|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| Social Brand Monitor | | | |
| Mit SAP HANA | | | |
|  | | | |
| Ihr Ansprechpartner: | | | |
| Christian Günther | +49 1511 |  | 2240942 |  |
|  | | | |
|  | | | |

|  |
| --- |
|  |

Inhaltsverzeichnis

1 Management Summary 4

1.1 Projektziele 4

1.2 Use Case – Social Brand Monitor 4

1.3 Use Case – Competition Monitor 4

2 Mockup 5

2.1 Usability Konzept 5

2.2 Visualisierung 6

3 Lösungsarchitektur 10

3.1 High Level Lösungs-Architektur und Prozessablauf 10

3.1.1 Prozessablauf 10

3.2 Komponenten der Architektur 11

3.3 Visualisierung 12

3.3.1 Fragen 12

3.4 Input Stream Verarbeitung 12

3.4.1 Fragen 12

3.5 Datenmodell 12

3.5.1 Fragen 12

3.6 Semantische Erkennung 12

3.6.1 Fragen 13

3.7 Analytics (Filterkonzept) 13

3.7.1 Fragen 13

3.8 Konfiguration 13

3.8.1 Fragen 14

3.9 Interfacing 14

3.9.1 Fragen 14

4 Umsetzung mit SAP HANA 15

4.1 Visualisierung 16

4.2 Input Stream – Abfrage Social Network 16

4.3 Datenmodell 16

4.4 Semantische Erkennung 16

4.5 Analytics oder Filterkonzept 16

4.6 Konfiguration 18

4.7 Interfacing 18

5 Referenzen 19

Abbildungsverzeichnis

Abbildung 1 – Struktur des UI 5

Abbildung 2 – Kartenansicht 6

Abbildung 3 – Detailansicht – Grafisch 7

Abbildung 4 – Detailansicht – Tabellarisch 8

Abbildung 5 – Detailansicht – Posts 9

Abbildung 6 – High Level Solution Architecture 10

Abbildung 7 – Bestandteile der Lösung 11

Abbildung 8 – Text Analysis und Analytics 13

Abbildung 9 – Systemarchitektur mit SAP HANA 15

Abbildung 10 – SAP HANA Elemente 16

Abbildung 12 – Graphendarstellung von Text 17

Abbildung 12 – Architektur Graphen-Engine in HANA 17

# Management Summary

## Projektziele

Das Ziel des Gesamtprojektes ist es schnell und einfach Daten aus Social Media Plattformen auf dem iPad zu visualisieren.

Es sollen zwei Anwendungsfälle auf dem iPad bereitgestellt werden:

1. **Social Brand Monitor**

Der Social Brand Monitor zeigt die Stimmung von Internetnutzern auf Facebook, Twitter und Co. gegenüber Firmen und ihren Produkten, indem die Tweets und Posts auf ihre jeweilige Meinung (positiv oder negativ) hin untersucht und nach verschiedenen Kriterien ausgewertet werden.

1. **Competition Monitor**

Der Competition Monitor vergleicht die Anzahl an Likes auf Facebook eines Unternehmens mit denen ihrer Mitbewerber und stellt diese als Graphen dar.

## Use Case – Social Brand Monitor

Der Nutzer hat eine 3D Kartenansicht auf. In dieser wird über Farben die Meinung (positiv = grün oder negativ = rot) in den einzelnen Regionen dargestellt. Zu dieser ersten Dimension kommt nun die 2. Die Menge der Blasen stellt die Menge an Äußerungen in Relation zu anderen Regionen dar. Über diese Sicht kann der Benutzer die Stimmung in einzelnen Regionen einsehen und bis auf einzelne Tweets oder Facebook-Posts navigieren.

## Use Case – Competition Monitor

Der Competition Monitor dient dem Vergleich der Stimmung (gemessen anhand Likes) auf Facebook des Unternehmens und direkten Konkurrenten. Hierfür werden Graphen herangezogen deren Datengrundlage sowohl global als auch von Nutzer einschränkbar auf Regionen sein sollen.

# Mockup

Ausgehend von den Anforderungen und dem Use Case, werden in diesem Kapitel das Usability-Konzept und der Mockup einer möglichen Oberfläche dargestellt.

## Usability Konzept

Bei der Erarbeitung eines Bedienkonzeptes wurde davon ausgegangen, dass es im Wesentlichen zwei Handlungsebenen innerhalb der Benutzung der App gibt. Zum einen ist dies die Definition des lokalen, inhaltlichen und zeitlichen Rahmens, in welchem sich die Informationen bewegen sollen. Dies betrifft die Auswahl eines geografischen Gebietes, gegebenenfalls die Eingabe von Keywords, die die Suche einschränken sollen und die Einschränkung auf einen bestimmten Zeitraum.

Die andere und Haupt-Handlungsebene ist die Rezeption der entsprechenden Ergebnisse.

Während es gilt, für die Auswahl des zeitlichen, inhaltlichen und räumlichen Rahmens, eine möglichst effiziente und einfache Form zu finden, sollten für die Handlungsebene der Rezeption Möglichkeiten geschaffen werden, die entsprechenden Inhalte in unterschiedlicher Form und entsprechend der jeweiligen Intention gefiltert darzustellen.

