**Twitteranalyser**

Dokumentation

621.252 PR Software Engineering II  
LV-Leiter: Dipl.Ing. Dr. Heinz Pozewaunig  
SS 2014

Inhaltsverzeichnis

Dokumentation 1

Inhaltsverzeichnis 2

1. Einleitung 3

2. Projektorganisation 4

3. Zeiterfassung 5

*3.1 Sprintverlauf* 7

*3.2 Stundenübersicht Sprints* 8

4. Product Backlog 9

*4.1. SOLL Product Backlog* 9

*4.2. IST Product Backlog* 10

5. Kosten 11

6. Qualitätsplan 12

7. Technische Produktbeschreibung 13

*7.1. Architektur* 13

*7.2 Controllers* 14

*7.3 Views* 14

*7.4 Apimanagement* 14

*7.5 Sprachen* 14

8. Risikoanalyse 15

9. Anhang 16

1. Einleitung

Im Rahmen der Lehrveranstaltung ‚Software Engineering II‘ wurde eine Twitteranalyse entwickelt, wobei als Softwareentwicklungsmodell *Scrum* Anwendung fand. Das Projekt, an dem acht Personen gearbeitet haben, wurde in drei Monaten, vom 18. März bis zum 18. Juni 2014 realisiert. Genaue Angaben zu den Kosten für das Produkt finden sich in Kapitel 5.  
  
Wie beim Kundengespräch am 18.03.2014 von Ihnen erwähnt, sollte es möglich sein, dass ein Mitarbeiter in der Marketingabteilung ein Stimmungsbild einer bestimmten Veranstaltung durch die Sammlung von ausgewählten Tweets abrufen kann. Der Mitarbeiter sollte im Programm eine Veranstaltung anlegen, eine Veranstaltung auswählen, löschen, ändern oder speichern können. Es soll ein Zeitraum definiert werden können, bis wann die Tweets ausgewählt werden sollen. Die so erhaltene Sammlung soll analysiert, auf einer Landkarte dann die positiven und negativen Stimmungen angezeigt werden können und über die firmeninterne Analyseseite abrufbar sein.

Diese Anforderungen wurden alle in das Projekt aufgenommen, in fünf Sprints aufgeteilt und schließlich erfolgreich erledigt.

2. Projektorganisation

Das Scrum-Entwicklungsteam besteht neben dem Scrum Master Herrn Dipl.Ing. Dr. Heinz Pozewaunig aus acht Mitglieder deren Aufgabenbereiche in Tabelle 1 ersichtlich werden. Die Stärken und Schwächen hat jeder selbst für sich angegeben. Wie man aus der Tabelle ersehen kann, hat sich jeder eine oder mehrere Aufgaben passend zu seinen Stärken gewählt.

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Stärken** | **Schwächen** | **Aufgaben** |
| Kesselbacher Max | Dokumentieren, Programmieren | Keine Angabe | Benutzer Dokumentation, User Tests |
| Kammel Johannes | Entwicklung, Projektmanagement, Logik | Dokumentation, Präsentation | Entwicklung, Logik, Datenbank, GUI |
| Krumpholz Maximilian | Testen, Präsentieren, Dokumentieren | Programmieren | Unit Tests, Projektmanagement, Präsentation |
| Mertens Henrik | Dokumentation, Testen | Programmieren, Technik | Unit Tests, Präsentation |
| Moser Manfred | Programmieren, Geduld, selbständiges Arbeiten, Arbeitstempo | Mangelnde Kritikfähigkeit | Entwicklung, Logik, Datenbank, GUI |
| Mößlacher Corinna | Dokumentieren | Programmieren | Techn.Dokumentation, Datenbank,GUI |
| Ressmann Andreas | Lösungsorientiert, Programmieren, Datenbank | User Dokumentation | Projektmanagement, SetUp, Datenbank, GUI |
| Steinkellner Rosemarie | Dokumentieren, Analysieren | Programmieren | Dokumentation, Kalkulation |

Tabelle 1: Projektmitglieder

3. Zeiterfassung

Im Weiteren wird der zeitliche Ablauf des Projekts beschrieben. Für jeden Sprint wurde ein Sprintziel definiert, das es zu erreichen galt.

