

LEHRSTUHL FÜR
ALLG. BWL UND WIRTSCHAFTSINFORMATIK
UNIV.-PROF. DR. HERBERT KARGL

Schwickert, Axel C.; Wendt, Peter

Web Site Monitoring

Teil 2: Datenquellen, Web-Logfile- Analyse, Logfile-Analyzer

ARBEITSPAPIERE WI
Nr. 07/2000

Schriftleitung:
PD Dr. habil. Axel C. Schwickert

Information

- Reihe:** Arbeitspapiere WI
- Herausgeber:** Univ.-Prof. Dr. Axel C. Schwickert
Professur für BWL und Wirtschaftsinformatik
Justus-Liebig-Universität Gießen
Fachbereich Wirtschaftswissenschaften
Licher Straße 70
D – 35394 Gießen
Telefon (0 64 1) 99-22611
Telefax (0 64 1) 99-22619
eMail: Axel.Schwickert@wirtschaft.uni-giessen.de
<http://wi.uni-giessen.de>
- Bis Ende des Jahres 2000 lag die Herausgeberschaft bei:
- Lehrstuhl für Allg. BWL und Wirtschaftsinformatik
Johannes Gutenberg-Universität Mainz
Fachbereich Rechts- und Wirtschaftswissenschaften
Welderweg 9
D - 55099 Mainz
- Ziele:** Die Arbeitspapiere dieser Reihe sollen konsistente Überblicke zu den Grundlagen der Wirtschaftsinformatik geben und sich mit speziellen Themenbereichen tiefergehend befassen. Ziel ist die verständliche Vermittlung theoretischer Grundlagen und deren Transfer in praxisorientiertes Wissen.
- Zielgruppen:** Als Zielgruppen sehen wir Forschende, Lehrende und Lernende in der Disziplin Wirtschaftsinformatik sowie das IuK-Management und Praktiker in Unternehmen.
- Quellen:** Die Arbeitspapiere entstanden aus Forschungsarbeiten, Diplom-, Studien- und Projektarbeiten sowie Begleitmaterialien zu Lehr- und Vortragsveranstaltungen des Lehrstuhls für Allg. Betriebswirtschaftslehre und Wirtschaftsinformatik Univ. Prof. Dr. Herbert Kargl an der Johannes Gutenberg-Universität Mainz.
- Hinweise:** Wir nehmen Ihre Anregungen und Kritik zu den Arbeitspapieren aufmerksam zur Kenntnis und werden uns auf Wunsch mit Ihnen in Verbindung setzen.
Falls Sie selbst ein Arbeitspapier in der Reihe veröffentlichen möchten, nehmen Sie bitte mit dem Herausgeber (Gießen) unter obiger Adresse Kontakt auf.
Informationen über die bisher erschienenen Arbeitspapiere dieser Reihe und deren Bezug erhalten Sie auf dem Schlußblatt eines jeden Arbeitspapiers und auf der Web Site des Lehrstuhls unter der Adresse <http://wi.uni-giessen.de>

Arbeitspapiere WI Nr. 7/2000

- Autoren:** Schwickert, Axel C.; Wendt, Peter
- Titel:** Web Site Monitoring – Teil 2: Datenquellen, Web-Logfile-Analyse, Logfile-Analyzer
- Zitation:** Schwickert, Axel C.; Wendt, Peter: Web Site Monitoring – Teil 2: Datenquellen, Web-Logfile-Analyse, Logfile-Analyzer, in: Arbeitspapiere WI, Nr. 7/2000, Hrsg.: Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Universität: Mainz 2000.
- Kurzfassung:** Aufgrund der steigenden Wettbewerbsrelevanz werden Unternehmen zu nicht unerheblichen Investitionen in ihre Web-Präsenzen veranlasst. Die Web Site entwickelt sich damit zu einem Kostenfaktor, der durch eine überzeugende Nutzenbegründung zu rechtfertigen ist. Für eine Nutzenbegründung ist es notwendig, den Nutzen zu ermitteln, den eine Web Site stiftet. Eine grundlegende Voraussetzung für die Nutzenermittlung ist es, festzustellen, wofür die Web Site durch wen und in welchem Ausmaß genutzt wird. Diese Nutzungsermittlung und eine darauf aufbauende Nutzungsanalyse führt somit erst zu einer fundierten Nutzenbegründung. In diesem Zusammenhang wird die Ermittlung und die Analyse der Web-Site-Nutzung unter dem Begriff „Web Site Monitoring“ (WSM) zusammengefaßt. Arbeitspapier Nr. 6/2000 definierte zunächst die zentralen Begriffe „Web Site“ und „Monitoring“ und verdeutlichte dessen Positionierung im Unternehmen. Aus welchen Quellen die Daten der Web-Site-Nutzung gewonnen werden, ist Gegenstand des vorliegenden Arbeitspapiers WI Nr. 7/2000: Electronic Mail, Web-Formulare, Logfiles von Web Servern, Netzwerküberwachung und Datenquellen außerhalb der Web Site. Eine zentrale Datenquelle sind die von Web Servern erzeugten Logfiles. Einige typische Software-Produkte zur Logfile-Analyse werden abschließend vorgestellt.
- Schlüsselwörter:** Monitoring, Electronic Mail, Web-Formulare, Logfile, Logfile-Analyzer, Web Site, Electronic Business, Electronic Commerce

Inhaltsverzeichnis

1	Nutzungsanalyse durch „Web Site Monitoring“	3
2	Datenquelle „Electronic Mail“	4
3	Datenquelle „Web-Formulare“	5
4	Datenquelle „Web-Logfiles“	8
4.1	Web Server und Logfiles	8
4.2	Inhalte und Meßgrößen von Logfiles	9
4.3	Technische Probleme bei der Nutzung von Logfiles	12
4.4	Standardisierungsbemühungen	16
4.5	Datenschutzrechtliche Rahmenbedingungen	16
4.6	Verarbeitungsverfahren von Logfiles.....	18
5	Datenquelle „Netzwerküberwachung“	19
6	Datenquellen außerhalb der Web Site	21
7	Zusammenfassender Überblick der Datenquellen	22
8	Software-Produkte zur Logfile-Analyse	23
8.1	Kategorisierung der Software-Produkte.....	23
8.2	FastStats Analyzer V2.74.....	25
8.3	Funnel Web 4.0	25
8.4	Summary 1.5	26
8.5	SurfStats Log Analyzer Professional Edition	27
8.6	SurfTrace Logfile Analyzer 2.11 Prof. Edition.....	28
8.7	WebSuxess 4 Enterprise	28
8.8	WebTrends Professional Suite	29
	Literaturverzeichnis	31

1 Nutzungsanalyse durch „Web Site Monitoring“

In den vergangenen Jahren hat sich die Bedeutung einer eigenen Internet-Präsenz für Unternehmen deutlich verändert. In der Anfangsphase war die Entwicklung einer Web Site vor allem durch die Chancenbewertung dieser neuartigen Technologie bzw. der zunehmenden Präsenz von Wettbewerbern motiviert, doch mit rasant steigenden Nutzerzahlen signalisierte das Internet seine hohe Wettbewerbsrelevanz als ein zusätzliches Kommunikations- und Vertriebsmedium.¹ Der Stellenwert einer Internet-Präsenz läßt sich daran erkennen, daß inzwischen der überwiegende Teil der Groß- und mittelständischen Unternehmen über eine eigene Web Site verfügt. Eine Besonderheit im Vergleich zu traditionellen Medien besteht dabei in der Interaktionsfähigkeit zwischen Web-Site-Anbieter und -Adressat.² Mit der Möglichkeit zum Dialog kann ein Anbieter individuell auf Bedürfnisse der Nutzer reagieren und sich so z. B. einen Wettbewerbsvorteil verschaffen.³

Aufgrund der steigenden Wettbewerbsrelevanz werden Unternehmen zu nicht unerheblichen Investitionen veranlasst. Die Web Site entwickelt sich damit zu einem Kostenfaktor, der durch eine überzeugende Nutzenbegründung zu rechtfertigen ist.⁴ Für eine Nutzenbegründung ist es notwendig, den Nutzen zu ermitteln, den eine Web Site stiftet. Eine grundlegende Voraussetzung für die Nutzenermittlung ist es, festzustellen, wofür die Web Site durch wen und in welchem Ausmaß genutzt wird. Diese Nutzungsermittlung und eine darauf aufbauende Nutzungsanalyse führt somit erst zu einer fundierten Nutzenbegründung. Im Zusammenhang des vorliegenden Arbeitspapiers wird die Ermittlung und die Analyse der Web-Site-Nutzung unter dem Begriff „Web Site Monitoring“ (WSM) zusammengefaßt. Dazu sind Methoden und Werkzeuge notwendig, die die Nutzung einer Web Site dokumentieren und deren Analyse ermöglichen.

Das diesem vorhergehende Arbeitspapier Nr. 6/2000 definierte zunächst die zentralen Begriffe „Web Site“ und „Monitoring“ und verdeutlichte die Positionierung des Web Site Controlling (WSC) im Unternehmen. Dabei stand die Einbindung in den übergeordneten Planungs- und Entwicklungsprozeß des Web Site Engineering (WSE) im Vordergrund. Web Site Monitoring wurde dabei als Bestandteil eines unternehmerischen Web Site Controlling eingeordnet und mit Handlungsebenen, Zielfelder, Controlling-Bereiche, Adressaten und Untersuchungsbereichen des Web Site Monitoring vorgestellt.

-
- 1 Vgl. Heuer, Kai R.; Wilken, Markus: Ansätze zur Datengewinnung für das Controlling im Online-Marketing, in: Handbuch Marketing-Controlling, Hrsg.: Zerres, Michael P., Berlin et al.: Springer 2000, S. 309.
 - 2 Vgl. McKenna, Regis: Marketing in Echtzeit, in: Erfolg im E-Business, Hrsg.: Tapscott, Don, München, Wien: Hanser 2000, S. 170.
 - 3 Vgl. Link, Jörg: Zur zukünftigen Entwicklung des Online Marketing, in: Wettbewerbsvorteile durch Online Marketing, Hrsg.: Link, Jörg, 2., überarbeitete und erweiterte Auflage, Berlin et al.: Springer 2000, S. 3.
 - 4 Vgl. Schwickert, Axel C.; Beiser, Armin: Web Site Controlling, in: Arbeitspapiere WI, Nr. 7/1999, Hrsg.: Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Universität Mainz 1999, S. 3.

Bisher verfügbare Software-Lösungen zur Unterstützung des Web Site Monitoring beschränken sich zumeist auf die Erfassung, Aggregation und Präsentation von technischen Nutzungsdaten. Aus welchen Quellen diese Daten gewonnen werden, ist Gegenstand des vorliegenden Arbeitspapiers WI Nr. 7/2000: Electronic Mail, Web-Formulare, Logfiles von Web Servern, Netzwerküberwachung und Datenquellen außerhalb der Web Site. Eine zentrale Datenquelle sind die von Web Servern erzeugten Logfiles. Einige typische Software-Produkte zur Logfile-Analyse werden abschließend vorgestellt.

Zur Informationsversorgung, Entscheidungsunterstützung und Nutzenbegründung im Management eines Unternehmens (einer Web Site) müssen die erfassten Daten in ihren wirtschaftlichen Kontext eingeordnet und zu aussagefähigen Kennzahlen verdichtet werden.⁵ Web Site Monitoring dient somit der Gewinnung von betriebswirtschaftlich verwertbaren Controlling-Kenngrößen zu einer Web Site. Derartige Controlling-Kenngrößen werden in Arbeitspapier Nr. 8/2000 zusammengetragen.

2 Datenquelle „Electronic Mail“

Die wohl häufigste Form der Kommunikation im Internet ist die „electronic mail“ bzw. eMail. Bei der Definition des WSM wurde der Internet-Dienst WWW als Basistechnologie einer Web Site bezeichnet. eMail ist nicht Bestandteil des WWW, sondern ein eigenständiger Internet-Dienst. Dennoch ist eMail als Datenquelle für das WSM zu erwähnen, da beide Dienste sowohl nach dem Empfinden der Nutzer, als auch durch die funktionale Gestaltung gängiger Internet-Software eng miteinander verbunden sind.⁶ So kann eine eMail nicht nur durch ein Mail-Programm generiert werden, sondern auch durch eine entsprechende Funktion in einer Web Site.

Die Datenübertragung beim Senden einer Nachricht geschieht mit Hilfe des Übertragungsprotokolls SMTP. Der Absender einer eMail sendet diese an seinen Mail-Server (Message Transfer Agent, MTA). Von dort wird sie über weitere MTAs zum Mail-Server des Empfängers weitergeleitet, wo sie bis zum Abruf zwischengespeichert wird. Wenn der Empfänger sein Mail-Programm aktiviert, startet dieses eine Anfrage an den Mail-Server und ruft vorliegende Nachrichten unter Verwendung der Übertragungsprotokolle POP oder IMAP ab.

Eine eMail enthält außer dem Nachrichtentext einen Kopfbereich (Header), dem folgende Angaben entnommen werden können:⁷

5 Vgl. Reichmann, Thomas; Fritz, Burkhard; Nölken, Dirk: EIS-gestütztes Controlling: Schnittstelle zwischen Controlling und Informationsmanagement, in: Handbuch Informationsmanagement, Hrsg.: Scheer, August-Wilhelm, Wiesbaden: Gabler 1993, S. 469.

6 Vgl. Wilde, Erik: World Wide Web, Berlin et al.: Springer 1999, S. 498.

7 Vgl. Köhntopp, Marit; Köhntopp, Kristian: Datenspuren im Internet, in: Computer und Recht, 4/2000, S. 254. Vgl. auch Pils, Manfred: Electronic Business im Blickfeld der Sensiblen Informations- und Kommunikationssysteme, in: Internet und Intranet – Auf dem Weg zum Electronic Business, Hrsg.: Höller, Johann; Pils, Manfred; Zlabiger, Robert, 2., neubearb. und erw. Aufl., Berlin et al.: Springer 1999, S. 294 f.

