

Leitfaden für die Auswertung der Logfiles

(sG): siehe Glossar am ende des Dokuments

"Als Logfile (dt. Logdatei) wird eine Datei bezeichnet, in der Prozesse, die in einem Computer- oder Netzwerksystem ablaufen, protokolliert werden. Eine andere Bezeichnung für Logfile ist daher auch Protokolldatei. Logfiles liefern wichtige Daten zur Analyse von Netzwerken oder Zugriffen auf einen Webserver bzw. eine Website."

Header.....	1
Unique visitors per day - including spiders.....	2
Requested files.....	2
Operating systems.....	2
Browsers.....	2
Time distribution.....	3
Referring sites.....	3
Glossar.....	3

Header

Der Header fasst viele Informationen kurz zusammen.

Total Requests: Gibt an wie viele Anfragen sind **insgesamt** auf den Server(sG) eingegangen.

Valid Requests: Gibt an wie viele Anfragen **erfolgreich** auf den Server eingegangen sind.

Failed Requests: Gibt an wie viele Anfragen **unerfolgreich** auf den Server eingegangen sind.

Init. Proc. Time: Gibt an wie lange die Loganalyse gedauert hat.

Unique Visitors: Von wie vielen IP-Adressen(sG) wurde zugegriffen.

Unique Files: Gibt an wie viele unterschiedliche URLs(sG) aufgerufen wurden.

Excl. IP Hits: Gibt an wie oft explizit ausgeschlossene(Excluded) IP-Adressen(sG) versucht haben auf den Server zuzugreifen.

Referrers: ?

Unique 404: Gibt an wie viele unterschiedliche den Server einen 404 Fehler Statuscode(sG) als Antwort auf eine Anfrage gesendet haben.

Static Files: Gibt an wie viele Dateien mit Endungen, wie .txt, .JPG, .CSS, .JS, .PNG etc. verwendet wurden.

Log Size: Gibt an wie groß die analysierten Logdateien sind.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den ganzen Server(sG).

Unique visitors per day - including spiders

Hier werden pro Tag Hits, Visitor, Bandwidth aufgelistet. Standard Sortierung nach Datum.

Hits: Gibt an wie viele Zugriffe es an diesem Tag gab.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Achtung: da die Webseite hinter einem Proxy steht, können einzelne Besucher nicht immer voneinander unterschieden werden. Der angegebene Wert ist also ein Minimum-Wert. Der Wert für Hits (einzelne Zugriffe) stimmt immer.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Requested files

Hier werden pro angefragte URL(sG) Hits, Visitor, Bandwidth, Method, Protocol aufgelistet. Standard Sortierung nach Hits.

Hits: Gibt an wie viele Zugriffe gab es an diesem Tag.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Method: Mit welcher HTTP(sG) Methode(sG) auf den Server zugegriffen wurden.

Protocol: Gibt an welche Version des HTTP verwendet wurde.

Operating systems

Hier werden pro Betriebssystem (sG) Hits, Visitors und Bandwith aufgelistet. Standard Sortierung nach Hits.

Hits: Gibt an wie viele Zugriffe es an diesem Tag gab.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Browsers

Hier werden pro Browser (sG) Hits, Visitors und Bandwith aufgelistet. Standard Sortierung nach Hits.

Hits: Gibt an wie viele Zugriffe es an diesem Tag gab.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Time distribution

Hier werden pro Länge des Besuches(nach Stunden) Hits, Visitors und Bandwith aufgelistet. Standard Sortierung nach Hits.

Hits: Gibt an wie viele Zugriffe es an diesem Tag gab.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Referring sites

Hier werden pro URL(sG) die einen Link oder Verweisung auf eine Museumsseite haben Hits, Visitors und Bandwith aufgelistet. Standard Sortierung nach Hits.

Hits: Gibt an wie viele Zugriffe es an diesem Tag gab.

