DIGITALE GESELLSCHAFT

WIE DAS INTERNET FUNKTIONIERT

Eine Anleitung für Entscheidungsträger und Interessierte

EUROPEAN DIGITAL RIGHTS

Diese Broschüre soll Entscheidungsträgern und Interessierten einen Überblick über das Internet und Internettechnologien geben. Sie erklärt einige Schlüsseltechnologien des Internets leicht und in verständlicher Sprache. Wir hoffen, dass sie im Dschungel des Computerjargons eine wertvolle Hilfe sein wird.

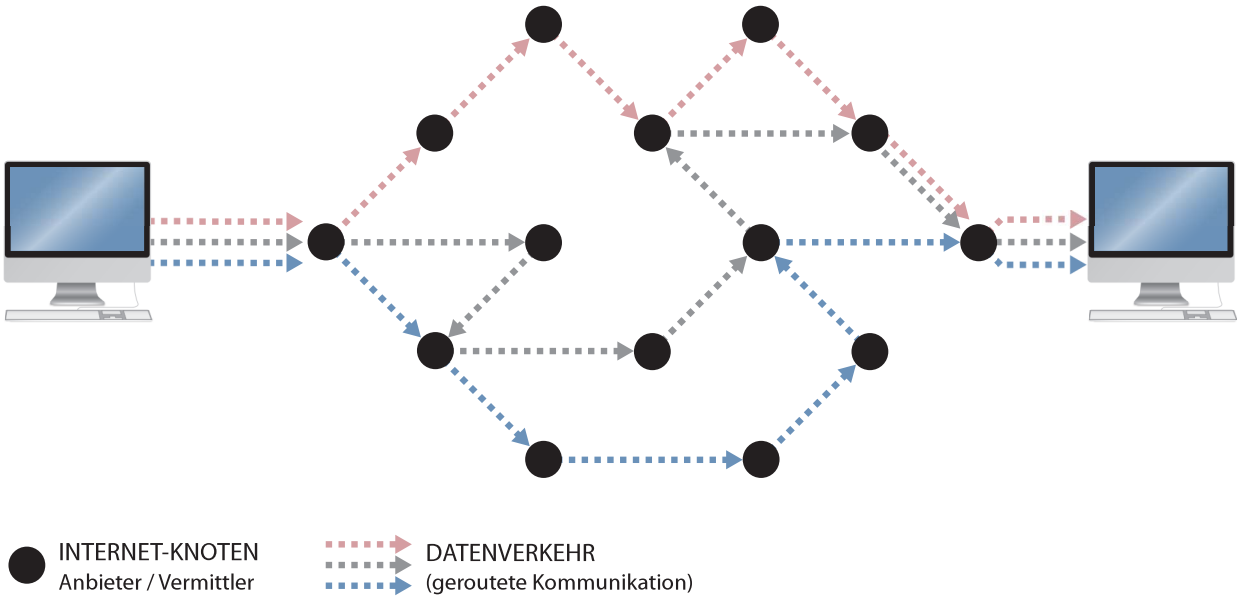
INHALT:

Die englischsprachige Vorlage wurde von Mitarbeitern und Aktivisten von European Digital Rights unter Beteiligung des Digitale Gesellschaft e.V. geschrieben. Das Design stammt von CtrlSPATIE. Übersetzung von Andreas Müller und Kilian Froitzhuber.

Der Druck dieser Broschüre konnte mit Spenden an den Digitale Gesellschaft e.V. finanziert werden. Unterstützen Sie uns bitte mit Ihrer Spende: http://digitalegesellschaft.de/spenden/

DAS INTERNET

EIN NETZWERK AUS COMPUTERNETZWERKEN



Das Internet ist ein globales System aus miteinander verbundenen Computer-netzwerken.

Sobald zwei elektronische Geräte (z.B. Computer) verbunden sind und miteinander kommunizieren können, werden sie Teil eines Netzwerks. Das Internet besteht aus weltweiten Verbindungen solcher Netzwerke, die Firmen, Regierungen oder Privatpersonen gehören. Jedes Gerät kann nun mit jedem anderen kommunizieren.

Damit diese Kommunikation funktioniert, müssen sich die Geräte gegenseitig verstehen können. Im Internet ist das möglich, weil alle Geräte dieselbe Sprache, dasselbe Protokoll benutzen, namentlich das „Internet Protocol (IP)", und zwar ohne Einschränkungen physischer, technischer oder nationaler Natur. Das Internet Protocol bildet die Basis für alle anderen Kommunikationssysteme im Internet.

Daten per Internet Protocol zu senden ist etwa damit vergleichbar, dass einzelne oder mehrere Seiten eines Buches jeweils in Briefumschlägen per Post verschickt würden. Auf allen Umschlägen steht die gleiche Absender- und Empfängeradresse. Auch wenn manche Briefe per Luftpost und andere mit dem Schiff verschickt werden, kommen alle früher oder

später am Ziel an und man kann das Buch dort wieder vollständig zusammensetzen. Es ist dabei egal, ob Seite 1 oder 47 zuerst beim Empfänger eintrifft.

Im Internet wird auch der Inhalt der Umschläge durch bestimmte Konventionen bzw. Protokolle (also ausgehandelte Formate) festgelegt. Für jede Form der Kommunikation gibt es ein eigenes Protokoll. Einige Beispiele für auf dem Internet Protocol basierende Protokolle sind:

- SMTP (Simple Mail Transfer Protocol) für das Versenden von E-Mails
- HTTP für den Zugriff auf Webseiten
- BitTorrent für „Peer-to-peer"-Kommunikation (Eine Methode, um Dateien mit einer großen Gruppe von Menschen zu teilen)

Jeder kann jederzeit eigene Protokolle definieren und über das Internet verwenden, wenn sie auf dem Internet Protocol aufbauen. Anders gesagt: Solange die Adresse auf dem Umschlag im Standardformat steht, ist die einzige Grenze der menschliche Erfindergeist. Diese technische Offenheit macht das Internet zu dem globalen Phänomen, als das wir es heute kennen.

Jede Einschränkung der Offenheit reduziert sein Entwicklungspotenzial. Der universelle Gebrauch eines einzigen Protokolls für alle Kommunikationsformen bringt einige Vorteile mit sich. Die Geräte, die den Datenstrom transportieren (sog. Router), müssen nicht für jede Kommunikationsform neu programmiert werden - sie müssen nicht einmal wissen, was für eine Art von Daten sie transportieren, solange alle das Internet Protocol benutzen. Genau wie ein Postbote müssen sie lediglich auf den Umschlag schauen, um die Post auszuliefern. Es spielt keine Rolle, ob im Umschlag eine Rechnung oder ein Liebesbrief steckt, außer für den Empfänger natürlich.

Diese Technik führt zu:

- Innovationsmöglichkeiten hinsichtlich neuer Anwendungen und Protokolle
- einer im Design verankerten Privatsphäre, da niemand den Inhalt der Kommunikation kennen muss außer Absender und Empfänger
- einem flexiblen und schnellen Datenfluss

Im Kern liefert das Internet nur eine einzige hochflexible Dienstleistung: Daten von einem Gerät auf ein anderes zu transportieren, egal um welche Geräte es sich handelt, wo sie stehen, wie sie mit dem Internet verbunden sind und egal um welche Daten es sich handelt.

