Web Analytics Tools

Stand: 21.06.2019

[1 Ziele des Dokuments 2](#_Toc9599959)

[2 Überblick 2](#_Toc9599960)

[3 Selbst-entwickeltes Web-Analytics Tool 3](#_Toc9599961)

[3.1 Vorteile 3](#_Toc9599962)

[3.2 Nachteile 3](#_Toc9599963)

[4 Piwik / Matomo 3](#_Toc9599964)

[4.1 Vorteile 3](#_Toc9599965)

[4.2 Nachteile 3](#_Toc9599966)

[5 Google Analytics (GA) 4](#_Toc9599967)

[5.1 Vorteile 4](#_Toc9599968)

[5.2 Nachteile 4](#_Toc9599969)

[6 Splunk 5](#_Toc9599970)

[6.1 Vorteile 5](#_Toc9599971)

[6.2 Nachteile 5](#_Toc9599972)

[7 Zusammenfassung / Vergleich 6](#_Toc9599973)

# Ziele des Dokuments

In diesem Dokument wird beschrieben, welche Möglichkeiten es gibt, Benutzeraktionen in unseren Portalen / Anwendungen zu analysieren, d.h. zu erkennen, welche Nutzer welche Funktionen wie nutzen.

# Überblick

Folgende Fragestellungen sollten von einem Web-Analytics Tool beantwortet werden:

1. Wie leicht / aufwendig kann man das integrieren?
2. Welche Infrastruktur braucht man dafür?
3. Preise.
4. Wie hilft einem das Produkt bei der Aufbereitung der Aufzeichnungen?

Die bekanntesten Möglichkeiten / Web-Analytics-Tools[[1]](#footnote-1):

1. Selbst-entwickeltes Tool  
   (basierend auf tool-spezifischen Aufrufen, Netzwerk-Protokoll-Analyse und/oder Server-Log-Analysen)
2. Piwik / Matomo (Open Source)
3. Google Analytics (gratis)
4. Splunk (kostenpflichtig mit eingeschränkter freien Version)

Neben diesen bekannten Tools gibt es kleinere Anbieter (meist kostenpflichtig), die auf Nischen-Anforderungen spezialisiert sind (z.B. Mindlab mit Individuallösungen; Odoscope für die Visualisierung von Benutzerpfaden; Webtrekk mit Marketing-Analysen und Echtzeit-Auswertungen).

Auf rein **serverseitiger Log-Analyse** fokussierten Tools sind weniger datenschutzrelevant, aber teilweise weniger aussagekräftig (z.B. ELK[[2]](#footnote-2) mit Kibana-Dashboards).

Die folgenden Seiten geben eine Kurzbeschreibung dieser Tools und bieten eine Zusammenfassung der Vor- und Nachteile.

# Selbst-entwickeltes Web-Analytics Tool

Es besteht die Möglichkeit, die zur Analyse benötigten Daten selbst zu sammeln, auswerten und graphisch (bzw. als Reports) aufbereiten.

## Vorteile

Dies ist die einzige Möglichkeit, die es ermöglicht, alle server- **und** client-seitige Informationen, und evtl. Log-Datei-Analyse zusammenzuführen.  
Auch nicht-traditionelle, (sondern auf Ajax-Requests und client-seitiges Routing basierende) Anwendungen können einfach analysiert werden.

## Nachteile

Entwicklungsaufwand ist höher.

Bestimmte Funktionen, wie Kampagne oder Customer-Journey‘s, sind schwer abbildbar.

Nach einem eventuellen Umzug in eine Public Cloud Umgebung werden manche Funktionen evtl. so nicht mehr nutzbar sein.

# Piwik / Matomo

Das **PHP**-basierte Web-Analytics Tool Matomo (ehemals Piwik) ist in der Kategorie Open Source am weitesten verbreitet.

Die Installation ist einfach, vorausgesetzt, eine MySQL/MariaDB Datenbank und ein Apache/Nginx Webserver sind vorhanden.

Unsere Web-Anwendung kann JavaScript-Code einbinden, der das Nutzerverhalten automatisch auswertet und an den Server weiterleitet. Auch der manuelle Aufruf der JavaScript-Analysefunktionen ist möglich (z.B. beim Absenden eines Formulars oder nach client-seitigem Routing).

