Internet bietet auch Zugang zu Informationen, deren Beschaffung sonst sehr teuer, umständlich oder sogar unmöglich ist. Praktisch zu allen Themen stehen im Internet Informationen kostenlos zur Verfügung.

Hier einige Beispiele:

- * Patentrecherchen in Deutschland, ganz Europa oder weltweit
- * Diskussionsforen zu allen denkbaren Themen
- * Handelsregistrauszüge, andere Firmeninformationen, Konkurrenzanalysen
- * deutsche Gesetze, Gerichtsurteile und Grundsatzentscheidungen
- * Telefonnummernsuche, z.B. nach Name oder Branche
- * Wörterbücher, Lexika
- * Fahrplanauskunft inklusive Umsteigen, Fahrtdauer, Preis und Reservierung
- * aktuelles Fernsehprogramm
- * Musikinformationen von Abba bis Zappa einschließlich Video oder Hörprobe
- *  [Diskussionsgruppen](#) zu fast allen vorstellbaren Themen (z.B. "Sammler von Blechspielzeug", "Direct Marketing", "Tariferhöhung Telekom", "Body Painting" u. v. a.)
- * alles, was mit Computern zu tun hat
- * Forschungsergebnisse aus praktisch allen Wissenschaftsgebieten

Diese Informationen liegen natürlich nicht "auf der Straße", sondern man muss nach ihnen suchen. Dafür gibt es verschiedene Möglichkeiten.

- * Man kennt eine Internetadresse (oder liest sie z. B. in einer Zeitung/Zeitschrift) zu einem bestimmten Thema und kommt über Links auf dieser Seite an die gewünschten Informationen.
- * Man kennt eine Firma/Hochschule/Institution, die das Thema bearbeitet. Unter der Internetadresse dieser Organisation findet man einen Link zu den gesuchten Informationen. (Mittlerweile bieten viele dieser Web-Seiten Suchfunktionen.)
- * Man befragt einen Auskunftsdienst im Internet.

Man unterscheidet zwei Arten von Auskunftsdiensten:

- * Kataloge, die redaktionell bearbeitete Informationen anbieten
- * Suchmaschinen, die die Informationen unbewertet (und oft auch unstrukturiert) weitergeben.

3.6.2 Kataloge

In den  [Katalogen](#) werden die Web-Seiten von Redakteuren geprüft, bewertet und katalogisiert. Der Vorteil dieses Verfahrens besteht darin, dass es eine nachvollziehbare Struktur ergibt. Allerdings ist der Aufwand dafür nicht unerheblich, sodass der Umfang dieser Kataloge beschränkt ist.

Bekannte Kataloge sind u. a.:

 www.dmoz.de,  www.dmoz.org  www.yahoo.de,  www.yahoo.com  www.web.de 
www.lycos.de,  www.lycos.com

3.6.3 Suchmaschinen

Unter  [Suchmaschinen](#) versteht man Programme (Crawler), die in regelmäßigen Abständen (etwa alle 2

Monate) ihnen bekannte Web-Seiten und darauf gefundene Links durchsuchen und die auf diese Art ermittelten Informationen in Datenbanken für öffentliche Recherchen zur Verfügung stellen.

Man bezeichnet auch die Web-Seiten, auf denen diese Programm bereitgestellt werden, sowie die Organisationen, die diese Seiten betreiben, als Suchmaschinen.

Natürlich kann es keine Suchmaschine geben, die alle Seiten des Internets enthält, schon allein deshalb, weil die Erfassung aller Seiten mehrere Jahre in Anspruch nähme und wegen der ständigen Veränderungen trotzdem nie abgeschlossen werden könnte. Man sollte also durchaus verschiedene Suchmaschinen zu einem Thema befragen, denn jedes dieser Suchprogramme kennt andere Seiten und wertet deren Inhalt mit anderen Methoden aus.

Stellt man einer Suchmaschine mittels Web-Formular eine Frage, so ist dies letztlich eine Datenbankabfrage: Kennt diese Datenbank Seiten, die das gesuchte Stichwort enthalten? Alle Datenbanken, auch die der Suchmaschinen, erlauben auch komplexere Abfragen. Die Regeln zur Formulierung solcher Anfragen werden in der Regel auf den Hilfeseiten der Suchmaschinen erklärt.

Bekannte Suchmaschinen sind u. a.:

 www.google.de,  www.google.com
 www.hotbot.com
 www.excite.de,  www.excite.com
 www.fireball.de
 www.altavista.de,  www.altavista.com

Diese Suchmaschinen sind thematisch nur grob oder überhaupt nicht sortiert und haben den Nachteil, dass sie nicht absolut aktuell sein können. Deshalb gibt es spezielle Kataloge und Suchmaschinen, die bestimmte Web-Seiten häufiger indizieren oder speziellen Themen gewidmet sind.