Diese Offenheit und Flexibilität ist der Hauptgrund für die Innovationen und den demokratischen und wirtschaftlichen Erfolg des Internets.

"Diese Offenheit und Flexibilität ist der Hauptgrund für die Innovationen und den demokratischen und wirtschaftlichen Erfolg des Internets."

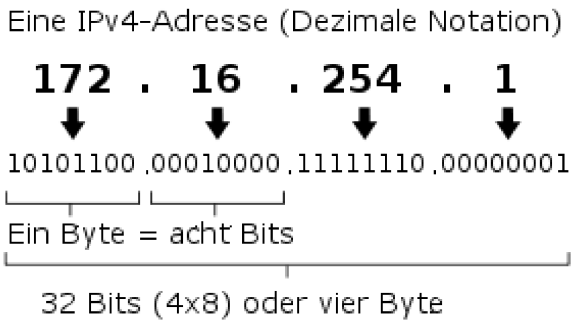
DIE IP-ADRESSE

EINE DIGITALE ADRESSE

Eine IP-Adresse ist eine aus Zahlen bestehende Adresse, die jedem mit dem Internet verbundenen Gerät zugewiesen wird.

In vielen Fällen können IP-Adressen benutzt werden, um Personen oder Unternehmen zu identifizieren, die über einen Service-Provider ein Gerät mit dem Internet verbunden haben. In anderen Fällen - vor allem bei Firmennetzwerken sowie bei öffentlichen oder ungeschützten Drahtlosverbindungen und mobilem Internet - kann eine Handlung im Internet nicht immer einer bestimmten Person zugeordnet werden. Da an herkömmlichen Internetanschlüssen häufig nur eine öffentliche IP-Adresse für die Verbindungen vieler angeschlossener Personen genutzt wird, wird über diese Adresse nur eine Gruppe von Personen statt eines Individuums identifizierbar. Daher ist es oft schwierig oder unmöglich, anhand der IP-Adresse zu bestimmen, wer welche Internetseiten oder Dienste aufgerufen hat.

IP-Adressen sind andererseits in manchen Kontexten personengebunden - dann müssen sie als schützenswerte, personenbezogene Daten behandelt werden.



"Die IP-Adresse identifiziert nicht immer denjenigen, der eine digitale Spur hinterlassen hat"

01 Aufgrund der Knappheit der derzeitigen Generation von IP-Adressen wird es immer üblicher dass eine IP-Adresse geteilt wird – z.B. von allen Computern in einem Büro.

DAS DOMAIN NAME SYSTEM (DNS)

DAS TELEFONBUCH DES INTERNETS

Sobald man eine Internetseite ins Netz stellt, ist sie auch über die IP-Adresse des Servers, auf dem sie gespeichert ist, erreichbar (zum Beispiel hat digitalegesellschaft.de bei Erstellung dieses Artikels die Adresse 46.4.67.52). Leider sind IP-Adressen für den Menschen nicht einfach zu merken. Zudem wäre es nicht praktisch, sie zur Identifizierung einzelner Online-Ressourcen zu benutzen: Internetdienste wechseln ihre IP-Adresse gelegentlich (zum Beispiel wenn sie zu einem neuen Anbieter wechseln).

Da die Benutzung von IP-Adressen also weder praktisch noch benutzerfreundlich ist, wurden die Domain-Namen (wie zum Beispiel „digitalegesellschaft.de") erfunden. Das globale „Domain Name System" (DNS) funktioniert in etwa wie ein Telefonbuch fürs Internet.

Wenn man den Namen der Domain (also z.B. „digitalegesellschaft.de") einer Website eintippt, fragt der Computer das Domain Name System unsichtbar und automatisch nach der IP-Adresse des Servers, auf dem die Seite gespeichert ist. Wenn Sie also „http://digitalegesellschaft.de" im Browser eingeben, erkennt der Computer automatisch die IP 46.4.67.52 und sendet eine Anfrage an diese Adresse, um die Seite zu empfangen.

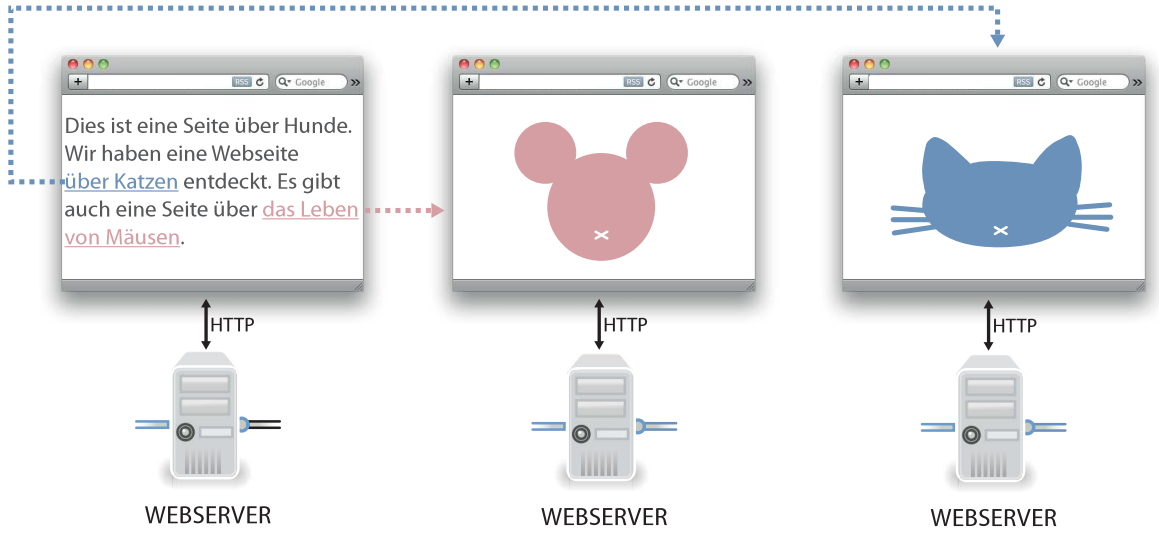
Das System, mit dem Domainnamen abgefragt werden, funktioniert hierarchisch. Sobald Sie „http://digitalegesellschaft.de" eintippen, verbindet sich Ihr Computer zuerst mit einem DNS-Server, um nach der Adresse zu fragen. Dieser Server wird standardmäßig von Ihrem Internet-Provider betrieben, der Zugriff auf alternative Server ist jedoch ebenfalls möglich.

Wenn kurz vor Ihnen schon jemand nach digitalegesellschaft.de gefragt hat, erinnert sich der DNS-Server daran und gibt Ihnen die richtige IP-Adresse zurück. Wenn der letzte Aufruf schon länger her oder die Domain noch unbekannt ist, wendet er sich an die nächsthöhere Instanz, auf der wieder der gleiche Vorgang abläuft. Die höchste Autorität sind hier die 13 „root server" (Stammverzeichnisse), die die Daten aller DNS-Server sammeln. Diese 13 Server sind sehr widerstandsfähig. Sie haben so viel Kapazität, dass sie selbst dann weiterlaufen, wenn sie groß angelegten Angriffen ausgesetzt sind (sog. DDoS, distributed denial of service).

02 Wenn Ihr Computer erst kürzlich auf http://digitalegesellschaft.de zugegriffen hat, dann kennt er die IP-Adresse bereits und braucht sie nicht mehr beim Provider abzufragen.

