Quickstart Guide

Author:Thomas MikschaE-Mail:thomas.mikscha@gmail.comLast modified:1/26/2014 2:00:50 PM

Requirements

- Windows 7
- .NET Framework 4.0
- IIS 7.0 Webserver
- Microsoft SQL Server 2008 R2
- Microsoft SQL Server Management Studio (optional)
- LinqToTwitter
- Log4net
- JSON.net
- Twitter API key

Installation and Deployment

Database

1. Create an empty database called aic13demo

Seite auswählen 🊰 Allgemein	🔄 Skript 🔻 🚺 Hilfe						
🚰 Optionen 🚰 Dateigruppen	Datenbankname:		aic13demo	aic 13demo			
	Besitzer:		<standard></standard>				
	Volltextindizieru	ng verwende	n				
	Datenbankdateien	:	-				
	Logischer Name	Dateityp	Dateigruppe	Anfangsgröße (MB)	Automatische Vergrößerun		
	aic13demo	Zeilen	PRIMARY	3	Um 1 MB, unbeschränkte		
	aic13demo_log	Protokoll	Nicht zutreffend	1	Um 10 Prozent, unbeschrä		
Verbindung							
Verbindung Server: localhost'sqlexpress							
Verbindung Server: localhost/sqlexpress Verbindung: WIN-B9JSU3N5D9N/tom							
Verbindung Server: localhost'sqlexpress Verbindung: WIN-B3JSU3N5D9N'tom Verbindungseigenschaften, anzeigen							
Verbindung Server: localhost'sqlexpress Verbindung: WIN-B3JSU3N5D9N'tom WIN-B3JSU3N5D9N'tom Webindungseigenschaften, anzeigen Status							
Verbindung Server: localhost\sqlexpress Verbindung: Wih-PBJSU3V5D9N\tom Verbindungseigenschaften, anzeigen Status Beret	<	n		Hinzufügen	Entfemen		
Verbindung Server: locahostvajexpress Verbindung: WIN-B9JSU3N5D9Ntom ⊮ Verbindungseigenschaften. anzeigen Status Beret Beret	<	п		Hinzufügen	• Entfemen		

2. Create a new user login on the SQL server e.g. aic13demo

Anmeldung - Neu		
Seite auswählen	🔄 Skript 🔻 🖪 Hilfe	
Algemein Serverrollen		
Penutzerzuordnung	Anmeldename:	aic13demo Suchen
Status	Windows-Authentifizierung	
_	SQL Server-Authentifizierung	
	Kennwort:	•••••
	Kennwort bestätigen:	•••••
	Altes Kennwort angeben	
	Altes Kennwort:	
	Kennwortrichtlinie erzwingen	
	Ablauf des Kennworts erzwingen	
	Benutzer muss das Kennwort bei der nä	chsten Anmeldung ändem
	Zugeordnet zu Zertifikat	· · · · · · · · · · · · · · · · · · ·
	Zugeordnet zu asymmetrischem Schlüssel	
Verbindung	Zu Anmeldeinformationen zuordnen	▼ Hinzufügen
Server: localhost\sqlexpress	Zugeordnete Anmeldeinformationen	Anmeldeinfor Anbieter
Verbindung: WIN-B9JSU3N5D9N\tom		
Verbindungseigenschaften anzeigen		
Status		Entfemen
Bereit	Standarddatenbank:	master 💌
.445 4	Standardsprache:	<standard></standard>
		OK Abbrechen
_		

