[Ü1] Allgemeines

Das World Wide Web ist ein über das Internet abrufbares System von elektronischen Hypertext-Dokumenten, sogenannten Webseiten, welche mit HTML beschrieben werden. Sie sind durch Hyperlinks untereinander verknüpft und werden im Internet über die Protokolle HTTP oder HTTPS übertragen. Die Webseiten enthalten meist Texte, oft mit Bildern und grafischen Elementen illustriert. Häufig sind auch Videos, Tondokumente oder Musikstücke eingebettet.

Umgangssprachlich wird das World Wide Web oft mit dem Internet gleichgesetzt,[1] es ist jedoch jünger und stellt nur eine von mehreren möglichen Nutzungen des Internets dar. Andere Internetdienste wie E-Mail, IRC oder SSH sind nicht in das World Wide Web integriert.

Zum Aufrufen von Inhalten aus dem World Wide Web wird ein Webbrowser benötigt, der z. B. auf einem PC oder einem Smartphone läuft. Mit ihm kann der Benutzer die auf einem beliebigen, von ihm ausgewählten Webserver bereitgestellten Daten herunterladen und auf einem geeigneten Ausgabegerät wie einem Bildschirm oder einer Braillezeile anzeigen lassen. Der Benutzer kann dann den Hyperlinks auf der angezeigten Webseite folgen, die auf andere Webseiten verweisen, gleichgültig ob diese auf demselben Webserver oder einem anderen gespeichert sind. So ergibt sich ein weltweites Netz aus Webseiten. Das Verfolgen der Hyperlinks wird auch als „Surfen im Internet“ bezeichnet.

Mit dem sogenannten Web 2.0 wurden ab etwa den 2000er Jahren Webseiten populär, deren Inhalt der Nutzer nicht nur wie etwa bei Nachrichten-Seiten passiv ansehen, sondern selbst ändern und ergänzen kann, z. B. um eigene Inhalte zu veröffentlichen oder mit anderen Nutzern zu kommunizieren. Dazu zählen Blogs als private Meinungsseiten, von einer losen Autorengemeinschaft geschaffene Seiten nach dem Wiki-Prinzip und Soziale Netzwerke wie Foren. Serverseitige Techniken und (Skript-)Sprachen, die diese Interaktivität umsetzen, sind vor allem CGI, Python, ASP, Apache Wicket, JSF, ColdFusion, Ruby, PHP und SSI. Zu clientseitigen Techniken, die z. B. über Filter die Inhalte individualisieren, gehören unter anderem CSS, JavaScript oder Java, wobei Java hauptsächlich zur plattformneutralen Ausführung von Programmen dient, die oft als Webanwendungen über das Internet geladen werden und mit internetbasierenden Datenbanken (z. B. SAP-Clients) kommunizieren. Mit der Interaktivität wurde der Einsatz von Suchmaschinen möglich, die die bis dato vorhandenen Webverzeichnisse ergänzten und bis heute weitgehend verdrängten.

Mit der zunehmenden Komplexität von Formaten, Protokollen und Techniken entstanden neue Berufsbilder, wie z. B. Webdesigner und Mediamatiker. Zu ihren Aufgaben gehört neben der Programmierung von Inhalten auch die Auswertung von Nutzerverhalten im Rahmen der Logdateianalyse.

Das WWW wurde unter Weiterentwicklung bekannter ähnlicher Konzepte 1989 von Tim Berners-Lee und Robert Cailliau am europäischen Forschungszentrum CERN in Genf entwickelt.[2][3] Berners-Lee entwickelte dazu das HTTP-Netzwerkprotokoll und die Textauszeichnungssprache HTML. Zudem programmierte er den ersten Webbrowser und die erste Webserver-Software. Er betrieb auch den ersten Webserver der Welt auf seinem Entwicklungsrechner vom Typ NeXTcube. Das Gesamtkonzept wurde der Öffentlichkeit 1991 unter Verzicht auf jegliche Patentierung oder Lizenzzahlungen zur freien Verfügung gestellt, was erheblich zur heutigen Bedeutung beitrug.