DAS WORLD WIDE WEB (WWW)

VERLINKTE INFORMATION



Das World Wide Web baut auf dem relativ jungen Protokoll HTTP auf, das wiederum auf dem Internet Protocol (IP) basiert. HTTP steht für HyperText Transfer Protocol und wurde für das Herunterladen von Hypertext-Dokumenten (die wir unter dem Namen „Webseiten" kennen) und das Senden einiger Basisinformationen an den Server entworfen.

Homepages werden mit der „Formatierungs-sprache" HTML (HyperText Markup Language) erstellt. Die Regeln für diese Sprache werden vom „World Wide Web Consortium" (W3C) erstellt und definieren bestimmte Schlüsselwörter, die sich auf Schriftsatz und Layout einer Homepage auswirken. Wenn zum Beispiel Text fett (engl. strong) dargestellt werden soll, steht ein „" davor und ein „" dahinter.

Auch wenn es inzwischen mehrere Versionen dieses Standards gibt (HTML 5 ist die neueste), wird HTML kontinuierlich weiterentwickelt. Jeder kann sich an dem Entwicklungsprozess beteiligen. Auch wenn neue Standards festgelegt werden, steht die Verwendung von HTML unter keiner Lizenz und ist kostenlos. Der Vorteil ist, dass alle Computersysteme HTML-Anweisungen verstehen - so kann jeder die Sprache kostenlos benutzen und sich sicher sein, dass die Homepage auf jedem Gerät angezeigt wird. Das Netz (und die Welt) wäre viel ärmer, wenn jeder für das Schreiben von Websites in verschiedenen Sprachen für verschiedene Computersysteme bezahlen müsste.

Dieser offene und freie Charakter von HTML ist entscheidend, wenn es darum geht, die Kompatibilität von Homepages auf allen möglichen Geräten (z.B. Desktop-Computer, Mobiltelefone, Tablets, Laptops usw.) zu garantieren. Eine richtige Umsetzung des HTML-Standards garantiert auch Sehbehinderten den

Zugang zu Websites. Wenn das nicht so wäre, könnten Vorleseprogramme nicht wissen, wie sie auf die Homepages zugreifen sollen.

Websites werden auf Maschinen veröffentlicht, die man Webserver nennt. Ein Webserver ist ein Computer, der immer unter der gleichen IP-Adresse zu erreichen ist (wie auf Seite 6 beschrieben). Normalerweise sind viele Domainnamen (z.B. www.edri.org und digitalegesellschaft.de) unter einer IP-Adresse zu erreichen, weil sie auf einem Webserver gespeichert ("gehostet") sind. So kann ein einzelner Webserver viele verschiedene Homepages anbieten. Im Falle kommerzieller Webhosting-Anbieter sogar mehrere hundert vollkommen unabhängige Seiten auf einem einzelnen Webserver.

Versuche, einzelne Webseiten auf Basis ihrer IP-Adresse zu blockieren, haben daher immer fatale Nebenwirkungen für alle anderen Seiten, die auf dem gleichen Server gespeichert sind.

Zusätzlich zu HTTP gibt es noch eine sichere Variante namens HTTPS. Normale HTTP-Verbindungen (und zwar sowohl Up- als auch Downloads) sind nicht verschlüsselt. Jeder, der Zugriff auf die Netzwerkkabel oder irgendwelche Geräte zwischen Endnutzer und Webserver hat oder sich in Reichweite eines verwendeten Funknetzwerks (WLan) befindet, hat vollen Zugriff auf alle hin- und hergesendeten Informationen.

HTTPS fügt dieser Verbindung eine Verschlüsselungstechnik hinzu, so dass theoretisch nur der Computer des Nutzers und der Webserver Informationen entschlüsseln können. Dieses System basiert auf Vertrauen: Der Autor einer Website beauftragt eine vertrauenswürdige Partei, ihm ein persönliches digital unterzeichnetes Zertifikat auszustellen, das die Identität des Verfassers bestätigt. Das funktioniert so ähnlich wie die Wachssiegel vergangener Jahrhunderte.

SPRACHE, DIE VON BROWSER-HERSTELLERN UND DEM WORLD WIDE WEB CONSORTIUM ENTWICKELT WIRD

W3C®

Dieser Text wird FETT geschrieben

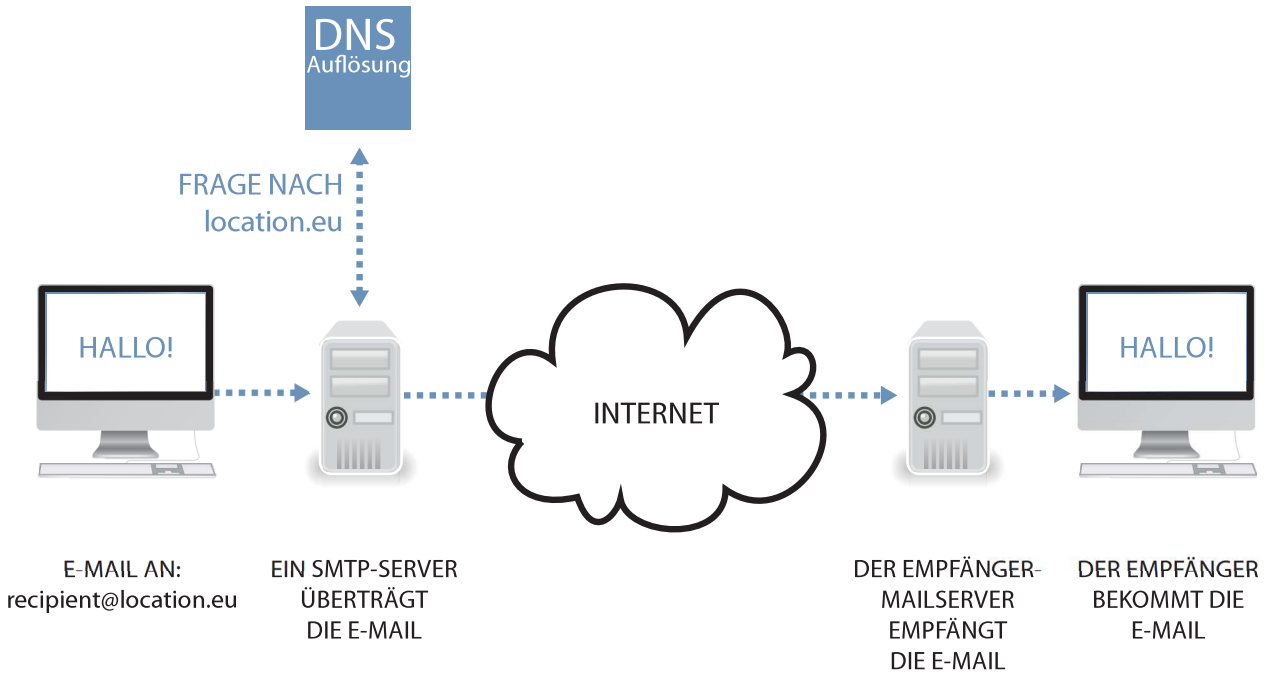
WIE ES DIE ENTWICKLER BENUTZEN

WAS SIE SEHEN

Wenn ein Nutzer einen Browser (so wie Internet Explorer, Firefox, Chrome, Safari oder Opera) installiert, wird diesem vom Hersteller eine Liste vertrauenswürdiger Zertifizierungsstellen mitgeliefert. Auf deren Integrität muss der Nutzer sich verlassen können - was zugleich die Schwachstelle von https ist: Wenn sich nur eine der Dutzenden Stellen als nicht vertrauenswürdig herausstellt, hat das https-Sicherheitskonzept ein riesengroßes Loch.