3. Add the user to the database and assign the roles "public" and "db_owner"

Seite auswählen	Skript 🔻 🖪 Hilfe		
Allgemein			
Serverrollen			
Sichen ungefähige Elemente	Benutzer, die dieser Anmeldung	zugeordnet sind:	2 1 1 1
Status	Zuord Datenbank	Benutzer	Standardschema
	AIC13		
	☑ aic13demo	aic13demo	
	master		
	model		
	msdb		
	tempdb		
	Gastkonto aktiviert für: aic 1	3demo	
Verbindung	Gastkonto aktiviert für: aic 13	3demo e für:aic13demo	
/erbindung	Gastkonto aktiviert für: aic 13 Mitgliedschaft in Datenbankrolle db_accessadmin db_backuragenter	3demo e für: aic13demo	
Verbindung Server: localhost\sqlexpress	Gastkonto aktiviert für: aic 13 Mitgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_backupoperator	3demo 9 für: aic13demo	
Verbindung Server: localhost\sqlexpress Verbindung:	Gastkonto aktiviert für: aic 13 Mitgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_datareader db_datawriter	3demo • für: aic 13demo	
Verbindung Server: localhostisqlexpress Verbindung: WIN-B9JSU3N5D9N\tom	Gastkonto aktiviert für: aic 13 Mitgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_datareader db_datawriter db_ddladmin	3demo • für: aic 13demo	
Verbindung Server: localhost\sqlexpress Verbindung: WIN-B3SU3N5D9N\tom ■2 Verbindungseigenschaften	Gastkonto aktiviet für: aic 1: Mitgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_datareader db_datawriter db_datamin db_datareader db_datareader db_datareader	3demo 9 für: aic 13demo	
Verbindung Server: localhost\sqlexpress Verbindung: WIN-B9JSU3N5D9N\tom 型 Verbindungseigenschaften anzeigen	Gastkonto aktiviert für: aic 1: Mtgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_datareader db_datawnter db_deladamin db_denydatareader db_denydatareader db_downer	3demo 9 für: alc13demo	
Verbindung Server: localhost\sqlexpress Verbindung: WIN-B9JSU3N5D9N\tom IV-trbindungseigenschaften_ anzeigen Status	Gastkonto aktiviert für: aic 1: Mitgliedschaft in Datenbankrolle db_backupoperator db_datareader db_datareader db_datawriter db_denydatareader db_denydatareader db_denydatawriter db_securityadmin	3demo 9 für: aic13demo	
Verbindung Server: localhost\sqlexpress Verbindung: WIN-B9JSU3N5D9N\tom Verbindungseigenschaften, anzeigen Status Bereit	Gastkonto aktiviert für: aic 1 Mitgliedschaft in Datenbankrolle db_accessadmin db_backupoperator db_datavräter db_datavräter db_datavräter db_denydatareader db_denydatareader db_denydatavräter db_securityadmin v public	3demo 9 für: aic13demo	
Verbindung Server: localhost\sqlexpress Verbindung: WIN-BSJSU3N5D9N\tom Verbindungseigenschaften_ anzeigen Status Bereit	Gastkonto aktiviert für: aic 1 Mitgliedschaft in Datenbankrolk db_backupoperator db_datareader db_datawriter db_datawriter db_denydatareader db_denydatawriter db_secuntyadmin dy_public	3demo 9 für: aic13demo	

4. Open the Microsoft SQL Server Management Studio, connect to the database, load the CreateTables.sql script and execute it in order to create the table structure.

atei Bearbeiten Anzeigen Abfrage Debuggen Extras Fen	ster Co	ommunity Hilfe			
Neue Abfrage 👔 🕞 🧀 🖂 🚟 🗕					
aic13demo + Ausführen + 🗸	3 🖷 🕻	1294 時間的1日日注意人。			
iekt-Explorer	μ×	SOL Query3 sol - localhost\\tom (59))	= ×	Eigenschaften	* † X
rbinden • 💷 💷 🐨 🕵		USE [AIC13DEMO]	î	Aktuelle Verbindun	arnarameter -
In the standard sector (COL Constraints of COL) with DOLCUDA EDGANA		GO	<u>^</u>	Actuene verbindun	gsparameter -
Distanbankan	om		E	2.	
Jatenbanken Sustemdatenbanken		☐ IF EXISTS (SELECT * FROM dbo.sysobjec	S WHERE	Aggregatstatus	
		BEGIN		Name	localhost\sqlexpress
Datenbankdiagramme		ALIER TABLE [dbo].[Customers] DROP CON	SIRAINI	Startzeit	22.01.2014 20:43:19
🖃 🧰 Tabellen				Status	Geöffnet
🛞 🧰 Systemtabellen		GO		Verbindungsfehle	
dbo.Customers				Verstrichene Zeit	00:00:00.0520000
🗑 🛄 Spalten		☐ IF EXISTS (SELECT * FROM dbo.sysobjec	S WHERE	Zeitpunkt der Fert	22.01.2014 20:43:19
III 🚞 Schlüssel		BEGIN		Zurückgegebene	0
🗉 🚞 Einschränkungen		ALTER TABLE [dbo].[Customers] DROP CON	STRAINT	Verbindung	
🗉 🚞 Trigger		- END		Verbindungsname	localhost\sqlexpress
🗄 🚞 Indizes				Verbindungsdeta	ails
🗉 🚞 Statistik		GO		Ablaufverfolgung	
🗉 💷 dbo.StatusTweet				Anmeldename	WIN-B9JSU3N5D9N\1
dbo.TwitterDataset		E IF EXISIS (SELECI * FROM db0.3930D]ec	CS WHERE	Anzeigename	localhost\sglexpress
🗉 🚞 Sichten		ALTER TABLE (dbo) (Customers) DROP CON	STRAINT	Servername	localhost\sglexpress
🕞 🚞 Synonyme		- END		Serverversion	10.0.2531
🐵 🚞 Programmierbarkeit				SPID	57
🛞 🚞 Service Broker	=	GO		Verbindungsstatu	Geöffnet
🛞 🚞 Speicher			-	Verstrichene Verb	00.00.00.0520000
🗉 🧰 Sicherheit		< III.	F	Zeitnunkt der Ver	22.01.2014.20-43-10
🗉 🔰 aic13demo		Ta Meldungen		Zeitpunkt der Ver	22.01.2014 20:43:10
🗉 🧰 Datenbankdiagramme		Befehl(e) wurde(n) erfolgreich abgeschlossen.		Zurückgegebene	0
🖃 🦲 Tabellen				Zuruckgegebene	0
Systemtabellen					
dbo.Customers					
dbo.Status I weet					
dbo. I witterDataset					
i ichten					
Synonyme					
Septice Programmerbarket					
Speicher					
Greeneit					
Sicherheit					
Anmeldungen					
##MS PolicyEventProcessingl onin##					
A ##MS PolicyTealExecutionLogin##					
A AIC13					
M NT SERVICE\MSSOLSSOLEXPRESS					
A NT-AUTORITÄT\SYSTEM					
a sa					
VORDEFINIERT\Benutzer				Name	
WIN-B9JSU3N5D9N\tom	-	•		Der Name der Verbir	ndung.
	•	ess (10.0 WIN-B9JSU3N5D9N\tom (57) aic13demo 00:00:	0 Zeilen		