3.6.4 Metasuchmaschinen

Neben den oben beschriebenen Suchmaschinen gibt es noch die  [Metasuchmaschinen](#). Dabei handelt es sich um Suchmaschinen, die Suchmaschinen abfragen, also Suchanfragen gleichzeitig an mehrere andere Suchmaschinen weiterleiten. Meistens werden die Ergebnisse bereinigt, also von doppelten Fundstellen befreit.

Einige Metasuchmaschinen:

 www.metacrawler.de,  www.metacrawler.com
 www.metager.de
 www.metaspinner.com
 www.vivisimo.com
 www.suchen.com

Soll ihre Firmen- oder private Webseite von Suchmaschinen erfasst werden, so müssen Sie die Internetadresse der Webseite bei der/den Suchmaschinen anmelden. Dies geschieht in einem WEB-Formular auf einer weiteren Seite der Suchmaschine.

3.6.5 Website bei Suchmaschinen anmelden

Soll Ihre Firmen- oder private Web-Seite von Suchmaschinen erfasst werden, so müssen Sie die Internet-Adresse der Web-Seite bei den gewünschten Suchmaschinen anmelden. Dies geschieht in der Regel über ein Web-Formular auf einer gesonderten Seite der Suchmaschinen-Website. Eine solche Anmeldung ist für jede einzelne Suchmaschine erforderlich.

Mehr detaillierte Informationen zum Thema Suchmaschinen finden Sie auf der Website  www.suchfibel.de.

3.7 Sicherheit im Internet

3.7.1 Gefahren des Internets

Das Internet ist kein reiner Segen, sondern hat auch seine Schattenseiten. Hierzu gehört zum Beispiel die unermessliche Informationsvielfalt ("Information Overkill"), in der man sich nicht immer leicht zurechtfindet. Ein weiteres Problem besteht darin, dass der Wahrheitsgehalt der angebotenen Informationen häufig nicht erkennbar oder überprüfbar ist.

Im Zusammenhang mit der Informationsvielfalt besteht auch die Gefahr, dass der Anwender sein eigentliches Suchziel aus den Augen verliert, während er sich von Link zu Link hangelt. Dieses Phänomen wird häufig mit "verloren im Internet" (**lost in cyberspace**) beschrieben.

Die  [Internet-Sucht](#) ist mittlerweile eine anerkannte  [Krankheit](#). Dabei verlieren die Süchtigen ihr Geld an die Telefongesellschaften, weil sie ihre gesamte Freizeit mit **Chatten, Spielen** oder **Surfen im Internet** verbringen.

Auf andere Art gefährlich sind die Möglichkeiten, das Internet dafür zu missbrauchen, die Gewohnheiten der Anwender auszuspionieren. Dafür werden Cookies (vgl. oben) zwischen Web Server und Browser ausgetauscht. Selbst der Industriespionage dient das Internet bereits als Hilfsmittel.

Aus dem Internet stammen derzeit die meisten [Computerviren](#). Traurige Berühmtheit haben die Viren, [Trojanischen Pferde](#),  [Hoaxes](#) usw. erlangt, die per E-Mail über das Internet verbreitet werden.

Viren, auch Makroviren, sind unter Linux zurzeit noch kein  [Problem](#). Allerdings wurden bereits Viren entwickelt, die auch unter Linux funktionieren. Da diese jedoch die Grenzen der Benutzerrechte in der Regel nicht überspringen können, sollte man zum Surfen im Internet **nie** als [Superuser mit dem Benutzernamen root](#) angemeldet sein. Das Bundesamt für Sicherheit in der Informationstechnik bietet online eine  [Broschüre](#) zum Thema Viren.

In letzter Zeit haben sich die so genannten **0190-Dialer** unter Windows als Gefahrenquelle erwiesen. Dabei werden Programme zur Installation angeboten, die einen schnelleren und besseren Zugang zu einem bestimmten Web-Angebot versprechen. Genehmigt der Anwender das Herunterladen dieser Software, so wird auf seinem Rechner ein Programm gespeichert und ausgeführt, welches das Einwahlprogramm von Windows so umkonfiguriert, dass der Internet-Zugang über eine 0190-Nummer erfolgt. Auf die meist sehr hohen Gebühren für die Nutzung dieser Nummern wird in der Regel nicht hingewiesen, und die Anwender merken häufig erst zu spät, dass ihnen Online-Gebühren in Höhe von mehreren Tausend Euro entstehen. Erfolgt der Internet-Zugang nicht über eine Modem-/ISDN-Verbindung, sondern über eine Ethernet-, DSL- oder Linux-Anbindung, so hat dieser Wahlmechanismus normalerweise keine Chance.