E-MAIL UND E-MAIL-SICHERHEIT

POST IN DER DIGITALEN GESELLSCHAFT



Wenn man in einem Mailprogramm oder auf einer Webmail-Seite eine E-Mail geschrieben hat, wird sie zuerst via SMTP an einen Ausgangs-Server übertragen. Von dort wird sie von einem E-Mail-Server zum nächsten wiederum via SMTP übertragen, und zwar so lange, bis sie beim Ziel-Server angekommen ist.

E-Mail-Server finden das Ziel einer E-Mail durch eine Abfrage im oben beschriebenen Domain Name Service (DNS). Hier ist auch die Information eingetragen, welcher Server E-Mails für welche Domain annimmt. Die Domain ist dabei immer der Teil der E-Mail-Adresse hinter dem @-Zeichen.

Ist die E-Mail beim Zielserver, der alle eingehenden E-Mails für diese Domain verwaltet, angekommen, bleibt sie dort gespeichert bis der Empfänger sie löscht. Einige Mailprogramme erledigen dies automatisch, sobald die E-Mails auf den PC oder das Smartphone des Empfängers heruntergeladen wurden.

E-MAIL-SICHERHEIT

E-Mails können von Dritten abgefangen werden, während sie von einem E-Mail-Server zum nächsten unterwegs sind. Es gibt zwei Wege, das zu verhindern: Sichere Kommunikation zwischen den Mailservern oder die Verschlüsselung

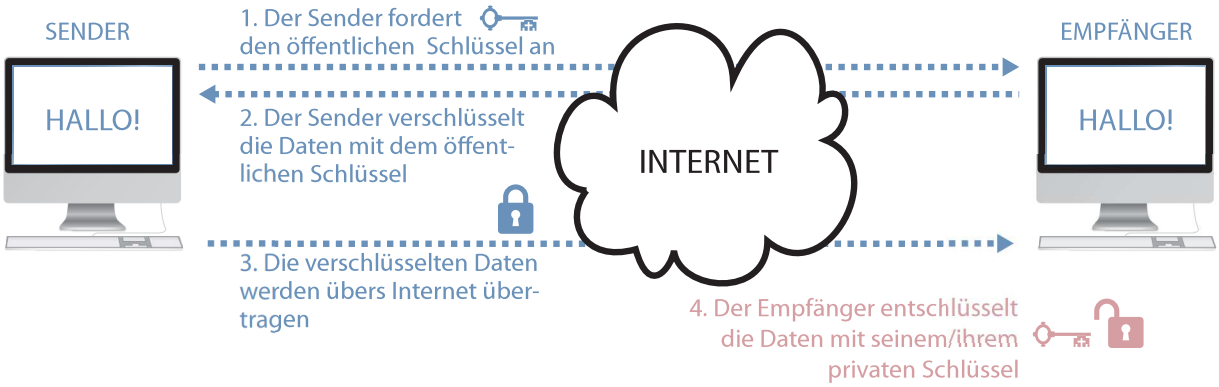
des Inhalts der E-Mails. Die Absicherung der Kommunikation zwischen E-Mail-Servern funktioniert auf die gleiche Weise wie das oben beschriebene HTTPS.

Eine Schwachstelle bei E-Mails ist, dass Ihr Computer nicht direkt mit dem Zielserver kommunizieren kann. Dadurch könnten Ihre E-Mails immer noch abgefangen werden, wenn eine der Zwischenstationen keine verschlüsselte Verbindung verwendet.

Deswegen ist es besser, die Nachricht selbst zu verschlüsseln. Eine beliebte und kostenlos verfügbare Methode dafür ist PGP (Pretty Good Privacy), das auch unter den Namen OpenPGP und GPG verfügbar ist.

VERSCHLÜSSELUNG

PRIVATSPHÄRE IM ÖFFENTLICHEN NETZ



Wie kann ein Benutzer eine vertrauliche Nachricht so verschicken, dass sie vor neugierigen Blicken geschützt bleibt? Wenn Sie einen Brief verschicken, könnte er abgefangen, geöffnet, gelesen und wieder versiegelt werden, ohne dass Sie es bemerken. Auch Telefongespräche können mitgehört werden.

Die rasante Entwicklung der Kryptographie im 20. Jahrhundert folgte der Verbreitung von Computertechnologie. Computer ermöglichten nicht nur die sehr schnelle Verschlüsselung elektronischer Nachrichten, sondern auch das sehr schnelle Knacken (“Cracken”) der bisher benutzten Verschlüsselungsverfahren.

Verschlüsselung ist kein Allheilmittel und garantiert keine vollkommene Sicherheit. Eine häufige Herangehensweise, Verschlüsselungen auszutricksen, ist die Nachricht abzufangen noch bevor sie verschlüsselt wird. Dies geschieht zum Beispiel mit einem versteckten „Trojanischen Pferd” auf dem Computer des Verfassers, das alle Tastaturanschläge auf dem PC oder Handy der Betroffenen mitprotokolliert. Nachrichten müssen immer vollständig übertragen werden, um ihre Verschlüsselung zu schützen. Ansonsten kann der Inhalt manipuliert werden, und zwar auch ohne den Schlüssel selbst zu kennen. Die meisten Verschlüsselungswerkzeuge kümmern sich deshalb automatisch um die Vollständigkeit der Nachrichten.

Das folgende Bild zeigt die Stufen der sog. Public-Key-Verschlüsselung, die mit einem Schlüsselpaar aus öffentlichem und privatem Schlüssel funktioniert.

1. Der Sender fordert eine Kopie des öffentlichen Schlüssels beim Empfänger der Nachricht an.

2. Eine geeignete Software verschlüsselt die Nachricht an den Empfänger mit dessen öffentlichem Schlüssel.

3. Die Nachricht wird verschickt und sieht für Außenstehende auf den ersten Blick aus wie „Kauderwelsch”.

4. Der Empfänger entschlüsselt die Nachricht mit einer Kombination aus öffentlichem und privatem Schlüssel.

DEEP PACKET INSPECTION

WERFEN WIR EINEN BLICK IN IHREN INTERNETVERKEHR

Daten werden im Internet in sogenannten Paketen (packets) verschickt. Jedes Paket hat eine Kopfzeile (header), in der Herkunft und Zielort stehen, wie bei einem normalen Postpaket. Diese Informationen sind nötig, damit die Internetknoten den im Moment besten Weg für die Daten finden können.

Früher schaute das Netzwerk nur auf Herkunft und Empfänger eines Pakets. Aber mit der rasanten Zunahme an schädlichen Aktivitäten entschieden die Eigentümer der Netzknoten, die Details jedes Pakets näher zu betrachten, um „sichere" Pakete von solchen unterscheiden zu können, die zu einer Hacker- oder „Denial-Of-Service"-Attacke gehören.