- Adressen von Sender und Empfänger
- Betreffzeile
- Identifikationsnummer der Nachricht
- Sendedatum und -uhrzeit
- Technische Infrastruktur des Absenders (Betriebssystem, Mail-Programm)
- Sendepfad der Nachricht

Diese Headerinformationen enthalten nur in geringem Maße verwertbare Hinweise zur Gestaltung der Web Site. Die Angaben zur technischen Ausstattung der Nutzer können für eine Abstimmung der technischen Anforderungen zur vollwertigen Nutzung der Web Site verwendet werden. Wesentlich wertvoller sind die gesammelten eMail-Adressen. Hier bieten sich Ansprechpartner, die aktiv an den Web-Site-Anbieter herangetreten sind und durch ihre eMail zum Dialog auffordern. Die genauen Anregungen und Bedürfnisse erschließen sich erst aus der inhaltlichen Untersuchung von eMails, weshalb der Inhalt als wichtigste Informationsquelle anzusehen ist. In ihm artikulieren die Nutzer u. a. ihre Zufriedenheit mit dem Web-Site-Angebot. Ein erhebliches Problem bei der Auswertung stellt die Tatsache dar, daß die Informationen im Text einer eMail in unstrukturierter Form vorliegen und ihre Analyse üblicherweise nur unter Verwendung personeller Ressourcen möglich ist. Ansatzpunkte zur Automatisierung bieten Text-Mining-Systeme, die Dokumente nach Schlagworten filtern und kategorisieren.⁸ Der Anwender erhält damit eine thematische Vorauswahl an eMails, die dann einer weitergehenden Analyse bedürfen.

Die Kommunikation per eMail zählt zu den reaktiven Verfahren der Datenerhebung, da die Initiative zur Übertragung vom Nutzer ausgeht. Abgesehen von der Möglichkeit einer automatischen Empfangsbestätigung wird die Auswertung zeitversetzt vorgenommen. Durch die inhaltliche Vielfalt von eMails lassen sich die Verwendungsmöglichkeiten für das WSM nicht genau abgrenzen. Deutlicher Vorteil dieser Datenquelle ist die Gewinnung von Adressdaten, die einem Web-Site-Anbieter eine Nutzergruppe aufzeigen, welche für eine planmäßige Datenerhebung, etwa im Rahmen einer Umfrage, herangezogen werden kann.

3 Datenquelle „Web-Formulare“

Der Einsatz von Web-Formularen zur Erhebung von Nutzerdaten erfolgt auf Initiative des Web-Site-Anbieters. Hierbei kann der Anbieter seinen Informationsbedarf in strukturierter Form anzeigen und sich auf die Erfassung relevanter Daten beschränken. Für den Entwurf von Formularen stehen eigene HTML-Befehle zur Verfügung, mit denen Eingabe- und Auswahlfelder in ein HTML-Dokument eingebunden werden. Die Übertragung von Daten erfolgt nach Betätigung eines Sendebuttons im Dokument, der einen URL aufruft, welcher ein serverseitiges Programm (z. B. ein CGI-Skript) ausführt. Ge-

8 Vgl. Gentsch, Peter: Business Intelligence: Aus Daten systematisch Wissen entwickeln, in: Electronic Business und Knowledge Management – Neue Dimensionen für den Unternehmenserfolg, Hrsg.: Scheer, August-Wilhelm, Heidelberg: Physica-Verlag 1999, S. 181 f.

ringe Datenmengen werden mit der GET-Methode⁹ übertragen, bei der die eingegebenen Werte als Abfrage-String an den URL angehängt werden und auch im Adressfenster des Browsers sichtbar sind. Für umfangreichere Eingaben eignet sich dagegen die POST-Methode, die die Parameter in einem eigenen Datenpaket an den Server schickt.¹⁰ Zudem ist auch eine Weiterleitung an eine eMail-Adresse möglich.

Das Einsatzgebiet von Formularen ist zweigeteilt: Einerseits die Identifikation von anonymen Web-Site-Nutzern und andererseits die zielgerichtete Datenerhebung durch Umfrageaktionen. Bei der Beobachtung von Nutzungsvorgängen ist die Person des Nutzers zunächst unbekannt. Erst mit der Angabe von Name und Adresse lassen sich Nutzer eindeutig identifizieren. Die Angabe kann freiwillig geschehen oder durch eine Transaktion auf der Web Site veranlaßt werden, wie etwa einer Bestellung von Waren. Einige Web Sites erlauben den Zugang zu bestimmten Bereichen nur, wenn sich der Nutzer vorher identifiziert hat, um so beispielsweise Adressdaten potentieller Kunden zu sammeln. Grundlage einer Identifikation sind zunächst persönliche Daten, die durch Zuordnung demographischer und psychographischer Angaben zu umfassenden Persönlichkeitsprofilen weiterentwickelt werden können. Anhand dieser Merkmale lassen sich Nutzer in Zielgruppen einordnen. Web Sites, die Einnahmen durch Werbung erzielen, können ihren Werbekunden damit die Erreichbarkeit einer bestimmten Zielgruppe nachweisen.¹¹

In den meisten Fällen erwartet der Nutzer für die Preisgabe seiner Daten eine adäquate Gegenleistung. Zwar verweisen die Nutzer oft auf den Schutz ihrer Privatsphäre, doch der eigentlich störende Faktor ist die ungleiche Verteilung des zusätzlichen Nutzens, der aus der Weitergabe persönlicher Daten resultiert.¹² Anreize zur Datenübermittlung werden in Form von monetären Vergütungen, kostenlosen Waren (Teaser) oder Zugang zu bestimmten Informationen und Downloads geboten. Die Datenerfassung in einem Formular wird nur dann vom Nutzer akzeptiert, wenn sie für ihn einen Nutzen verspricht, begründbar ist und eine transparente Verwendung aufweist. Für die Gestaltung von Formularen empfiehlt sich daher, möglichst wenige Eingaben zwingend erforderlich zu machen, aber die optionale Eingabe zusätzlicher Informationen zuzulassen. Die Angabe fehlerhafter Daten kann absichtlich oder unabsichtlich erfolgen und den verfügbaren Datenbestand verringern bzw. verfälschen. Für absichtliche falsche Angaben sind zwei Hauptgründe zu nennen: Entweder besteht der Nutzer auf der Wahrung seiner Anonymität oder das Formular suggeriert einen lästigen Arbeitsaufwand, der möglichst schnell umgangen wird.¹³

9 GET bzw. POST bezeichnen hier die Anfragemethoden des HTTP-Protokolls vom Client an einen Server. Vgl. Janetzko, Dietmar: Statistische Anwendungen im Internet, München: Addison-Wesley-Longman 1999, S. 21.

10 Vgl. Wilde, Erik: World Wide Web, a. a. O., S. 262.

11 Vgl. Zierl, Marco: Anreize bieten: So erfahren Sie mehr über Ihre Kunden, in: Internet Professionell, 9/2000, S. 46.

12 Vgl. Hagel III., John; Rayport, Jeffrey F.: Die kommende Schlacht um Kundendaten, in: Erfolg im E-Business, Hrsg.: Tapscott, Don, München, Wien: Hanser 2000, S. 183.

13 Vgl. Zierl, Marco: Anreize bieten: So erfahren Sie mehr über Ihre Kunden, a. a. O., S. 50.

Die Verwendung der Web Site als Medium zur Datenerhebung im Rahmen von Umfragen weist im Vergleich zu konventionellen Verfahren deutliche Unterschiede auf. Vorteile sind die geringe Beanspruchung zeitlicher, personeller und finanzieller Ressourcen bei der Durchführung sowie die Erreichbarkeit eines großen Adressatenkreises, aus dem sich die Teilnehmer selbst rekrutieren.¹⁴ Eine solche Selbstrekrutierung kann allerdings zu einem Ergebnis führen, das nicht repräsentativ für die Gesamtheit der Web-Site-Nutzer ist. Sofern hohe Ansprüche an die wissenschaftliche Methode der Datenerhebung bestehen, erhöht sich der Aufwand bei der Erhebung durch planmäßige Teilnehmerselektion, Maßnahmen zur Vermeidung von Mehrfachteilnehmern und Plausibilitätsprüfungen von Eingaben.

Bei der Erfassung personenbezogener Daten sind rechtliche Aspekte des Datenschutzes zu beachten. Maßgeblich für Datenerhebungen im Internet ist das „Informations- und Kommunikationsdienste-Gesetz“ (IuKDG).¹⁵ Personenbezogene Daten werden im Bundesdatenschutzgesetz (BDSG) als „Einzelangaben über persönliche oder sachliche Verhältnisse einer bestimmten oder bestimmbarer natürlichen Person“¹⁶ definiert. Dazu zählen auch grundlegende Bestandsdaten wie Name, Adresse oder Geburtsdatum. Ob eine eMail-Adresse einen Personenbezug zuläßt, ist umstritten.¹⁷ Die Erhebung personenbezogener Daten ist nach dem Gesetz nur zulässig, wenn sie zur Erfüllung eines Vertragsverhältnisses zwischen Nutzer und Web-Site-Anbieter, bspw. bei der Bestellung von Waren, notwendig ist, oder der Nutzer ausdrücklich einwilligt. Für die Identifikation von Nutzern und nicht-anonyme Befragungen bedeutet dies, daß der Nutzer über Zweck und Umfang der Erhebung und Verarbeitung seiner Daten informiert werden muß. Die Weitergabe der Daten an Dritte erfordert ebenfalls eine Einwilligung des Nutzers. Durch eine transparente Darstellung der Datenverwendung erreicht der Anbieter eine rechtliche Absicherung und zugleich eine höhere Akzeptanz und Zufriedenheit unter den Nutzern.

Die Vorgehensweise bei der Datengewinnung mit Hilfe von Web-Formularen beinhaltet mehrere Phasen. Am Anfang steht der Entwurf eines Formulars, für den bereits ein HTML-Editor ausreicht. Datenerfassung und -verwaltung werden meistens durch den Einsatz relationaler Datenbanksysteme oder die Integration in ein Data Warehouse unterstützt. Zur Analyse können die Daten in statistische Anwendungen (z. B. von SPSS oder SAS) eingelesen werden, die auch eine numerische oder graphische Ergebnispräsentation bieten. Neben der Unterstützung des Datenerhebungsprozesses durch Standardanwendungen existieren zahlreiche Programme (z. B. Perception von Question

14 Vgl. Janetzko, Dietmar: Statistische Anwendungen im Internet, a. a. O., S. 143 f.

15 Vgl. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie: Gesetz zur Regelung der Rahmenbedingungen für Informations- und Kommunikationsdienste, Beschluß des Bundestages vom 13. Juni 1997, Online im Internet unter <http://www.iid.de/rahmen/iukdgbt.html>.

16 Vgl. Bundesdatenschutzgesetz § 3 Abs. 1, in: Gola, Peter; Schomerus, Rudolf: BDSG – Bundesdatenschutzgesetz mit Erläuterungen, 6. Aufl., München: Beck 1997, S. 6.

17 Vgl. Höller, Johann: Die rechtlichen Rahmenbedingungen für Electronic Business, in: Internet und Intranet – Auf dem Weg zum Electronic Business, Hrsg.: Höller, Johann; Pils, Manfred; Zlabiger, Robert, 2., neubearb. und erw. Aufl., Berlin et al.: Springer 1999, S. 375 f.

Mark oder Infopoll)¹⁸, die auf Anforderungen und Probleme von Datenerhebungen im Internet abgestimmt sind und den Prozeß teilweise oder vollständig begleiten. Formulare verfügen über ein hohes Automatisierungspotential, da Format, Struktur und Umfang der Eingaben im Vorfeld der Erhebung abgeschätzt werden können. Die Erfassung auf elektronischem Wege ermöglicht zudem einen Transfer ohne Medienbrüche in die Systeme zur Verarbeitung der Daten.

Web-Formulare zählen zu den reaktiven Datenquellen. Der Anbieter ist auf die Kooperation der Nutzer angewiesen, die eine Datenübermittlung zwar bewußt, aber nicht immer völlig freiwillig vornehmen. Obwohl einige Umfragen dem Nutzer nach seiner Teilnahme einen Einblick in das Umfrageergebnis gewähren und auch den Anbieter über den aktuellen Stand einer laufenden Umfrage informieren, zählen Erhebungen durch Web-Formulare eher zu den zeitversetzten Verfahren. Informationen über Nutzerstruktur oder Nutzungstrends lassen sich erst dann gewinnen, wenn die Daten einer größeren Teilnehmeranzahl vorliegen. Damit entsteht eine zeitliche Differenz zwischen der Phase der Datenerfassung und den Entscheidungen, die als Reaktion auf die gewonnenen Informationen getroffen werden. Die Verwendung von Formularen im Rahmen des WSM bietet ein flexibles Erhebungsinstrument für aktuelle Fragestellungen. Von einer Identifikation und Segmentierung der Web-Site-Nutzer kann vor allem das Marketing profitieren und zielgerichtete Maßnahmen zur individuellen Kundenbetreuung ergreifen. In Kombination mit beobachtbaren Nutzungsdaten, z. B. aus Logfiles, können damit Persönlichkeitsprofile entwickelt werden, die Bestands- und Verhaltensdaten einbeziehen.

4 Datenquelle „Web-Logfiles“

4.1 Web Server und Logfiles

Web Server haben die Aufgabe, auf Anfrage von Web Clients Dateien (HTML- und sonstige Dateien) zur Anzeige zur Verfügung zu stellen. Um die Zugriffe auf die bereitgestellten Dateien nachvollziehbar zu machen, zeichnet jede Web-Server-Software Anfragen von Clients in einem oder mehreren Logfiles (i. d. R. Textdateien) auf. Als Logfiles werden Dateien bezeichnet, in denen eingetretene Ereignisse automatisch protokolliert werden.¹⁹ Dem Betreiber eines Web Servers liegen damit Protokolle vor, die die Beanspruchung einer Web Site und von Web-Site-Teilbereichen objektiv abbilden.

Der Aufruf einer Internet-Seite basiert auf den Übertragungsverfahren des Internet-Protokolls HTTP. Dabei gibt der Nutzer auf der Client-Seite in einem Internet-Browser

18 Vgl. Question Mark Corp.: Question Mark Computing – World leaders in computerized testing and assessment. Online im Internet unter <http://www.qmark.com>. Vgl. Infopoll Inc.: Internet Resources for Survey and Polling Software. Online im Internet unter <http://www.infopoll.com/default.htm>. Einen Überblick über die Leistungsfähigkeit von Anwendungsprogrammen zur Erhebung von Daten im WWW bietet: Janetzko, Dietmar: Statistische Anwendungen im Internet, a. a. O., S. 285-329.

19 Vgl. Heinrich, Lutz J.; Roithmayr, Friedrich: Wirtschaftsinformatik-Lexikon, 6., vollständig überarbeitete und erweiterte Auflage, München: Oldenbourg 1998, S. 194, 331.

die Adresse (URL) eines gewünschten Dokumentes an. Der Browser veranlaßt die Herstellung einer Verbindung zu demjenigen Web Server, auf dem das Dokument vorliegt und sendet eine Anfrage (Request) zur Übertragung. Der Server sendet das Dokument bzw. dessen Inhalte an die Adresse des Nutzers (Response) und protokolliert die Übertragung im Logfile.²⁰ Um den physischen Übertragungsweg nicht unnötig zu belasten, wird die Verbindung zwischen Server und Client in der HTTP-Version 1.0 nach jeder Request/Response-Interaktion wieder getrennt.²¹ Das Logfile zeigt daher keinen zusammenhängenden Nutzungsvorgang, sondern eine chronologische Auflistung von bearbeiteten Anfragen. In der aktuellen Version 1.1 ermöglicht HTTP auch persistente Verbindungen, doch die meisten Übertragungen erfolgen noch nach der älteren Version 1.0, um Übertragungsprobleme aufgrund älterer Software-Ausstattung zu vermeiden.