- 5. Finally the Twitter dataset (corpus) from <u>http://cs.stanford.edu/people/alecmgo/trainingandtestdata.zip</u> must be imported into the table "TwitterDataset"
- 6. (Optional) Database backup: The deployment folder contains a complete dump of the sentiment database already containing the training data (corpus), so the database can be restored using this dump without having to create it from scratch using the scripts. (File: aic13_database_dump.zip)

Webserver

1. Copy the folder "SentimentAnalyis" containing the webservice and webfrontend to your webserver

📔 « Projects	▶ Bitbucket_Repos ▶ aic13 ▶ Deployment	▶ 20130122 ▶ Sentimen	tAnalysis 🕨	
en 🔻 🛛 In Bib	liothek aufnehmen 🔻 🛛 Freigeben für 🔻	Brennen Neuer Ord	ner	
iten	Name	Änderungsdatum	Тур	Größe
ktop	鷆 bin	22.01.2014 21:05	Dateiordner	
vnloads	🗈 packages.config	12.11.2013 20:54	XML Configuratio	1 KB
tzt besucht	SentimentAnalysis.aspx	22.01.2014 19:31	ASP.NET Server Pa	5 KB
	脑 SentimentAnalysisWS.asmx	20.11.2013 23:17	ASP.NET Web Ser	1 KB
heken	📄 settings.xml	24.11.2013 11:07	XML-Dokument	1 KB
er	🗈 Web.config	22.01.2014 19:36	XML Configuratio	2 KB
umente				

2. Next, open the IIS Management Console, click on "Default website" and map the webapplication by creating a new "Web application"

→ (WIN-B9JSU3N5D9N → Sites → Default Web Site	: •	
Datei Ansicht ?		
/erbindungen	Default Web Site Startseite	Aktionen
MIN-B9JSU3N5D9N (WIN-B9JSU3N5D9N\tom) Anwendungspools	Filter: • 🔐 Start - 🖓 Alle anzeigen Gruppieren nach:	Berechtige Site beart
 Sites Default Web Site 		Bindunger
🕞 💮 SentimentAnalysis	.NET-AutorisierunNET-Fehlerseiten .NET-Globalisierung .NET-Kompilierung	Anwendur Virtuelle V anzeigen
		Website verw
	Aver-ventaderise Anwendungseinst Computerschutssei	 Neu starte Starten
	Saitap und Sitzungszustand SMTD-E-Mail Verbindungszeich	Beenden Website o
	Seven und Sitzungszustahd SMTP-E-Mail Verbindungszeich Steuerelemente	• *:80 (http)

3. Specify an alias where the webapplication should be accessed later on

Anwendung hinzufügen		8 23
Sitename: Default Web Site Pfad: /		
Alias:	Anwendungspool:	
aic13demo	DefaultAppPool	Auswählen
Beispiel: Vertrieb		
Physikalischer Pfad:		
C:\inetpub\wwwroot\SentimentAna	alysisDemo	
Pass-Through-Authentifizierung		
Verbinden als Einstellunge	n testen	
	ОК	Abbrechen

- 4. Ensure that the Framework is set to .NET 4.0
- 5. Open the settings.xml file and specify your credentials for accessing the Twitter API and your SQL server

📔 C:\P	rojects\Deployment\201401261256\SentimentAnalysisService\settings.xml - Notepad++
Datei	Bearbeiten Suchen Ansicht Kodierung Sprachen Einstellungen Makro Ausführer
6) 🗄 🖻 🕞 🕼 🎝 🖌 🛍 🜔 🔵 🗲 🏙 🍢 🍳 👒 🖫 🔂 🎫 1 🏢
📄 settir	igs xml 🔀
1	xml version="1.0" encoding="utf-8"?
2	<pre>□<configuration></configuration></pre>
3	<pre>twitter></pre>
4	<accesstoken></accesstoken>
5	<accesstokensecret></accesstokensecret>
6	<consumerkey></consumerkey>
7	<consumersecret></consumersecret>
8	<accountname></accountname>
9	<pre><password></password></pre>
10	-
11	<pre>database></pre>
12	<dbserver>LOCALHOST\SQLEXPRESS</dbserver>
13	<dbname>AIC13</dbname>
14	<user>AIC13</user>
15	<pre><password></password></pre>
16	-
17	L