Nach dem ersten Sprint (Tabelle 2) sollte bereits eine minimale Version des Twitteranalyser existieren. Dieses Sprintziel wurde nicht ganz erreicht, da ein Kollege überraschend das Team verlassen hatte. Daher konnte das Ergebnis aus Sprint I Veranstaltungen anlegen gemeinsam mit Veranstaltungen anzeigen und bearbeiten erst am Ende von Sprint II (Tabelle 3) erfolgreich erreicht werden. Ziel von Sprint III und IV (Tabelle 4 und 5) war es, Tweets anzeigen und zu einer Analyse zusammenfassen zu können. Außerdem sollte eine Übersicht der Tweets sowie eine multilinguale Ausgabe der Keywords möglich sein. Alle diese Ziele wurden erfolgreich umgesetzt. Im finalen Sprint V (Tabelle 6) wurde das Ziel bereits angelegte Filter löschen können und die mögliche Darstellung der Analyse in Diagrammen ebenfalls erfolgreich umgesetzt.

Legende zu Tabellen 2 - 6: T = Abk. für Twitteranalyser, 1 = Sprint, a = Sprintbereich

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **ID** | **User Story ID** | **Sprint Ziel** |
| T1a | 1 | Als MM möchte ich eine Veranstaltung anlegen, sodass das Sammeln der Tweets beginnen kann |
| T1b | 2\_1 | Als MM möchte ich eine vorhandene Veranstaltung anzeigen. |
| T1c | 2\_2 | Als MM möchte ich eine vorhandene Veranstaltung bearbeiten. |

Tabelle 2: Sprint 1

|  |  |  |
| --- | --- | --- |
| **ID** | **User Story ID** | **Sprint Ziel** |
| T1a | 1 | Als MM möchte ich eine Veranstaltung anlegen, sodass das Sammeln der Tweets beginnen kann |
| T1b | 2\_1 | Als MM möchte ich eine vorhandene Veranstaltung anzeigen. |
| T1c | 2\_2 | Als MM möchte ich eine vorhandene Veranstaltung bearbeiten. |
| T2a | 3\_2 | Als <MM> will ich Tweets nach Anlegen einer Veranstaltung zur |
|  |  | Verfügung haben, sodass ich später eine Sentimentanalyse fahren kann |
| T2b | 3\_4 | Als <MM> will ich alle gesammelten Tweets einer Veranstaltung |
|  |  | anzeigen, sodass ich eine Übersicht bekomme |

Tabelle 3: Sprint 2

|  |  |  |
| --- | --- | --- |
| **ID** | **User Story ID** | **Sprint Ziel** |
| T3a | 3\_1 | Als <MM> will ich eine positiv/negativ Sentiment Analyse der Tweets erhalten |
| T3b | 4\_1 | Als <MM> will ich die vorhandenen Tweets einschränken, |
|  |  | sodass nur relevante Tweets analysiert werden |

Tabelle 4: Sprint 3

|  |  |  |
| --- | --- | --- |
| **ID** | **User Story ID** | **Sprint Ziel** |
| T4a | 5\_1 | Als <MM> will ich bereits angelegte Filter auch löschen können. |
| T4b | 3\_3 | Als <MM> will ich eigene Keywords in beliebiger Sprache |
|  |  | für die positiv / negativ Semtimentanalyse bestimmen können. |
| T4c | 2\_1 | Als <MM> möchte ich eine Mindestanzahl von zu sammelden Tweets pro Veranstaltung angeben, damit ich bewerten kann, ob das Analyseergebnis sinnvoll ist. |
| T4d | 4\_3 | Als <MM> will ich einzelene Tweets manuell löschen, sodass diese nicht wieder analysiert werden. |
| T4e | 4\_2 | Als <MM> will ich einzelene Tweets manuell für einen Filter ausblenden, |
|  |  | sodass diese nicht analysiert werden. |

Tabelle 5: Sprint 4

|  |  |  |
| --- | --- | --- |
| **ID** | **User Story ID** | **Sprint Ziel** |
| T5a | 5 | Als <MM> will ich zu einer bestimmten Veranstaltung gehörende Daten als .csv exportieren. |
| T5b | 3\_2 | Als <MM> will ich verschiedene Darstellungen der analysierten Daten. |