4.2 Inhalte und Meßgrößen von Logfiles

Logfiles treten in einer Vielzahl von Formaten auf, die sich nach Art und Reihenfolge der enthaltenen Angaben unterscheiden. Ein standardisiertes Format ist das Common Logfile Format (CLF), das die meisten Web Server neben proprietären Formaten unterstützen. Jeder Eintrag eines CLF-Logfiles beinhaltet folgende Angaben:

- IP-Adresse oder Domain-Name des Nutzers
- Identifizierung des Client-Rechners
- Authentifizierung des Nutzers
- Datum und Uhrzeit der Anfrage
- Anfragebefehl des Clients
- Angeforderte Datei
- Verwendetes Übertragungsprotokoll
- Dreistelliger Status-Code der Übertragung
- Übertragungsvolumen in Byte

Die IP-Adresse ist eine Ziffernfolge mit einer Länge von 32 Bit, die Rechner im Internet eindeutig identifiziert. Jeder IP-Adresse ist auch ein Domain-Name zugeordnet, der die Lokalisation eines Rechners durch das Domain Name System (DNS) verdeutlicht. Eine Zuordnung beider Daten wird über eine Anfrage an DNS-Server vorgenommen, die Verzeichnisse über IP-Adressen und zugehörige Domain-Namen bereithalten.²² Über die IP-Adresse werden die einzelnen Eintragungen im Logfile zu Nutzungsvorgängen kombiniert. Tritt eine IP-Adresse in kurzem zeitlichem Abstand wiederholt im Logfile auf, ist mit relativ hoher Wahrscheinlichkeit davon auszugehen, daß beide Einträge auf denselben Nutzer zurückzuführen sind. Weitere Angaben zur Identifizierung bzw. Au-

20 Vgl. Köhntopp, Marit; Köhntopp, Kristian: Datenspuren im Internet, a. a. O., S. 250 f.

21 Vgl. Wilde, Erik: World Wide Web, a. a. O., S. 56.

22 Vgl. Köhntopp, Marit; Köhntopp, Kristian: Datenspuren im Internet, a. a. O., S. 248.

thentifizierung erscheinen nur, wenn dies auf der Seite des Clients notwendig wurde. Öffentliche Web-Site-Bereiche verzichten meist auf Zugangsüberprüfungen, wodurch diese Eintragungen im Logfile fehlen. Der Status-Code ist eine Information, die nicht von der Client-Seite übermittelt, sondern vom Web Server erzeugt wird. Darin protokolliert der Server, wie die gestellte Anfrage bearbeitet wurde.²³ Beispielsweise wird für eine fehlerfreie Bearbeitung der Code „200“ verzeichnet. Bekannt ist auch die Meldung „404“, die einem Nutzer anzeigt, daß das angeforderte Dokument auf dem Server nicht gefunden wurde.

Außer den Eintragungen im CLF-Logfile können weitere Informationen festgehalten werden. Die wichtigsten davon sind die Angabe des Referrers und des Agents. Ein Referrer-Eintrag gibt die Adresse der Seite an, die im Browser des Nutzers zuletzt aufgerufen wurde.²⁴ Mit diesen Angaben kann der Weg eines Nutzers durch die Pages einer Web Site nachvollzogen werden, da jeder Logfile-Eintrag eine Verbindung zur vorherigen Seite enthält. Zudem lassen sich diejenigen Einstiegspunkte der Web Site entdecken, die durch Links von anderen Seiten, Werbebanner oder Suchmaschinen entstehen. Bei der Verarbeitung von Suchbegriffen hängen Suchmaschinen die Begriffe oft als Parameter eines Abfrage-Strings an den URL des verarbeitenden Programms an, so daß diese auch im Referrer-Eintrag der gesuchten Seite erscheinen.

Im Agent-Eintrag stehen Angaben zur Software-Ausstattung des Clients, darunter Typ und Version von Browser und Betriebssystem. Aus diesen Angaben ergibt sich ein Bild der technischen Ausstattung der Nutzer. Die Web Site sollte dementsprechend so gestaltet sein, daß der überwiegende Teil der Nutzer die Seite ohne Darstellungsprobleme aufrufen kann. Referrer und Agent werden zum Teil in eigenen Logfiles aufgezeichnet oder mit dem CLF-Logfile kombiniert, was als Combined-Log-Format bezeichnet wird. Ein Ausschnitt aus einem Logfile im Combined-Log-Format wird in Abbildung 1 dargestellt. Weiterhin verdeutlicht die Abbildung auch die genauen Inhalte eines Logfiles am Beispiel eines einzelnen Eintrags.

Die Einbindung aktiver Programmkomponenten, z. B. durch Javascript in einer Web Site, kann dem Anbieter weitere Informationen verschaffen,²⁵ bspw. über die Aktivierung von Javascript, installierte Plug-Ins oder die benutzte Monitorauflösung. Eine ausführliche Liste erfaßbarer Daten zeigt der URL <http://www.anonymizer.com> im Internet.²⁶

23 Vgl. Janetzko, Dietmar: Statistische Anwendungen im Internet, a. a. O., S. 172.

24 Vgl. Heindl, Eduard; Maier, Karin: Der Webmaster – Praktische Realisierung der Internetpräsenz, 2., aktualisierte und erweiterte Auflage, München: Addison-Wesley 2000, S. 185.

25 Vgl. Heindl, Eduard; Maier, Karin: Der Webmaster – Praktische Realisierung der Internetpräsenz, a. a. O., S. 184.

26 Vgl. Anonymizer, Inc.: Anonymizer.com – Online Privacy Services, Online im Internet unter [http://www.anonymizer.com/#Here's what WE know about YOU](http://www.anonymizer.com/#Here's%20what%20WE%20know%20about%20YOU), 12.10.2000.

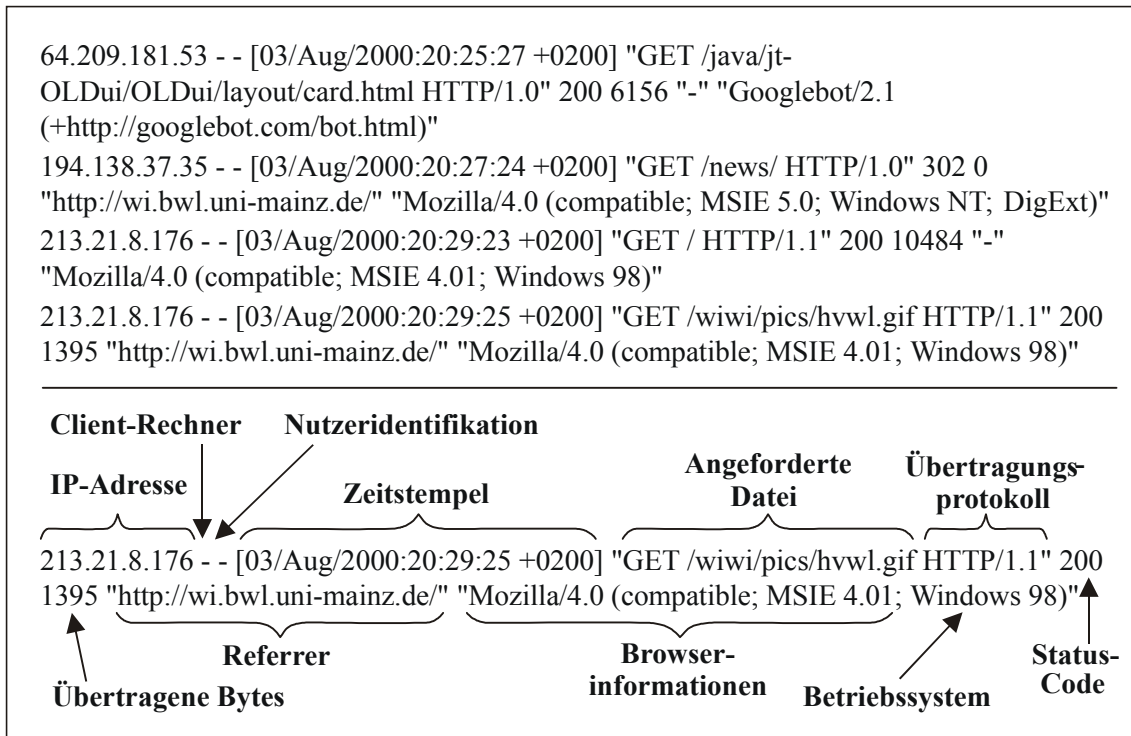


Abb. 1: Bestandteile eines Logfiles im Combined-Log-Format

Der Erfolg einer Internet-Präsenz wird oft an der Anzahl von Zugriffen auf die Web Site gemessen. Sowohl im Vergleich der Nutzungsintensität verschiedener Web Sites als auch zur unternehmensinternen Bewertung der Web-Site-Aktivität ist es notwendig, sich auf eindeutig definierte Meßgrößen zu einigen. Dabei ist zu unterscheiden zwischen objektiv meßbaren und betriebswirtschaftlich interessanten Größen. Ein Unternehmen ist daran interessiert, aussagefähige Zahlen wie die Anzahl der Nutzer festzustellen, verfügt aber in seinen Logfiles nur über eine Aufzeichnung abgerufener Dateien. Den Zusammenhang zwischen beiden Größen verdeutlicht eine hierarchische Gliederung der Web-Site-Aktivität, die folgende Größen definiert:²⁷

- User (Nutzer, Besucher)
- Visits (Sessions, Besuche)
- Page Views (Page Impressions)
- Hits (abgerufene Dateien)

Als User bzw. Nutzer wird eine individuelle natürliche Person bezeichnet, die auf eine Web Site zugreift. Ein Visit ist ein zeitlich zusammenhängender Nutzungsvorgang eines Nutzers. Handelt es sich dabei um einen wiederkehrenden Nutzer, so verursacht dieser im Zeitablauf mehrere Visits. Jeder Visit besteht wiederum aus einem oder mehreren Page Views, d. h. einer Abfolge von einzelnen HTML-Seiten. Dabei kann eine Seite mehrere Objekte beinhalten, die als eigenständige Dateien auf dem Server vorliegen,

²⁷ Vgl. Cutler, Matt; Sterne, Jim: E-Metrics: Business Metrics For The New Economy, Cambridge (MA): NetGenesis Corp. 2000, S. 16. Online im Internet: <http://www.netgen.com/emetrics/#Click> here to download the E-Metrics Whitepaper, 12.10.2000.

z. B. der Seitentext, Bilder oder Videodateien. Jeder Aufruf einer einzelnen Datei wird als Hit bezeichnet und in einem Eintrag des Logfiles protokolliert, so daß die Darstellung einer Seite im Browser mehrere Hits verursachen kann. In Abbildung 2 wird der Zusammenhang der definierten Größen graphisch verdeutlicht.

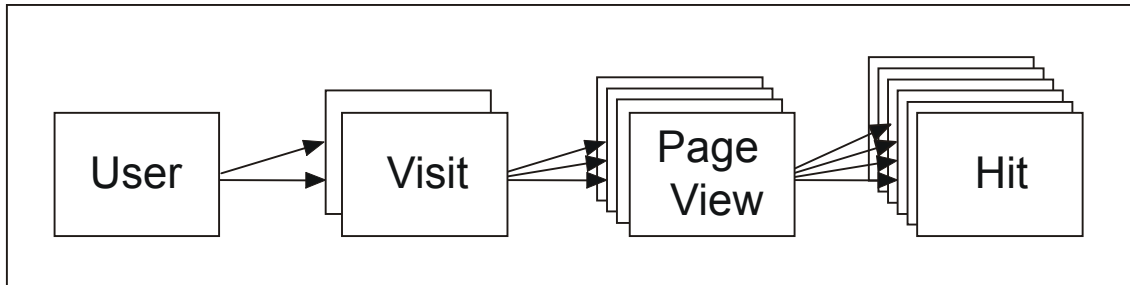


Abb. 2: Hierarchie der Web-Site-Aktivität

Eine verbreitete Meßgröße zur Bestimmung der Werbeeffizienz ist die sogenannte Klickrate (AdClicks).²⁸ Sie beschreibt denjenigen Anteil an Nutzern, der auf die Präsentation eines Werbeträgers, bspw. eines Banners, durch Betätigen des Hyperlinks reagiert. Zur Berechnung der Klickrate werden die Zugriffszahlen der werbetragenden und der beworbenen Seite benötigt. Liegt eine dieser Seiten auf einem fremden Web Server, ist die Kooperation des Besitzers bei der Ermittlung der Zugriffszahlen notwendig.

Von Interesse sind für einen Web-Site-Anbieter nicht nur absolute Nutzungszahlen, sondern auch die Verweildauern der Nutzer auf einzelnen Seiten. Diese sind mit Logfiles schwierig festzustellen, da Logfiles in ihrem Zeitstempel keine Zeiträume sondern nur diskrete Zeitpunkte aufzeichnen. Verweildauern lassen sich durch die Zeitdifferenzen der Logfile-Einträge ermitteln, was jedoch Ungenauigkeiten mit sich bringt. Es werden weder die Übertragungszeiten berücksichtigt, noch Informationen über die Verweildauer auf der letzten betrachteten Seite gewonnen. Eine exakte Zeitbestimmung ist durch die Einbindung eines JavaScript-Programms in die Web Site möglich, das eine Zeitmessung auf der Client-Seite vornimmt. Dies erfordert aber, daß der Browser des Nutzers Javascript unterstützt und diese Option nicht vom Nutzer deaktiviert wurde.²⁹

4.3 Technische Probleme bei der Nutzung von Logfiles

Bedingt durch die einfache Konzeption von Logfiles und die Architektur des Internet entstehen technische Probleme, welche die Qualität und Quantität des Datenbestandes beeinflussen. Eine Beurteilung von Informationen, die auf Logfiles basieren, sollte deshalb folgende Problembereiche nicht außer Acht lassen:

28 Vgl. Janetzko, Dietmar: Surfer im Visier, in: c't, 20/1999, S. 87.

29 Vgl. Dastani, Parsis: Online Mining, in: Wettbewerbsvorteile durch Online Marketing, Hrsg.: Link, Jörg, Berlin et al.: Springer 2000, S. 245.