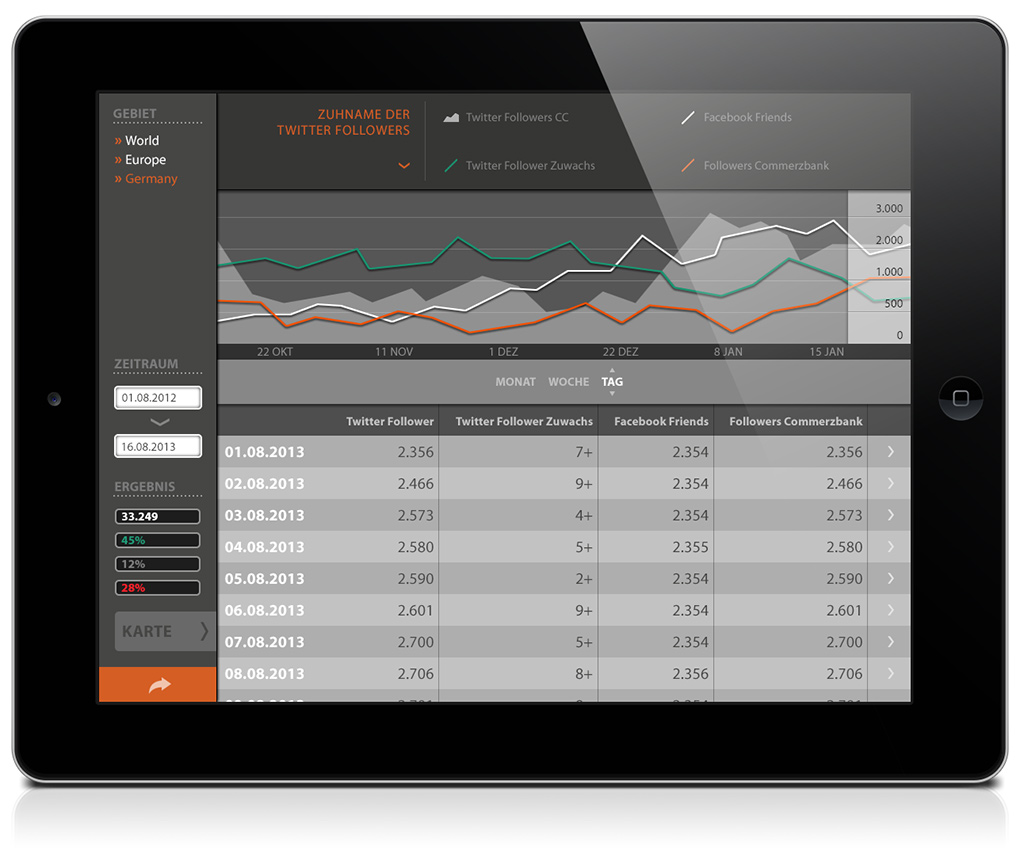


Abbildung – Struktur des UI

Der Screen ist in zwei Bereiche geteilt: auf der linken Seite sind globale Informationen und Einstellungsmöglichkeiten untergebracht:

* Wechsel zwischen Karten- und Detail-Ansicht
* Derzeitige lokale Eingrenzung als Text
* Eingabemöglichkeit von Keywords
* Dargestellter Zeitraum mit Einstellungsmöglichkeit
* Schlüsselkennzahlen für den eingestellten Bereich
* Exportmöglichkeit der Daten des eingestellten Bereiches

## Visualisierung

Die größere, rechte Seite dient der Darstellung der eigentlichen Informationen. Nach dem Start wird hier eine Karte angezeigt, in der geografisch navigiert werden kann. Je nach Darstellung werden Marker auf der Karte abgebildet, die eine bestimmte Kennzahl anhand ihrer Größe für ein bestimmtes Gebiet visualisieren (z.B. Gesamtanzahl der Twitter- und Facebook-Posts innerhalb des eingegrenzten Zeitrahmens). Das Gebiet mit der größten Anzahl wird dabei mit 25 Punkten dargestellt und repräsentiert 100%. Die Darstellung der anderen Gebiete erfolgt prozentual zu diesem. Die differenzierte Darstellung innerhalb eines Markers durch unterschiedliche Farben (Grün, Rot, Grau) wird der prozentuale Anteil an positiven, negativen und neutralen Posts visualisiert. Durch diese Darstellung erhält man einen schnellen Überblick, welches Gebiet das potentiell stärkste ist und in welchem Verhältnis die anderen Gebiete zu diesem stehen.



Abbildung – Kartenansicht

Mit einem Tap auf einen Marker wird an dieses Gebiet heran gezoomt: Es erscheinen jetzt Unterbereiche für dieses Gebiet.

Wurde ein Gebiet ausgewählt, kann in die Detail-Ansicht in die gewechselt werden, indem man oben links auf das orangene Switch-Symbol tippt Vortex:Users:chris:Desktop:Bildschirmfoto 2013-12-13 um 10.20.51.png – der rechte Bereich des Screens wird ausgetauscht und zeigt hier detaillierte Informationen an. Diese Darstellung kann nun auf unterschiedliche Weise angepasst werden:

**Anpassung der Werte-Details**

Unten rechts besteht die Möglichkeit zwischen den Darstellungen „TAG“, „WOCHE“ und „MONAT“ zu wechseln. Ein Wechsel in diesem Bereich bewirkt eine jeweilige Analyse der Daten bezogen auf einen Tag, eine Woche oder einen Monat. Wird z.B. bei der Darstellung in der Monatsansicht ein signifikanter Anstieg eines Wertes festgestellt, erhält man durch den Wechsel zur Wochen- oder Tagesansicht eine genauere Information darüber, wann dieser Anstieg stattgefunden hat.

**Darstellung als Grafik oder als Tabelle**

Direkt über der Einstellung der Werte-Details besteht die Möglichkeit zwischen einer grafischen und einer tabellarischen Darstellung. Ein Tap auf das entsprechende Icon, wechselt die Ansicht, wobei alle anderen Einstellungen erhalten bleiben.



Abbildung – Detailansicht – Grafisch

**Definieren der dargestellten Werte**

Ganz oben im Anzeigebereich wird eingestellt, welche Werte überhaupt angezeigt werden sollen. Dieses Konzept sieht vor, dass gleichzeitig maximal 4 Werte dargestellt und untereinander verglichen werden können. Die farbliche Kodierung dieser Werte spiegelt sich in den grafischen Darstellungen wider. Mit einem Tap auf einen Wert wird eine Liste aller verfügbaren Werte angezeigt, aus der beliebig ausgewählt werden kann. Es ist natürlich auch möglich, eine Auswahl leer zu lassen, wenn beispielsweise nur zwei Werte verglichen werden sollen.