Die weltweit erste Webseite info.cern.ch wurde am 6. August 1991 veröffentlicht.[4] Eine Nachbildung dieser Seite ist über nachfolgenden Link erreichbar: info.cern.ch

Das WWW führte zu umfassenden, oft als revolutionär beschriebenen Umwälzungen in vielen Lebensbereichen, zur Entstehung neuer Wirtschaftszweige und zu einem grundlegenden Wandel des Kommunikationsverhaltens und der Mediennutzung. Es wird in seiner kulturellen Bedeutung, zusammen mit anderen Internet-Diensten wie E-Mail, teilweise mit der Erfindung des Buchdrucks gleichgesetzt.

[Ü1] Geschichte

[Ü2] Name

In Berners-Lees erstem Projektentwurf vom März 1989 hieß das Web noch Mesh (englisch Geflecht).[11] Der Name wurde aber schnell verworfen, da er zu sehr an Mess (englisch Unordnung) erinnert. Die folgenden Benennungsversuche Mine of Information (englisch Informations-Mine) oder The Information Mine hatten keinen Bestand, da die Abkürzungen MOI (französisch ich) und TIM zu egozentrisch wirkten. Außerdem war eine Mine ein nur teilweise geeignetes Bild, da man aus ihr bloß etwas herausholen kann, das Web dagegen sowohl Informationen liefern als auch mit ihnen befüllt werden sollte.

Schließlich legte Berners-Lee sich auf Web und World Wide Web fest (angelehnt an den im frühen 20. Jahrhundert geprägten Ausdruck World Wide Wireless für die damals aufkommende weltweite Radiokommunikation),[12][13] obwohl er von Kollegen gewarnt wurde, dass die im Englischen und Französischen zungenbrecherische Abkürzung WWW den Projekterfolg gefährden würde. Web erschien ihm als Bild besonders passend, da es in der Mathematik ein Netz von Knoten (englisch Nodes) bezeichnet, von denen jeder mit jedem verbunden sein kann.[14]

[Ü2] Entwicklung

Das Web entstand 1989 als Projekt an der Forschungseinrichtung CERN, in der Nähe von Genf auf schweizerischem und französischem Gebiet liegend, an dem Tim Berners-Lee ein Hypertext-System aufbaute.[5] Die Idee hierzu stellte er erstmals am 12. März 1989 in der Forschungseinrichtung vor.[6] Das Konzept wurde von dem Belgier Robert Cailliau mit entworfen. Das ursprüngliche Ziel des Systems war es, Forschungsergebnisse auf einfache Art und Weise mit Kollegen auszutauschen. Eine Methode dafür war das „Verflechten“ von wissenschaftlichen Artikeln – also das Erstellen eines Webs. In Berners-Lees eigenen Worten:

„Das World Wide Web ist eine großräumige Hypermedia-Initiative zur Informationsbeschaffung mit dem Ziel, den allgemeinen Zugang zu einer großen Sammlung von Dokumenten zu erlauben.“

– Tim Berners-Lee

Das dem Hypertext zugrunde liegende Konzept stammt von früheren Entwicklungen ab, wie Ted Nelsons Projekt Xanadu,[7] Vannevar Bushs „memex“-Maschinenidee und dem Note Code Project.

Das World Wide Web unterscheidet sich von damaligen Hypertext-Systemen (Note Code benutzte beispielsweise eine einfache und lesbare Syntax und semantische Deskriptoren). Das WWW benötigt nur unidirektionale Links statt bidirektionaler, was es ermöglicht, einen Link auf eine Ressource zu setzen, ohne dass deren Besitzer eingreifen muss. Zudem baut das World Wide Web – anders als andere Protokolle wie HyperCard oder Gopher – auf einem freien Protokoll auf, was die Entwicklung von Servern und Clients ohne Beschränkungen durch Lizenzen möglich machte. Tim Berners-Lee machte das World Wide Web-Projekt am 6. August 1991 mit einem Beitrag zur Newsgroup alt.hypertext öffentlich und weltweit verfügbar.[8]