- Datenspeicherung in Caches
- Abgrenzung von Visits
- Nutzeridentifikation
- Ein- und Ausstiegspunkte der Web Site
- Virtuelle Nutzer (Spider, Bots, Crawler)
- Verwendung von Frames

Caches sind Speicher, die Daten temporär zwischenlagern, um den Zugriff bei einer erneuten Anforderung zu beschleunigen. Im Internet-Verkehr wird so die Auslastung von Verbindungen reduziert.³⁰ Je nach Lokalisation lassen sich Caches auf der Client-Seite, auf der Server-Seite oder auf einem Rechner in der Übertragungskette unterscheiden. Eine solche Zwischenstation kann beispielsweise ein Proxy-Server sein, der die Schnittstelle zwischen dem Internet und einem Intranet bildet und Requests aus dem internen Bereich weiterleitet. Fordert ein Nutzer ein HTML-Dokument an, antwortet der Web Server mit der Übertragung und hält dies im Logfile fest. Auf dem Sendeweg wird im Cache des Proxy-Servers bzw. des Clients eine temporäre Kopie abgelegt. Bei einer erneuten Anforderung desselben Dokuments sucht der Browser zunächst im lokalen Cache und startet nur einen neuen Request, falls das Dokument dort nicht vorliegt. Genauso gibt der Proxy-Server das bei ihm zwischengelagerte Dokument zurück und leitet den Request nicht an den Web Server weiter. Als Resultat erhält der Nutzer nur eine Kopie und der Web Server registriert keine erneute Anfrage (siehe Abbildung 3). In ein Dokument kann zwar eine Anmerkung eingefügt werden, die das Caching unterbindet, doch dies würde dem Sinn des Cache-Prinzips nicht gerecht und eine höhere Belastung der Internet-Verbindungen bedeuten.

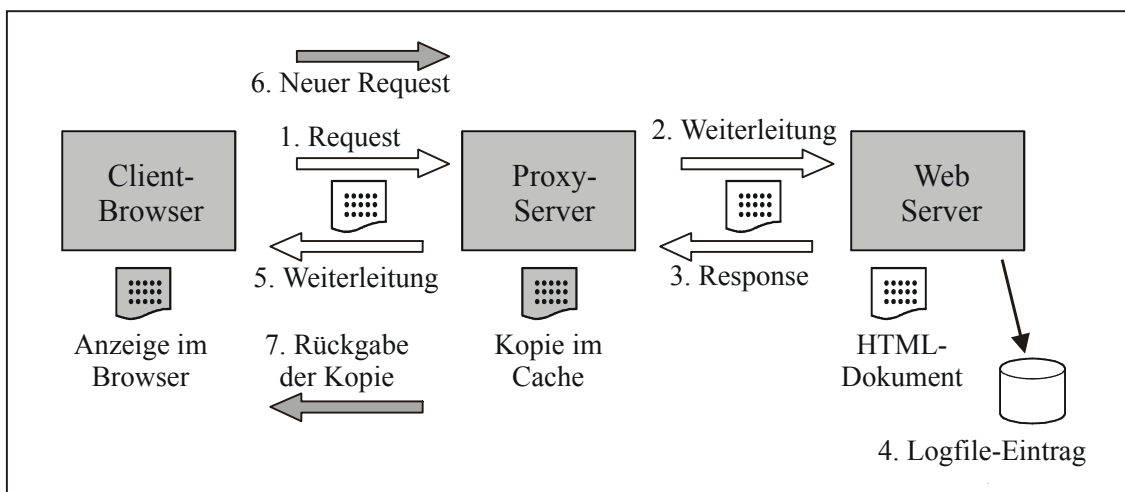


Abb. 3: HTTP Caching

Probleme bei der eindeutigen Abgrenzung von Visits resultieren ebenfalls aus den Eigenheiten von HTTP. Da HTTP nach jeder Datenübertragung die Verbindung zum

30 Vgl. Wilde, Erik: World Wide Web, a. a. O., S. 108.

Client unterbricht und beim nächsten Request wieder neu herstellt, ist ein Nutzungsvorgang nur durch die Auflistung einzelner Zugriffe dokumentiert. Daß es sich um einen zusammenhängenden Vorgang handelt, läßt sich anhand der zeitlichen Nähe einzelner Einträge mit hoher Wahrscheinlichkeit vermuten. In der Praxis werden zur Abgrenzung von Visits Zeitfenster zwischen fünf Minuten und einer halben Stunde vorgegeben,³¹ nach deren Ablauf ein Zugriff auch von derselben IP-Adresse als neuer Besuch der Web Site angesehen wird.

Die Identifikation des Nutzers erfolgt im Logfile durch die IP-Adresse seines Rechners. Ob es sich dabei immer um dieselbe Person handelt, ist aber unsicher, denn in Unternehmen oder Universitäten können mehrere Nutzer den gleichen Rechner verwenden. Auch hier können Proxy-Server die Datenerhebung beeinflussen. An der Schnittstelle zum Internet leiten sie Requests der Clients weiter und versehen sie mit ihrer eigenen IP-Adresse.³² Greifen also mehrere Personen über den gleichen Proxy-Server auf eine Web Site zu, werden ihre Anfragen unter einer IP-Adresse protokolliert. Weiterhin können IP-Adressen auch dynamisch vergeben werden, d. h., ein Nutzer erhält bei jedem Besuch im Internet eine neue Adresse. Verbreitet ist dies vor allem bei großen Providern wie AOL oder T-Online, die für eine große Anzahl an Kunden nur über einen begrenzten Vorrat an IP-Adressen verfügen. In den Logfile-Einträgen können daher mehrere Nutzer mit derselben Adresse oder ein Nutzer unter verschiedenen Adressen auftreten.³³

Eine andere Methode zur Nutzeridentifikation bietet der Einsatz der Cookie-Technologie. Cookies sind kleine Textdateien, die ein Web Server neben einer angeforderten Seite übermittelt und auf der Festplatte des Clients in der Datei cookies.txt speichert.³⁴ Sie enthalten eine eindeutige Identifikationsnummer und die Angabe der Domain, die als Empfänger des Cookies bestimmt ist. Bei jeder erneuten Anfrage wird das Cookie an den Web Server zurückübertragen, so daß dort die eindeutige Identifikation des anfragenden Browsers über mehrere Schritte eines Nutzungsvorgangs möglich ist.³⁵ Durch eine Zeitangabe wird die Lebensdauer des Cookies festgelegt. Ein Nutzer kann daher auch bei wiederholtem Besuch einer Domain Cookiedaten übermitteln und wiedererkannt werden. Der beschriebene Einsatz dieser Technologie schränkt die Anonymität von Nutzern in hohem Maße ein.³⁶ Deshalb verfügen Browser über eine Option, die Annahme von Cookies zu verweigern oder den Benutzer vorher zu informieren. Zudem kann der Nutzer das Cookie-File von seiner Festplatte entfernen. Die Wirksamkeit von Cookies hängt damit von der Kooperationsbereitschaft der Nutzer ab. Eine eindeutige Identifikation ist nur durch eine Anmeldung bei jedem Nutzungsvorgang zu erreichen.

31 Vgl. Bensberg, Frank; Weiß, Thorsten: Web Log Mining als Marktforschungsinstrument für das World Wide Web, in: Wirtschaftsinformatik, 5/1999, S. 429 und net.Genesis Corporation: Build A World Wide Web Commerce Center, New York et al.: John Wiley & Sons 1996, S. 278.

32 Vgl. Garfinkel, Simson; Spafford, Gene: Web Security & Commerce, Cambridge et al.: O'Reilly & Associates 1997, S. 89 f.

33 Vgl. net.Genesis Corporation: Build A World Wide Web Commerce Center, a. a. O., S. 276.

34 Vgl. Garfinkel, Simson; Spafford, Gene: Web Security & Commerce, a. a. O., S. 90.

35 Vgl. Wilde, Erik: World Wide Web, a. a. O., S. 136.

36 Vgl. Garfinkel, Simson; Spafford, Gene: Web Security & Commerce, a. a. O., S. 92.

Eine Web Site ist normalerweise nicht vom übrigen Internet isoliert, sondern verfügt über Kontaktpunkte in Form von Hyperlinks zu Web Sites anderer Anbieter. Die Betätigung eines Links von Suchmaschinen, Werbeflächen oder Kooperationspartnern stellt damit einen Moment dar, an dem ein Nutzer die Web Site betritt oder wieder verläßt. Die Analyse dieser Kontaktpunkte gibt Anhaltspunkte darüber, was ein Nutzer beim Betreten erwartete bzw. wo und warum er das Interesse an der Web Site wieder verlor. Besonders bei Werbeflächen ist die Messung der Kontaktanzahl wichtig, da sich oft der Preis der Werbefläche daraus ableitet. Angaben über die Einstiegspunkte werden im Logfile in den Referrer-Einträgen festgehalten. Wieviele Nutzer eine Seite über einen Link wieder verlassen, ist nicht direkt aus dem Logfile ersichtlich, da der Link einen Request an einen fremden Server aktiviert. Der letzte Eintrag ist daher beim Betreten der Seite erstellt worden, die den Link enthält. Eine Lösung dieses Problems läßt sich durch das Einfügen zusätzlicher Seiten erreichen, die selbst keinen Inhalt anzeigen, sondern zuerst ein Zählprogramm aktivieren und dann unmittelbar auf die Zielseite weiterverbinden.³⁷

Datenspuren in Logfiles müssen nicht immer von realen Personen stammen. Nach der Anmeldung bei einer Suchmaschine wird der Inhalt automatisch indiziert. Dazu dienen Suchprogramme, die als Spider, Crawler oder Bots bezeichnet werden. Diese virtuellen Nutzer hinterlassen charakteristische Spuren in den Agent-Einträgen und IP-Adressen der Logfiles, was eine Filterung der Nutzungszahlen zuläßt.³⁸

Der Inhalt von Logfiles wird nicht nur durch technische Rahmenbedingungen bestimmt, sondern auch durch die Strukturierung der Web Site. Beim Einsatz von Frames besteht eine Seite aus mehreren Dateien. Neben dem Frameset, das Größe und Anordnung der übrigen Frames bestimmt, sind dies Inhalts-, Navigations- oder Werbeframes. Für den Benutzer ist normalerweise nur der Inhaltsframe von Interesse; die anderen Frames dienen nur als Hilfsmittel zur übersichtlichen Gestaltung. Dennoch werden diese bei jedem Serverabruf wieder im Logfile vermerkt und damit als Teil des Nutzungsvorgangs dargestellt. Um die Dokumentation auf die relevanten Inhalte zu beschränken, können die Seiten im Quellcode mit einer Markierung versehen werden, die den Typ des jeweiligen HTML-Dokuments anzeigt.³⁹ Eine Eintragung im Logfile erfolgt nur, wenn die Datei als Seiteninhalt markiert ist, während Framesets oder Navigationsframes nicht protokolliert werden.

37 Vgl. Greenspun, Philip: Datenbankgestützte Web-Sites, München, Wien: Hanser 1998, S. 106 f.

38 Vgl. Puscher, Frank: Logfiles richtig lesen!, in: Internet World, 1/2000, S. 102. Vgl. auch Ideenreich.com: Suchmaschinen - Robots, Spider, Crawler, exotische Browser, Online im Internet unter <http://www.ideenreich.com/dr-rob.shtml>, 12.10.2000.

39 Vgl. Reichelt, Dirk: Zahlen bitte! – So funktioniert das IVW-Verfahren, in: Internet Professionell, 9/2000, S. 58.

4.4 Standardisierungsbemühungen

Aufgrund unterschiedlicher Definitionen von Meßgrößen und den technischen Problemen bei der Auswertung von Logfiles ist bei einem Vergleich der Web-Site-Aktivität verschiedener Anbieter eine Einigung auf ein standardisiertes Meßverfahren notwendig. Für den deutschsprachigen Raum wird ein solches Verfahren von der Informationsgemeinschaft zur Feststellung der Verbreitung von Werbeträgern e. V. (IVW) angeboten. Bei den angeschlossenen Anbietern werden die Meßgrößen Page Impressions (entspricht Page Views) und Visits monatlich erhoben und auf der Web Site der IVW⁴⁰ veröffentlicht.

Gemeldet sind bei der IVW ausschließlich Web Sites, die Werbung Dritter zulassen und durch die Bekanntgabe ihrer Zugriffszahlen potentiellen Werbekunden einen Vergleich der Reichweite ermöglichen. Die Teilnehmer des IVW-Verfahrens implementieren eine spezielle Zählsoftware auf ihrem Web Server, welche die Erfassung der Meßgrößen gemäß den IVW-Kriterien überwacht. Für die Zählung sind die Seiten in Content Pages, Werbeseiten, Navigationsseiten und die Homepage (Startseite eines Angebots) zu klassifizieren. Bei der Durchführung werden nur die Content Pages und die Homepage berücksichtigt. Zusätzlich wird in jede Seite eine ein Pixel große, unsichtbare Graphik eingefügt, die bei jedem Seitenzugriff neu geladen und nicht in einem Cache zwischengespeichert wird. Durch die geringe Größe der Graphik von nur 43 Byte entsteht nur eine geringe Mehrbelastung der Übertragungswege. Ein Eintrag in das Logfile erfolgt bei jedem Aufruf der Graphik. Zur Auswertung stellt die IVW ein eigenes Tool bereit, das Page Impressions und Visits nach Definition der IVW ermittelt. Das IVW-Verfahren bietet weder eine tiefgreifende Analyse der Web-Präsenz, noch exakte Zahlen des Nutzungsverhaltens. Dafür lassen sich einheitliche Zugriffszahlen für die angeschlossenen Web Sites ermitteln, die das quantitative Werbepotential anzeigen.⁴¹

4.5 Datenschutzrechtliche Rahmenbedingungen

Auch bei der Verarbeitung von Nutzungsdaten aus Logfiles sind die datenschutzrechtlichen Vorschriften des IuKDG zu beachten. Das Gesetz geht bei der Erfassung von Daten durch Telediensteanbieter⁴² von folgenden Grundsätzen aus:⁴³

40 Vgl. Informationsgesellschaft zur Feststellung der Verbreitung von Werbeträgern e. V.: Willkommen bei der IVW, Online im Internet unter <http://www.ivw.de>.

41 Vgl. Reichelt, Dirk: Zahlen bitte! – So funktioniert das IVW-Verfahren, a. a. O., S. 58.

42 Teledienste sind nach § 2 Abs. 1 des Teledienstegesetzes als „Informations- und Kommunikationsdienste, die für eine individuelle Nutzung von kombinierbaren Daten (...) bestimmt sind und denen eine Übermittlung mittels Telekommunikation zugrundeliegt“ definiert. Vgl. Engel-Flechsig, Stefan: „Teledienstedatenschutz“ – Die Konzeption des Datenschutzes im Entwurf des Informations- und Kommunikationsdienstegesetzes des Bundes, in: Datenschutz und Datensicherheit, 1/1997, S. 10.