Die deutsche  [Bundesregierung](#) plante 2003 wegen der anhaltenden Probleme mit 0190-Dialern die entstehenden Kosten bei 0190-Nummern auf einen bestimmten Betrag pro Minute, sowie die Dauer einer solchen Verbindung auf maximal eine Stunde zu begrenzen. Seit dem 01.01.2006 sind 0190-Nummern für Mehrwertdienste durch 0900-Nummern ersetzt worden. Sie bieten eine kostenfreie Ansage, wie viel die angewählte Verbindung kostet, sowie eine Kostenbegrenzung: Eine 0900-Verbindung darf maximal zwei Euro pro Minute oder 30 Euro pro Verbindung kosten.

3.7.2 Wie kann man sich schützen?

3.7.2.1 Generell

- * Lokale Script-Sprachen deaktivieren.
- * Schutz vor Cookies einrichten ( [siehe Anleitung](#)).
- * Möglichst mit einem Rechner surfen, der nicht in ein [Netzwerk](#) eingebunden ist.
- * Zum Surfen nach Möglichkeit keine Windows-Plattform verwenden, da praktisch alle Viren nur unter Windows bzw. Microsoft Office funktionieren.
- * Regelmäßig manuell oder automatisch einen [Virensan](#) durchführen.
- * Möglichst über einen Proxyserver auf das Internet zugreifen.
- * Einsatz einer [Firewall](#)

3.7.2.2 E-Mail-Verkehr

- * Mails von unbekannten Absendern nicht öffnen, ansehen oder lesen.
- * Mails mit dubiosem, nichts sagendem Betreff löschen.
- * Mails, in denen Sie nicht direkt als Adressat genannt sind, löschen.
- * Mails nach Möglichkeit verschlüsseln, z.B. mittels GnuPG.
- * Ein- und ausgehende Mails auf Viren prüfen.

3.7.2.3 Download aus dem Internet

- * Lesen Sie alle Meldungen, bevor Sie sie bestätigen. Damit schützen Sie sich besonders vor Dialern.
- * Lassen Sie nicht zu, dass ein Programm eine Installation eigenständig ausführt, sondern geben Sie immer selbst an, was und wo gespeichert werden soll ("als [Datei](#) speichern"). Probieren Sie dies eventuell auf einem separaten Rechner aus.

3.7.2.4 Anmelden auf Rechnern

Verwenden Sie bei einer Anmeldung über das Internet nicht telnet, sondern unbedingt ein verschlüsselndes Programm wie z.B. ssh oder  [putty](#).

3.7.3 Firewall

Die Einrichtung einer **Firewall** ist in einer Firmenumgebung unerlässlich. Aber auch private Surfer sollten sich mit dieser Technik vertraut machen.

Eine [Firewall](#) ("Brandschutzmauer") kann den Datenstrom mit Hilfe eines Regelwerkes kontrollieren. So kann der Anwender z. B. festlegen, dass Daten über bestimmte Ports nicht angenommen oder versendet werden dürfen, Pakete mit Protokollfehlern verworfen werden müssen oder keine Verbindung von außen nach innen hergestellt werden darf. Die Einrichtung und Pflege eines solchen Programms ist nicht unbedingt einfach, insbesondere wegen der durch die Firewall erzwungenen Einschränkungen. So sind unter Umständen bestimmte Web-Seiten nicht mehr erreichbar, wenn der Web-Server auf einem anderen Port läuft.

Hier sei noch einmal darauf hingewiesen, dass Daten im Internet normalerweise im Klartext übertragen werden. Insbesondere bei E-Mails sollte man diese Tatsache nicht vergessen. Da es keine Person/Institution gibt, die für alle Vorgänge im Internet verantwortlich ist, und die Daten in der Regel über Rechner in verschiedenen Ländern laufen, ist ein umfassender Datenschutz weder machbar noch durchzusetzen. Daher ist die Verschlüsselung wichtiger Informationen unverzichtbar. Für die Verschlüsselung stehen verschiedene Programme zur Verfügung, z. B.  [Pretty Good Privacy](#) (pgp bzw das freie [GnuPG](#)) oder auch das vom  [Bundesamt für Sicherheit in der Informationstechnik](#) bereitgestellte Programmpaket.

Vorsicht ist geboten. Sie sind nun gewarnt!

Für weitere Informationen zum Thema Sicherheit von Computern empfiehlt sich das Kapitel [Grundlagen Sicherheit](#) im SelfLinux Tutorial.