Zusätzlich zu dieser Auswahl kann für einen Wert auch noch ein anderer Zeitraum als der definierte gewählt werden – z.B. der aktuelle Zeitraum minus eine Woche. Damit können gleiche Werte über einen unterschiedlichen Zeitraum innerhalb einer Grafik bzw. einer Tabelle verglichen werden (Wochen-, Monats- oder Jahresvergleich).

Jede Kombination von angezeigten Werten entspricht einem „Werte-Set“. Eine bestimmte, vordefinierte Anzahl dieser Sets liefert das System mit. Der Name des jeweils angezeigten Sets wird rechts oben angezeigt. Ein Tap darauf öffnet eine Liste aller bereits definierten Sets aus der eine beliebige Kombination ausgewählt werden kann. Zusätzlich besteht die Möglichkeit eigene Kombinationen zu speichern. Damit wird ein flexibles System zur Verfügung gestellt, welches die unterschiedlichsten Anforderungen erfüllen kann.

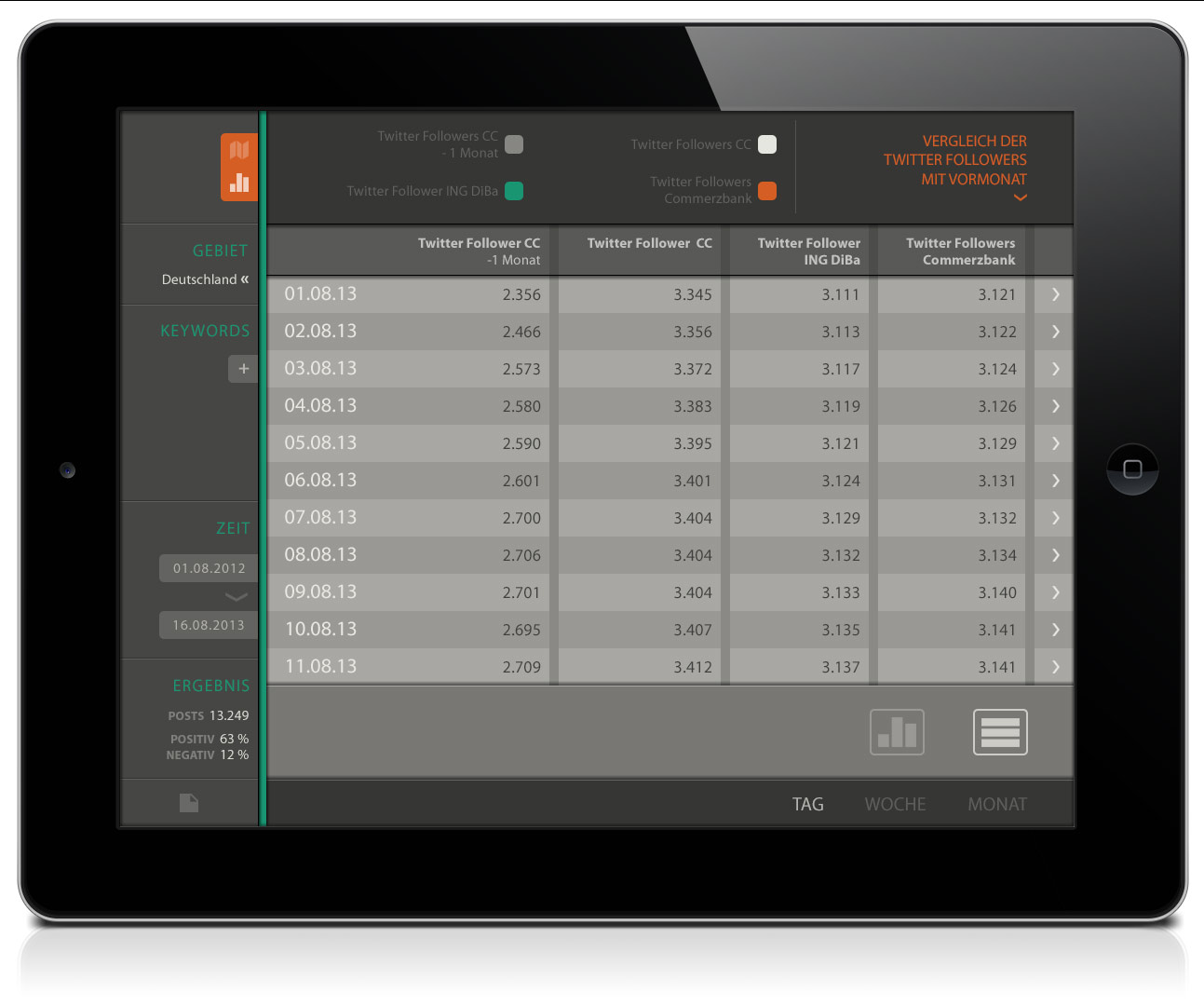


Abbildung – Detailansicht – Tabellarisch

**Grafiken, Tabellen und Posts**

Im mittleren Bereich werden die Werte über die Zeit zum einen grafisch dargestellt. In der Tages-Ansicht erfolgt dies als Liniendiagramm – in der Wochen- und Monatsansicht als Balkendiagramm. Innerhalb der Diagramme kann sowohl gescrollt als auch gezoomt werden. So erhält man sowohl eine Darstellung des gesamten Zeitraumes, der eingestellt ist, als auch Ansichten bestimmter Abschnitte daraus.



Abbildung – Detailansicht – Posts

Die Tabellenansicht gibt eine tabellarische Aufstellung der definierten Werte. Je nach Auswahl werden hier pro Zeile Tages-, Wochen- oder Monatswerte angezeigt. In dieser Tabellenansicht können mit einem Tap auf den Pfeil am Ende jeder Zeile sämtliche Posts für diesen Zeitraum aufgerufen werden. Die Darstellung dieser Posts erfolgt im gleichen Anzeigebereich. Sie können nach entsprechenden Kriterien gefiltert werden – z.B. danach ob es ein positiver oder negativer Post ist.

Mit einem Tap auf „zurück“ gelangt man zur vorhergehenden Tabellenansicht.