43 Vgl. Engel-Flechsig, Stefan: „Teledienstedatenschutz“ – Die Konzeption des Datenschutzes im Entwurf des Informations- und Kommunikationsdienstegesetzes des Bundes, a. a. O., S. 9 f. Vgl. auch Seidel, Uwe M.: Das neue Datenschutzrecht der Telediensteanbieter, in: Wirtschaftswissenschaftliches Studium, 12/1998, S. 636 ff.

- Keine Verarbeitung personenbezogener Daten ohne Notwendigkeit dazu
- Zweckbindung der Verarbeitung an die Erbringung von IuK-Diensten
- Transparente Darstellung der Datenverwendung
- Technische Sicherung der Nutzeranonymität
- Kontrolle durch eine unabhängige Instanz

Die Verwendung personenbezogener Daten zu Marketingzwecken oder zur Gestaltung von Web Sites ist ausschließlich mit Einwilligung des Nutzers zulässig.⁴⁴ Um Nutzungsdaten in Logfiles verwenden zu dürfen, ist darauf zu achten, daß die Daten anonymisiert vorliegen und keinen Personenbezug aufweisen. Andernfalls wären diese Nutzungsdaten sofort nach Ende der Nutzung wieder zu löschen,⁴⁵ es sei denn, sie werden für Abrechnungszwecke benötigt. Bei kostenlos nutzbaren Web-Angeboten wäre damit sogar das Führen von Logfiles unzulässig.⁴⁶

Zur technischen Wartung und Weiterentwicklung der Web Site sind Logfiles mit anonymen Einträgen in den meisten Fällen ausreichend. Fehler, die bei der Nutzung auftreten oder benötigte Übertragungskapazitäten in einem bestimmten Zeitraum lassen sich auch unabhängig von der Zuordnung zu individuellen Nutzern erfassen. Die darauf basierenden Anpassungsmaßnahmen betreffen nur selten einzelne Nutzer, sondern eher allgemeine Veränderungen der technischen Gestaltung einer Web Site. Im Marketingbereich ist hingegen die Verbindung erhobener Nutzungsdaten mit vorliegenden Bestandsdaten von Kunden von hohem Interesse. Gelingt die Verbindung von objektiven Verhaltensdaten und Demographika, entstehen detaillierte Persönlichkeitsprofile, die es erlauben, Kunden individuell anzusprechen und zu betreuen. Die Erstellung von Persönlichkeitsprofilen ist nach dem Gesetz nur bei Verwendung von Pseudonymen erlaubt; eine Zusammenführung personenbezogener Daten ist unzulässig.⁴⁷ Liegt eine wirksame Einwilligung des Nutzers in die Erstellung eines Persönlichkeitsprofils vor, ist das informationelle Selbstbestimmungsrecht gegenüber der gesetzlichen Schutzfunktion des Verbotes abzuwägen. Hier bewertet das Datenschutzrecht das Schutzbedürfnis des Nutzers als höherwertig, da dieser auch bei Einwilligung keine Möglichkeit hat, sich über Entwicklung und Richtigkeit des Profils zu informieren. Persönlichkeitsprofile sind daher nur unter erheblichen rechtlichen Bedenken anzufertigen.⁴⁸

Insgesamt erweist sich das Datenschutzrecht für Telediensteanbieter als erheblich restriktiver als für Anbieter mit konventionellen Kommunikations- und Vertriebswegen, bspw. beim Postversand von Waren, und bildet so ein schwerwiegendes Hindernis für

44 Vgl. Baeriswyl, Bruno: Data Mining und Data Warehousing: Kundendaten als Ware oder geschütztes Gut?, in: Recht der Datenverarbeitung, 1/2000, S. 8.

45 Vgl. Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie: Gesetz zur Regelung der Rahmenbedingungen für Informations- und Kommunikationsdienste, a. a. O., Artikel 2 § 6 Abs. 2 Nr. 1.

46 Vgl. Höller, Johann: Die rechtlichen Rahmenbedingungen für Electronic Business, a. a. O., S. 376.

47 Vgl. Engel-Flehsig, Stefan: „Teledienstedatenschutz“ – Die Konzeption des Datenschutzes im Entwurf des Informations- und Kommunikationsdienstegesetzes des Bundes, a. a. O., S. 14.

48 Vgl. Wittig, Petra: Die datenschutzrechtliche Problematik der Anfertigung von Persönlichkeitsprofilen zu Marketingzwecken, in: Recht der Datenverarbeitung, 2/2000, S. 61 f.

Marketing-Aktivitäten.⁴⁹ In der Praxis ist zu erwarten, daß diese restriktive Rechtslage absichtlich oder unbeabsichtigt wenig Beachtung erfährt.⁵⁰ Dies beruht überwiegend auf der Unwissenheit der Nutzer und den wirtschaftlichen Interessen in Verbindung mit einem geringen Unrechtsbewußtsein auf Seiten der Anbieter.

4.6 Verarbeitungsverfahren von Logfiles

Logfiles liegen zunächst als Textdateien auf dem Server vor. Eine Verarbeitung und Auswertung kann auf dem gleichen Rechnersystem erfolgen, was aber zu einer erhöhten Beanspruchung der Server-Ressourcen führt. Bei Servern, die eine große Anzahl an Requests beantworten müssen, ist es daher sinnvoller, die Logfiles in festen Zeitintervallen auf ein anderes Rechnersystem zu übertragen und dort die Analyse durchzuführen. Zur Weiterverarbeitung existieren zahlreiche Programme, die Logfile-Daten unterschiedlicher Formate organisieren, auswerten und präsentieren (siehe dazu Kapitel 8). Verwaltet werden die Daten in relationalen Standard-Datenbanken oder proprietären Datenbankformaten der Analyseprogramme. Deutliche Unterschiede gibt es in der Anzahl und Qualität der erstellten Berichte, die sowohl graphisch als auch tabellarisch präsentiert werden. Ausgereifte Produkte beherrschen neben der Darstellung von Berichten in Form von HTML-Seiten auch den Export in weitere Datenformate. Führende Anbieter von Logfile-Analyseprogrammen sind die Firmen Webtrends und für den deutschsprachigen Raum Exody mit ihrem Produkt WebSuxess.⁵¹ Komplexe und dementsprechend kostspielige Anwendungssysteme zur eBusiness-Analyse, wie z. B. Accrue Insight oder net.Analysis,⁵² erlauben die Integration weiterer interner und externer Daten. Die Nutzungsdaten können damit zum inhaltlichen Bestandteil eines Data Warehouse werden. Mit Hilfe von Data-Mining-Verfahren wird der Datenbestand dort auf das Vorhandensein impliziter Strukturen analysiert.

Da Logfiles automatisch vom Web Server angelegt werden, ist diese Form der Datenerhebung ein nicht-reaktives Verfahren. Selbst wenn der Nutzer über diese Technik informiert ist, hat er keine Möglichkeit, die Einträge im Logfile zu beeinflussen, ohne sein Nutzungsverhalten dabei zu verändern. Die Analyse von Logfiles erfolgt in den meisten Fällen zeitversetzt. Um Veränderungen des Nutzungsverhaltens feststellen zu können,

49 Vgl. Imhof, Ralf: One-to-One-Marketing im Internet – Das TDDSG als Marketinghindernis, in: Computer und Recht, 2/2000, S. 110, 116.

50 Vgl. Baeriswyl, Bruno: Data Mining und Data Warehousing: Kundendaten als Ware oder geschütztes Gut?, a. a. O., S. 10.

51 Vgl. WebTrends Corp.: Welcome to WebTrends Corporation, Online im Internet unter <http://www.webtrends.com/default.htm> und Exody E-Business Intelligence GmbH: Exody – WebSuxess, Online im Internet unter <http://www.exody.net/ger/index.cfm>. Für einen Vergleich der Leistungsfähigkeit von zehn gängigen Logfile-Analyseprogrammen vgl. Thorbrügge, Marco: Erfolgreich im Web, in: PC Professionell, 7/2000.

52 Vgl. Accrue Software, Inc.: eBusiness Analysis and Accrue Insight - White Paper / Part II of II, Fremont (CA): Accrue Software, Inc. 1999, S. 2. Download im Internet unter <http://www.accrue.com/products/whiteprs.html> #Accrue Software, Inc. Vgl. auch net.Genesis Corp.: net.Analysis - An Architecture for Acquiring E-Business Intelligence, Cambridge (MA): net.Genesis Corp. 1998, S. 5. Download im Internet unter <http://www.netgen.com/forms/whitepaper/> #Architecture Whitepaper.

beinhalten Logfiles die Daten konstanter und damit vergleichbarer Zeitintervalle (Tage, Wochen oder Monate). Die Berichtserstellung und darauf basierende Handlungsmaßnahmen erfolgen daher ebenfalls zeitversetzt. Um eine zeitnahe Informationsversorgung zu gewährleisten, verwenden einige Analyseprogramme, z. B. LiveStats von Media-House Software,⁵³ ungenutzte Prozessorkapazitäten für die Verarbeitung von Logfiles. Die Verarbeitung erfolgt ständig als Hintergrundprozeß, wodurch jederzeit der Zugriff auf aktuelle Nutzungsdaten möglich ist.

Die Logfile-Analyse bildet zur Zeit den Hauptbestandteil der Aktivitäten des Web Site Monitoring. Dabei ist allerdings zu beachten, daß diese Datenquelle sehr nahe an der technischen Ebene des Betriebs einer Web Site anzusiedeln ist. Für die Bewertung der qualitativen und quantitativen Serverleistung geben die gesammelten Daten ein objektives Bild ab, doch Rückschlüsse auf Nutzerverhalten und Nutzerinteressen enthalten immer ein beträchtliches Maß an Unsicherheit.

5 Datenquelle „Netzwerküberwachung“

Die Überwachung des Netzwerkverkehrs stellt eine noch größere Annäherung an die technischen Grundlagen einer Web Site dar als die Verwendung von Logfiles. Nachrichten werden im Internet in Form von Datenpaketen übertragen, die an ihrem Bestimmungsort, d. h. einem Rechnersystem, wieder zusammengesetzt werden. Bei der Beobachtung der Übertragung kommen sogenannte Packet Sniffer zum Einsatz, die den Strom von Paketen im Netzwerkverkehr passiv beobachten und protokollieren, ohne die Datenpakete selbst zu verändern.⁵⁴ Implementiert werden diese Werkzeuge auf einem eigenständigen Rechnersystem zwischen dem zu beobachtenden Web Server und der Schnittstelle zum Internet, bzw. zu einem internen Netzwerk.⁵⁵ Die Protokollierung von Daten auf einem eigenen Rechner, der dem Web Server vorgeschaltet ist, macht das Führen von Logfiles auf dem Server überflüssig, so daß dieser seine Ressourcen voll zur Beantwortung eingehender Requests einsetzen kann. Für die Datenerhebung und Datenanalyse steht ein eigenes System zur Verfügung, das Daten in Echtzeit auswertet. Durch die Verlagerung der Monitoring-Aktivität ist es nicht notwendig, Logfiles vom Server zu übertragen oder direkt dort auszuwerten. Die Positionierung von Packet Sniffern erlaubt zudem die Beobachtung des Web Servers und erfaßt Daten, die diesem verborgen bleiben. Außer allen Informationen, die auch in einem Logfile festgehalten werden können, lassen sich zusätzlich folgende Daten dokumentieren:⁵⁶

53 Vgl. MediaHouse Software Inc.: Statistics Server LiveStats – Website Statistics Software, Online im Internet unter <http://www.mediahouse.com/statisticsserver>, 12.10.2000.

54 Vgl. Gibson, Steve: OptOut – How to Watch Spyware Watching You!, Online im Internet unter <http://grc.com/oo/packetsniff.htm>, 12.10.2000.

55 Vgl. WildPackets Inc.: Analyzing Complex Web Environments, S. 8. Online im Internet unter <http://www.wildpackets.com/support/support.html> #Analyzing Complex Web Environments, 12.10.2000.

56 Vgl. Wildpackets Inc.: Analyzing Complex Web Environments, a. a. O., S. 6. Vgl. auch Accrue Software, Inc.: eBusiness Analysis and Accrue Insight – White Paper / Part I of II, Fremont (CA): Accrue Software, Inc. 1999, S. 7. Online im Internet unter <http://www.accrue.com/products/whiteprs.html> #Accrue Software Inc.

- Steuerungssignale an den Server, z. B. Abbruch der Übertragung nach Betätigen des Stop-Buttons beim Client
- Tatsächliches Übertragungsvolumen in Byte
- Übertragungsvolumen einzelner Applikationen
- Antwortzeit und -verhalten des Servers
- Übertragungszeit zum Client

Damit erhält der Web-Site-Anbieter aufschlußreiche Informationen über die Leistungsfähigkeit und Ressourcenbeanspruchung seiner Netzwerk-Architektur. Verwendung finden diese Informationen im Rahmen der Kapazitätsplanung oder bei der effizienteren Gestaltung der Web Site.

Der Einsatz von Instrumenten zur Netzwerküberwachung weist auch eine Reihe von Nachteilen auf.⁵⁷ Die Notwendigkeit, einen Packet Sniffer vor dem Web Server zu installieren, erfordert eine Zugriffsmöglichkeit auf diesen Netzwerk-Bereich. Unternehmen, die den Betrieb ihrer Web Server Internet Service Providern (ISP) überlassen, können dieses Verfahren nur mit der Zustimmung des ISP anwenden. Der Ausfall eines Packet Sniffers führt zum Verlust der Protokolldaten im betreffenden Zeitraum, wobei die Funktionsfähigkeit des Web Servers davon völlig unbeeinflusst sein kann. Um eine kontinuierliche Datenerfassung zu gewährleisten, ist daher die Funktion des Packet Sniffers regelmäßig zu überprüfen, was den Aufwand für Pflege und Instandhaltung der Web Site erhöht. Nachteilig wirkt sich zudem die zunehmende Dynamisierung von Web Seiten aus. Dynamisierte Pages präsentieren bei ihrem Aufruf kein fest vorgegebenes HTML-Dokument, sondern werden vom Web Server unter Verwendung aktueller Parameter für jede Anfrage eigens generiert. Kriterien, die den Verarbeitungsprozeß innerhalb des Web Servers beeinflussen können, sind Präferenzen oder Eigenschaften des jeweiligen Nutzers und Faktoren, die vom Web-Site-Anbieter abhängen. Beispielsweise kann ein Web Server je nach Herkunft eines Nutzers mit einer Seite in dessen Landessprache reagieren oder ein Shop-System in einer Web Site nur solche Artikel anzeigen, bei denen noch ein ausreichender Lagerbestand vorrätig ist. Durch die Trennung vom Server ist ein Packet Sniffer nicht in der Lage, die Entscheidungskriterien für die Auswahl einer bestimmten Seite zu dokumentieren, sondern kann nur das übertragene Ergebnis registrieren.