Bestimmte Netzwerksicherheitsprogramme, genannt „Firewalls", können zum Beispiel ein einzelnes Paket blockieren, das von einem bestimmten Absender zu einem bestimmten Empfänger für eine bestimmte Anwendung unterwegs ist. Wenn Sie diese Kriterien richtig einsetzen, können Sie so alle eingehenden Daten zu ihrem Büronetzwerk blockieren, da sie der Öffentlichkeit keine Internet-Dienste vom Büro aus zur Verfügung stellen. Und Sie können dabei immer noch alle anderen Facetten des Internets nutzen, indem Sie die Pakete durchlassen, die von Ihrem Büro nach Außen gehen.

Irgendwann werden Sie vielleicht auf die Idee kommen, in Ihrem Büro einen Webserver zu installieren, um Dokumente zu veröffentlichen. Dann müssten Sie Ihre Firewall so konfigurieren, dass eingehende Anfragen durchgeleitet werden, aber nur die, die Webseiten anfordern. Hier gibt es aber einige Angriffsmöglichkeiten durch Pakete, die für eine Firewall harmlos aussehen. Nur auf Basis von Informationen wie Herkunft und Zielort kann sie nicht unterscheiden, ob ein Paket harmlos ist oder nicht.

Netzwerktechniker kamen bald auf die Idee, dass es besser wäre, ein bisschen „tiefer" in die Pakete zu schauen, um Angriffe zu erkennen. Theoretisch ist das ganz einfach – die „Kopfzeilen" sind nur durch die Formatierung vom eigentlichen Inhalt getrennt. Wir müssen uns also nur die nächsten paar Bytes nach denen anschauen, die wir sowieso schon analysieren. Oder wir gehen noch tiefer und betrachten den gesamten Datenblock im Paket.

Geräte, die das taten, wurden anfangs „Eindringlingsverhinderungssysteme" (Intrusion Prevention Systems, IPS) genannt und bald in die

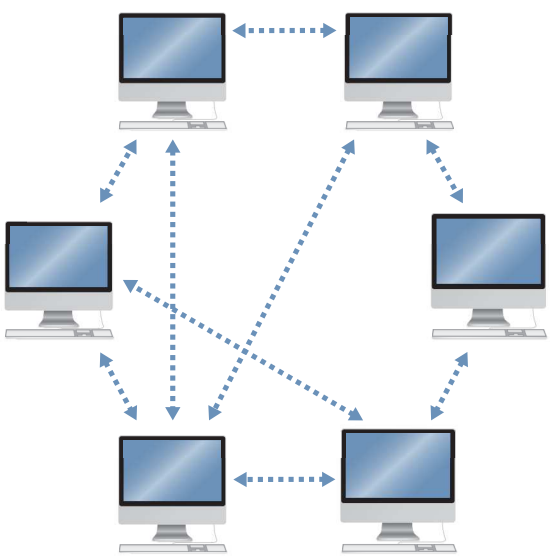
meisten Netzwerkgeräte implementiert. Solange das nur zur Abwehr von Hackerangriffen diente, gab es darüber keine Kontroversen.

Mit der Zeit bemerkten aber Regierungen, Internetanbieter und Netzwerkbetreiber, dass sie mit dieser Technik, Deep Packet Inspection (DPI) genannt, viel mehr Kontrolle über die Daten der Netzwerknutzer bekommen konnten, als das vorher möglich war. DPI-Techniken sind bereits bei der Kriminalitätsbekämpfung (Überwachung, Blockaden usw.), beim Erstellen von Marketingprofilen und beim zielgenauen Platzieren von Werbeanzeigen im Einsatz und sollen bald auch bei der Durchsetzung von Urheberrechten benutzt werden.

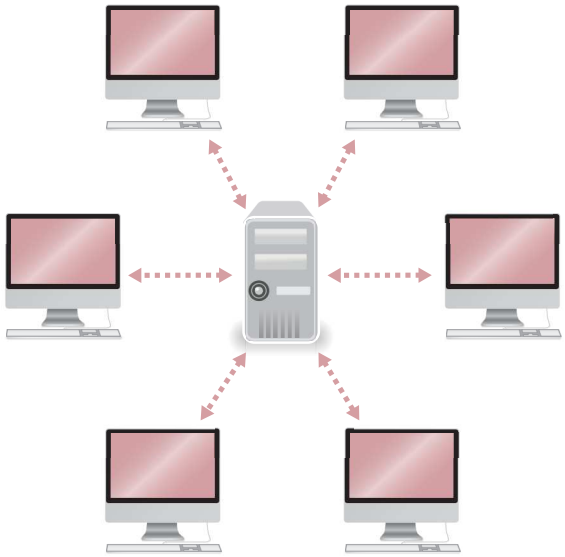
Von der Nutzerseite aus können DPI-Techniken durch Verschlüsselung umgangen werden. Die „tiefen" Inhalte eines verschlüsselten Pakets bleiben so für den Betreiber einer Deep Packet Inspection verschleiert.

PEER-TO-PEER

VON MIR ZU IHNEN OHNE JEMANDEN DAZWISCHEN



PEER TO PEER
SYSTEM VON KNOTEN OHNE
ZENTRALE INFRASTRUKTUR



ZENTRALISIERT
SERVERBASIERTES NETZWERK
(NICHT PEER-TO-PEER)

Peer-to-Peer-Netzwerke (wörtl: Kollege zu Kollege) bestehen aus Computern (Webserver oder Heim-PCs), die zugleich an einer bestimmten Form der Kommunikation teilnehmen. Jeder Computer ("peer") kann mit anderen Computern kommunizieren, es gibt keinen Unterschied zwischen Konsument und Produzent, Client und Server. Peer to Peer heißt einfach, dass viele Computer mit vielen anderen kommunizieren.

Im Internet werden für Peer-to-Peer-Anwendungen eigene Peer-To-Peer-Protokolle verwendet, die wiederum auf dem IP-Protokoll basieren.

Peer-to-Peer-Netzwerke haben eine Reihe von Vorteilen: Ein Fehler an einem Punkt im Netzwerk kann das Netz nicht zusammenbrechen lassen, da es keine Zentralinstanz gibt. In einem „Einer-zu-Vielen"-Netz versagt das Netz, wenn der „Eine" ausfällt. In einem „Viele-zu-Vielen"-Netz ist der Gesamtschaden dagegen minimal. Diese Netze können auch sehr einfach wachsen, weil jeder neue Teilnehmer auch selbst neue

Ressourcen (also Bandbreite, Speicherplatz und Rechenleistung) für das Netzwerk mitbringt. Peer-to-Peer-Netze brauchen keine Verwaltung (Administration), da es keine zentrale Stelle gibt.

Sie garantieren ihren Benutzern Freiheit. Nicht nur die teilnehmenden Rechner sind gleichberechtigt, auch ihre Benutzer sind es.

Eine der wichtigsten Aufgaben eines Peer-to-Peer-Netzwerkes ist es, das Netzwerk zu organisieren und Ressourcen im Netz zur Verfügung zu stellen.

Bis zu einem gewissen Grade sind E-Mail-Server ein frühes Beispiel für Peer-to-peer-Anwendungen. Per SMTP-Protokoll kann ein E-Mail-Server eine Mail an jeden anderen Server schicken. Auch das Domain Name System kann mehrere Server ansprechen, die jeweils die E-Mail einer bestimmten Domain verarbeiten können, was die Zuverlässigkeit erhöht.