Tabelle 6: Sprint 5

## *3.1 Sprintverlauf*

Die folgende Tabelle 7 bietet eine Übersicht über den gesamten Projektverlauf. Die Aufgabenbereiche sind dabei den dafür zuständigen Personen zugeteilt. Die grauen Balken zeigen an, über welchen Zeitraum diese Tätigkeiten erledigt wurden. Weiße Felder im Sprint bedeuten, dass für diese Aufgabe keine Zeit aufgewendet wurde.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Aufgabe*** | ***Person*** | ***Sprint I*** |  | ***Sprint II*** |  | ***Sprint III*** |  | ***Sprint IV*** |  | ***Sprint V*** |  |
| Technische Dokumentation | Mößlacher |  |  |  |  |  |  |  |  |  |  |
| Benutzer Dokumentation | Kesselbacher |  |  |  |  |  |  |  |  |  |  |
| Dokumentation | Steinkellner |  |  |  |  |  |  |  |  |  |  |
| Unit Tests | Krumpholz |  |  |  |  |  |  |  |  |  |  |
|  | Mertens |  |  |  |  |  |  |  |  |  |  |
| User Tests, User Manuel | Kesselbacher |  |  |  |  |  |  |  |  |  |  |
| Multilingualität | Steinkellner |  |  |  |  |  |  |  |  |  |  |
|  | Mertens |  |  |  |  |  |  |  |  |  |  |
|  | Mößlacher |  |  |  |  |  |  |  |  |  |  |
| Logik | Ressmann |  |  |  |  |  |  |  |  |  |  |
| Design | Ressmann |  |  |  |  |  |  |  |  |  |  |
| GUI | Ressmann |  |  |  |  |  |  |  |  |  |  |
|  | Kammel |  |  |  |  |  |  |  |  |  |  |
|  | Moser |  |  |  |  |  |  |  |  |  |  |
|  | Mößlacher |  |  |  |  |  |  |  |  |  |  |
| Sentiment, Twitter | Kammel |  |  |  |  |  |  |  |  |  |  |
|  | Moser |  |  |  |  |  |  |  |  |  |  |
| Datenbank | Moser |  |  |  |  |  |  |  |  |  |  |
|  | Kammel |  |  |  |  |  |  |  |  |  |  |
|  | Mößlacher |  |  |  |  |  |  |  |  |  |  |
|  | Ressmann |  |  |  |  |  |  |  |  |  |  |
| Filter-anpassung | Kammel |  |  |  |  |  |  |  |  |  |  |
|  | Moser |  |  |  |  |  |  |  |  |  |  |
|  | Ressmann |  |  |  |  |  |  |  |  |  |  |
| Darstellung Analyse | Kammel |  |  |  |  |  |  |  |  |  |  |
|  | Ressmann |  |  |  |  |  |  |  |  |  |  |
|  | Moser |  |  |  |  |  |  |  |  |  |  |

Tabelle 7: Übersicht Projekt

## *3.2 Stundenübersicht Sprints*

In der folgenden Tabelle 8 finden sich die Stundenanzahlen, die pro Sprint in das Projekt investiert wurden und die gesamte Stundenanzahl aller Mitarbeiter, die sich auf 678 Stunden beläuft. Man sieht, dass der zweite und der fünfte Sprint am längsten gedauert haben, was wohl durch den Mehraufwand in Sprint II durch das Verlassen des Teams eines Mitarbeiters zu erklären ist und in Sprint V wurde mehr Zeit auf die abschließenden Dokumentationen und die Präsentation verwendet. Tabelle 9 zeigt dazu den Zeitaufwand pro Sprint in Prozent.