Der Einsatz von Packet Sniffern ist auch mit sicherheitstechnischen Bedenken verbunden. Der Datenstrom in einem Netzwerksegment besteht nicht ausschließlich aus Paketen, die für den zu beobachtenden Server bestimmt sind. Durch die Netzwerküberwachung ist eine Protokollierung aller Daten möglich, die ein Netzwerksegment als physischen Übertragungsweg nutzen.⁵⁸ Packet Sniffer können deshalb auch eingesetzt werden, um sensible Daten, wie z. B. Passwörter, auszuspähen.

57 Vgl. net.Genesis Corp.: net.Analysis - An Architecture for Acquiring E-Business Intelligence, a. a. O., S. 10.

58 Vgl. Gibson, Steve: OptOut – How to Watch Spyware Watching You!, a. a. O.

Die Netzwerküberwachung stellt wie die Verwendung von Logfiles ein nicht-reaktives Verfahren dar. Erhobene Daten können in Echtzeit ausgewertet werden, wodurch Informationen zeitnah zur Verfügung stehen. In besonderem Maße profitiert die Datenverarbeitung von der Netzwerküberwachung. Durch die Zuordnung von Übertragungsmengen zu einzelnen Servern bzw. Clients kann die effektive Netzbelastung einzelner Komponenten bestimmt werden. In einer Intranet-Umgebung liegt damit z. B. eine Basis zur Bestimmung von Kostengrößen für die Beanspruchung von Netzwerkressourcen vor.⁵⁹

6 Datenquellen außerhalb der Web Site

Die bisher erwähnten Quellen setzen immer eine elektronische Datenübertragung zwischen Client und Server voraus. Zur Weiterentwicklung der Web Site können aber auch Daten verwendet werden, die sich bereits im internen Bestand des Unternehmens befinden oder von externen Informationsanbietern bezogen werden. Zusätzliche interne Daten liegen oft schon elektronisch gespeichert vor, so daß die Integration mit den Web-Site-Nutzungsdaten und die Weiterverarbeitung ohne hohen Übertragungsaufwand möglich ist. Einen Beitrag zur Analyse und Entwicklung der Web-Präsenz können folgende interne Datenquellen leisten:

- Kundendatenbanken
- Kundenhistorien
- Kundenservice
- ERP-Systeme
- Logistik-Informationssysteme
- Konfigurationsdaten der IuK-Systeme
- Kostenrechnungsdaten des Web-Site-Betriebs

Der Einbezug von Kundendaten faßt Informationen aus anderen Kommunikationskanälen mit denen der Web Site zusammen. Daraus kann ein vollständigeres Bild des Kunden entwickelt und der Einsatz von Marketing-Instrumenten auf die individuellen Bedürfnisse abgestimmt werden. Ähnlich stellt sich die Situation für Unternehmensprozesse im Intra- und Extranet dar. Die Web-Technologie bietet hier einen allgemeinen, flexiblen Standard, der Schnittstellenprobleme reduziert. Eine vollständige Betrachtung der Unternehmensprozesse und des Unterstützungspotentials der Web Site ist nur durch die Verbindung von Daten aus unterschiedlichen Quellen möglich, z. B. durch Integration von traditionellen Kommunikationsmöglichkeiten und Daten des Rechnungswesens in die Beobachtung logistischer Verflechtungen eines Unternehmens oder die Analyse des Informationsbedarfs von Mitarbeitern bzw. Abteilungen und dessen Abdeckung durch das Informationsangebot der Web Site. Auch die Bewertung der technischen Leistungsfähigkeit läßt sich durch die Integration zusätzlicher interner Daten, z. B. Infor-

59 Vgl. WildPackets Inc.: Analyzing Complex Web Environments, a. a. O., S. 9.

mationen über Eigenschaften von Komponenten der Web-Site-Architektur oder subjektive Erfahrungsberichte von Anwendern, unterstützen. Die objektive Beurteilung der Web Site erfordert den Vergleich der tatsächlichen, meßbaren Leistung mit den Kapazitäten der technischen Infrastruktur.

Externe Daten sind mit einem höheren Aufwand bei Beschaffung und Integration verbunden. Die folgenden Datenquellen beinhalten externe Informationen, die für das Web Site Monitoring von Nutzen sein können:

- Studien von Marktforschungsinstituten
- Statistisches Datenmaterial öffentlicher Institutionen
- Adressdaten von Listbrokern
- Mikrogeographische Daten
- Presseberichte über das Unternehmen
- Berichte über Potentiale neuer Technologien

Mit diesen externen Quellen verschafft sich ein Unternehmen ein genaueres Verständnis bestehender oder potentieller Kunden bzw. Märkte. Eigene Beobachtungen der Web-Site-Nutzer werden mit Marktsegmenten im Internet oder der Gesamtheit der Internet-Nutzer verglichen, wodurch sich typische Eigenschaften der bisherigen Nutzer erkennen lassen. Diese Informationen dienen der Prognose von Kundengruppen, die für eine profitable Kundenbeziehung in Frage kommen und direkt oder durch die attraktive Gestaltung eines entsprechenden Web-Site-Angebots angesprochen werden. In Berichten über das Unternehmen können Kommentare zur Web-Präsenz, Produktbewertungen oder Imageinformationen enthalten sein, die auf die Unternehmenspräsentation über die Web Site rückwirken. Eine aktuelle und qualitativ hochwertige Web Site sollte im Weiterentwicklungsprozeß auch technologische Neuerungen auf ihr Nutzenpotential überprüfen. Die Nutzbarkeit der Web Site muß dabei aber für den überwiegenden Teil der Nutzer gewährleistet bleiben.

Nach Auswahl und Beschaffung relevanter Daten werden diese gegebenenfalls in IuK-Systeme transferiert und dort aufbereitet. Eine Speicherung erfolgt gemäß der Struktur der Daten in geeigneten Systemen zur Datenverwaltung, wie z. B. relationalen Datenbanken, Systemen zur Dokumentenverwaltung oder einem Data Warehouse. Die Integration zusätzlicher interner bzw. externer Datenquellen ist notwendig, weil sich Nutzerinteressen und potentielle Wettbewerbsvorteile gegenüber den Web Sites von Konkurrenten in allen Informationskanälen eines Unternehmens artikulieren können.

7 Zusammenfassender Überblick der Datenquellen

Anschließend werden die besprochenen Datenquellen des Web Site Monitoring anhand ihrer wichtigsten Merkmale vergleichend gegenübergestellt. Dabei beinhaltet der Überblick in Abbildung 4 nur diejenigen Erhebungsinstrumente im einzelnen, die auf eine Datenübertragung per Internet zurückgreifen, da sich die sonstigen externen und internen Datenquellen in Struktur und Inhalt zum Teil erheblich unterscheiden.

Datenquellen des Web Site Monitoring				
Merkmal	eMail	Web-Formular	Logfile	Netzwerküberwachung
Initiator der Datenerhebung	Nutzer Web-Site-Anbieter	Web-Site-Anbieter	Web-Site-Anbieter	Web-Site-Anbieter
Reaktivität	Reaktiv	Reaktiv	Nicht-reaktiv	Nicht-reaktiv
Objektivität der Informationen	Fehlinformation möglich	Fehlinformation möglich	Objektive Messung	Objektive Messung
Freiwilligkeit der Übermittlung	Freiwillig	Teilweise freiwillig Zugangsberechtigung als Anreiz	Unfreiwillig	Unfreiwillig
Inhaltliche Flexibilität	Variabler Inhalt	Flexibel Zielgerichtete Erhebung	Begrenzt	Begrenzt
Analysezeitpunkt	Zeitversetzt	Zeitversetzt	Zeitversetzt	Echtzeit-Analyse
Analyseverfahren	Mitarbeiter Text-Mining-Systeme	Speicherung in Datenbanken Statistische Verfahren	Speicherung in Datenbanken Statistische Verfahren Logfile-Analyseprogramme	Speicherung in Datenbanken Statistische Verfahren Packet Sniffer
Automatisierungspotential	Niedrig	Hoch	Hoch	Hoch

Abb. 4: Datenquellen des Web Site Monitoring

8 Software-Produkte zur Logfile-Analyse

8.1 Kategorisierung der Software-Produkte

Für die Ermittlung von Nutzungsdaten der Web-Präsenz aus Logfiles existiert bereits ein reichhaltiges Software-Angebot. Nachfolgend wird eine Auswahl kommerzieller Lösungen zur Logfile-Analyse vorgestellt. Das erhältliche Angebot an Software läßt sich in die drei Kategorien Freeware-Tools, Logfile-Analyzer und eBusiness-Analyseprogramme unterteilen.

Freeware-Tools werden hier nicht berücksichtigt, da sie meist aus Eigenbedarf an Analyseprogrammen von Web-Site-Entwicklern oder Programmierern entwickelt und danach der Öffentlichkeit zur Weiterverwendung zur Verfügung gestellt werden. Daraus resultiert, daß diese Programme oft in hohem Maß auf die funktionalen Anforderungen und technischen Systemvoraussetzungen des Entwicklers abgestimmt sind. Andere Nut-

zer von Freeware-Tools müssen daher einige Nachteile und Risiken in Kauf nehmen, darunter z. B.:

- Erhöhter Anpassungsaufwand an bestehende Systemarchitekturen
- Eingeschränkt nutzbarer Funktionsumfang
- Unzureichende optische Gestaltung
- Fehlerhafter Programmcode
- Kein Produktsupport
- Mangelhafte Dokumentation

Diese Produkte scheinen deshalb für die Einbindung in ein professionelles und systematisches Web-Site-Monitoring wenig geeignet, bieten aber möglicherweise für private Web-Site-Betreiber oder Web Sites von geringer geschäftlicher Relevanz eine preiswerte Alternative zu kommerziellen Analyseprogrammen.

Ebenfalls nicht in der Auswahl vorgestellter Programme enthalten sind komplexe eBusiness-Analysertools, da diese auf sehr umfangreiche und komplexe Web-Präsenzen abzielen. Ihr Einsatz erfordert eine zeit- und ressourcenintensive Anpassung an bereits vorhandene Anwendungssysteme. Die Integration in die bestehende Systemarchitektur umfaßt dabei auch die Anbindung an unternehmenseigene Datenbanksysteme, die zur Speicherung und Verwaltung umfangreicher Datenbestände aus der Nutzung der Web Site verwendet werden. Die Entscheidung über Auswahl und Einsatz eines solchen Systems liegt nicht zuletzt auch aufgrund des hohen Anschaffungspreises nicht im Ermessen einer für die Web Site verantwortlichen Person, sondern ist im Bereich der strategischen Entscheidungen einzuordnen. Für den Betreiber einer Web Site von geringem bis mittlerem Umfang sind Produkte dieser Kategorie deutlich überdimensioniert.

In den folgenden Kapiteln werden einige typische und verbreitete Produkte aus der Gruppe der kommerziellen Logfile-Analysertools in Frage. Die Anbieter dieser Software versuchen, eine Vielzahl professioneller Web-Site-Betreiber zu erreichen und bieten Programme, die Logfiles übersichtlich tabellarisch und graphisch darstellen. Die Auswahl innerhalb dieser Gruppe beschränkt sich auf Programme, die als kostenlose Testversion im Internet angeboten werden.⁶⁰ Weiteres Kriterium war, daß Installation und Bedienung des Programms einfach vorzunehmen ist, ohne laufende Anwendungen zu beeinflussen. Die getesteten Programme generieren ihre Berichte innerhalb eigenständiger Windows-Anwendungen oder als eine Reihe von HTML-Seiten, die in einem Web Browser dargestellt werden.

Grundlage der Analyse war ein Logfile, die Nutzungsdaten einer Web Site für einen Zeitraum von 77 Tagen enthielt. Der Umfang dieses Files betrug ca. 475.000 Einträge (ca. 80 MB).

60 Die Web-Adressen werden beim jeweiligen Produkt angegeben. Die Nutzungsdauer dieser Testversionen ist zeitlich auf 14 Tage bis 4 Wochen befristet.

8.2 FastStats Analyzer V2.74

Anbieter:	Mach 5 Enterprises
Web Site:	http://www.mach5.com/
Preis:	99,95 \$

FastStats hat sich seinen Namen wirklich verdient. Auch umfangreiche Logfiles werden in kürzester Zeit eingelesen und verarbeitet. Durch die übersichtliche Benutzeroberfläche, die große Ähnlichkeiten zu Windows-Standardanwendungen aufweist, ist die Nutzung des Programms intuitiv erfassbar und schnell verständlich. Wie im Windows-Explorer werden die einzelnen Berichtsgruppen in Form einer Baumstruktur im linken Fensterbereich angezeigt, während im Hauptfenster der jeweils gewählte Bericht zu sehen ist. Die Berichte sind sowohl in tabellarischer, als auch in graphischer Form verfügbar, wobei allerdings beide Darstellungsarten nicht gleichzeitig, sondern nur wechselweise im Fenster angezeigt werden können. Inhaltlich bietet das Programm die üblichen Informationen zu Besuchern, Zugriffsstatistiken, Referrer-Angaben und den häufigsten Fehlern. Anzahl und Flexibilität der Berichte ist allerdings recht gering. Der Nutzer wird nicht einer Vielzahl von Graphiken und Tabellen gegenübergestellt, sondern erhält die wichtigsten Daten der Web Site präsentiert, ohne interessante Sachverhalte weiter analysieren zu können.

Die Möglichkeiten des Datenexports sind ebenfalls eingeschränkt. Neben der Darstellung als HTML-Seite können Berichte als Word-Dokument ausschließlich in tabellarischer Form gespeichert werden.

Besonderes Merkmal ist die Fähigkeit, die Struktur einer Web Site als Baumdiagramm darzustellen. Gerade bei der Verfolgung von Klickpfaden ist diese Funktion eine nützliche Orientierungshilfe.

FastStats bietet ein leicht erlernbares und leistungsfähiges Programm, das vor allem für schnelle Informationen während Betrieb und Weiterentwicklung der Web Site geeignet ist. Verbesserungspotentiale existieren aber noch in der Berichtsgenerierung, insbesondere wenn diese in unterschiedlichen Datenformaten oder gedruckter Form benötigt werden.

8.3 Funnel Web 4.0

Anbieter:	Active Concepts Inc.
Web Site:	http://www.activeconcepts.com/
Preis:	Ab 349,- \$

Auf den ersten Blick ist die Benutzeroberfläche von Funnel Web etwas gewöhnungsbedürftig. Nach Programmstart öffnet sich zunächst nur ein Fenster mit fünf Buttons. Erst bei deren Betätigung gelangt der Benutzer zu weiteren Fenstern, die neben der Hilfefunktion Einstellmöglichkeiten des Programms, eine Funktion zur Automatisierung der

Analyseaufgaben, einen Generator für individuelle Berichte oder die Anzeige der automatisch erstellten Berichte enthalten. Die Berichte werden als HTML-Seiten in einem Browser dargestellt. Durch die Verwendung von Frames ist der Browser in einen Navigationsbereich und einen Darstellungsbereich unterteilt. Erwähnenswert ist die gelungene optische Gestaltung der aufbereiteten Daten. Wo andere Programme lediglich bunte und einfach gestaltete Graphiken erstellen, zeigt Funnel Web Diagramme, die Präsentationsqualität besitzen.