FILESHARING

Peers in Filesharing-Netzwerken wissen nicht automatisch die IP-Adressen von anderen Peers im Netz und sie wissen auch nicht, welche Dateien (bzw. welchen Teil einer Datei) ein Peer hat. Dieses Problem wird typischerweise durch einen Prozess gelöst, bei dem Peers Informationen darüber mitteilen, welche Dateien auf anderen Peers angeboten werden. Dateien werden mittels sogenannter „Hash Keys" identifiziert, die im Grunde genommen so etwas wie eindeutig zuzuordnende Fingerabdrücke einzelner Dateien sind.

Verteilte Hash-Tabellen (Distributed Hash Tables, DHT) ermöglichen es Peers nachzuschauen, auf welchen anderen Peers eine gewünschte Datei liegt.

Nutzer von Peer-to-Peer-Netzwerken brauchen einen Weg, die Hash-Fingerabdrücke der benötigten Dateien zu bekommen.

Einige werden auf Webseiten veröffentlicht, zum Beispiel um die neueste Version von Ubuntu Linux herunterzuladen. Es gibt auch Wörterbücher, die mit einer Übersetzung von menschlicher Sprache in Hash-Keys eine Suche nach Dateien in Peer-to-Peer-Netzen ermöglichen.

Websites wie zum Beispiel thepiratebay.org und mininova.org bieten solche Wörterbücher an. Trotzdem können die Hash-Fingerabdrücke von Dateien auch über E-Mail, Chats oder Soziale Netze verbreitet werden – es gibt keine zentrale Stelle im System.

Inzwischen gibt es auch Peer-to-Peer-Netze, die ihren Nutzern Anonymität garantieren.

VERHALTENS-BASIERTE WERBUNG

JETZT WIRD'S PERSÖNLICH

Verhaltensbasierte Werbung, auch verhaltensbasiertes Targeting genannt, ist eine Technik, die darauf beruht, die Aktivitäten der Internetnutzer aufzuzeichnen und zu verfolgen. Sie wird benutzt, um Profile von Internetnutzern zu erstellen. Durch den Einsatz dieser Technik kann Werbung effizienter angezeigt werden, da diese bei korrektem Profil auf die Interessen des Nutzers zugeschnitten und somit für ihn von höherer Relevanz ist.

Verhaltensbasierte Werbung hat ein leicht verständliches Prinzip: Sobald ein Nutzer zum ersten Mal eine Seite (zum Beispiel über Fußball) besucht, wird eine kleine Datei (ein sog. „Cookie") im Webbrowser (wie z.B. Internet Explorer, Firefox oder Chrome) platziert. Üblicherweise enthält eine Webseite Inhalte aus verschiedenen Quellen. Zum Beispiel stammen Text und Bilder einer Seite von der Adresse, die Sie eingetippt haben, zusätzliche Inhalte wie z.B. Werbung wird aber von anderen Quellen, die unter Umständen nichts mit der Seite zu tun haben, nachgeladen. Jedes Mal wenn Inhalte geladen werden, können Daten aus dem Cookie von Ihrem Computer mitgeschickt werden.

Für verhaltensbasierte Werbung enthalten Cookies eine eindeutige Identifikationsnummer. Liest ein User also später einen Artikel über Autos, kann die Verhaltensmarketingfirma Annahmen über Nutzer machen, die Artikel über Autos und Fußball lesen. In unserem Beispiel wäre so eine einfache Annahme, dass dieser User am besten auf Bierwerbung anspricht. Eine weitere Annahme wäre, dass es keine gute Idee ist, Werbung für Autoversicherungen anzuzeigen, weil der User wahrscheinlich ein junger Mann ist.

Je mehr Webseiten, die wie die meisten Zeitungen und viele andere Seiten Teil eines Verhaltensmarketingnetzes sind, ein Nutzer besucht, desto mehr Daten sammeln sich in seinem Profil an. So kann in relativ kurzer Zeit ein sehr detailliertes Profil angelegt werden – und die Identifizierbarkeit der Daten steigt, auch wenn sie in der Theorie anonym angelegt werden.

Große Datenmengen können die Größe einer Gruppe von Menschen, die einem bestimmten Suchmuster entsprechen, auf einige wenige Individuen reduzieren. Vor einigen Jahren veröffentlichte ein Suchmaschinenbetreiber ein solches Set „anonymer" Daten, die beim Suchen

angefallen waren. Journalisten konnten aus diesen anonymen Informationen Einzelpersonen identifizieren. Das zeigt, dass „anonym" eben doch nicht anonym ist. Man weiß auch nicht, ob weitere Daten aus anderen Quellen für verhaltensbasierte Werbung benutzt werden. Viele im verhaltensbasierten Marketing tätige Unternehmen, z.B. Google oder Yahoo!, bieten gleichzeitig andere Dienstleistungen wie z.B. eine Internetsuche an.

Wenn man diese Datenbanken zusammenlegen würde, entstünden riesige Mengen persönlicher Daten, die relativ einfach echten Menschen zugeordnet werden könnten.

Es wird davon ausgegangen, dass verhaltensbasiertes Marketing eine der Triebfedern des Erfolgs der Onlinemarketingindustrie in den letzten Jahren ist. Die Technik selbst wird auch zu anderen Zwecken eingesetzt, zum Beispiel um Nutzern Nachrichten anzuzeigen, die sie interessieren.

User werden nicht nach ihrer Zustimmung zur Verabeitung ihrer persönlichen Daten gefragt. Das Argument der Werbeindustrie ist, dass diese Art der Aufzeichnung im Interesse der Nutzer sei, weil sie ihnen hilft, nur „relevante" Werbung zu bekommen. Sie praktizieren eine sogenannte „Opt-Out"-Lösung, die von manchen gefordert wird um die Standards der elektronischen Privatsphäre zu gewährleisten. „Opt-Out" bedeutet, dass die Informationen eines jeden Nutzers solange aufgezeichnet werden, bis dieser sich gegenteilig äußert.

Die Kernfragestellung ist hier, ob die „Cookie-Einstellungen" im Browser, die selten standardmäßig auf „Privat" stehen, schon eine Einwilligung seitens der User bedeuten. Die Europäische Datenschutzbehörde sagt, dass dem nicht so ist. Viele Internetnutzer wissen weder etwas von Cookies noch wie sie die Einstellungen dazu ändern. Auch technisch gesehen ist der Vorschlag der Werbeindustrie schwierig umzusetzen, da „Opt-Out"-Lösungen nicht alle werbetreibenden Firmen umfassen. Außerdem benutzt das Einverständnis-Entziehungs (= Opt-Out)-System selbst Cookies: Löscht man alle Cookies, wird auch die Einverständnisverweigerung gelöscht.

Zudem bieten moderne Browser und Plug-Ins (wie z.B. Flash) zusätzlich zum herkömmlichen Cookie viele andere Möglichkeiten, Daten zu speichern und auszulesen. Diese sind für den durchschnittlichen Nutzer schwer zu verwalten und teilweise nicht in den Cookie-Einstellungen des Browsers aufgeführt.

Zurzeit werden europäische Bürger durch ein Gesetz auf europäischer Ebene geschützt, bleiben de facto allerdings ungeschützt, da der Wille fehlt, dieses Gesetz umzusetzen.