6. Then the webservices can be accessed using the URL http://localhost/aic13demo/SentimentAnalysisWS.asmx

a balant has in all land at half at	
🧧 tmikscha / aic13-sc 🗙 🗋 SentimentAnalysis' 🗙 🈏 (2) Twitter 🛛 🙁 🗋 Real-time HTML Ec 🗙 📘	
← → C [] localhost/SentimentAnalysis/SentimentAnalysisWS.asmx	☆ 〓
🗰 Apps 🕒 Development Tasks: 🐰 Google Accounts	
SentimentAnalysisWS	
Folgende Vorgänge werden unterstützt. Eine ausführliche Definition finden Sie in der Dienstbeschreibung.	
ClassifyTweet	
<u>CreateSentimentAnalysiskeport</u>	
<u>DoPreProcessing</u>	
DoSentimentAnalysis	
<u>GetSentimentReport</u>	
GrabTweets	
DreDrocescTweet	
<u>RegisterSentimentSearch</u>	

7. The GUI for registering new keywords for sentiment analysis should be accessible using the URL http://localhost/aic13demo/SentimentAnalysis.aspx

localhost/aic13dem × yaic	Btmik (aicl3tmi ×) 📴 tmikscha / aicl3-sei ×) Neuer Tab 🛛 ×	
← → C □ localhost/aic1	.3demo/SentimentAnalysis.aspx	☆ =
🔛 Apps 🕒 Development Tasks:	8 Google Accounts	
Company ID:		
Company ID.		
Company Name:		
E-Mail:		
Keyword:		
Algorithm:	Simple Classifier (Bayes) 🔻	
Limit:	100 •	
	Register Reset Report	

Windows Service

1. Copy the folder "SentimentAnalyisService" to the desired target directory

Organisieren 🔻 🛛 In Bibl	iothek aufnehmen 🔻 🛛 Freigeben für 🔻	Brennen Neuer Ord	Iner	· · · · · · · · · · · · · · · · · · ·
🔆 Favoriten	Name	Änderungsdatum	Тур	Größe
🥅 Desktop	SentimentAnalysisService.exe	22.01.2014 19:57	Anwendung	16 KB
\rm Downloads	LinqToTwitter.dll	10.11.2013 21:35	Anwendungserwe	534 KB
📃 Zuletzt besucht	🚳 log4net.dll	10.11.2013 21:35	Anwendungserwe	292 KB
	Newtonsoft.Json.dll	10.11.2013 21:35	Anwendungserwe	446 KB
对 Bibliotheken	StemmersNet.dll	10.11 <mark>.</mark> 2013 21:35	Anwendungserwe	107 KB
📔 Bilder	🚳 SVM.dll	22.01.2014 19:57	Anwendungserwe	65 KB
Dokumente	🚳 Tools.dll	22.01.2014 19:57	Anwendungserwe	44 KB
🚮 Git	TweetClassifier.dll	22.01.2014 19:57	Anwendungserwe	15 KB
🁌 Musik	S TweetGrabber.dll	22.01.2014 19:57	Anwendungserwe	9 KB
🛃 Videos	TweetReportingService.dll	22.01.2014 19:57	Anwendungserwe	6 KB
	TweetTokenizer.dll	22.01,2014 19:57	Anwendungserwe	21 KB
🍣 Heimnetzgruppe	InstallUtil.InstallLog	24.11.2013 11:30	INSTALLLOG-Datei	7 KB
	SentimentAnalysisService.InstallLog	24,11,2013 11:30	INSTALLLOG-Datei	13 KB
🖳 Computer	SentimentAnalysisService.InstallState	24.11,2013 11:30	INSTALLSTATE-D	8 KB
🕌 Lokaler Datenträger	📋 params.txt	19.01. <mark>2014 18:24</mark>	Textdokument	3 KB
	📋 results.txt	19.01.2014 19:20	Textdokument	1 KB
📭 Netzwerk	results1.txt	19.01.2014 18:47	Textdokument	1 KB
	SentimentAnalysis.log	29.11.2013 07:28	Textdokument	2.847 KB
	📋 stopwords.txt	19.01.2014 15:40	Textdokument	1 KB
	📋 svmTest.txt	19.01.2014 19:20	Textdokument	1 KB
	svmTrainingSet.txt	19.01 <mark>.</mark> 2014 19:19	Textdokument	327 KB
	📋 tokenList.txt	19.01.2014 19:19	Textdokument	261 KB
	TweetGrabberConsole.log	24.11.2013 10:45	Textdokument	6 KB
	RegisterService.bat	24,11,2013 10:24	Windows-Batchda	1 KB
	SentimentAnalysisService.exe.config	24.11.2013 10:49	XML Configuratio	1 KB
	📄 log4net.xml	10.11.2013 21:35	XML-Dokument	1.482 KB
	Newtonsoft.Json.xml	10.11.2013 21:35	XML-Dokument	459 KB
	settings.xml	14.11.2013 22:10	XML-Dokument	1 KB
	P Report.xslt	22.01.2014 20:00	XSLT Stylesheet	9 KB