Die Anzahl der generierten Berichte ist zufriedenstellend, doch auch hier macht sich eine mangelnde Detaillierbarkeit der präsentierten Daten bemerkbar. Als einziges Programm im Test weist Funnel Web einen Loyalitätsindex aus. Abhängig von der Besucheranzahl und weiteren einstellbaren Faktoren werden den Nutzern der Web Site Loyalitätswerte zugeordnet.

Ihre Stärken zeigt die Software in den reichhaltigen Möglichkeiten zur Automatisierung. Die Verarbeitung von Logfiles kann zu festen Terminen voreingestellt werden, das Programm ist fernbedienbar und kann auch als Windows-NT-Service installiert werden. Zudem simuliert Funnel Web eine Echtzeitanalyse der Logfile-Daten. In einstellbaren Zeitintervallen werden die inzwischen erfolgten Ergänzungen des Logfiles eingelesen und verarbeitet.

Das Programm eignet sich vor allem für die regelmäßige Informationsbereitstellung am Bildschirm, da ein Export in verbreitete Formate wie z. B. in Word- oder Excel-Dateien nicht möglich ist.

8.4 Summary 1.5

Anbieter:	Summary.net
Web Site:	http://www.summary.net/
Preis:	249,- \$

Summary faßt sämtliche Berichte in Form von HTML-Dateien zusammen. Die Anzeige der Daten und die Konfiguration des Programms läuft vollständig im Web Browser ab. Ausgehend von einer Hauptseite, auf der alle Berichtstypen aufgeführt sind, präsentiert die Software ihre Auswertungen sehr übersichtlich. Das umfangreiche Angebot erstellter Berichte entschädigt dabei den Benutzer etwas für die sehr statische Darstellung der analysierten Daten, die keine detaillierten Nachforschungen zuläßt. Die graphische Präsentation ist sehr spärlich gehalten und beschränkt sich auf ein nicht veränderbares Säulendiagramm zu jeder tabellarischen Darstellung. Auch die Änderungsmöglichkeiten der Voreinstellungen zur Berichtsgenerierung, sowie die Konfigurationseinstellungen des Programms bieten nicht viele Möglichkeiten zur Variation.

Vorteile bietet das Programm durch die Vielzahl erstellter Berichte. Unter den Angaben zur Software-Ausstattung finden sich z. B. auch Häufigkeitsverteilungen zur Verwendung von Java, Cookies oder Plugins.

Summary kann seinen Ursprung aus der Kategorie der Free- bzw. Sharewareprogramme nicht ganz verbergen. Es stellt eine gelungene technische Lösung dar, die Logfiles sehr schnell verarbeitet und eine Vielzahl ungewöhnlicher Berichte bietet, doch in der optischen Gestaltung und der Anwenderfreundlichkeit zeigen sich noch zahlreiche Verbesserungsmöglichkeiten.

8.5 SurfStats Log Analyzer Professional Edition

Anbieter:	SurfStats
Web Site:	http://www.surfstats.com/index.htm
Preis:	295,- \$

SurfStats weist einen schwerwiegenden Nachteil auf, der sämtliche positiven Ansätze des Programms in den Hintergrund stellt: Der Zeitbedarf zur Verarbeitung von Logfiles ist außerordentlich hoch. Beim Einlesen eines Logfiles stehen zwei Modi zur Auswahl. Entweder werden nur Einträge beachtet, die einem definierbaren Auswahlfilter entsprechen (z. B. bestimmte Zeiträume oder Dateien) oder die gesamten Informationen des Files werden verarbeitet. Für die detaillierte Analyse des Testfiles von 80 MB benötigte SurfStats etwa eine halbe Stunde Rechenzeit, während andere Programme das File in deutlich weniger als einer Minute verarbeiteten. Ursache dieses immensen Zeitbedarfs ist, daß die Software das ursprüngliche Textfile zunächst in ein eigenes Datenbankformat überträgt. Auch nach Einlesen des Logfiles benötigt SurfStats beim erstmaligen Aufruf eines Berichts wiederum etwa eine Minute, bis aus der Datenbank die Berichtsdaten ermittelt und angezeigt werden. Nach der Erstellung eines Berichts wird dieser bei weiteren Zugriffen aber ohne Zeitverzögerung angezeigt.

Demgegenüber hebt sich das Programm auch durch einige außergewöhnliche Ideen von anderen Logfile-Analysern ab. Zur Gestaltung der Berichte bietet SurfStats eine Palette von Graphikoptionen, mit denen der Benutzer Farben, Darstellung und Hintergrund von Diagrammen verändern und den Berichten ein qualitativ hochwertiges Erscheinungsbild verleihen kann.

Eine weitere Besonderheit stellt die Mehrsprachigkeit des Programms dar, das seine Berichte in insgesamt sieben Sprachen präsentieren kann. In deutscher Sprachausgabe finden sich zwar einige Schreibfehler und holperige Satzkonstruktionen, doch die Beschriftungen bzw. Erläuterungen können in einem Editor nachbearbeitet werden, so daß eine fehlerfreie Beschriftung mit etwas Nachbearbeitung zu erreichen ist.

Die Benutzeroberfläche ist sehr übersichtlich gestaltet. Der Berichtsbereich ähnelt optisch der Windows-Hilfefunktion und weist die übliche Trennung in Navigations- und Reportbereich auf.

Die Stärken des Programms liegen damit nicht im eigentlichen Bereich der Analyse von Logfiles; hingegen sind die Gestaltungsmöglichkeiten und die Mehrsprachigkeit Funk-

tionen, die dieses Produkt aus einer Menge gleichartiger Programme hervorheben. Durch die geringe Verarbeitungsgeschwindigkeit verliert SurfStats jedoch viel von seiner Konkurrenzfähigkeit.

8.6 SurfTrace Logfile Analyzer 2.11 Prof. Edition

Anbieter:	Surftrace Internet Software
Web Site:	http://www.surftrace.de/
Preis:	199,- DM

Auch SurfTrace benötigt für die Auswertung größerer Logfiles außergewöhnlich lange, da Logfile-Daten zunächst in ein proprietäres Datenformat übertragen werden. Neben dem Zeitbedarf ist aber auch die Ressourcenbeanspruchung durch Surftrace ungewöhnlich hoch, wodurch sämtliche auf dem System laufende Anwendungen beeinträchtigt werden.

Sowohl inhaltlich, als auch in ihrer Benutzerfreundlichkeit kann die Software nicht mit den Konkurrenzprodukten mithalten. Die Navigation ist umständlich zu handhaben, optisch bietet das Programm wenig ansprechende Tabellen und Diagramme und die Auswahl an Berichten ist vergleichsweise gering. Die Software ist darüber hinaus nur zur Nutzung am Bildschirm zu verwenden, da eine Exportfunktion der Berichte in gebräuchliche Formate nicht zur Verfügung steht.

Positiv fällt auf, daß dem Benutzer eine Einführungstour durch das Programm angeboten wird. Innerhalb weniger Minuten lassen sich die Grundfunktionen von SurfTrace erlernen. Die qualitativen Mängel kann die Software dadurch aber nicht ausgleichen.

8.7 WebSuxess 4 Enterprise

Anbieter:	Exody E-Business Intelligence GmbH
Web Site:	http://www.exody.net/ger/index.cfm
Preis:	880,12 DM

WebSuxess ist das erfolgreichste deutschsprachige Produkt unter den Logfile-Analyseprogrammen. Die Software generiert nicht nur hochwertige Berichte, sondern ist auch auf Besonderheiten, wie z. B. das IVW-Verfahren abgestimmt. Über eine Schnittstelle zur Software der IVW können von WebSuxess die standardisierten Kennzahlen Page Views und Visits ausgewiesen werden.

Die Benutzeroberfläche ist verständlich aufgebaut und leicht handhabbar. Für Anfänger bietet WebSuxess einen Assistenten, der die wichtigsten Funktionen kurz vorstellt. Ebenfalls verfügbar sind Anleitungen, mit denen der Anwender schnell eigene Analysen generieren kann.

Die Berichte sind übersichtlich in sieben Gruppen (Zusammenfassung, zeitliche Nutzungsverteilung, Seitenbereiche, Nutzerstatistik, Navigation, Kampagnen und Browserstatistik) angeordnet. Nützliche Zusatzfunktionen erleichtern die Arbeit mit dem Datenmaterial. So sind z. B. die URLs der Web Site kommentierbar, wodurch kryptische Adressen der eigenen Web Site mit Klartexten bezeichnet werden können. Die Web-Adressen sind durch Hyperlinks unterlegt, so daß die entsprechende Seite zu ihren Informationen aufgerufen werden kann. In der Benutzerstatistik besteht die Möglichkeit, Namen aus den Adressen der Nutzer herauszufiltern, wodurch das Verhalten bestimmter Personen oder Organisationen isoliert beobachtet werden kann. WebSuxess beschränkt sich auch nicht darauf, nur Typ und Version der genutzten Browser anzuzeigen, sondern erstellt auch Statistiken zu deren Fähigkeiten wie z. B. der Verwendung von Frames.

Die graphische Präsentation von Daten läßt sich in Darstellungsart und -tiefe nach den Wünschen des Anwenders einrichten. Die voreingestellte Darstellungsart (z. T. auch abhängig vom Datenmaterial) ist nicht immer besonders aussagekräftig, so daß eine nachträgliche Einstellung der Graphiken teilweise nötig ist.

Insgesamt bietet WebSuxess eine große Anzahl an Berichten in einer ausgereiften Benutzeroberfläche. Klare Struktur und hohe Verarbeitungsgeschwindigkeit ermöglichen auch ungeübten Anwendern einen schnellen Einstieg in die Analyse von Logfiles.

8.8 WebTrends Professional Suite

Anbieter:	WebTrends Corp.
Web Site:	http://www.webtrends.com/default.htm
Preis:	799,- \$

WebTrends ist in Bezug auf die Flexibilität im Umgang mit unterschiedlichen Datenformaten das vielseitigste der hier besprochenen Programme. Es unterstützt sowohl eine große Anzahl einzulesender Logfile-Formate, als auch viele Exportformate, darunter unter anderem auch den Export als Word- oder Excel-Datei.

Komfortabel in der Bedienung wird die Software durch zahlreiche Möglichkeiten zur Automatisierung. WebTrends kann als Windows-NT-Service eingerichtet werden, über einen Scheduler Logfiles automatisiert auswerten und diese Files auch per FTP übertragen.

Dargestellt werden die Berichte von WebTrends ähnlich wie bei den meisten anderen Programmen in einem Fenster mit Navigations- und Berichtsbereich, wobei die Auswahl an verfügbaren Informationen sehr umfangreich ist. Jeder Bericht besteht aus einer graphischen Darstellung, wobei der Diagrammtyp nicht veränderlich ist, einem Tabeleinteil, sowie einem Bereich, der die dargestellten Informationen erläutert. Dieser Hilfetext enthält Definitionen, Interpretationsansätze der dargestellten Daten und Hinweise auf Programmeinstellungen, welche die Berichtsgenerierung beeinflussen. Damit wird das präsentierte Datenmaterial auch schnell für Anwender verständlich, die wenig

Kenntnisse im Bereich Logfile-Analyse besitzen. Einziger Kritikpunkt an den Berichten ist das Auftreten einiger Darstellungsfehler im Test. Dabei wurde die Sortierung von Daten nach Häufigkeiten tabellarisch nicht korrekt dargestellt. Davon abgesehen ist die Berichtserstellung äußerst flexibel handhabbar. Die Darstellung läßt sich an die Vorstellungen des Anwenders anpassen und ist auch in deutscher Sprache verfügbar.

WebTrends bietet ein ausgereiftes Produkt, das sowohl Anfängern durch zahlreiche Hilfen und die Verträglichkeit mit einer Vielzahl an Datenformaten zu empfehlen ist, als auch Fortgeschrittene durch das breite Angebot verfügbarer Informationen zufriedenstellt.

Literaturverzeichnis

- Accrue Software, Inc.: eBusiness Analysis and Accrue Insight - White Paper / Part II of II, Fremont (CA): Accrue Software, Inc. 1999. Online im Internet unter <http://www.accrue.com/products/whiteprs.html> #Accrue Software, Inc., 12.10.2000.
- Anonymizer, Inc.: Anonymizer.com – Online Privacy Services, Online im Internet unter <http://www.anonymizer.com/> #Here's what WE know about YOU, 12.10.2000.
- Baeriswyl, Bruno: Data Mining und Data Warehousing: Kundendaten als Ware oder geschütztes Gut?, in: Recht der Datenverarbeitung, 1/2000, S. 6-11.
- Bensberg, Frank; Weiß, Thorsten: Web Log Mining als Marktforschungsinstrument für das World Wide Web, in: Wirtschaftsinformatik, 5/1999, S. 426-432.
- Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie: Gesetz zur Regelung der Rahmenbedingungen für Informations- und Kommunikationsdienste, Beschluß des Bundestages vom 13. Juni 1997. Online im Internet unter <http://www.iid.de/rahmen/iukdgbt.html>, 12.10.2000.
- Cutler, Matt; Sterne, Jim: E-Metrics: Business Metrics For The New Economy, Cambridge (MA): NetGenesis Corp. 2000. Online im Internet: <http://www.netgen.com/emetrics/#Click> here to download the E-Metrics Whitepaper, 12.10.2000.
- Dastani, Parsis: Online Mining, in: Wettbewerbsvorteile durch Online Marketing, Hrsg.: Link, Jörg, Berlin et al.: Springer 2000, S. 235-259.
- Engel-Flechsigt, Stefan: „Teledienstedatenschutz“ – Die Konzeption des Datenschutzes im Entwurf des Informations- und Kommunikationsdienstegesetzes des Bundes, in: Datenschutz und Datensicherheit, 1/1997, S. 8-16.
- Garfinkel, S.; Spafford, G.: Web Security & Commerce, Cambridge: O'Reilly & Ass. 1997.
- Gentsch, Peter: Business Intelligence: Aus Daten systematisch Wissen entwickeln, in: Electronic Business und Knowledge Management – Neue Dimensionen für den Unternehmenserfolg, Hrsg.: Scheer, August-Wilhelm, Heidelberg: Physica-Verlag 1999, S. 167-195.
- Gibson, Steve: OptOut – How to Watch Spyware Watching You!, Online im Internet unter <http://grc.com/oo/packetsniff.htm>, 12.10.2000.
- Gola, Peter; Schomerus, Rudolf: BDSG – Bundesdatenschutzgesetz mit Erläuterungen, 6. Aufl., München: Beck 1997.
- Greenspun, Philip: Datenbankgestützte Web-Sites, München, Wien: Hanser 1998.
- Hagel III., John; Rayport, Jeffrey F.: Die kommende Schlacht um Kundendaten, in: Erfolg im E-Business, Hrsg.: Tapscott, Don, München, Wien: Hanser 2000, S. 182-195.
- Heindl, Eduard; Maier, Karin: Der Webmaster – Praktische Realisierung der Internetpräsenz, 2., aktualisierte und erweiterte Auflage, München: Addison-Wesley 2000.
- Heinrich, Lutz J.; Roithmayr, Friedrich: Wirtschaftsinformatik-Lexikon, 6., vollständig überarbeitete und erweiterte Auflage, München: Oldenbourg 1998.
- Heuer, Kai R.; Wilken, Markus: Ansätze zur Datengewinnung für das Controlling im Online-Marketing, in: Handbuch Marketing-Controlling, Hrsg.: Zerres, Michael P., Berlin et al.: Springer 2000, S. 309-319.
- Höllner, Johann: Die rechtlichen Rahmenbedingungen für Electronic Business, in: Internet und Intranet – Auf dem Weg zum Electronic Business, Hrsg.: Höllner, Johann; Pils, Manfred; Zlabinger, Robert, 2., neubearb. und erw. Aufl., Berlin et al.: Springer 1999, S. 351-384.
- Ideenreich.com: Suchmaschinen - Robots, Spider, Crawler, exotische Browser, Online im Internet unter <http://www.ideenreich.com/dr-rob.shtml>, 12.10.2000.

