

Website-Monitoring: Durch Auswertung von Nutzerdaten die eigene Homepage optimieren

Nutzer einer Website hinterlassen dort, ob gewollt oder ungewollt, eine Fülle von Daten, die der Website-Betreiber nutzen kann - zur Verbesserung der Website, zur Wirkungskontrolle von Kommunikationsmaßnahmen, zur Personalisierung von Marketingmaßnahmen, zur Optimierung seines Produktangebots oder auch über eine Vermarktung der gewonnenen Nutzerdaten, also den Verkauf an externe Unternehmen.

Gewinnung von Nutzerdaten

Nutzerdaten können über folgende prinzipielle Maßnahmen gewonnen werden:

- **Protokolldateien des Webservers:** Alle Zugriffe auf eine Homepage laufen über den **Webserver** des Websitebetreibers. Alle Zugriffe auf diesen Webserver werden von diesem in einer **Server-Log-Datei** protokolliert. Inhalte des Server-Log sind i.d.R. (der Administrator des Webservers kann dies konfigurieren): Name/IP-Adresse des Besuchers, Benutzeridentifikation (falls möglich), Datum und Zeitpunkt des Zugriffs, Datei, auf die zugegriffen wird, Status der Übertragung (erfolgreich oder nicht), übertragene Bytes. Aus diesen Daten lässt sich eine Fülle von **Mess- und Bewertungskriterien** für den Erfolg einer Website ableiten; man unterscheidet
 - **Kontaktkennzahlen** wie Reichweite (Anzahl der Besucher oder Besuche bzw. Visits), Visit Duplication (Anzahl der Mehrfachbesucher), Frequenz (Häufigkeit der Visits eines Nutzers)
 - **Interaktivitätskennzahlen** wie Verweildauer, Zeit zwischen zwei Visits, Visit-Tiefe (Besuch von Nachfolgeseiten).

Daraus kann man u.a. folgende interessante Informationen ableiten:

- Welche Seiten werden am häufigsten aufgerufen?
 - Aus welchen Ländern kommen die Nutzer?
 - Gibt es Wiederholungsnutzer?
 - Welches Tages-/Wochen-/Monats-/Jahresprofil weisen die Nutzungen auf?
 - Von welchen Seiten wird auf welche gesprungen?
 - Welche Links auf einer Seite werden am häufigsten betätigt?
 - Gibt es "tote" Seiten?
 - Von welchen Seiten aus verlassen die Nutzer unser Angebot?
 - Wie lange verweilen die Nutzer auf einer Seite?
- **Cookies:** Cookies sind Datensätze, die ein Browser wie Firefox oder der Internet-Explorer auf **Anweisung eines Webservers** (des Website-Betreibers) beim Besuch einer Seite **auf dem Computer des Besuchers** in einer Textdatei speichert (cookies.txt). Beim nächsten Besuch der Website versucht der Webserver, die von ihm gesetzten Cookies wieder auszulesen. Cookies können für den Website-Besucher den Komfort beim Surfen steigern, indem sie eine Personalisierung der Website zulassen; Transaktionen in Online-Shops sind mit Cookies einfacher und schneller durchführbar. Cookies können allerdings ein **Datenschutzrisiko** sein. Mit ihnen können Benutzer leicht identifiziert und Verhaltensprofile von angelegt werden. Wenn sie nicht nur an den Server, der das Cookie gesetzt hat, sondern auch an andere Webserver übertragen werden, gelangen Nutzerdaten an Dritte - **ohne Wissen des Websitebesuchers**. In den Sicherheitseinstellungen der Browser findet man daher Funktionen zum Löschen und Verwalten von Cookies, die das verhindern können.

- **Explizite Benutzerregistrierung:** Vor dem Zugriff auf die Seite werden vom Besucher Zugangsdaten (Benutzerkennung, Passwort) angefordert, die er über eine Registrierung erhält. Damit erhält man über die einzelnen Besucher nicht nur Verhaltensdaten, sondern darüber hinaus gehende **persönliche Daten**. Zur Registrierung muss ein Besucher aber erst motiviert werden, beispielsweise durch Teilnahme an einem Gewinnspiel, durch individuelle Erstellung eines „unverbindlichen Angebots“ oder bei der Bestellung in Online-Shops.

Verhindern der Datensammlung und Auswege für Website-Betreiber

Die Möglichkeiten von Website-Besuchern, sich gegen eine Sammlung ihrer Daten zu wehren, hängen von der Methode ab, wie sie gewonnen werden sollen: Die Sammlung von Server-Log-Daten ist durch Benutzer prinzipiell nicht zu verhindern, kann jedoch **verfälscht** werden: Eine interessante Möglichkeit, das eigene "Surfverhalten" zu verbergen, bietet die **Anonymisierung** beispielsweise über <http://www.anonymizer.com>. Solche Anonymizer fordern die Site anonym an und leiten sie an den Benutzer weiter. Dessen Browser stellt sie dann so dar, als ob er sie selber abgeholt hätte. Von hier kann man problemlos anonym weitersurfen, denn die Anonymizer verändern jeden Link in den übertragenen Webseiten so, dass er automatisch anonym abgerufen wird. Ein Portal mit zahlreichen Links zum Thema Anonymisierung sowie zu Anonymizer-Softwaretools findet man unter <http://www.epic.org/privacy/tools.html>.