Im Anhang finden Sie eine detaillierte Aufstellung der Sprints mit Zuteilung der Mitarbeiter und den Aufgaben.

|  |  |
| --- | --- |
| *Sprint* | *Stunden* |
| Sprint I | 121 |
| Sprint II | 186 |
| Sprint III | 108 |
| Sprint IV | 85 |
| Sprint V | 178 |
| *Gesamt* | 678 |

Tabelle 8: Zeitaufwand pro Sprint

Tabelle 9: Zeitaufwand pro Sprint

4. Product Backlog

## *4.1. SOLL Product Backlog*

Die nachfolgende Tabelle 10 stellt die ursprüngliche Liste mit Anforderungen, die nach dem Kundengespräch aufgestellt wurde, dar.

|  |  |  |
| --- | --- | --- |
| **Beschreibung** | **User Story Points** | **Priorität** |
|  |  |  |
| Anlegen einer Veranstaltung | **3** | high |
| Webanwendung - Detailansicht | 1 | high |
| Parameteränderung | 3 | low |
| Auswertung | 3 | high |
| Filterobjekterstellung |  | high |
| Funktionalität: Löschen | 2 | medium |
| Filterobjektspeicherung |  | high |
| Funktionalität: Ausblenden | 7 | medium |
| Export | **2** | low |
| Darstellung | 5 | high |
| Benachrichtigung |  | low |

Tabelle 10: SOLL Product Backlog

## *4.2. IST Product Backlog*

Daraus wurde dann der folgende Product Backlog (Tabelle 11) erstellt, der für die weiteren Sprints zur Abarbeitung bestimmt war. Den jeweiligen Anforderungen sind Aufgabenbereiche zugeteilt mit Prioritäten von 10 = nicht so wichtig bis 100 = sehr wichtig. Zudem ist angegeben, in welchem Sprint (siehe Legende) die Tätigkeit erfüllt wurde.

Legende zu Tabelle 11: T = Abk. für Twitteranalyser, 1 = Sprint, a = Sprintbereich

|  |  |  |  |
| --- | --- | --- | --- |
| **ID** | **Sprint** | **Beschreibung** | **Priorität** |
| T1a | 1 | Als <MM> will ich eine Veranstaltung anlegen, | 100 |
|  |  | sodass das Sammeln der Tweets beginnt |  |
| T1b | 2-1 | Als <MM> will ich eine vorhandene Veranstaltung anzeigen | 60 |
| T1c | 2-2 | Als <MM> will ich eine vorhandene Veranstaltung bearbeiten | 55 |
| T2a | 3-2 | Als <MM> will ich das Tweets nach Anlegen einer Veranstaltung | 99 |
|  |  | zur Verfügung haben, so dass ich später eine Sentimentanalyse |  |
|  |  | fahren kann |  |
| T2b | 3-4 | Als <MM> will ich die gesammelten Tweets einer Veranstaltung | 55 |
|  |  | alle anzeigen, sodass ich ein Übersicht bekomme |  |
| T3a | 3-1 | Als <MM> will ich eine positiv/negativ Sentiment Analyse der Tweets | 40 |
|  |  | erhalten |  |
| T3b | 4-1 | Als <MM> will ich die vorhandenen Tweets einschränken, | 30 |
|  |  | sodass nur relevante Tweets analysiert werden |  |
| T4a | 5-1 | Als <MM> will ich bereits angelegte Filter auch löschen können. | 30 |
| T4b | 3-3 | Als <MM> will ich eigene Keywords in beliebiger Sprache für die | 27 |
|  |  | positiv / negativ Semtimentanalyse bestimmen können. |  |
| T4c | 2-1 | Als <MM> möchte ich eine Mindestanzahl von zu sammelden Tweets | 26 |
|  |  | pro Veranstaltung angeben, damit ich bewerten kann, |  |
|  |  | ob das Analyseergebnis sinnvoll ist. |  |
| T4d | 4-3 | Als <MM> will ich einzelene Tweets manuell löschen, sodass diese | 20 |
|  |  | nicht wieder analysiert werden. |  |
| T4e | 4-2 | Als <MM> will ich einzelene Tweets manuell für einen Filter | 19 |
|  |  | ausblenden, sodass diese nicht analysiert werden. |  |
|  | 5 | Als <MM> will ich zu einer bestimmten Veranstaltung gehörende | 15 |
|  |  | Daten als .csv exportieren. |  |
|  | 3-2 | Als <MM> will ich verschiedene Darstellungen der analysierten Daten. | 16 |
|  | 1-5 | Als <MM> will ich über Tweets informiert werden. | 10 |

Tabelle 11: IST Product Backlog

5. Kosten

Nachfolgend findet sich eine Aufstellung der Kosten für das Projekt (siehe auch die Aufwandsschätzung vom 05.05.2014):

Die Kostenschätzung mittels Function-Point Verfahren ergibt 54 UFP und 47,52 AFP.

