

Data Deluge

Welche Informationen wie gesammelt werden...

Doris Köberl

Nicole Kowatsch

„Wir ertrinken in Informationen,
aber hungern nach Wissen.

John Naisbitt (*1930), amerik. Prognostiker

Web Analytics

- Gezielt Informationen zu sammeln und auszuwerten
- Website Optimierung
- Zielgruppen Definierung
-

Logfile-Analysen

Jeder Dateiaufruf (Hit) wird aufgezeichnet.

- Aufrufdatum und Zeit
- URL
- IP-Adresse
- Verwendete Browser und Betriebssystem (User Agent)
- Cookie
- Herkunft des Aufrufs (Referer)
- Ergebnis des Aufrufes (Status)
- Abgeschlossene Downloads und Übertragungsrate
- Leichte Wiederaufbereitung historischer Daten



Nachteile:

- Caching, Proxy
- Framesets, Java Script, Flash, Ajax
- Besucheridentifikation
- Suchmaschinen-Crawler und Sonden
- Verteilte Systeme
- Zu wenig Informationen und hoher Aufwand

Page Tagging

- Verwendete Browser, eingestellte Browsersprache
- Aktivierte Funktionen wie Java und Cookies
- Mausbewegung, Position des Zeigers, Mausklicks
- Tastatureingaben im Browser und Cursor-Position
- Bildschirmauflösung
- Abspieldauer bzw. Unterbrechung von Multimedia-Daten
- Besuchte Links
- Verweildauer
- uvm.

Nachteile:

- Seiten ohne Page Tags
- Suchmaschinen-Crawler und Sonden
- Web Beacon
 - Bsp. Ad-Blocker der Banner von Drittsites ausblendet
- RSS-Feeds
- Abgeschlossene Downloads und Übertragungsrate

Traffic Quellen

Wie sind die Besucher auf die Website gelangt?

- Analyse und Optimierung umliegender Verknüpfungen
- Suchmaschinen-Keywords
- Direktzugriff, Lesezeichen
-

Favoriten

Social Bookmarks (<http://del.icio.us>)

- <http://del.icio.us/rss?url?url=http://www.ebay.de/>
- <http://badges.del.icio.us/feeds/json/url/data?hash=51fc89b9f7f31f7d420334c5f9f091dc>

```
[{"hash":"51fc89b9f7f31f7d420334c5f9f091dc","title":"eBay Deutschland \u2013  
Der weltweite Online-  
Marktplatz","url":"http:\:\/\/www.ebay.de\/","total_posts":3414,"top_tags":{"shopping":813,"ebay":758,"shop":331,"einkaufen":316,"auktion":299,"auction":235,"auktionen":170,"german":90,"kaufen":71,"verkaufen":69}}]
```