# Lösungsarchitektur

## High Level Lösungs-Architektur und Prozessablauf

Das folgende Schaubild skizziert die grundlegende Lösungs-Architektur, stellt aber keine exakte System-Architektur dar – diese wird in Kapitel 4 ab Seite 15 definiert. Aufgezeigt werden dagegen die einzelnen Komponenten mit ihrem jeweiligen Aufgabengebiet.

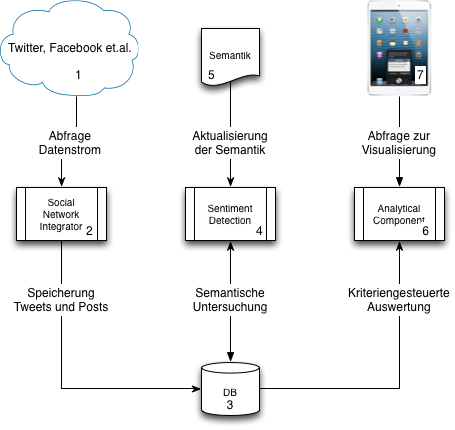


Abbildung – High Level Solution Architecture

### Prozessablauf

1. Aktuelle Tweets und Nachrichten werden in den sozialen Netzwerken gepostet
2. Der Datenstrom der sozialen Netzwerke wird periodisch abgefragt
3. Speicherung der Posts in der Datenbank
4. Semantische Untersuchung der Daten zur Bestimmung positiver oder negativer Tendenz
5. Periodische oder bedarfsgesteuerte Aktualisierung der Semantik
6. Auswertung der gespeicherten Daten nach vorgegebenen Kriterien, z.B. Region
7. Visualisierung der ausgewerteten Daten in der iPad App (siehe Mockup)

## Komponenten der Architektur

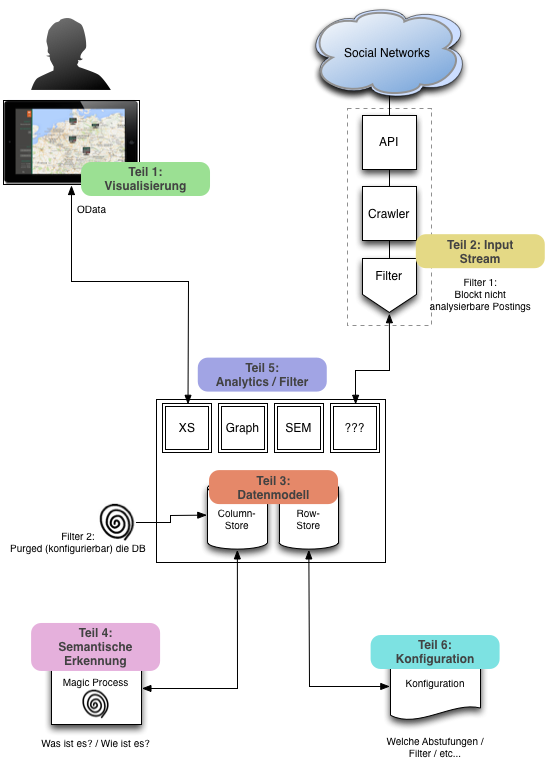


Abbildung – Bestandteile der Lösung

1. Teil 1 Visualisierung
2. Teil 2 Input Stream-Verarbeitung
3. Teil 3 Datenmodell
4. Teil 4 Semantische Erkennung
5. Teil 5 Analytics – Filterkonzept
6. Teil 6 Konfiguration
7. Teil 7 Interfacing – HANA zu UI