In Deutschland ist Matomo nach Google Analytics das meistverwendete Webanalytik-Werkzeug.

## Vorteile

Die Installation und Pflege sind einfach. Die Dashboards / Auswertungen und Reports sind informativ.

Auch fortschrittliche Funktionen wie Kampagne, Referreranalyse und Ecommerce-Auswertungen sind möglich.

Clients für Android und iOS sind verfügbar.

Die Software ist gratis.

## Nachteile

Pflege der fachspezifischen erweiterten Funktionen (z.B. Konfiguration der Ecommerce-Einstellungen) erfordert eine Einarbeitung.

Um serverseitige Analyse-Informationen bereitzustellen, müssen die spezifischen API-Aufrufe implementiert werden.

Eigene serverseitige Infrastruktur (Webserver, Datenbank) muss aufgebaut werden.

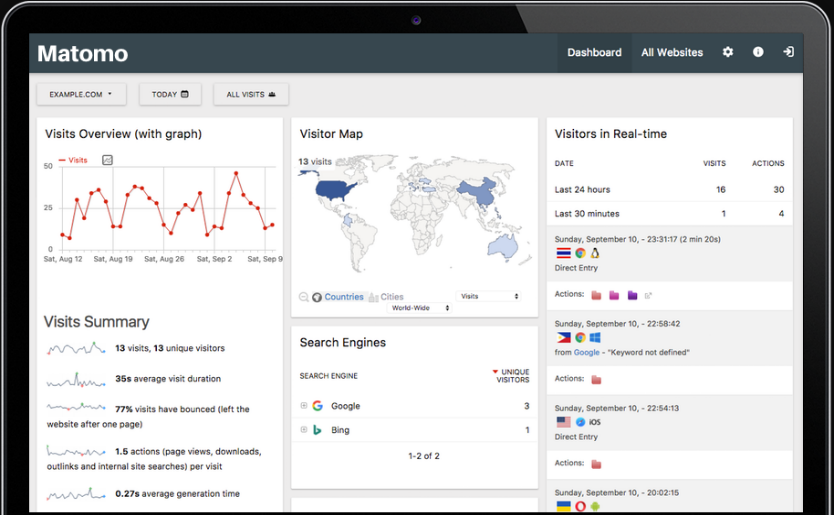


Abbildung: Matomo Dashboard

# Google Analytics (GA)

GA ist das weltweit gesehen meist verwendete Web-Analytics Tool, das sogar geräte-übergreifende Analyse ermöglicht.

Die Verknüpfung mit weiteren Google-Diensten (AdWords, AdSense, Google-Webmaster-Tools) ist möglich; die automatisch durchgeführte google-interne Analyse ermittelt demographische Daten.

Die Anwendung kann JavaScript-Code einbinden, die das Nutzerverhalten automatisch auswertet und (für unsere Anwendung transparent) an die Google-Server weiterleitet.

Der Einsatz von GA in Deutschland ist aus mehreren datenschutz-relevanten Gründen umstritten (nicht-transparente Datennutzung durch Google auf US-amerikanischen Servern inkl. Verknüpfung mit weiteren Google-Informationen und KI-Analyse).

## Vorteile

Keine serverseitige Infrastruktur nötig.

Die Nutzung ist frei, auch für kommerzielle Zwecke (wenn man auf AdWords u. ä. verzichtet).

Die durch Google vorgenommene Auswertungen / Reports sind professionell (aber nur begrenzt anpassbar).

Durch die Nutzung der JavaScript-APIs sind Aktionen benutzerspezifisch aufrufbar (im Falle von client-seitigem Routing).

Bei Nutzung weiterer Google-Dienste können die Informationen aller Google-Tools zusammengeführt werden.

## Nachteile

Die Nutzung ist datenschutztechnisch umstritten.

Serverseitige Informationen oder Log-Auswertung sind nicht out-of-the-box möglich, können aber mittels SDK selbst entwickelt werden[[3]](#footnote-3).

Google Analytics ist auf maximal zehn Millionen Aktionen pro Monat **limitiert**.