- Imhof, Ralf: One-to-One-Marketing im Internet – Das TDDSG als Marketinghindernis, in: Computer und Recht, 2/2000, S. 110-116.
- Infopoll Inc.: Internet Resources for Survey and Polling Software. Online im Internet unter <http://www.infopoll.com/default.htm>, 12.10.2000.
- Informationsgesellschaft zur Feststellung der Verbreitung von Werbeträgern e. V.: Willkommen bei der IVW, Online im Internet unter <http://www.ivw.de>, 12.10.2000.
- Janetzko, D.: Statistische Anwendungen im Internet, München: Addison-Wesley 1999.
- Janetzko, Dietmar: Surfer im Visier, in: c't, 20/1999, S. 86-92.
- Köhntopp, M. und K.: Datenspuren im Internet, in: Computer und Recht, 4/2000, S. 248-257.
- Link, Jörg: Zur zukünftigen Entwicklung des Online Marketing, in: Wettbewerbsvorteile durch Online Marketing, Hrsg.: Link, Jörg, 2. Aufl., Berlin et al.: Springer 2000, S. 1-34.
- McKenna, Regis: Marketing in Echtzeit, in: Erfolg im E-Business, Hrsg.: Tapscott, Don, München, Wien: Hanser 2000, S. 167-181.
- MediaHouse Software Inc.: Statistics Server LiveStats – Website Statistics Software, Online im Internet unter <http://www.mediahouse.com/statisticsserver>, 12.10.2000.
- net.Genesis Corp.: net.Analysis - An Architecture for Acquiring E-Business Intelligence, Cambridge (MA): net.Genesis Corp. 1998. Online im Internet unter <http://www.netgen.com/forms/whitepaper/#Architecture> Whitepaper, 12.10.2000.
- net.Genesis Corporation: Build A World Wide Web Commerce Center, New York et al.: John Wiley & Sons 1996.
- Pils, Manfred: Electronic Business im Blickfeld der Sensiblen Informations- und Kommunikationssysteme, in: Internet und Intranet – Auf dem Weg zum Electronic Business, Hrsg.: Höllner, J.; Pils, M.; Zlabiger, R., 2. Aufl., Berlin et al.: Springer 1999, S. 285-305.
- Puscher, Frank: Logfiles richtig lesen!, in: Internet World, 1/2000, S. 100-103.
- Question Mark Corp.: Question Mark Computing – World leaders in computerized testing and assessment. Online im Internet unter <http://www.qmark.com>, 12.10.2000.
- Reichelt, Dirk: Zahlen bitte! – So funktioniert das IVW-Verfahren, in: Internet Professionell, 9/2000, S. 56-58.
- Reichmann, T.; Fritz, B.; Nölken, D.: EIS-gestütztes Controlling: Schnittstelle zwischen Controlling und Informationsmanagement, in: Handbuch Informationsmanagement, Hrsg.: Scheer, August-Wilhelm, Wiesbaden: Gabler 1993, S. 463-489.
- Schwickert, A. C.; Beiser, A.: Web Site Controlling, in: Arbeitspapiere WI, Nr. 7/1999, Hrsg.: Lehrstuhl für Allg. BWL und Wirtschaftsinformatik, Johannes Gutenberg-Univ. Mainz 1999.
- Seidel, Uwe M.: Das neue Datenschutzrecht der Teledienstanbieter, in: Wirtschaftswissenschaftliches Studium, 12/1998, S. 635-642.
- Thorbrügge, Marco: Erfolgreich im Web, in: PC Professionell, 7/2000, S. 186-201.
- WebTrends Corp.: Welcome to WebTrends Corporation, Online im Internet unter <http://www.webtrends.com/default.htm> und Exody E-Business Intelligence GmbH: Exody – Web-Success, Online im Internet unter <http://www.exody.net/ger/index.cfm>, 12.10.2000.
- Wilde, Erik: World Wide Web, Berlin et al.: Springer 1999.
- WildPackets Inc.: Analyzing Complex Web Environments. Online im Internet unter <http://www.wildpackets.com/support/support.html>, 12.10.2000.
- Wittig, Petra: Die datenschutzrechtliche Problematik der Anfertigung von Persönlichkeitsprofilen zu Marketingzwecken, in: Recht der Datenverarbeitung, 2/2000, S. 59-62.
- Zierl, Marco: Anreize bieten: So erfahren Sie mehr über Ihre Kunden, in: Internet Professionell, 9/2000, S. 46-51.

Bisher erschienen

Stand: Dezember 2000 – Den aktuellen Stand der Reihe erfahren
Sie über unsere Web Site unter <http://wi.uni-giessen.de>

Nr. 1/1996	Grundlagen des Client/Server-Konzepts.....	Schwicker/Grimbs
Nr. 2/1996	Wettbewerbs- und Organisationsrelevanz des Client/Server-Konzepts.....	Schwicker/Grimbs
Nr. 3/1996	Realisierungsaspekte des Client/Server-Konzepts	Schwicker/Grimbs
Nr. 4/1996	Der Geschäftsprozeß als formaler Prozeß - Definition, Eigenschaften, Arten	Schwicker/Fischer
Nr. 5/1996	Manuelle und elektronische Vorgangsteuerung.....	Schwicker/Rey
Nr. 6/1996	Das Internet im Unternehmen - Neue Chancen und Risiken	Schwicker/Ramp
Nr. 7/1996	HTML und Java im World Wide Web.....	Gröning/Schwicker
Nr. 8/1996	Electronic-Payment-Systeme im Internet.....	Schwicker/Franke
Nr. 9/1996	Von der Prozeßorientierung zum Workflow-Management - Teil 1: Grundgedanken, Kernelemente, Kritik	Maurer
Nr. 10/1996	Von der Prozeßorientierung zum Workflow- Management - Teil 2: Prozeßmanagement und Workflow	Maurer
Nr. 11/1996	Informationelle Unhygiene im Internet.....	Schwicker/Dietrich/Klein
Nr. 12/1996	Towards the theory of Virtual Organisations: A description of their formation and figure.....	Appel/Behr
Nr. 1/1997	Der Wandel von der DV-Abteilung zum IT-Profitcenter: Mehr als eine Umorganisation.....	Kargl
Nr. 2/1997	Der Online-Markt - Abgrenzung, Bestandteile, Kenngrößen	Schwicker/Pörtner
Nr. 3/1997	Netzwerkmanagement, OSI Framework und Internet SNMP	Klein/Schwicker
Nr. 4/1997	Künstliche Neuronale Netze - Einordnung, Klassifikation und Abgrenzung aus betriebswirtschaftlicher Sicht	Strecker/Schwicker
Nr. 5/1997	Sachzielintegration bei Prozeßgestaltungsmaßnahmen.....	Delnef
Nr. 6/1997	HTML, Java, ActiveX - Strukturen und Zusammenhänge.....	Schwicker/Dandl
Nr. 7/1997	Lotus Notes als Plattform für die Informationsversorgung von Beratungsunternehmen.....	Appel/Schwaab
Nr. 8/1997	Web Site Engineering - Modelltheoretische und methodische Erfahrungen aus der Praxis	Schwicker
Nr. 9/1997	Kritische Anmerkungen zur Prozeßorientierung	Maurer/Schwicker
Nr. 10/1997	Künstliche Neuronale Netze - Aufbau und Funktionsweise	Strecker
Nr. 11/1997	Workflow-Management-Systeme in virtuellen Unternehmen	Maurer/Schramke
Nr. 12/1997	CORBA-basierte Workflow-Architekturen - Die objektorientierte Kernanwendung der Bausparkasse Mainz AG	Maurer
Nr. 1/1998	Ökonomische Analyse Elektronischer Märkte.....	Steyer
Nr. 2/1998	Demokratiopolitische Potentiale des Internet in Deutschland	Muzic/Schwicker
Nr. 3/1998	Geschäftsprozeß- und Funktionsorientierung - Ein Vergleich (Teil 1)	Delnef
Nr. 4/1998	Geschäftsprozeß- und Funktionsorientierung - Ein Vergleich (Teil 2)	Delnef
Nr. 5/1998	Betriebswirtschaftlich-organisatorische Aspekte der Telearbeit	Polak
Nr. 6/1998	Das Controlling des Outsourcings von IV-Leistungen	Jäger-Goy
Nr. 7/1998	Eine kritische Beurteilung des Outsourcings von IV-Leistungen.....	Jäger-Goy
Nr. 8/1998	Online-Monitoring - Gewinnung und Verwertung von Online-Daten.....	Guba/Gebert
Nr. 9/1998	GUI - Graphical User Interface.....	Maul
Nr. 10/1998	Institutionenökonomische Grundlagen und Implikationen für Electronic Business.....	Schwicker
Nr. 11/1998	Zur Charakterisierung des Konstrukts "Web Site".....	Schwicker
Nr. 12/1998	Web Site Engineering - Ein Komponentenmodell.....	Schwicker
Nr. 1/1999	Requirements Engineering im Web Site Engineering – Einordnung und Grundlagen.....	Schwicker/Wild
Nr. 2/1999	Electronic Commerce auf lokalen Märkten	Schwicker/Lüders
Nr. 3/1999	Intranet-basiertes Workgroup Computing	Kunow/Schwicker
Nr. 4/1999	Web-Portale: Stand und Entwicklungstendenzen.....	Schumacher/Schwicker
Nr. 5/1999	Web Site Security.....	Schwicker/Häusler
Nr. 6/1999	Wissensmanagement - Grundlagen und IT-Instrumentarium.....	Gaßen
Nr. 7/1999	Web Site Controlling.....	Schwicker/Beiser
Nr. 8/1999	Web Site Promotion	Schwicker/Arnold
Nr. 9/1999	Dokumenten-Management-Systeme – Eine Einführung	Dandl
Nr. 10/1999	Sicherheit von eBusiness-Anwendungen – Eine Fallstudie	Harper/Schwicker
Nr. 11/1999	Innovative Führungsinstrumente für die Informationsverarbeitung	Jäger-Goy
Nr. 12/1999	Objektorientierte Prozeßmodellierung mit der UML und EPK	Dandl
Nr. 1/2000	Total Cost of Ownership (TCO) – Ein Überblick.....	Wild/Herges
Nr. 2/2000	Implikationen des Einsatzes der eXtensible Markup Language – Teil 1: XML-Grundlagen.....	Franke/Sulzbach
Nr. 3/2000	Implikationen des Einsatzes der eXtensible Markup Language – Teil 2: Der Einsatz im Unternehmen	Franke/Sulzbach
Nr. 4/2000	Web-Site-spezifisches Requirements Engineering – Ein Formalisierungsansatz	Wild/Schwicker
Nr. 5/2000	Elektronische Marktplätze – Formen, Beteiligte, Zutrittsbarrieren	Schwicker/Pfeiffer
Nr. 6/2000	Web Site Monitoring – Teil 1: Einordnung, Handlungsebenen, Adressaten.....	Schwicker/Wendt
Nr. 7/2000	Web Site Monitoring – Teil 2: Datenquellen, Web-Logfile-Analyse, Logfile-Analyzer	Schwicker/Wendt
Nr. 8/2000	Controlling-Kennzahlen für Web Sites.....	Schwicker/Wendt
Nr. 9/2000	eUniversity – Web-Site-Generierung und Content Management für Hochschuleinrichtungen.....	Schwicker/Ostheimer/Franke

Bestellung (bitte kopieren, ausfüllen, zusenden/zufaxen)

Adressat: Professur für BWL und Wirtschaftsinformatik
 Fachbereich Wirtschaftswissenschaften
 Licher Straße 70
 D – 35394 Gießen
 Telefax: (0 641) 99-22619

Hiermit bestelle ich gegen Rechnung die angegebenen Arbeitspapiere zu einem Kostenbeitrag von DM 10,- pro Exemplar (MwSt. entfällt) zzgl. DM 5,- Versandkosten pro Sendung.

Nr.	An
1/1996	
2/1996	
3/1996	
4/1996	
5/1996	
6/1996	
7/1996	
8/1996	
9/1996	
10/1996	
11/1996	
12/1996	

Nr.	An
1/1997	
2/1997	
3/1997	
4/1997	
5/1997	
6/1997	
7/1997	
8/1997	
9/1997	
10/1997	
11/1997	
12/1997	

Nr.	Anz
1/1998	
2/1998	
3/1998	
4/1998	
5/1998	
6/1998	
7/1998	
8/1998	
9/1998	
10/1998	
11/1998	
12/1998	

Nr.	Anz
1/1999	
2/1999	
3/1999	
4/1999	
5/1999	
6/1999	
7/1999	
8/1999	
9/1999	
10/1999	
11/1999	
12/1999	

Nr.	Anz
1/2000	
2/2000	
3/2000	
4/2000	
5/2000	
6/2000	
7/2000	
8/2000	
9/2000	

Absender:

Organisation _____

Abteilung _____

Nachname, Vorname _____

Straße _____

Plz/Ort _____

Telefon _____ Telefax _____ eMail _____

Ort, Datum _____ Unterschrift _____