## Visualisierung

### Fragen

**Welche Technologie?**

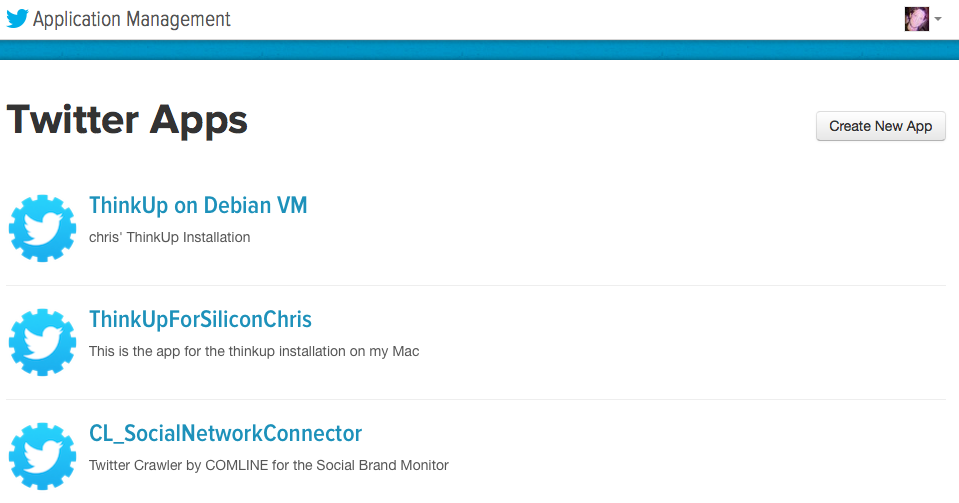
**Wie fragt es ab?**

## Input Stream Verarbeitung Python

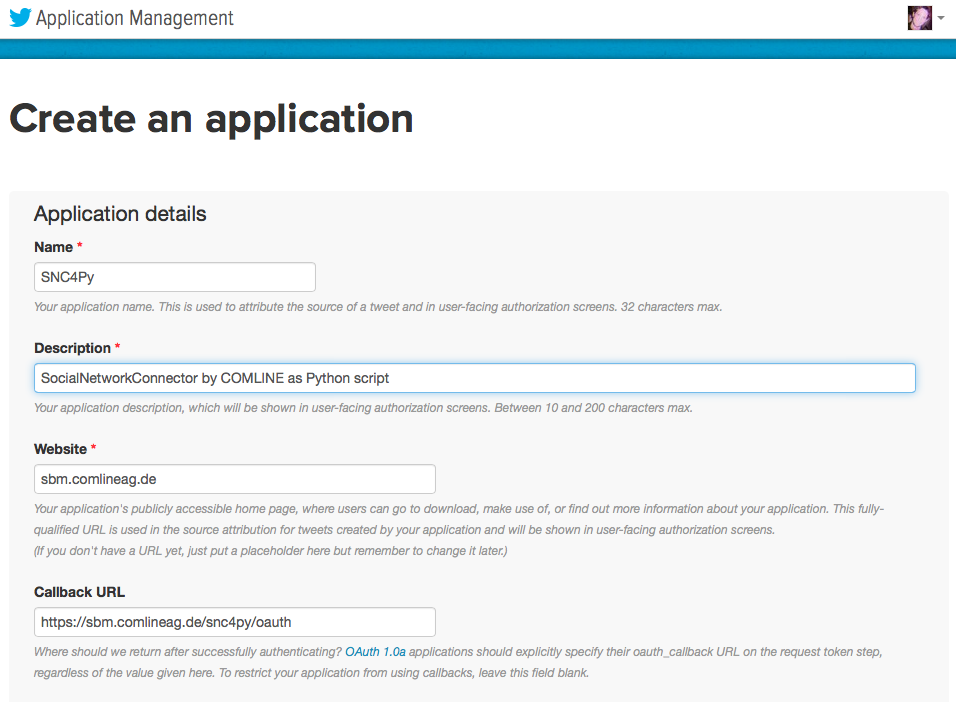
Die Twieets von Twitter werden in der die Input-Stream Verarbeitung (hier als Python Anwendung SNC4Py ausgeführt) geladen und in die Datenbank geschrieben..