- 2. Again, add your credentials to the settings.xml which can be copied from the web application
- 3. In order to run the sentiment analysis process, the service must be registered. This can be either done by executing the RegisterService.bat file located in the service directory or by running the following commands from an administrator console

```
%windir%/Microsoft.NET/Framework/v4.0.30319/InstallUtil /u
SentimentAnalysisService.exe
%windir%/Microsoft.NET/Framework/v4.0.30319/InstallUtil
SentimentAnalysisService.exe
```

4. Then you should see a new entry in your service console

LOKal) Dienste (Lokal)	Eigenschaften von Sentin	eantApplysis (Lakalas Computer)
	Name Resch	nerovnalysis (cokaler computer)
SentimentAnalysis	Allgemein Anmelden V	Wiederherstellung Abhängigkeiten
Den Dienst starten	Parental Controls Dieser	
	Peer Name Resolu Aktivi Dienstname: Senti	imentAnalysis
	Anzeigename: Senti	imentAnalysis
	Peernetzwerkident Bietet	
	Plug & Play Ermog Beschreibung:	- -
	DNDD Committee Disease	
	Proceeding	D
	Programmkompa Dieser C.\Projects\Bitbucket_	_Repos \aic 13 \SentimentAnalysis Service \bin \Debug
	RAS-Verbindungs Verwa Starttyp: Man	nuell
	Remotedesktopal Ermot	,
	Remoteprozedura Der Kr <u>Unterstützung beim Kon</u>	figurieren der Startoptionen für Dienste
	Richtlinio zum Ent Läczter Disasteleten Base	
	Richtlinie zum Ent Lasst Dienststatus: Been Pouting und PAS Pietet	Idet
	PDC Endpunktrup Löct P Starten	Beenden Anhalten Fortsetzen
	Sekundäre Appel Attivity wenn der Dienst von hie	er aus gestartet wird.
	Sentiment Analysis	-
	Server Unter	
	Server für Threads Bietet	
	Shellbardwareerke Zeigt	
		OK Abbrehen Obernehme
	<	OK Abbiechen Obernening

TweetGrabberConsole

- For testing and debugging or simply running the collection of tweets manually, the TweetGrabberConsole can be used. The console offers better capabilities to monitor the whole sentiment analysis process compared to the windows service
- 2. Copy the folder TweetGrabberConsole to the desired target folder

In Bib	liothek aufnehmen ▼ Freigeben für ▼ B	rennen Neuer Ord	ner	3== ▼ □	0
Favoriten	Name	Änderungsdatum	Тур	Größe	
🧮 Desktop	TweetTokenizer.dll	22.01.2014 19:57	Anwendungserwe	21 KB	- 1
〕 Downloads	TweetReportingService.dll	22.01.2014 19:57	Anwendungserwe	6 KB	- 1
📃 Zuletzt besucht	TweetGrabberConsole.vshost.exe.manifest	17.03.2010 22:39	MANIFEST-Datei	1 KB	- 1
	TweetGrabberConsole.vshost.exe.config	21.01.2014 21:15	XML Configuratio	2 KB	- 1
🗃 Bibliotheken	TweetGrabberConsole.vshost.exe	22.01.2014 20:03	Anwendung	12 KB	- 1
📔 Bilder	TweetGrabberConsole.log	22.01.2014 20:03	Textdokument	10.282 KB	
Dokumente	TweetGrabberConsole.exe.config	21.01.2014 21:15	XML Configuratio	2 KB	- 1
🚮 Git	TweetGrabberConsole.exe	22.01.2014 19:57	Anwendung	12 KB	
🎝 Musik	🚳 TweetGrabber.dll	22.01.2014 19:57	Anwendungserwe	9 KB	
😸 Videos	TweetClassifier.dll	22.01.2014 19:57	Anwendungserwe	15 KB	- 1
	🚳 Tools.dll	22.01.2014 19:57	Anwendungserwe	44 KB	
👌 Heimnetzgruppe	📄 tokenList.txt	22.01.2014 20:00	Textdokument	100 KB	
	🚳 test.csv	19.11.2013 19:57	Microsoft Office E	105.487 KB	
Computer	svmTrainingSet.txt	22.01.2014 20:00	Textdokument	103 KB	1
🏭 Lokaler Datenträger	📄 svmTest.txt	19.01.2014 19:20	Textdokument	1 KB	- 1
	SVM.dll	22.01.2014 19:57	Anwendungserwe	65 KB	- 1
📮 Netzwerk	i stopwords.txt	19.01.2014 15:40	Textdokument	1 KB	- 1
	🚳 StemmersNet.dll	10.11.2013 21:35	Anwendungserwe	107 KB	- 1
	settings.xml.orig	12.11.2013 21:00	ORIG-Datei	1 KB	- 1
	📄 settings.xml	12.11.2013 21:00	XML-Dokument	1 KB	- 1
	results1.txt	19.01.2014 18:47	Textdokument	1 KB	- 1
	results.txt	19.01.2014 19:20	Textdokument	1 KB	- 1
	Report.xslt	22.01.2014 19:55	XSLT Stylesheet	9 KB	
	params.bt	19.01.2014 18:24	Textdokument	3 KB	
	🔮 Output.xml	11.11.2013 22:14	XML-Dokument	737 KB	
	🔮 Newtonsoft.Json.xml	10.11.2013 21:35	XML-Dokument	459 KB	
	Newtonsoft.Json.dll	10.11.2013 21:35	Anwendungserwe	446 KB	
	🔮 log4net.xml	20.01.2014 22:25	XML-Dokument	1.482 KB	
	🚳 log4net.dll	20.01.2014 22:25	Anwendungserwe	294 KB	
	LinqToTwitterCP.dll	02.11.2013 14:23	Anwendungserwe	531 KB	