Verweisende Websites

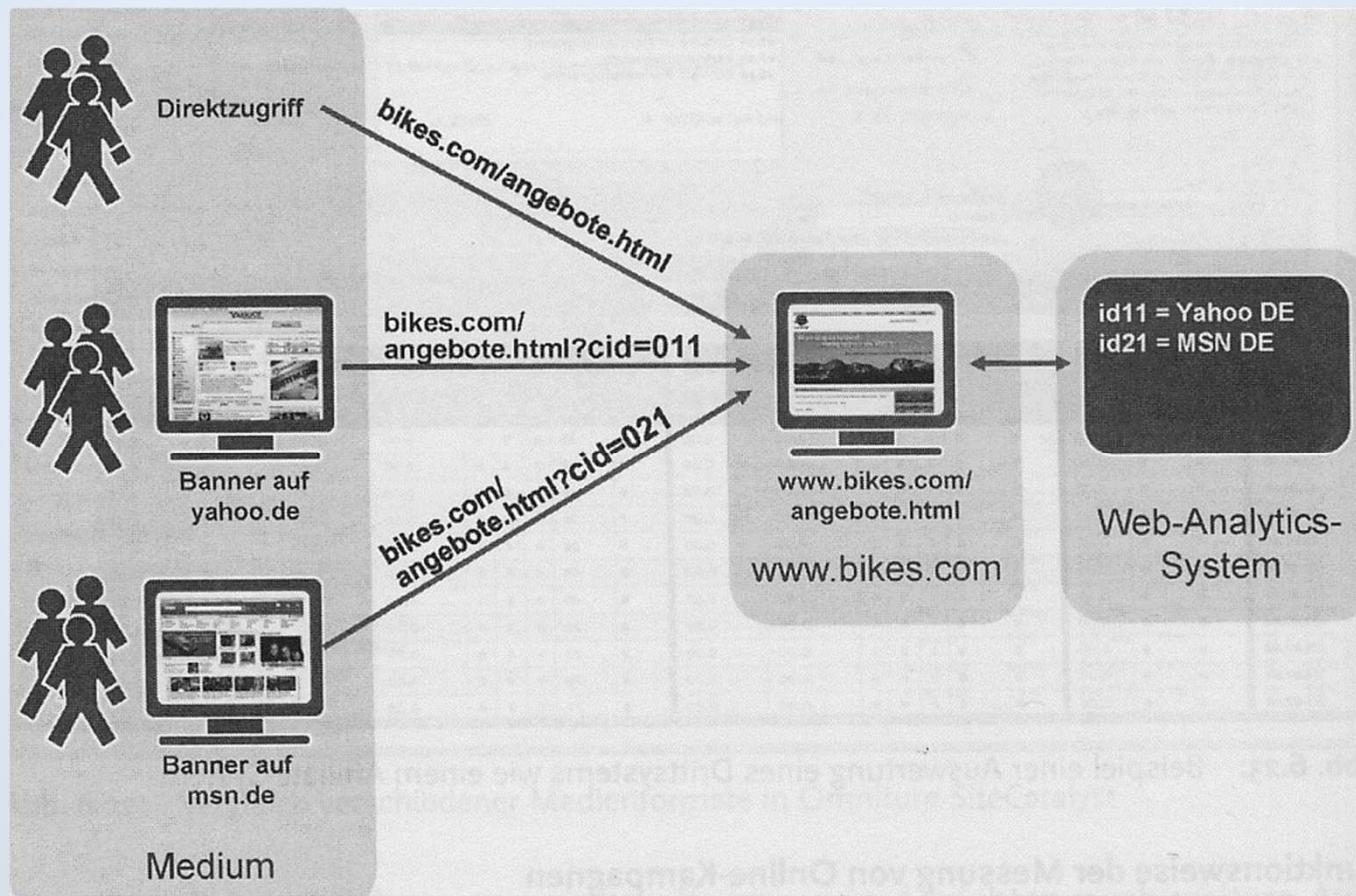
- von anderen Websites
- über Suchmaschinen

Verlinkungsgrad

- link:www.amazon.de
- www.touchgraph.com
- Page Rank (0-10 Relevanz der Seite)
zb: toolbar.google.co

Kampagnen

Mittels Tracking-Code (Hassler Marco, Web Analytics)



Traffic von Social Networks

www.bit.ly

The screenshot shows the bit.ly homepage. At the top, there is a navigation bar with the bit.ly logo, the tagline "Shorten, share, and track your links", and buttons for "Shorten", "Manage", "Sign In or Sign Up Now", and "Tools". Below the navigation bar, there is a main form titled "Shorten with bit.ly". It contains a text input field with the URL "http://student.cosy.sbg.ac.at/~nkowa/" and a "Shorten" button. Below this, it shows the shortened link "http://bit.ly/9QsKix" and a "Sign In to Share Or Customize" button. To the right of the form, there is a section titled "What's bit.ly?" with the text "The easiest way to make long links short, share and track them." and a "Sign Up for Free" button. There are also two small decorative icons: a yellow cartoon character and a blue gear.

Besucher

- Neue und wiederkehrende Besucher
 - Besuchertreue
 - Besucherfrequenz
 - Besucheraktivität
 - Browser, Bildschirmgröße
 - Herkunftsland und Region
 - Sprache
 - Geschlecht und Altersgruppe
-
-

Geschlecht und Alter

- Besucher Daten explizit preisgibt
- Microsoft AdCenters (<http://adlab.microsoft.com/Demographics-Prediction/DPUI.aspx>)

Demographics Prediction

This tool predicts a user's age, gender, and other demographic information, based on their online behavior, such as what queries they search online and what web sites they visit. It lets merchants learn more about their customers before bidding for keywords.