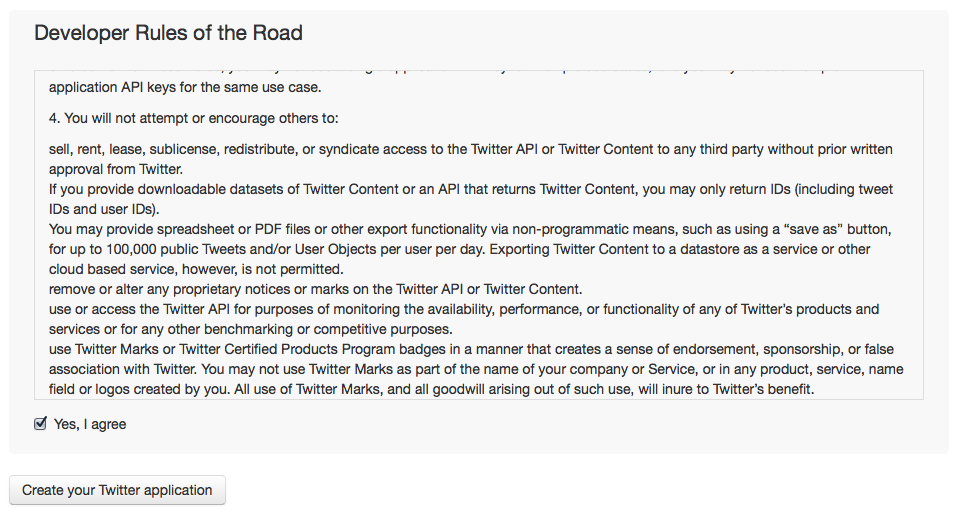
### Social Network Integration

Applikation auf Twitter registrieren: apps.twitter.com

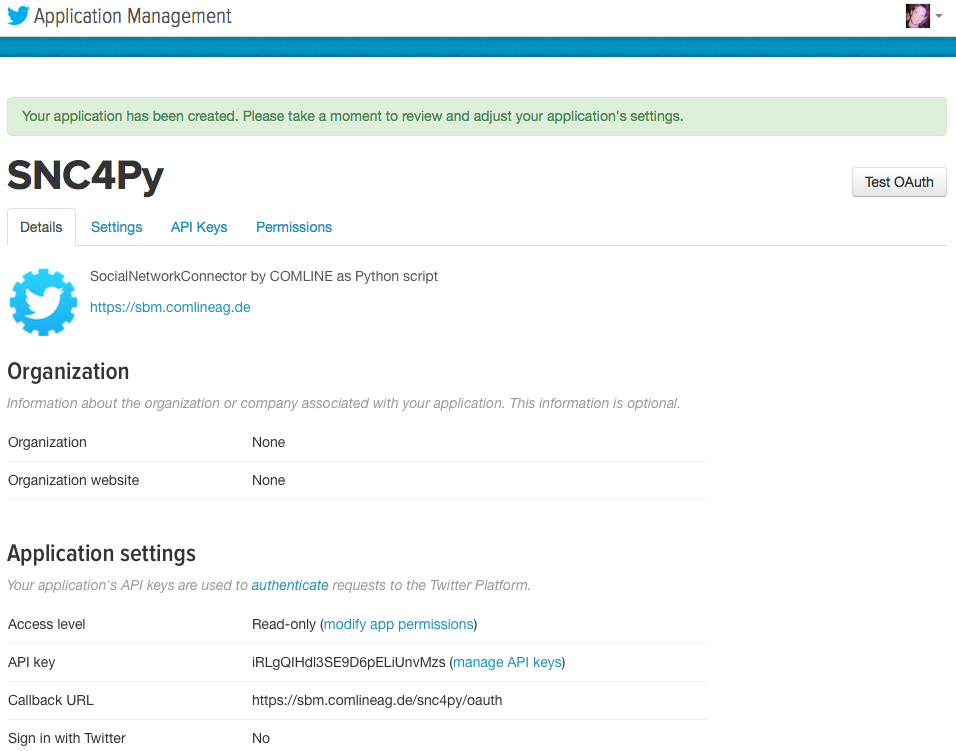


Create New App

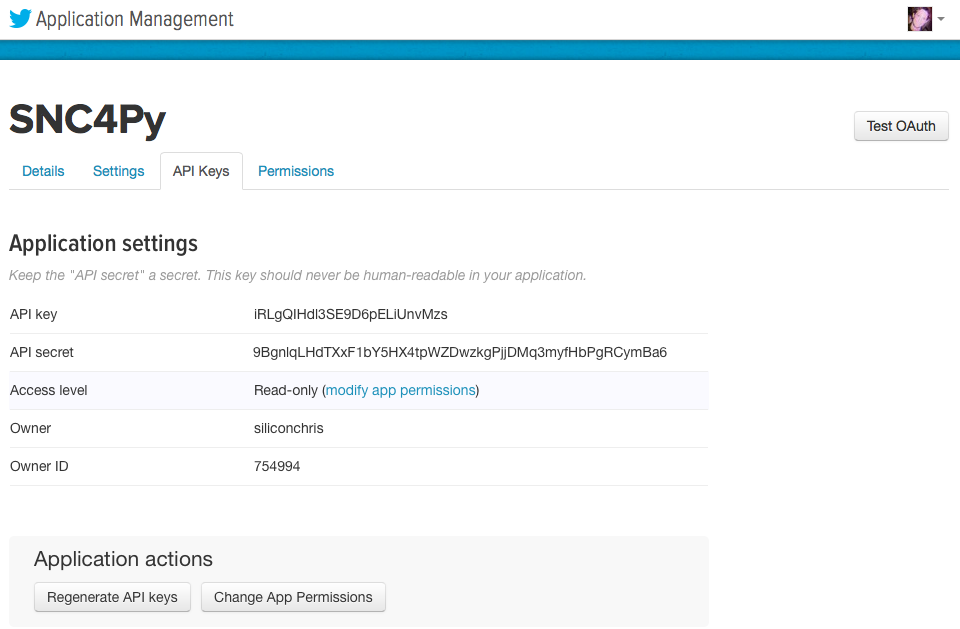




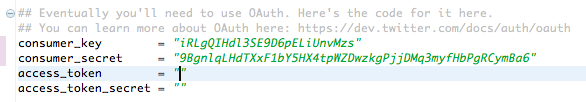
Die Application wurde erstellt.



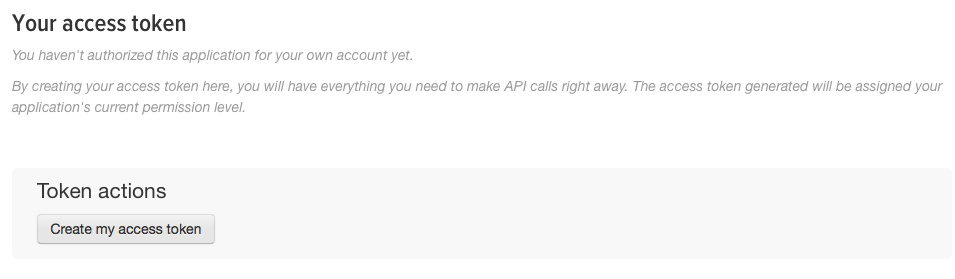
Generierte Keys kopieren



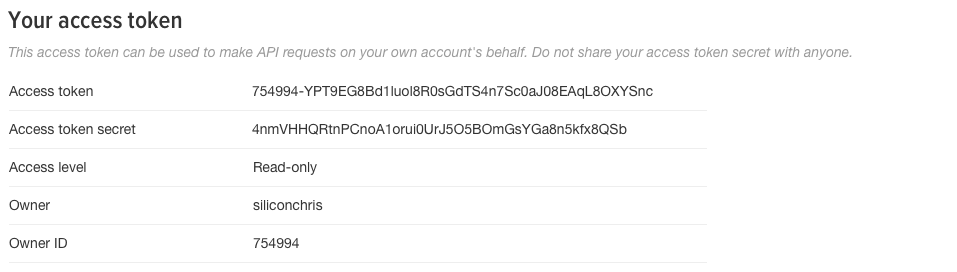
Einfügen der API Keys und secret (Consumer Key und Consumer Secret)



Access Tokens erzeugen:



Evtl. muss die Webseite ein paar mal refreshed werden, bevor sie angezeigt werden.



Access Token und Secret kopieren und ebenfalls in die OAuth-Methode des Programms einfügen:

## Input Stream Verarbeitung Java

Die Input-Stream Verarbeitung wird als Java-EE Anwendung im SocialNetworkConnector ausgeführt.

### Social Network Integration

Twitter API nutzen

Twitter-Crawler mit Input-Filter entwickeln

Jobgesteuerte Ausführung Crawler aufsetzen

### Crawler

Es wird ein generalisierter Crawler entwickelt. Von diesem wird jeweils eine spezialisierte Ableitung für jedes anzubindende Social Network abgeleitet. Der spezialisierte Crawler implementiert jeweils die API des Social Network.

### Fragen

**Welche Klassen/Methoden/Softwarekomponenten?**

## Datenmodell

### Fragen

**Wie sieht das Datenmodell aus?**

**Welche Tabellen brauchen wir?**

**Müssen wir im Datenmodell die Graph-Engine vorsehen?**

**Wie wird eine Abfrage gegen einen Graphen programmiert?**

## Semantische Erkennung

Im Rahmen der semantischen Analyse müssen für den Anwender zwei Fragen beantwortet werden:

***Was ist es?*** Ist ein Apple eine Frucht, oder ein Computer?

und

***Wie ist es?*** Findet der Schreibende Äpfel gut oder schlecht?

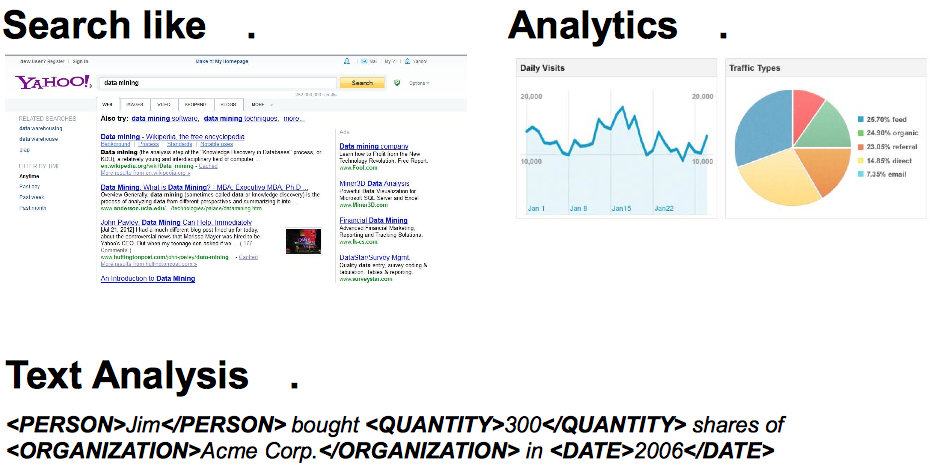


Abbildung – Text Analysis und Analytics

Um diese Fragen (Was und Wie ist es?) zu beantworten, muss die Struktur eines Textes vom System erfasst werden. Die Text-Analyse strukturiert einen Text nach **Core Entities** und **Domajn Facts**.

**Core Entities:**

Davey Jones was one of the Monkeys.

<PERSON> Davey Jones </PERSON> was one of the Monkeys.

**Domain Facts:**

I love your product.

I <STRONGPOSITIVESENTIMENT> love </SPS><TOPIC> your product </TOPIC>.

**Core Entities** sind also die Dinge der realen Welt; Personen, Produkte oder alles, *worüber* man redet. **Domain Facts** können als Eigenschaften einer Core Entity (Farbe, Geschmack o.ä.) verstanden werden. Sie können aber auch eine geäußerte Meinung, also *wie* man über das Objekt redet, sein.

### Fragen

**Welche Komponenten nutzen wir?**

**Wie soll diese arbeiten? Trigger/Batch/Job etc...**

## Analytics (Filterkonzept)

### Fragen

**Wie dynamisch muss die Analyse sein?**

**Welche Arten von Filter brauchen wir? GEO/Keywords etc...**

## Konfiguration

### Fragen

**Wie wird die Konfiguration aufgebaut? Dynamisch/Statisch etc...**

**Enthält diese auch semantische Erkennung?**

## Interfacing

### Fragen

**Welche Interfaces stellt HANA bereit?**

# Umsetzung mit SAP HANA

Im Folgenden wird die Umsetzung der avisierten Lösung auf Basis von SAP HANA und der bereits präsentierten iPad-Anwendung beschrieben.

Das folgende Schaubild veranschaulicht die grundsätzliche Systemarchitektur bei Umsetzung mit der SAP HANA Appliance[[1]](#footnote-1).

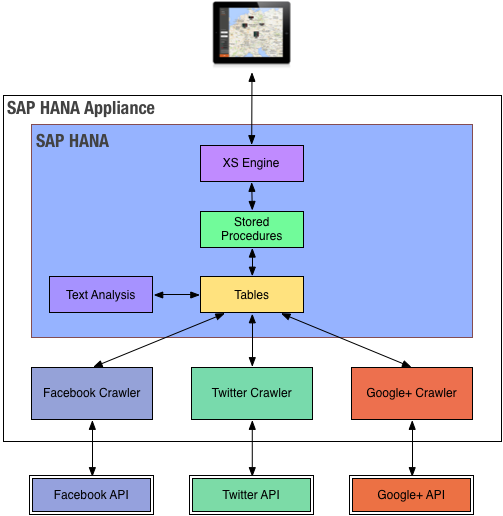


Abbildung – Systemarchitektur mit SAP HANA

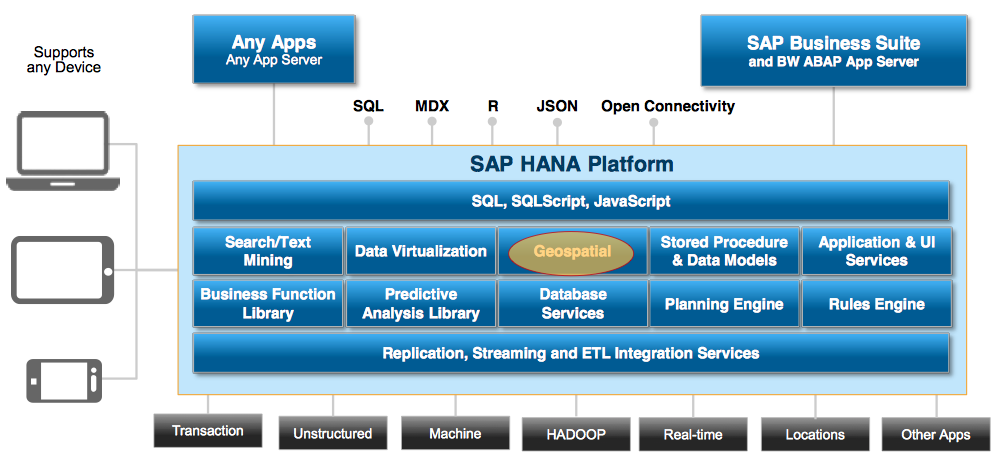


Abbildung – SAP HANA Elemente

## Visualisierung

## Input Stream – Abfrage Social Network

Auf der SAP HANA Appliance werden Crawler für jedes social network (Twitter, Facebook, Googl+ etc.) installiert, welche die API der jeweiligen Netzwerke ansprechen, um alle relevanten Daten periodisch abzufragen. Die erhaltenen Daten werden in Tabellen der SAP HANA Datenbank abgelegt.

## Datenmodell

## Semantische Erkennung

Die semantische Analyse ...

## Analytics oder Filterkonzept

Die Abfrage der relevanten Tweets und Posts aus der Datenbank wird auf Basis eines Graphensystems umgesetzt.