3. Modify the settings.xml by adding your credentials for the SQL server and the Twitter API

Usage

Overview of Architecture

The sentiment analysis process is basically split into three modules:

- A common database server where all collected tweets and jobs (e.g. keywords to analyze) are stored
- A webapplication that is used for registering new keywords and accessing the reports after the classification (=user frontend for interaction)
- TweetGrabberConsole/Windows Service for running the classification process



Generally speaking whole classification process consists of the following steps/phases

- Register a new search term or keyword for analysis (Webapplication)
- Collect, tokenize, classify and analyze the tweets (TweetGrabberConsole or Windows Service)
- Reporting using XSLT (Webapplication)

Register	Collect	Tokenize	Classify	Analyze	Report
 Prospective customer registers at ASPX webpage Website calls SOAP- based webservice for storing registration data Webservice adds records to database on MS SQL server 	TweetGrabber checks for new requests Tweets are collected using Twitter Streaming API Live capture of tweets for a certain period of time (number of tweets, limit specified at registration time) Collected tweets are stored in database	 TweetTokenizer performs stop word removal using a Snowballing algorithm Porter2 stemming algorithm will be applied Processed tweets are stored in database again 	 TweetClassifier processed the tokenized tweets by applying a classification algorithm Naive Bayes classifier is used Classification is stored in database 	TweetReportingService creates a report out of collected and calculated data Data exported as XML and transformed by XSLT into a HTML document HTML report is stored in database	 Customer can access the result of the sentiment analysis as a report (HTML) by calling a method provided by the SOAP- based webservice Webservice is called by ASPX-webpage which acts as the frontend for the customer

ASPX/SOAP WS Webserver	C# Windows Service Application Server	ASPX/SOAP WS Webserver
------------------------------	--	------------------------------

Usage of Components

1. Register a new search term that should be analyzed in the Twitter universe using the webapplication / - frontend e.g. Justin Bieber

localhost/SentimentAnaly ×	acl3tmik (aicl3tmik) auf × g tmikscha / aicl3-sentimer ×	
		= 2
E Development rasks		
Company ID:		
Company Name:	Justin Bieber	
E-Mail:	test@test.com	
Keyword:	Justin Bieber	
Algorithm:	Simple Classifier (Bayes) ▼	
Limit:	100 •	
	Register Reset Report	

2. After pressing the Register button, the search term will be added to the database for further processing by either the TweetGrabberConsole or the SentimentAnalysis Windows service