Das erste Web-Anzeigeprogramm, das eher ein Browser-Editor-Hybrid war, nannte Berners-Lee einfach „WorldWideWeb“. Er hatte es im Herbst 1990 auf einem NeXT-Computer geschrieben. Später benannte er es – um Verwechslungen mit dem World Wide Web (mit Leerzeichen) zu vermeiden – in „Nexus“ um. Es konnte damals nur Text anzeigen, aber spätere Browser wie Pei Weis Viola (1992) fügten die Fähigkeit Grafiken anzuzeigen hinzu. Marc Andreessen vom NCSA veröffentlichte im Jahre 1993 einen Browser namens „Mosaic für X“, der bald dem Web und auch dem gesamten Internet ungekannte Popularität jenseits der bisherigen Nutzerkreise und ein explosionsartiges Wachstum bescherte. Marc Andreessen gründete die Firma „Mosaic Communications Corporation“, später „Netscape Communication“. Mittlerweile können moderne Browser auch zusätzliche Merkmale wie dynamische Inhalte, Musik, Animationen und Videos wiedergeben.[9] Am 30. April 1993 gab das Direktorium des europäischen Kernforschungszentrums CERN das World Wide Web kostenlos für die Öffentlichkeit frei.[10]

[Ü1] Funktionsweise

Das WWW basiert auf drei Kernstandards:[15]

HTTP als Protokoll, mit dem der Browser Informationen vom Webserver anfordern kann.

HTML als Auszeichnungssprache, die festlegt, wie die Information gegliedert ist und wie die Dokumente verknüpft sind (Hyperlinks).

Uniform Resource Identifier (URI) als eindeutige Bezeichnung einer Ressource, die in Hyperlinks verwendet wird.

Folgende Standards kamen später dazu:

Cascading Style Sheets (CSS) beschreiben das Aussehen und die Anordnung der Elemente einer Webseite, womit der Inhalt von dessen Darstellung separiert wird.

Hypertext Transfer Protocol Secure (HTTPS) ist eine Weiterentwicklung von HTTP, bei dem das Protokoll TLS zwischen TCP und HTTP geschoben wird und in der Folge der komplette Datentransfer verschlüsselt wird.

Document Object Model (DOM) als Programmierschnittstelle für externe Programme oder Skriptsprachen (wie JavaScript) von Webbrowsern.

Das World Wide Web Consortium (W3C), das heute von Tim Berners-Lee, dem Erfinder des WWW, geleitet wird, und andere entwickeln den HTML- und den CSS-Standard; andere Standards kommen von der Internet Engineering Task Force, der ECMA und Herstellern wie Sun Microsystems. Nicht vom W3-Konsortium standardisiert ist JavaScript, die am weitesten verbreitete Skript- oder Makrosprache von Webbrowsern.

JavaScript ist eine Skriptsprache mit Anweisungen für den Browser, mit der Programme (Skripte) eingebettet werden können. Dadurch können Webseiten mit Hilfe des Document Object Models (DOM) dynamisch geändert werden. Skripte sind üblicherweise kleine Programme, können aber auch als Client Manager mit Hilfe des DOM die vollständige Kontrolle über die Anzeige übernehmen. Eine von Microsoft entwickelte Variante von JavaScript heißt JScript. Beide Sprachen sind ähnlich, allerdings nicht kompatibel zueinander. Diese Inkompatibilität war ein entscheidender Teil des sogenannten Browserkriegs.

Das WWW wurde und wird durch andere Techniken ergänzt. Schon früh wurden Bilder zur Illustration benutzt; man verwendete die Formate GIF für Grafiken und Animationen und JPEG für Fotos. Später wurde GIF mehr und mehr von PNG verdrängt, da für dessen Verwendung – im Gegensatz zu GIF – keine Lizenzgebühren anfielen.

Zudem konnten in Browsern zahlreiche weitere Dateitypen durch Browsererweiterungen, sogenannte Plug-ins, dargestellt werden. Dadurch ließen sich Multimediainhalte von Animationen bis hin zu Musik und Videos oder ganze Anwendungen wie zum Beispiel Versicherungsrechner oder Navigationsoberflächen darstellen. Beispielsweise ermöglichten Java-Applets das Einbetten von Programmen, die lokal auf dem Computer des WWW-Benutzers ablaufen und Flash wurde für interaktive Inhalte oder Animationen verwendet.

Mit der Einführung von HTML5 und weiteren standardisierten Techniken wurden Plug-ins schnell vom Markt verdrängt, da die Funktionen für Animationen und Multimediainhalte o. Ä. jetzt direkt in die Browser eingebaut waren und somit ohne externe Abhängigkeiten umgesetzt werden konnten. Aus Sicherheits- und Stabilitätsgründen deaktivieren die meisten größeren Browser meist große Teile bzw. die gesamte NPAPI-Schnittstelle, die für diese Plug-ins genutzt wurde. Weiterhin beliebt sind allerdings beispielsweise Formate wie PDF zum Anzeigen von Dokumenten, die meist mittlerweile auch durch in den Browser eingebaute PDF-Reader angezeigt werden können.