Weitere Schwierigkeiten beim Datensammeln könnten sein:

- **Caching/Mirroring:** Der Zugriff des Browsers erfolgt aus Performance-Gründen üblicherweise nicht auf die "Originalsite", die auf dem Server des Anbieters liegt, sondern auf **zwischengespeicherte Seiten** im Cache-Speicher (siehe <http://www.firefox-browser.de/wiki/Cache>) des Browsers, auf einen Proxy-Server oder auf eine Mirror-Site, auf der Websites gespiegelt werden (um damit beispielsweise robust gegenüber Serverausfällen zu sein).
- **Verbindung mit dem eigenen Internet-Provider über DHCP** (Dynamic Host Configuration Protocol; siehe http://de.wikipedia.org/wiki/Dynamic_Host_Configuration_Protocol): Die meisten Internet-Provider (wie beispielsweise t-online) bauen eine Datenübertragungs-Verbindung zwischen dem Endgerät des Surfers (Client) und dem Internet so auf, dass sie IP-Adressen (die wichtigsten Verbindungsinformationen) beim Aufbau der Verbindung **dynamisch zuweisen**, d.h. diese ändern sich bei jedem Verbindungsaufbau. Damit kann nicht mehr nachvollzogen werden, welcher Nutzer über mehrere Verbindungen hinweg auf bestimmte Webseiten zugegriffen hat.
- **NAT** (Network Address Translation; http://de.wikipedia.org/wiki/Network_Address_Translation): Wenn Benutzer beispielsweise über ein Firmen-Netzwerk auf die Website zugreifen, dann erfolgt dies wahrscheinlich über eine zwischengeschaltete Firewall. Die meisten Firewalls "maskieren" die Verbindungsdaten, indem sie nach außen jedem Benutzer eine einheitliche IP-Adresse zuordnen; dadurch kann von außen nur nachvollzogen werden, dass jemand aus dieser Organisation auf eine Website zugegriffen hat, jedoch nicht, welcher konkrete Benutzer dies war.
- **Benutzerverweigerung:** Bei jedem Browser kann eingestellt werden, dass Cookies nicht akzeptiert werden. Benutzer können eine Registrierung verweigern.

Auf der anderen Seite werden ständig **neue Methoden** entwickelt, um immer hochwertigere Benutzer-Daten zu erhalten. Ein Verfahren, mit dem Nutzer-Daten gesammelt und zur Individualisierung von Online-Werbung sowie zur Website-Optimierung ausgewertet werden, ist **Behavioral Targeting** (siehe http://de.wikipedia.org/wiki/Predictive_Behavioral_Targeting). Damit werden Website-Inhalte davon abhängig gemacht, welche Seiten die Nutzer zuvor aufgerufen haben und mit Befragungsdaten zu Soziodemographie, Produktinteressen und Lifestyle angereichert. Mit **Mouse-Tracking-Verfahren** (siehe beispielsweise <http://www.whiteboxx-research.de/methoden/mousetracking.html>) werden Mausbewegungen des Nutzers aufgezeichnet; anhand der Art, wie der Benutzer die Maus bewegt, kann man Schlüsse ziehen, welche Informationen ihn interessieren. Die gewonnenen Daten werden ergänzt um Logfile-Daten; daraus kann ein Informationsprofil des Nutzers einerseits und eine Bewertung von HTML-Seiten andererseits erstellt werden. Damit werden neuartige Navigationsverfahren möglich; je nach Mausbewegung erhält der Nutzer ganz bestimmte Informationen (siehe auch <http://www.media.mit.edu/~alockerd/papers/cheese.pdf>).

Analyse-Softwaretools

Zur Analyse der Website-Daten steht mittlerweile eine Vielzahl an **Software-Produkten** zur Verfügung (siehe beispielsweise <http://www.heise.de/software/download/o0g0s3l1k287>)- zum großen Teil als Open-Source-Software. Nahezu jeder Webspacer-Provider bietet dem Homepage-Betreiber Analyse-Statistiken an. Auch Google bietet mit Google Analytics einen solchen Dienst an (siehe <http://www.google.com/intl/de/analytics>) - wobei dann natürlich auch Google über die Benutzung der eigenen Website Bescheid weiß. Google bietet dazu unter dem Namen Google Analytics Individual Qualification (IQ; <http://www.google.com/intl/de/analytics/education.html>) auch eine kostenlose Online-Fortbildung an; damit kann man sich Know-How im Umgang und der Erstellung von Web-Analysenerwerben.

Auch Dienstleister bieten Website-Nutzungsstatistiken an; dazu werden i. d. R. **Web-Bugs** eingesetzt. Dies sind kleine, versteckte Bilder oder Codeteile in einer Webseite, die eine Verbindung zu einer anderen Site aufbauen, so dass dort ein Profil der Surfgeohnheiten erstellt werden kann (Third-Party-Cookie).

CEBIS hilft weiter

Unternehmen, die Informations- und Beratungsbedarf zu Einsatzmöglichkeiten des Internets für das eigene Marketing haben, können sich an CEBIS wenden. In **Veranstaltungen** und an **Beratertagen** können Unternehmen Hilfestellung durch kompetente Berater erhalten. Informieren Sie sich auf der CEBIS-Website über die entsprechenden Termine und melden Sie sich möglichst frühzeitig an.

Quelle und Copyright: Internetauftritt des Landkreises Neu-Ulm, <http://www.landkreis.neu-ulm.de>

Tipp des Monats November 2010