Kicrosoft SQL Server Management Studio													
Datei Bearbeiten Anzeigen Abfrage	Debu	iggen E	xtras Fenster	Community Hil	fe								
😫 🛄 Neue Abfrage 📑 📑 📑 📑	<u>,</u>	÷											
🖳 📴 master 👻 💡	Ausfü	ihren 🕨	- 🗸 🎲 🖻	- 🖃 🎌 🖷 🍳	a 🖾 🖏 🗉	g		Åå ₌					
Objekt-Explorer 🚽 🕂 🗙	S		4.sql - localhost	\tom (56)) Crea	teTables.sgl -	localho	st\\tor	n))*	₹ :	×	Eigenschaften	→ ‡ ×	
Verbinden - 💷 💷 🝸 🛒		/****	** Skript f	ür SelectTopN	IRows-Befe	hl au	s SSM	5 ****	**/		Aktuelle Verbind	ungsparameter 🔹	
	E	SELEC	T TOP 1000	[id]						^		angsparameter	
Iocalnost\sqlexpress (SQL Server 10.0.2			,[company_	name]							2 I 🖾		
🖃 🦲 Datenbanken			,[email_ad	dress]						=	Aggregatstatus		
Jystemdatenbanken			,[keyword]								Name	localhost\sqlexpres	
AlCIS AlCIS Datenbankdiagramme			,[status]								Startzeit	22.01.2014 21:47:32	
Taballan			,[stage]								Status	Geöffnet	
			, [CIASSIII	erj							Verbindungsfe	ł	
dho Customers			<pre>,[IICMITEDO []imit]</pre>	10]							Verstrichene Ze	00:00:00.0920092	
Spalten		FRC	M [AIC13].[dbol.[Custome	rsl						Zeitpunkt der F	22.01.2014 21:47:32	
E Constant										*	Zurückgegeber	4	
🗉 🧫 Einschränkungen	•								•		Verbindung		
🐨 🛄 Trigger		Ergebnis	se 🚮 Meldunge	n							Verbindungsna	localhost\sglexpres	
🗉 🥅 Indizes		id	company_name	email_address	keyword	status	stage	classifier	htmlreport		Verbindungsd	etails	
🗉 🧰 Statistik	1	136	justin bieber	asd	Justin Bieber	2	3	2	NULL		Ablaufverfolgu		
dbo.StatusTweet	2	137	Lady Gaga	weird@office.com	Lady Gaga	2	4	3	0xEFBBBF3C68746D6C2	07	Anmeldename	WIN-B9ISU3N5D9N	
😠 🔲 dbo.TwitterDataset	3	138	Justin Bieber	Justin Bieber	Justin Bieber	2	4	3	0xEFBBBF3C68746D6C2	07	Anzeigename	localbost\sqlexpres	
🕀 🚞 Sichten	4	139	Justin Bieber	test@test.com	Justin Bieber	1	1	3	NULL		Servername	localhost\sglexpres	
🕀 🚞 Synonyme											Serverversion	10.0.2531	
											Serversion	101012331	

- 3. Running the TweetGrabberConsole will start the classification process. The tool will query the database looking for open tasks and running the
 - a. Collection of tweets
 - b. Tokenization (stop word removal, stemming etc.)
 - c. Classification using the selected algorithm (Simple Bayesian classifier, Naïve Bayes or SVM)
 - d. Evaluation (Reporting/XSLT)

C:\Projects\Bitbucket_Repos\aic13\TweetGrabberConsole\bin\Debug\TweetGrabberConsole.exe	
#33 - ?Spread Your Wings? "Swaggy_xz" tweeted at 22.01.2014 21:03:42: RT @JBSour ce: People seem to forget that Justin Bieber is a human being just like the rest	*
of us. #33 - Follow Jasmine, ily "VILLEGASPORRA" tweeted at 22.01.2014 21:03:42: Foto: lovemisterbieber: Photo - Last chance to see Justin Bieber's tour here http://t. co/x0dUGOsHVD http://t.co/Zm8uQD2xiu	
H35 - Lujain "curlyromeox" tweeted at 22.01.2014 21:03:42: RT @birdcagewifi: "T he Wanted split"	
"Justin Bieber retiring"	Ξ
are you scared bc i am	
#36 - clara "thoughtdrew" tweeted at 22.01.2014 21:03:43: RT @selfiejustins: "Yo u're too obsessed with Justin Bieber" http://t.co/lfeaeA3gw3	
#37 - Djulayn "Djulscorputty" tweeted at 22.01.2014 21:03:43: Justin Bieber - H1 1 that matters ?? #38 - FOLLOW ME JUSTIN ? "kidrauhlninjax" tweeted at 22.01.2014 21:03:43: RT @se	
lfiejustins: "You're too obsessed with Justin Bieber" http://t.co/LfeaeA3gw3 #39 - Sam Everson "samevo2" tweeted at 22.01.2014 21:03:43: My psych teacher sai	
d justin Bieber is a horrible person because he spent 75 grand at a strip club . That's not a reason to hate lol #40 - Belieue, "Sonia MadiDamon" tweeted at 22 01 2014 21:03:43: "Divstinsnizza:	
dylan obrien, justin bieber & tyler posey aka my life" Same	-