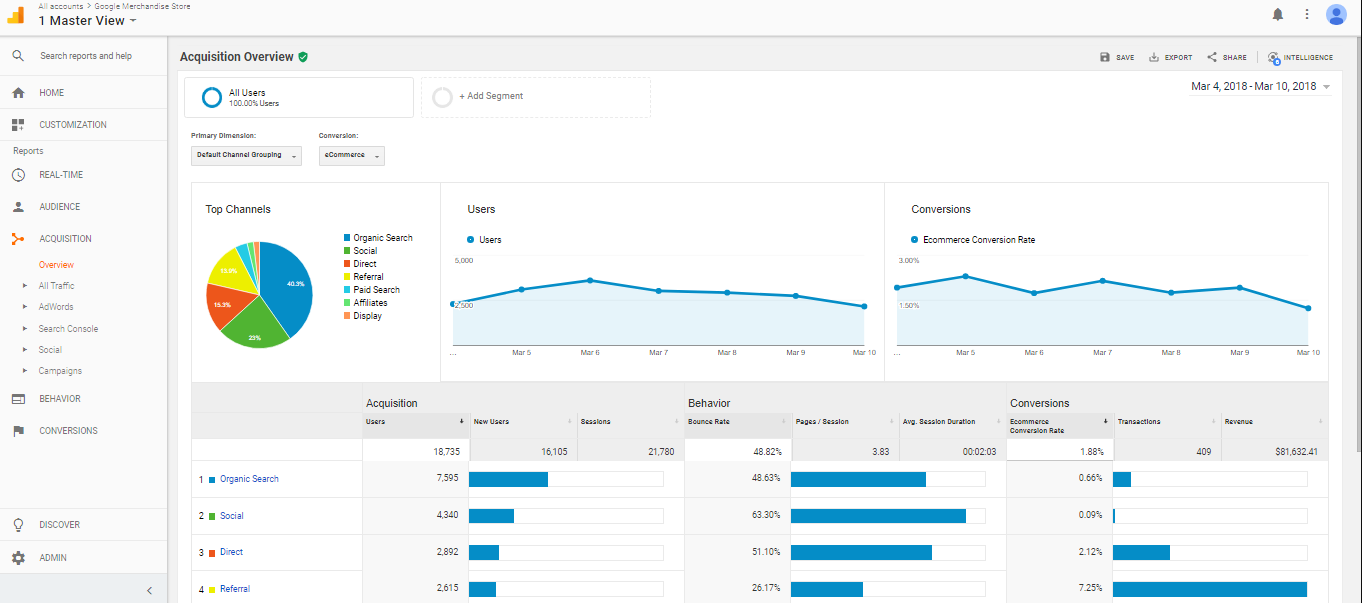


Abbildung: Google Analytics Dashboard

# Splunk

Splunk ermöglicht die Auswertung der Benutzerinteraktion basierend auf Log-Auswertung (es sind keine clientseitigen Änderungen notwendig) und deren intelligenter Konfiguration. Somit sind keine Änderungen an den einzelnen, zu analysierenden Applikationen nötig, außerdem tauchen viel weniger datenschutzrelevanten Bedenken auf.

Die Basisversion des Analyse-Werkzeugs ist frei, die Preise der Enterprise-Version sind nutzungsabhängig[[4]](#footnote-4).

Die Auswertung erfolgt auch mit Big Data- und Machine Learning Techniken und ist insbesondere geeignet für große Datenmengen und/oder für Apps in der Public Cloud (AWS).

Splunk bietet ein ganzes Ökosystem an Anwendungen, deren Kombination das gesamte Anwendungsmanagement (von Server-Monitoring, durch Cloud-Management bis User Behavior Analytics) abdeckt.

## Vorteile

Datenschutzbedenken gibt es wenig, da clientseitige Benutzeranalyse nicht erfolgt.

Die Auswertungen und deren Darstellung (Dashboards, Reports) sind fortschrittlich und vorbereitet auf große Datenmengen.

Durch die Kombination der Splunk-Tools (Enterprise, Cloud, Security, Service Intelligence, User Behavior Analytics) können Synergie-Effekte erzielt werden.

Aufwände durch eigene Entwicklungen können durch oft freie Zusatz-Tools und Addons minimiert werden[[5]](#footnote-5).

Bietet Zusatzfunktionen wie Alerting.

## Nachteile

Ist kostenpflichtig (wenn nicht die Basisversion benutzt wird).

Macht für größere Datenmengen und für Public-Cloud-Anwendungen mehr Sinn.

KI-gestützte Auswertungen bringen nicht immer genug aussagekräftige neue Informationen bzw. rechtfertigen die Kosten nicht immer.