03 http://www.edps.europa.eu/EDPSWEB/webdav/site/mySite/shared/Documents/Consultation/Opinions/2009/09-01-09_ePricacy_2_EN.pdf

DIE SUCH-MASCHINE

EIN INDEX FÜR DAS INTERNET

Im Internet navigiert man mithilfe von Hyperlinks, also Bildern oder Texten, die eine andere Webseite öffnen, wenn man daraufklickt.

Jeder Web-Autor kann zu jedem beliebigen Online-Inhalt verlinken. Durch die Praxis des Verlinkens helfen alle Internetuser mit, die Informationen in ein Netz aus verknüpften Ressourcen zu integrieren.

Das Internet selbst hat keinen zentralen Index, der über alle Inhalte auf dem Laufenden bleibt. Deswegen sind Suchmaschinen die wichtigste Plattform, die die Bedürfnisse der Internetnutzer erfüllt und es ihnen ermöglicht, effizient im Netz zu navigieren.

Es gibt verschiedene Arten von Suchmaschinen. Die wichtigste davon basiert auf sog. „Crawlern".

Die dabei verwendeten Programme werden „Crawler" oder auch Spinnen (Spider) genannt und schauen regelmäßig nach, was im Internet steht. Daraus bauen sie systematisch einen Index auf. Die Raffinesse und Effektivität eines Crawler-Programms bestimmt über die Größe und Aktualität dieses Index – beides Kennzahlen für die Qualität einer Suchmaschine. Einfach gesagt funktioniert ein Crawler folgendermaßen: Er folgt jedem Link einer Webseite, indiziert alle verlinkten Seiten, folgt wiederum allen Links dieser Seiten, indiziert alle Seiten, folgt jedem Link und so weiter.

Die wichtigste Aufgabe einer Suchmaschine ist die Verknüpfung der Suchanfrage eines Users mit den Informationen in ihrem Index. Typischerweise gibt dieser Vorgang eine sortierte Liste an Referenzen zurück. Diese Suchmaschinentreffer bestehen normalerweise aus einem Titel, ein paar Informationen und Hyperlinks, die auf die Webseiten verweisen, welche die Suchmaschine als möglicherweise relevant erkannt hat.

Neben diesen „natürlichen" Suchergebnissen (also den Seiten, die die Suchmaschine gefunden hat), binden kommerzielle Suchmaschinen sog. „gesponsorte" Ergebnisse ein, die von Werbetreibenden je nach Schlagwort ersteigert werden können.

Der Vorgang des Abgleichs zwischen Suche und Index ist höchst komplex und kommerzielle Suchmaschinenbetreiber halten die Algorithmen, mit denen die Reihenfolge der Suchergebnisse bestimmt werden, als Betriebsgeheimnis unter Verschluss.

Einer der berühmtesten dieser Algorithmen

ist der „PageRank" von Google. Er sagt die Relevanz einer Webseite nach einer Analyse ihrer Linkstruktur im Internet (das heißt vor allem, welche Seiten auf sie verlinken) voraus.

Andere wichtige Methoden der Verknüpfung des Informationsbedarfs eines Nutzers mit dem Suchindex sind die Analyse des Seiteninhalts oder der Abgleich mit Nutzungsdaten anderer Nutzer. Kommerzielle Suchmaschinen benutzen Cookies, um die Suchanfragen der Nutzer, ihre Klicks auf Links und mehr in personalisierter Form für lange Zeit zu speichern.

Eine „vertikale" oder spezialisierte Suchmaschine konzentriert sich auf einen bestimmten Typ von Information, zum Beispiel Reisen, Shopping, wissenschaftliche Artikel, Nachrichten oder Musik. Die großen kommerziellen Suchanbieter integrieren spezialisierte Suchmaschinen als Extra-Features in ihre Webseiten.

Eine Meta-Suchmaschine ist eine Suchmaschine, die keinen eigenen Index mit Suchergebnissen aufbaut, sondern die Suchergebnisse einer oder mehrerer Suchmaschinen benutzt. Ein „Verzeichnis" ist ein Link-Archiv, das nach verschiedenen Kategorien sortiert ist. Das Yahoo! Directory und das Open Directory Project sind bekannte Beispiele dafür.

CLOUD COMPUTING

DAS INTERNET WIRD IHR COMPUTER

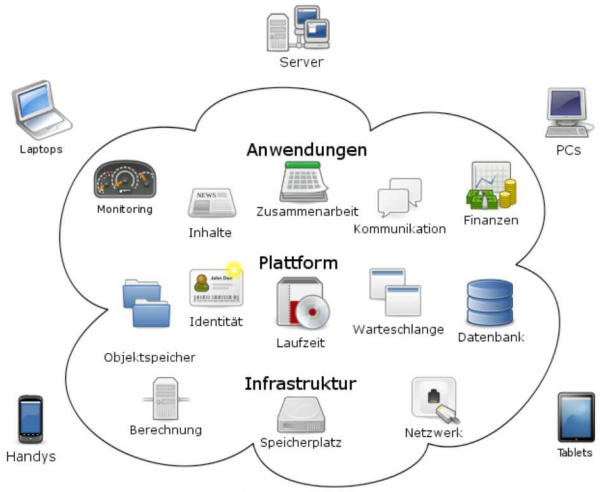
Cloud Computing wurde in letzter Zeit zu einem beliebten Werbeslogan. Das Konzept selbst ist nicht besonders neu, auch wenn in letzter Zeit eine Menge neuer Anwendungen dafür auf den Markt kamen.

In Netzwerkdiagrammen wird das Symbol der Wolke benutzt, um anzuzeigen, dass ein Netzwerk außerhalb des Netzwerks eines Benutzers angelegt ist. Cloud Computing bezieht sich also auf alle Computerdienste, die im Netzwerk und nicht auf dem Computer des Nutzers ablaufen.

Zu den ersten derartigen Dienstleistungen gehörten internetbasierte E-Mail-Dienste ("Webmailer"). Nutzer von Webmail können ihre Mails von jedem Gerät aus lesen, das mit dem Internet verbunden ist und nicht nur auf einem Computer. Zu den beliebtesten Webmail-Diensten gehören z.B. GMX, Hotmail und GMail.

Durch den stetigen Anstieg der Verbindungsgeschwindigkeit stieg die Zahl der verfügbaren Cloud-Dienste in den letzten Jahren exponentiell. Heutzutage ist es zum Beispiel möglich, im Internet riesige Datenmengen auf „virtuellen Festplatten" wie der von Microsoft Live abzulegen.

Auch Bürosoftware (z.B. Textverarbeitung



Cloud Computing

und Datenbanken) wird inzwischen online angeboten.

Googles Chrome-Betriebssystem geht einen Schritt weiter in diese Richtung. Durch die Benutzung des Chrome-Browsers als Basis soll die Cloud-Technik als Standard eingeführt werden. Das bedeutet, dass auf dem Computer nur minimale Software installiert ist, die sich voll auf die Verfügbarkeit von Internetsoftware verlässt – unter vielen Gesichtspunkten ein genau entgegengesetzter Ansatz zu traditionellen Computern, bei denen fast die gesamte Software auf der Festplatte gespeichert wird.

SOCIAL MEDIA

WO WIR UNS TREFFEN

Soziale Medien sind ein Set aus Online-Kommunikationswerkzeugen zur Erstellung und zum Austausch von durch Nutzer generierten Inhalten.