[Ü1] Dynamische Webseiten und Webanwendungen

Mit Hilfe der dynamischen WWW-Seiten kann das WWW als Oberfläche für entfernt ausgeführte Programme dienen: Ein Programm wird nicht mehr konventionell lokal auf dem Rechner gestartet, sondern wird durch einen Webbrowser betrachtet und bedient. Vorteilhaft ist hier, dass die Programme nicht mehr auf den einzelnen Rechnern verteilt sind und dort (dezentral) administriert werden müssen.

Dynamische Webanwendungen werden entweder am Webserver oder direkt im Browser ausgeführt.

Ausführen von Webanwendungen am Webserver: Der Inhalt wird durch in Skriptsprachen (wie PHP oder Perl) oder kompilierte Anwendungen (wie JSP, Servlets oder ASP.NET) geschriebene Webanwendungen erzeugt und an den Browser geliefert.

Dynamische Websites am Client: Der Browser erzeugt oder ändert Inhalt mittels JavaScript.

Gemischte Ausführung: Eine gemischte Ausführung stellt Ajax dar – hier sendet der Browser mittels JavaScript eine Anfrage, der vom Webserver bearbeitet wird und so dynamisch Teile der HTML-Struktur erneuert.

Nachteilig sind die begrenzten Ausdrucksmöglichkeiten von WWW-Seiten, so dass Programme in Form von Internetseiten im Allgemeinen nicht so einfach bedient werden können wie konventionelle Programme. Ein Trend, der versucht, beides in Einklang zu bekommen, sind Rich Internet Applications.

Zurzeit ist zu beobachten, dass immer mehr Dienste, die ursprünglich vom WWW getrennt waren und als eigenes Programm liefen, über das WWW angeboten werden und mittels eines Browsers genutzt werden können:

So wird Webmail oft als E-Mail-Client oder WebFTP als FTP-Client genutzt; Webforen ersetzen das Usenet und Webchats den IRC.

[Ü1] Kompatibilität und Zugänglichkeit

Oft führten Browser-Hersteller neue Möglichkeiten ein, ohne auf eine Standardisierung zu warten. Umgekehrt werden jedoch immer noch nicht alle Teile von Standards wie HTML oder CSS korrekt implementiert. Das führt zu Inkompatibilitäten zwischen bestimmten Webseiten und manchen Browsern. Besonders „hervorgetan“ durch solche Inkompatibilitäten hatte sich zu Beginn des Internet-Booms die Firma Netscape, heute vor allem das Unternehmen Microsoft mit seinem Internet Explorer.

Außerdem ging durch die Vielzahl der Ad-hoc-Erweiterungen von HTML ein wesentlicher Vorteil dieser Sprache verloren – die Trennung von Inhalt und Darstellung. Durch diese Trennung können die in HTML ausgezeichneten Inhalte optimal für das jeweilige Ausgabegerät aufbereitet werden. Die Spannweite an unterstützungswürdigen Ausgabegeräten reicht vom hochauflösenden Bildschirm eines Desktop-Computers, dem kleinen Bildschirm eines leistungsschwachen Mobiltelefons, über schwarz-weiße E-Book-Reader mit geringer Bildfrequenz bis hin zur Sprachausgabe mittels Screenreader und Steuerung mittels Braillezeile für Benutzer mit Sehschwierigkeiten.

Das W3C und andere Initiativen treiben daher die Entwicklung in die Richtung einer Standardisierung und browserübergreifenden Vereinheitlichung von HTML/XHTML und CSS voran, um diese Vorteile von HTML wiederzuerlangen. Auch der zunehmende Verzicht auf Plug-ins, wie Flash, unterstützen diesen Trend und führten zu einer besseren Barrierefreiheit. XHTML wurde später zugunsten von HTML5 aufgegeben, welches auch explizite Features zur Verbesserung der Barrierefreiheit und Maschinenlesbarkeit einbaute, die visuell nicht sichtbar sind.