Die Kostenschätzung mittels Application-Point Verfahren beträgt bei einem System mit 12 Screens, 4 Berichten mittlerer Komplexität, 6 3GL Modulen und ca. 25 % Wiederverwendung, eine gute Umgebung vorausgesetzt und wenn wir schätzen, dass die Mitarbeiter mittlere Erfahrung haben 104 AP’s und 17 NAP’s = 6 Personenmonate

Der Durchschnitt der beiden Varianten ergibt nun 13 Personenmonate.

Die Gesamtkosten errechnen sich wie folgt:

Stundensatz von 90 Euro x 678 Stunden 61.020,00 €

Nettorechnungsbetrag 61.020,00 €

+ 20 Prozent Umsatzsteuer 12.204,00 €

**Bruttorechnungsbetrag 73.224,00 €**

6. Qualitätsplan

Schon während der Entwicklung des Software-Produkts wurden regelmäßig Unittests durchgeführt, um eine hohe Qualität zu gewährleisten. Zudem wurde das Produkt von Mitarbeitern und externen Personen kontinuierlich auf Benutzerfreundlichkeit getestet. Im Anhang befinden sich Links zu den Ergebnissen dieser Tests. Außerdem ist dort eine Verlinkung zu den Code-Metriken zu finden.

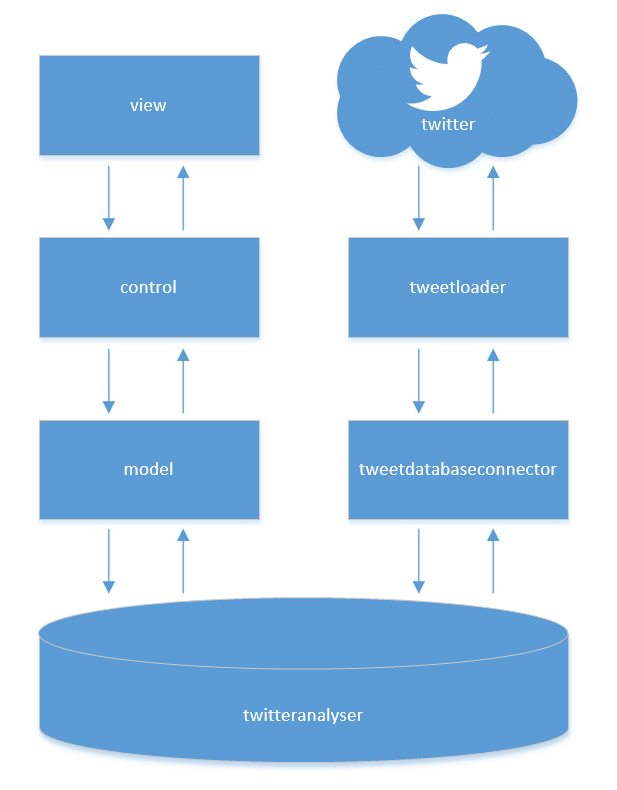
7. Technische Produktbeschreibung

Im Folgenden wird auf die technische Realisierung der Software näher eingegangen. Für nähere Details siehe auch die technische Dokumentation, die sich auf der CD befindet.

Die Software wurde mittels der objektorientierten Programmier- und Skriptsprachen PHP und Java realisiert.

Die Struktur des Projektes ist eine klassische MVC Architektur, realisiert mit Zend Framework (Version 1).