Soziale Medien unterscheidet sich fundamental von herkömmlichen Medien, da sie nicht nur Informationen vermitteln, sondern mit einer Person interagieren, während diese Informationen weitergibt. Diese Interaktion kann eine einfache Aufforderung zum Kommentieren sein, beispielsweise durch Klicken auf „like"- oder „unlike"-Links für jede Aktion eines anderen Users. Somit ist jeder Nutzer nicht nur Zuschauer, sondern auch Teil der Medien, da wiederum auch jeder andere User diese Kommentare lesen kann.

Die Menschen gewöhnen sich daran, auf das, was andere schreiben, reagieren zu können und ihren Standpunkt dazu auszudrücken. Dadurch vergrößert sich der Einfluss der Gesellschaft auf laufende Debatten. Die Zahl der Nutzer von sozialen Medien steigt Jahr für Jahr, ihr Einfluss wird stärker und stärker.

Jede Webseite, die es Besuchern erlaubt, mit anderen Besuchern in Kontakt zu treten, kann als Teil der Social Media gesehen werden. Sie können grob in sechs verschiedene Gruppen unterteilt werden:

1. Gemeinschaftsprojekte (z.B. Wikipedia), bei denen User durch das Erstellen oder Bearbeiten von Artikeln mitwirken können.
2. Blogs und Microblogs (z.B. Twitter).
3. Inhalts-Communities (z.B. YouTube oder Flickr), auf denen User durch das Teilen und Kommentieren von Photos, Videos und anderen Inhalten interagieren.
4. Soziale Netzwerke (z.B. Facebook, Myspace, StudiVZ, Google+), in denen User durch das Verknüpfen mit Freunden, das Kommentieren auf Profilen und den Beitritt zu und der Diskussion in Gruppen miteinander in Kontakt treten können.
5. Virtuelle Spielewelten (z.B. World Of Warcraft).
6. Virtuelle Welten (z.B. Second Life).

Ein wichtiges Thema ist der Schutz der Benutzer von sozialen Medien, vor allem der Schutz ihrer Privatsphäre. Obwohl Nutzer üblicherweise entscheiden können, ob sie Informationen veröffentlichen oder verstecken wollen, sind die Standardeinstellungen zum Schutz der Privatsphäre kontrovers diskutierte Punkte. Spezielle Maßnahmen zum Schutz von Kindern und Jugendlichen sind ebenfalls in der Diskussion. Auch haben bestimmte Seiten wie zum Beispiel Facebook bereits mehrfach die Privatsphäreneinstellungen ihrer User ohne Nachfrage abgeändert.

INTERNET GOVERNANCE

DIGITALE DEMOKRATIE

Die ersten Versuche, den Begriff „Internet Governance" zu definieren, wurden bei Vorbereitungstreffen zum UN-Weltgipfel zur Informationsgesellschaft unternommen. Eine erste allgemein akzeptierte Definition entwickelte die Arbeitsgemeinschaft zur „Digital Governance", einer aus vielen Interessengruppen zusammengesetzten Gruppe, die vom UN-Generalsekretär ins Leben gerufen wurde. Folgende Definition wurde in die Tunis-Agenda für die Informationsgesellschaft einbezogen:

"[Internet Governance ist die] durch Regierungen, den Privatsektor und die Zivilgesellschaft in ihren jeweiligen Rollen vorgenommene Entwicklung und Anwendung von einheitlichen Prinzipien, Normen, Regeln, Entscheidungsfindungsprozessen und Programmen, die die Evolution und die Benutzung des Internets formen."

Diese Definition hebt das Prinzip der Beteiligung verschiedener Interessengruppen bei der Diskussion von Themen hervor, die mit dem Internet zusammenhängen: die Teilnahme aller Akteure auf eine offene, transparente und verantwortliche Art und Weise.

Um dieses Ziel zu erreichen, wurde das „Internet Governance Forum" als Forum für Diskussionen über die Themen des öffentlichen Sektors gegründet, welche Schlüsselelemente der Internet Governance betreffen. Dieses Forum, das zwischen 2006 und 2011 bereits sechs Ausgaben hatte, bewirkte die Gründung ähnlicher Foren auf nationaler und regionaler Ebene (z.B. EuroDIG – der paneuropäische Dialog über Internet Governance). Es ist hervorzuheben, dass diese Gremien keine Entscheidungen treffen, sondern nur die politischen Richtlinien beeinflussen können.

Was gehört zur Internet Governance?

- Infrastruktur und Standardisierung
- Technische Themen, die den Betrieb des Internets betreffen: die Telekommunikationsinfrastruktur, Internetstandards und -dienste (z.B. das Internet Protocol oder das Domain Name System) und Inhalts- und Anwendungsstandards (z.B. HTML)
- Themen zum sicheren und stabilen Funktionieren des Internets: Cyber Security, Verschlüsselung, Spam
- Rechtliche Fragestellungen: Nationale und Internationale Gesetzgebung, die auf Internetthemen angewendet werden

kann (Urheberrechte, Privatsphäre und Datenschutz)

- Entwicklungsthemen: Die digitale Kluft, universeller Zugang zum Internet
- Wirtschaftliche Themen: E-Commerce, Steuern, elektronische Unterschriften und Zahlungsmethoden
- Soziokulturelle Probleme: Menschenrechte (Meinungsfreiheit und das Recht, Informationen zu suchen, zu teilen und weiterzugeben), Vorschriften zu Inhalten, Privatsphäre und Datenschutz, Mehrsprachigkeit und kulturelle Vielfalt, Bildung und Online-Sicherheit von Kindern

Wer nimmt an der Internet Governance teil?

- Regierungen: Sie arbeiten das Internet betreffende Richtlinien und Regelungen aus und implementieren diese.
- Der private Sektor: Internetanbieter (Internet Service Provider, ISP), Netzwerkanbieter, Registrierungsstellen für Domains (z.B. denic), Softwarehersteller und Anbieter von Inhalten.
- Die Zivilgesellschaft: Nichtregierungsorganisationen (NGOs), welche die Internetnutzer repräsentieren.
- Internationale Organisationen: die International Telecommunication Union, die UNESCO, das UN-Entwicklungsprogramm
- Die technische Gemeinschaft: die Internetgesellschaft, die Internet Engineering Task Force, das Internet Architecture Board, die ICANN (Internet Corporation for Assigned Names and Numbers)

Mehr Informationen:

Jovan Kurbalija, An Introduction to Internet Governance, Diplo Foundation, 2010

DIGITALE GESELLSCHAFT

EUROPEAN DIGITAL RIGHTS

Die englischsprachige Vorlage wurde von Mitarbeitern und Aktivisten von European Digital Rights unter Beteiligung des Digitale Gesellschaft e.V. geschrieben. Das Design stammt von CtrlSPATIE. Übersetzung von Andreas Müller und Kilian Froitzhuber.

Kontakt:
Digitale Gesellschaft e.V.
Schönhauser Allee 6/7, 10119 Berlin
info@digitalegesellschaft.de | @digiges auf Twitter
V.i.S.d.P. Markus Beckedahl

Der Druck dieser Broschüre konnte mit Spenden an den Digitale Gesellschaft e.V. finanziert werden. Unterstützen Sie uns bitte mit Ihrer Spende:

http://digitalegesellschaft.de/spenden/