Visitors: Gibt an wie viele unterschiedliche IP-Adressen(sG) an diesem Tag auf den Server zugegriffen haben.

Bandwidth: Gibt an wie viel Datenverkehrsmenge verschoben wurde über den Server.

Glossar

IP-Adresse:

„Eine Internet Protocol Address IP-Adresse (Kurzform für: Internet Protocol Address) ist eine Netzwerkadresse, die für jedes Gerät in einem Netzwerk nur einmal vergeben werden darf. Das ist notwendig, damit Datenpakete richtig adressiert und zugestellt werden können. Eine IP-Adresse ist das eindeutige Identifikationsmerkmal eines Computers, um den Standort im Internet zu definieren.“

URL(Uniform Resource Locator):

Eine URL identifiziert und lokalisiert eine Ressource (Server Seite). Jede Server Seite hat genau eine URL.

(Beispiel: <https://via.museumfuernaturkunde.berlin/wiki/Hauptseite>)

HTTP:

„Bei HTTP(Hypertext Transfer Protocol) werden die Anfragen an die Webseite sowie die Antworten der Seite festgehalten. Es ermöglicht die Datenübertragung zwischen den verschiedenen Netzwerken und sorgt für eine optimale Formatierung des dargestellten Inhalts.“

HTTP-Methode:

Die am meisten verwendete Methode wird bei uns immer GET sein. Mit der GET-Methode können Sie eine Ressource (zum Beispiel eine Datei) vom Server anfordern.

Die zweit meiste Methode wird die POST-Methode sein. Mit der POST-Methode können Sie große Datenmengen (wie Bilder oder HTML-Formular-Daten) zur weiteren Verarbeitung zum Server senden.

Die restlichen werden nicht so häufig verwendet. Aber Erklärung für alle Methoden:
<https://wiki.selfhtml.org/wiki/HTTP/Anfragemethoden>

Statuscodes:

Ein Statuscode gibt an, wie die Anfrage auf eine URL verlaufen ist.
Es gibt 4 wichtige Gruppen:

2xx Erfolgreiche Operation

Beispiel: 202 Die Anfrage wurde akzeptiert, wird aber zu einem späteren Zeitpunkt ausgeführt.

3xx Umleitung

Beispiel: 301 Die angeforderte Ressource steht ab sofort bereit (auch Redirect genannt). Die alte Adresse ist nicht länger gültig

4xx Client Fehler (Der Fehler liegt am Nutzer)

Beispiel: 404 Die angeforderte Ressource wurde nicht gefunden. Dieser Statuscode kann ebenfalls verwendet werden, um eine Anfrage ohne näheren Grund abzuweisen.

5xx Server Fehler (Nicht klar von den so genannten Client-Fehlern abzugrenzen. Die Ursache des Scheiterns der Anfrage liegt jedoch eher im Verantwortungsbereich des Servers)

Beispiel: 501 Die Funktionalität, um die Anfrage zu bearbeiten, wird von diesem Server nicht bereitgestellt. Ursache ist zum Beispiel eine unbekannte oder nicht unterstützte HTTP-Methode.

Betriebssystem:

„Ein Betriebssystem (englisch: Operating System) ist eine Software, die den Betrieb eines Computers steuert. Es ist im Prinzip eine Vermittlungsinstanz zwischen der vorhandenen Hardware, Software und dem User.“

Beispiele sind Windows, IOS, Linux

Browser:

Ein Browser ist ein Programm welches Webseiten korrekt anzeigen lassen kann.

Beispiele sind Internet Explorer, Mozilla Firefox, Google Chrome etc.

Server:

„Ein Server ist ein leistungsstarker, zentraler Netzwerkrechner. Er stellt seine Ressourcen für andere Computer bereit. Über den Server werden funktionale und infrastrukturelle Netzdienste realisiert. Hier können Daten, auf die mehrere Personen zugreifen müssen, gespeichert werden.“

Das betrifft alle Seiten die in den Logs festgehalten werden.