Type a query or a Web address (URL), select Query or URL, and then click Go.

Query examples: real estate, dodge caravan, dodge pickup, xbox, baby names.

URL examples: <http://www.harvard.edu>, <http://zone.msn.com>, <http://www.seomoz.org>

[▶ Learn More](#)

Related Tools

Detecting Online Commercial Intention
Predict customer intention to do a transaction based on their search queries or recently visited URLs

Explore MSN
Analyze publisher web site content and its audience for ad placement and targeting

www.amazon.de GO

Tool Input Query URL Reset Tool

URL: www.amazon.de

Gender: Female-oriented, with the following confidence:

 Male	:0.30
 Female	:0.70

Age: 35~49 Oriented with following distribution:

	General Distribution	Predicted Distribution
35~49	23.00%	37.04%
25~34	27.20%	22.22%
50+	13.20%	18.52%
18~24	26.80%	14.81%
<18	9.80%	7.41%

Besucherverhalten

- Besuchsdauer
- Besuchstiefe
- Navigationsverhalten
- Klickverhalten und Besucherfluss
- Beweggründe und Aufgaben von Besuchern

Inhalt

- Genutzte Inhalte
- Inhaltsgruppen
- Ein- und Ausstiegsseiten
- Attraktivität einer Seite
- Fehler und Ausfälle

Google Analytic

- wurde 2005 eingeführt
- frei bis zu 5 Mio. Zugriffen/Monat
- verwendet First-Party-Cookies
- Page Tags inkl. Datei Downloads, Mail-to Links, partielle Ausfüllhilfen, Flash, Ajax
- Datenspeicherung für 25 Monate

Wesentliche Hauptmerkmale

- Kampagnen und E-Commerce Bericht
- Integration mit AdWords und AdSense
- Goal Conversion
- Funnel Visualization
- Website Overlay
- Datenexport
- Motion Charts
- API
- Intelligenz
- Pivot Ansichten
-

Funktionsweise von Google Analytic

1. Aufruf der Website → Request

<http://www.google-analytics.com/ga.js>

- Einmalig für Sitzung
- 18 KB JavaScript Datei
- Durchführung Datenerhebungen
- Erzeugung „First-Party“-Cookies

First-Party Cookies

Cookie Name	Lebensdauer	Verwendungszweck
__utma	24 Monate	speichert Besucher- und Domain-Kennung
__utmb	Session	speichert Session-Kennung
__utmc	Session	speichert Session-Kennung; erlischt bei 30 Min. Inaktivität
__utmv	24 Monate	speichert Benutzerart
__utmz	6 Monate	speichert Kampagene-Variablen



Funktionsweise von Google Analytic

2. Seitenzugriff Google Data Collection Server über
1x1-Pixel Gif Aufruf (inkl. In-Page Aktionen)
3. Auswertungen nach 3-4 Stunden verfügbar

Was Google Analytics nicht kann

- Daten wiederaufbereiten
- Angebotsmanagement
- Echtzeitberichterstattung
- Importierung von Third-Party Cost Data
- Verfolgung von einzelnen Personen

Urchin

- wurde 2005 von Google gekauft
- läuft auf lokalen Webserver
- verarbeitet Server Logfiles
- Ressourcen muss man selbst bereitstellen
- kostenpflichtig
- ergänzt Google Analytics
 - historische Berichte über Besucher
 - Bandbreiten-Bericht
 - Login Berichte
 - Fehlerseite und Statuscode Berichte

Vorteile von Urchin

- Datenverarbeitung und Wiederaufbereitung
- Unlimitierte Datenspeicherung
- Wirtschaftsprüfung durch Dritte
- Intranet und Firewalls
- Datenbankzugriff

Google Analytics

Danke
für die Aufmerksamkeit!

...