## *7.1. Architektur*



Der Beginn der Anwendung ruft die Index.php Datei auf welche den Zend Bootloader für die primäre Initialisierung des Projekts ladet. Alle Anfragen, die über die Website empfangen werden, werden an die entsprechende Steuerung durch das Zend Framework weitergeleitet. Die Steuerung wird durch die Unterordner der Anwendung bestimmt. Alle Hauptmenüpunkte haben eine entsprechende Steuerung. Der Standard Controller ist der Index Controller.   
Alle Menüunterpunkte sind Aktionen im entsprechenden Controller. Jeder Klick auf einen Unterpunkt ruft eine andere Aktion in der Steuerung auf. Die Index Action wird aufgerufen, wenn der Controller aufgerufen wird. Die Anwendung nutzt Twitter-Bootstrap. Dies ist für die Gestaltung der grafischen Benutzeroberfläche und der Präsentation (z.B. der Datumsauswahl) verantwortlich.

## *7.2 Controllers*

Die Controller Logik ist in fünf Controller aufgeteilt:

Index Controller (Haupt Controller), API Controller, Error Controller, Map Controller und Statistics Controller

## *7.3 Views*

Die phtml-Dateien in "Views" sind für die Darstellung der Web-Seite eingesetzt. Sie enthalten neben den HTML-Code auch PHP-Code für die dynamischen Aspekte der Seite. Alle Nutzer-Interaktionen rufen eine Aktion im Controller hervor, zu der die View gehört. Für die Views "Analyse", "Event" und "Tweets" ist es der API-Controller. Die View "Ereignis" erzeugt eine Liste als Übersicht aller derzeit bestehenden Veranstaltungen. Für die Ansicht der Karte und der Statistik werden Google Maps und Google Visualisation für die Darstellung der Karte und Diagramme verwendet.

## *7.4 Apimanagement*

Die Dialoge sind durch die Datei "apimanagement.js" verwaltet. Alle Dialoge der Web-Anwendung werden unter Verwendung von EJS-Code generiert, der geladen und dynamisch über JQuery angezeigt wird. Das API Management sendet JSON API-Anfragen an den Controller. Der Controller sendet Antworten oder Validierungsfehler.

## *7.5 Sprachen*

Die CSV-Dateien enthalten die Texte, die auf der Webseite angezeigt werden.

*7.6 Models*  
Für alle Datenbanktabellen, die relevant für die Ansichten sind, existiert ein entsprechendes Modell.

8. Risikoanalyse

In der nachfolgenden Tabelle werden mögliche Risiken, die Wahrscheinlichkeit ihres Eintritts und Pläne zur Vermeidung des jeweiligen Risikos aufgelistet. Die Auswertung, der Durchschnitt von Wahrscheinlichkeit und Schweregrad, wird von 2 = geringes Risiko bis 4 = hohes Risiko definiert.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Risiko*** | ***Wahrschein-lichkeit*** | ***Schwere-grad*** | ***Auswertung*** | ***Gegenmaßnahmen*** |
| Personalausfall | Möglich | Moderat | 2 | Konfliktarbeit, Motivation der Mitarbeiter |
| Unrealistische Termin- und Budgetplanung | Wahrscheinlich | Hoch | 4 | Genaue Aufwands- und Kostenschätzung |
| Entwicklung falscher Funktionalitäten | Möglich | Hoch | 3 | Genaue Anforderungsanalyse, Kundengespräche protokollieren, stetiger Kundenkontakt, |
| Entwicklung einer falschen Benutzeroberfläche | Möglich | Hoch | 3 | Kundengespräche protokollieren, stetiger Kundenkontakt, genaue Anforderungsanalyse, Usability-Tests |
| Anforderungen ändern sich kontinuierlich | Möglich | Moderat | 2 | Leicht änderbare Software entwickeln |
| Nichterfüllen der gewünschten Anforderungen | Möglich | Hoch | 3 | Genaue Zeitplanung, gute Arbeitsaufteilung |
| Mangelnde Echtzeit-Performance | Möglich | Hoch | 3 | Kontinuierliche Tests, Messung und Bewertung |
| Datenverlust | Möglich | Schwer-wiegend | 4 | Backups, Versionsverwaltung mit Git |

Tabelle 12: Risikoanalyse

9. Anhang

CD

* Software Twitteranalyser
* Technische Dokumentation
* User Dokumentation
* Dokumentation
* Sprintablauf Detailansicht