C:\Projects\Bitbucket_Repos\aic13\TweetGrabberConsole\bin\Debug\TweetGrabberConsole.exe
Nipple . Nipple . #97 - paul walker "mymjnd" tweeted at 22.01.2014 21:04:07: RT @selfiejustins: "Y ou're too obsessed with Justin Bieber" http://t.co/LfeaeA3gw3 #98 - Follow Jasmine, ily "UILLEGASPORRA" tweeted at 22.01.2014 21:04:09: Foto: lovemisterbieber: Photo - Last chance to see Justin Bieber's tour here http://t. co/x0qUGOSHWD http://t.co/0cgOOKNcFp #98 - Cristiano Ronaldo. ♥ "Gnovas_PaulaCR7" tweeted at 22.01.2014 21:04:09: RT @InJustinWorld: Every Belieber wants to hear these 5 words, 8 syllables, 25 lett ers, one dream come true: "You are meeting Justin Bieber. #100 - 8 Jaque Beatriz ? "JBeatriz_s2" tweeted at 22.01.2014 21:04:10: Hey, you' re the best idol knew? I love you very much @justinbieber ♥ quarenta e quatro' #ArtistaDoSéculo Justin Bieber #101 - Alyssia Rodriguez "4_14_13_baby" tweeted at 22.01.2014 21:04:10: RT @robb ona3: Justin Bieber got 100,000 retweets for tweeting "Live life full". That's j ust 3 random words. I'm going to try now.
Nipple .
Step 2: Tokenization finished.
Step 3: Classification finished.
Step 4: Report generation finished.

4. Accessing the final report using the webfrontend again

localhost/Se	entim	entAnaly ×	🈏 aic13tmik	(aic13tmik) au	f 🗙 📴 tm	nikscha / aic13-sentim	ner ×						23		
← → C'		ocalhost/S	entimentAna	alysis/Senti	mentAnaly	sis.aspx						5	Ξ		
🔛 Apps 🕒 De	evelo	pment Tasks:.	<u>8</u> Google A	Accounts											
6		Freedl	Kaumand		Senti	ment		Sent	iment ((Weig	ghted)				
Company		cmail	Keyword	Pos	tive	Negative		Positive			Negative				
Justin Bieber	test	t@test.com	Justin Bieber	65.3465346	5346535%	34.6534653465346	65%	66.144408077122	59%	33.85	55591922877444%				
					Mainhain -				Terre						
User name	ð	Followers	Weig	hting	weighting %	Classification	A	Avg Followers	(Orig	t i.)	Text (Processed)				
		13041	7.852605	0329988735	785.26%	1.0000	166	0.7227722772277	1.000	io F	RT weheartjustin justin bieber is like a book he judg by the cover but nobodi read his stori i'v never seen a statement more t				
Lujain		10371	6.24487	1313337268	624.49%	0.0000	166	0.7227722772277	0.000	io F	RT birdcagewifi the want split justin bieber retir we'r slowli take over twitter are you scare bc i am				
?		10198	6.140699	8026625645 614.07		6.1406998026625645		1.0000	166	0.7227722772277	1.000		ariana is my queen life she is everyth to me she isn't famous like justin bieber or 1D but she is grow bigger and bigger		
maria		9823	5.91489	4504957283 591.49%		5.914894504957283		0.0000	166	0.7227722772277	0.000	0 3	you'r all that matter to me artistadoséculo justin bieber		
?follow me justin?		8061	4.85391	4.853910679472733		1.0000	1.0000 1660		1.000	io F	RT bieberloft i'm justin bieber and i'll make u question your sexual http://t.co/8aagrky0uu				
News.Stream		6971	4.19756	9947476048	419.76%	1.0000	166	0.7227722772277	1.000	o j	justin bieber smoke a control http://t.co/vh	a cuban cigar as singer camp say his ego Is out of invov7pav			
		5555	3.34492	9143340905	334.49%	0.0000	166	0.7227722772277	0.000		RT alexisbroci I just now omg I love you	met justin drew bieber I cant contain MY feel right with all MY heart justin http://t			

5. Another approach is to access all components using the webservice methods available under http://localhost/SentimentAnalysis/SentimentAnalysisWS.asmx



- Helper-Method: ClassifyTweet
 - Takes a string containing a preprocessed text and and integer for the desired classifier e.g. 1=NaïveBayes, 2=SVM and 3=SimpleClassifier
- Helper-Method: PreProcessTweet
 - Takes a string containing a tweet and the language of the tweet ("EN", "DE", etc.) and applied the pre-processing (stemming, stopword-removal, etc.)
- Helper-Method: GetSentimentReport
 - Returns the previously created HTML report containing the sentiment analysis for the specified customer/job id
- Step 1: RegisterSentimentSearch
 - Allows the registration of a new search term for sentiment analysis
 - Adds the "keyword search" as a job to the database for further processing
 - Returns the ID of the search job
- Step 2: GrabTweets
 - Starts the collection of tweets for the specified ID
 - Uses the Twitter streaming API
 - Adds the tweets to the database and returns the number of collected tweets
- Step 3: DoPreProcessing
 - Applies the stemming, stopword-removal, etc. on all collected tweets for the specified customer/job id
 - Returns an HTML message containing all processed tweets
- Step 4: DoSentimentAnalysis
 - Starts the classification for the specified customer/job id
 - Return a HTML string containing all tweets together with their classification
- Step 5: CreateSentimentAnalysisReport
 - Creates a HTML report of the collected tweets containing the results of the classification
 - XML data will be transformed to HTML using XSLT
 - HTML file is stored in database and returned to the client