

# Nutzungsstatistiken in Webarchiven Einblicke in die AWstats-Auswertung in SWBcontent

nestor-Expertengespräch „Bereitstellung und Nutzung  
archivierter Webseiten“

Stuttgart, 27.2.2013  
Stefan Wolf, BSZ

---

# Überblick

- Installation, Funktion, Logerstellung
- Welche Zahlen werden ermittelt?
- Aussagefähigkeit

## Was ist das – AWstats?

- eine freie Webanalyse-Software
- zur Auswertung von Logdateien, die Webserver auf Basis von Besucheranfragen erstellen.
- erzeugt Berichte als HTML-Seiten, die mit einem Browser betrachtet werden können und auf diese Weise beispielsweise direkt von der analysierten Webseite verlinkt werden können.

(vgl. <http://de.wikipedia.org/wiki/AWStats>)

# Eine Installation von SWBcontent

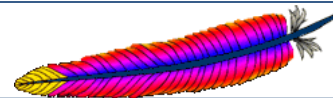
Software SWBcontent  
mit Datenbanken und  
Schnittstellen



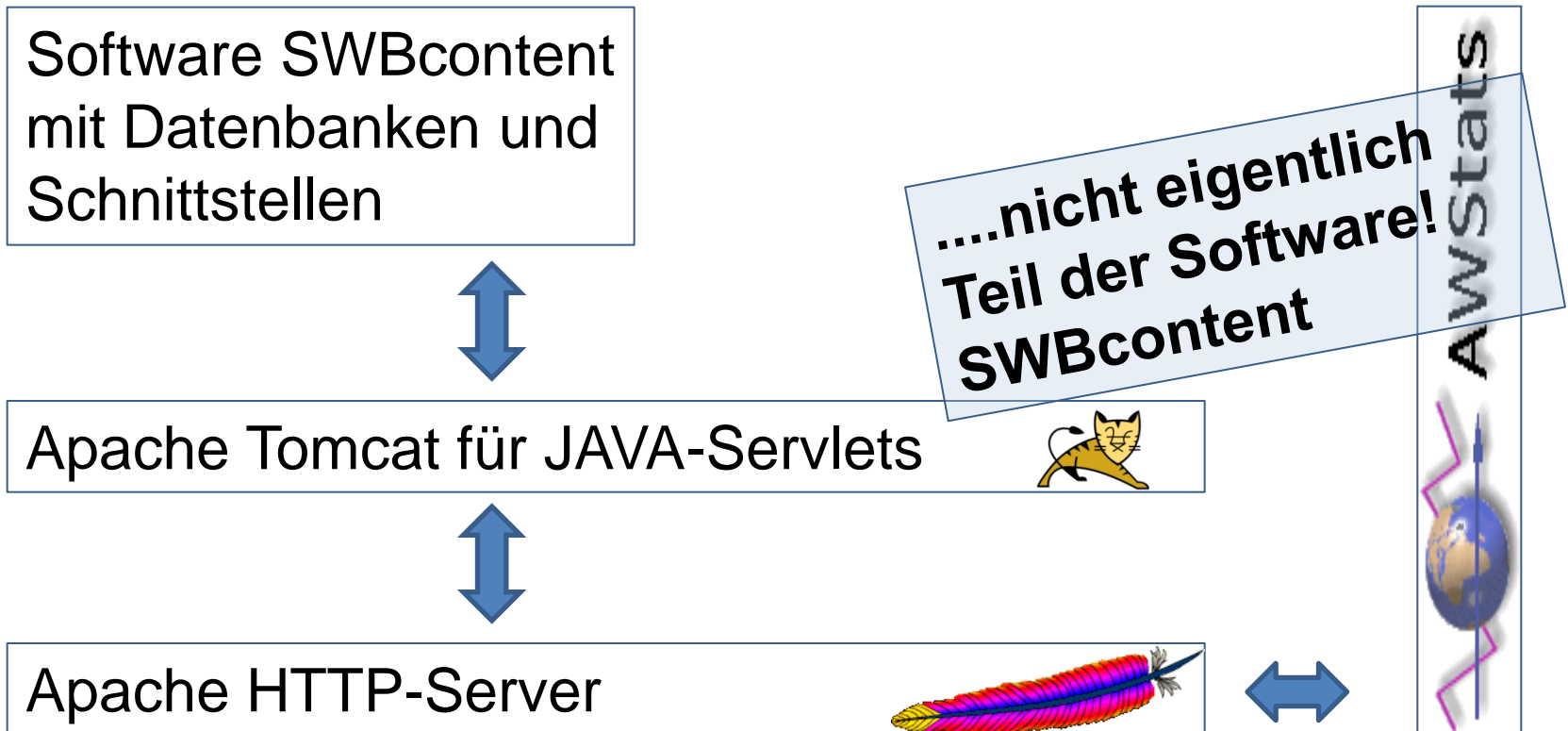
Apache Tomcat für JAVA-Servlets



Apache HTTP-Server



# Eine Installation von SWBcontent



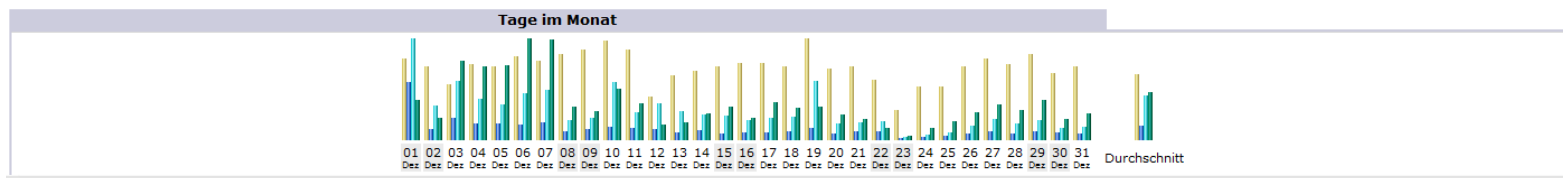
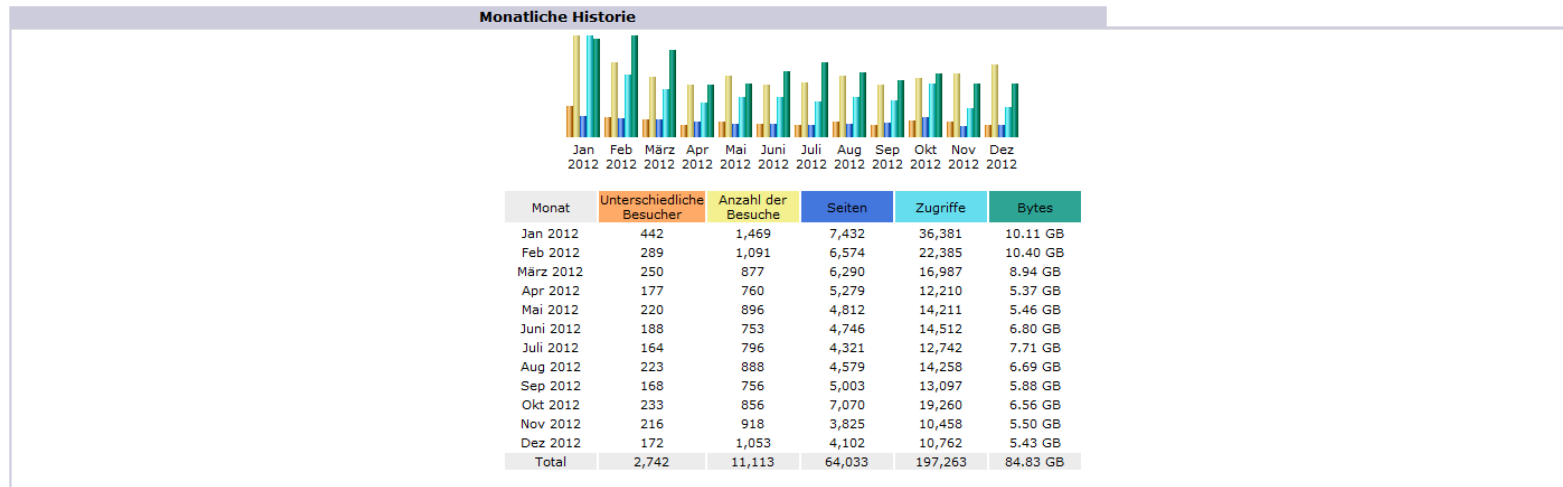
- In allen Installationen seit ca 2 Jahren mitlaufend
- Über URL, Passwort und Kennung für die Betreiber (Kunden) zugänglich
  - Für Betreiber: Auskunft über Nutzung der Installation
  - Für BSZ: Teil der Nutzungsmessung für die Webserver

# Statistik am Beispiel:

Zuletzt aktualisiert: 04.03.2013 - 14:15  
 Zeitraum: - Jahr - 2012 OK 

Zusammenfassung					
Zeitraum	Jahr 2012				
Erster Zugriff	01.01.2012 - 00:12				
Letzter Zugriff	31.12.2012 - 23:47				
	Unterschiedliche Besucher	Anzahl der Besuche	Seiten	Zugriffe	Bytes