Nutzung ist nicht einfach für Nicht-Techniker.

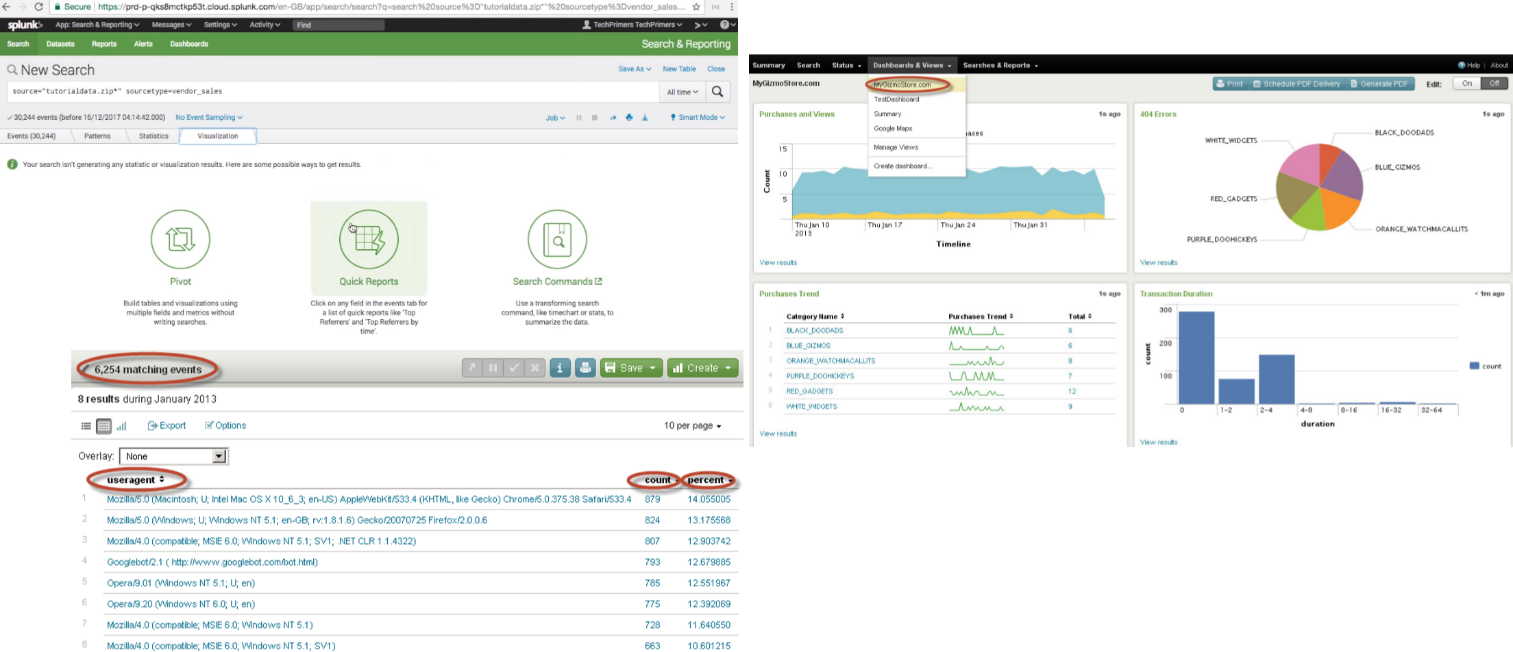


Abbildung: Splunk Query Analyse

# Zusammenfassung / Vergleich

Vergleich der hier vorgestellten Web-Analyse-Tools:



Vergleich der verschiedenen Web-Analytics-Tools

1. Liste bekannter Web-Analytics-Tools: <https://de.wikipedia.org/wiki/Liste_von_Web-Analyse-Software> [↑](#footnote-ref-1)
2. <https://www.elastic.co/de/elk-stack> [↑](#footnote-ref-2)
3. <https://developers.google.com/analytics/> , <https://github.com/googleanalytics> [↑](#footnote-ref-3)
4. <https://www.splunk.com/en_us/software/pricing/faqs.html#Splunk-Enterprise> [↑](#footnote-ref-4)
5. <https://splunkbase.splunk.com/apps/#/search/angular/> [↑](#footnote-ref-5)