

Seminarthemen

1 **BigData - Verarbeitung von verteilten Massendaten**

- Was ist bei der Verarbeitung von Massendaten zu beachten?
- Worin unterscheiden sich BigData-Datenbanken zum relationalen Datenbank-Konzept?
- Hadoop und HBase im Vergleich zur Integration von Hadoop in Cassandra
- Horizontale Skalierung durch das MapReduce-Framework von Hadoop
- Herausforderungen der Echtzeit-Verarbeitung (inkrementell) gegenüber Analysen über alle Daten

2 **NoSQL - Nicht-relationale Datenbanken**

NoSQL ist ein Oberbegriff für alle Datenbanken, die nicht relational sind. Dabei gibt es jedoch sehr unterschiedliche Konzepte, die sich dahinter verbergen. Darunter finden sich wiederum Gruppen, wie Dynamo, BigTable, Key-Value, Dokument- und Graphen-Datenbanken. Worin unterscheiden sich diese Konzepte und welche Probleme kann man mit ihnen jeweils lösen? Außerdem gibt es auch für einige NoSQL-Datenbanken durchaus die Möglichkeit SQL-Queries darauf anzusetzen.

3 **Cookieloses Web-Controlling**

Die Wiedererkennung von Besuchern einer Webseite ist ein essentieller Bestandteil für das Web-Controlling. Anhand der Information ob ein Besucher eine Webseite bereits besucht hat, lassen sich beispielsweise besuchsübergreifend Werbekampagnenkontakte erfassen und optimieren.

Die Erkennung wird momentan meist unter Zuhilfenahme von Browser-Cookies gelöst. Diese Nutzung könnte jedoch in Zukunft erschwert werden, da eine EU-Richtlinie vorsieht, dass Daten auf den Rechnern von Websitebesuchern nur nach deren Zustimmung abgelegt werden dürfen. Es gibt verschiedene alternative Möglichkeiten Besucher ohne Cookies wieder zu erkennen. Eine Möglichkeit ist die Ermittlung von Browser-Fingerprints. Da jeder Browser verschiedenen Informationen an einen Webserver übermittelt, lassen sich anhand der Kombination dieser Daten relativ eindeutige Besucher ermitteln. Ein anschauliches Beispiel um die Einzigartigkeit eines Browsers zu ermitteln bietet die Seite <https://panopticlick.eff.org/>

Ziel der Seminararbeit ist es, die Möglichkeiten der cookielosen Zählung zu analysieren

4 **Erkenntnisse aus dem Mouse-Tracking**

Modernes Web-Controlling bietet viele verschiedene Möglichkeiten zur Optimierung von Webseiten. Eine Variante ist das Mouse-Tracking. Es zeichnet unter anderem die Bewegung der Maus, das Scrollverhalten und den sichtbaren Bereich im Browser auf. Der Webseitenbetreiber hat so die Möglichkeit seinem Besucher quasi über die Schulter zu sehen und Probleme bei der Bedienung zu erkennen.

Die aufgezeichneten Sessions lassen zwar einzelne Probleme erkennen, jedoch liegt das größte Potential in der Aggregation dieser Daten um weitere Informationen zur Optimierung einer Webpräsenz zu erhalten. Da niemand die Zeit hat, tausende von Besucheraufzeichnungen zu analysieren, stellt sich die Frage: Wie kann aus der Masse der aufgenommenen Daten Erkenntnis abgeleitet werden?

5 **Verteilte Versionsverwaltung mit Git und Mercurial**

CVS und Subversion war gestern. Mittlerweile arbeiten moderne Software-Unternehmen mit Git oder Mercurial. Warum ist das so und was sind die Vorteile? Inwieweit spielt das Konfigurationsmanagement-Werkzeug eine zentrale Rolle im Software-Entwicklungsprozess?

6 **Wechselwirkungen von Werbekampagnen (CampaignControl)**

Wie beeinflussen sich unterschiedliche Werbekampagnen z.B. Branding- & Sale-Kampagnen. Entwickle ein Modell zur Berechnung des Lift-Offs.

7 Besucher-Clustering

Mittlerweile ist man in der Lage Massendaten zu erfassen und Bewegungen riesiger Besucherströme über unterschiedliche Websites zu messen. Wie können aus diesen Daten nun Erkenntnisse generiert werden? Ein Ansatz ist es, anhand bestimmter Merkmale Cluster von ähnlichen Besuchern zu generieren. Wie sehen die Algorithmen aus? Wie können aus den Daten, die beispielsweise mit etracker generiert werden, solche Cluster generiert werden.

8 Real-Time Bidding

RTB ist derzeit in der Online Marketing Branche eines der „heißesten“ Buzzwords. Doch was steckt dahinter und welche Rolle können klassische Webanalytics Anbieter wie etracker dabei bieten?

9 Die Zukunft des Webcontrolling

Was kommt nach den PI (Page Impressions) als nächste Kennzahl im Web-Controlling? Welche Kennzahlen sind im Web-Controlling in den nächsten Jahren die entscheidenden? Dieses Seminar-Thema ist im Schwerpunkt fachlicher Natur. D.h. es gilt die unterschiedlichen Zielgruppen für Web-Controlling zu analysieren und zu erarbeiten, welches die relevanten Kenngrößen sind.