gesehener Traffic *	<b>&lt;= 2,742</b> Der genaue Wert ist nicht in der Jahres-Ansicht verfügbar	<b>11,113</b> (4.05 Besuche/Besucher)	<b>64,033</b> (5.76 Seiten/Besuch)	<b>197,263</b> (17.75 Zugriffe/Besuch)	<b>84.83 GB</b> (8004.5 KB/Besuch)
nicht gesehener Traffic *			<b>29,518</b>	<b>49,220</b>	<b>276.89 MB</b>

\* Nicht gesehener Traffic ist Traffic, welcher von Robots, Würmern oder Antworten mit speziellem HTTP-Statuscode



# Leistungsumfang

- Besucher: nach Tag/Monat/Jahr  
nach Wochentagen/Uhrzeit
- Herkunft der Anfragen nach Ländern
- Suchmaschinenanfragen werden ausgeblendet
- Aufenthaltsdauer der Besucher
- Abgerufene Dateitypen
- Top 10 der Downloads, aufgerufenen Seiten, anfragenden Browser
- Woher die Besucher kamen
- Fehlercodes



# Zur Tragweite der Statistik

- Es gibt Nutzung!
  - mehr als wir selbst erwartet haben
- Wir haben
  - eine bedingte Auskunft über die Herkunft der Nutzung
  - eine bedingte Auskunft über das leitende Interesse
- Gegenfrage:
  - ist das bei „normaler“ Bibliotheksnutzung anders?
- Hinweis:
  - Die Statistik dient nicht z.B. der Relevanzmessung von Publikationen/archivierten Seiten

## Ergänzung: Mengenstatistik

- Für alle Installationen von SWBcontent wird auch eine Mengenstatistik erstellt, die Auskunft gibt über
  - Zahl der Titelaufnahmen
  - Zahl der „Objekte untergeordneter Hierarchie“ (Zeitschriftenteile, Zeitschnitte)
  - Belegtes Datenvolumen
- Struktur und Frequenz dieser Statistik wurde mit den Betreibern abgesprochen
- Veröffentlichung liegt in ihren Händen

Stefan Wolf  
stefan.wolf@bsz-bw.de  
07531/883050

Danke für die Aufmerksamkeit