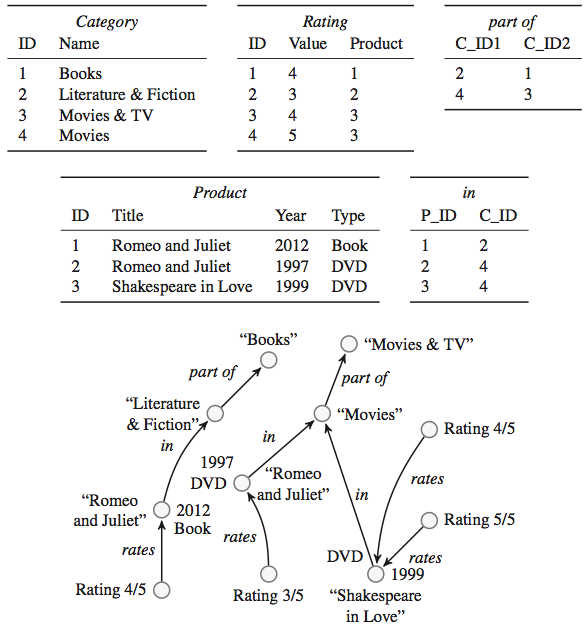


Abbildung – Graphendarstellung von Text

Innerhalb der HANA ist die Graphen-Engine wie folgt implementiert ...

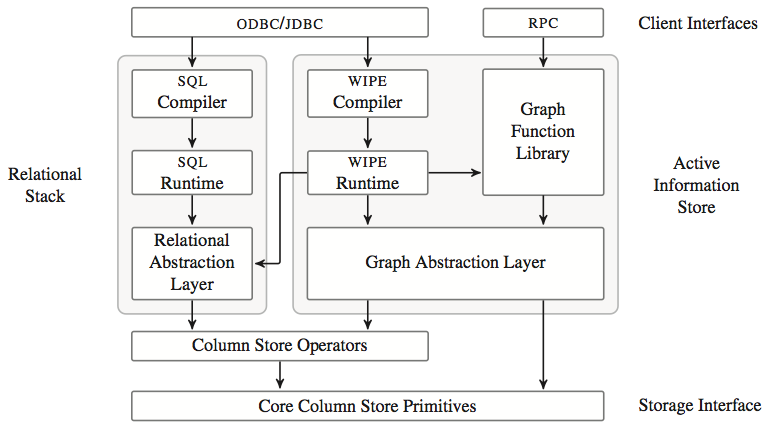


Abbildung – Architektur Graphen-Engine in HANA

## Konfiguration

## Interfacing

Die Tabellen werden über Stored Procedures, geschrieben in SQLScript, angesprochen. Die Stored Procedures werden über die XS-Services der HANA exponiert und stellen eine REST-Schnittstelle für die iPad-Anwendung bereit.

# Referenzen

**Text Search and Text Analysis with SAP HANA**

<http://www.saphana.com/servlet/JiveServlet/previewBody/3996-102-2-8477/Text_Search_and_Text_Analysis_with_SAP_HANA_%5B1%5D.pdf>

**Social Media Analysis with SAP HANA and Predictive Analysis**

<http://www.saphana.com/community/blogs/blog/2013/08/16/sharknado-social-media-analysis-with-sap-hana-and-predictive-analysis>

**Geo-Spatial Processing in SAP HANA**

<http://www.saphana.com/servlet/JiveServlet/previewBody/4299-102-1-9008/HANA_SPS07_NEW_Geospatial.pdf>

**The Graph Story of the SAP HANA Database**

<http://www.slideshare.net/Mahican/the-graph-story-of-the-sap-hana-database>

<http://subs.emis.de/LNI/Proceedings/Proceedings214/403.pdf>

**10 Tips for River development**

<http://scn.sap.com/community/hana-in-memory/blog/2014/01/20/10-tips-for-sap-river-developers>

**SAP Lumira Anwenderhandbuch**

<http://help.sap.com/businessobject/product_guides/vi01/de/lumS12_user_guide_de.pdf>

**Integration of Apache Hadoop with SAP HANA**

<http://events.sap.com/teched/en/session/8586>

**Building an End-to-End App from HANA via ABAP to SAP UI5**

<http://events.sap.com/teched/en/session/8907>

**SAP HANA Architecture**

<http://pi1.informatik.uni-mannheim.de/filepool/big-data-seminar-pdfs/5_2_The%20SAP%20HANA%20Database%20--%20An%20Architecture%20Overview..pdf>

**Large Graph-Analytics in the SAP HANA DB through summarization**

<http://event.cwi.nl/grades2013/16-Rudolf.pdf>

Versionsnachweis

|  |  |  |  |
| --- | --- | --- | --- |
| Ver. | Datum | Autor | Kommentar |
| 0.1 | 06.01.2014 | C. Günther | Initialversion mit Strukturvorgabe |
| 0.2 | 10.03.2014 | C. Günther | HANA Feature Graphanalyse aufgenommen |
| 0.3 | 12.03.2014 | C. Günther | Kapitel 3 überarbeitet |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

Dokumentenpflege/Genehmigung

|  |  |
| --- | --- |
| Erstellt: | Christian Günther, Chief Solution Architect |
| Geprüft: |  |
| Genehmigt: | t.b.d. |

|  |  |  |  |
| --- | --- | --- | --- |
| Erstellt: | Christian Günther, Chief Solution Architect |  | 12.03.2014 |
| Unterschrift/Datum | |
| Geprüft: | Stephan Huber, Account Manager |  | Datum eintr. |
| Unterschrift/Datum | |
| Genehmigt: | t.b.d. |  | Datum eintr. |
| Unterschrift/Datum | |

Copyright COMLINE AG 2013

Inhaltliche Änderungen vorbehalten. Alle Rechte an dieser Dokumentation, insbesondere das Recht der Vervielfältigung und Verbreitung sowie der Übersetzung, bleiben vorbehalten.

Kein Teil der Dokumentation darf in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein sonstiges Verfahren) ohne vorherige schriftliche Zustimmung der COMLINE AG reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

1. Für eine genauere Betrachtung der SAP HANA Appliance siehe [www.sap.com/hana](http://www.sap.com/hana) [↑](#footnote-ref-1)