

Chrillys Goldpreis-Report Mai 2018

von

Dr. Chrilly Donniger, Chef für Alles
Anni Donniger, Illustration&Lektorat.



Inhaltsverzeichnis:

1) Highland Analytica oder der Twitterbot Vanellus.

Der Chef für Alles hat einen Twitter-Bot gebaut und berichtet in diesem Beitrag über seine Erfahrungen und Resultate in diesem Metier.

2) König Zuckerbergs Glück ohne Ende.

Der Facebook Skandal als Sturm im Medien-Wasserglas.

3) Guantanamera, Spanisch-Deutsch:

Nach Meinung der Redaktion eines der schönsten Heimatlieder. Wie üblich geschrieben von einem Dichter im Exil.

4) Hört, sieht und riecht die Signale:

Ausgehend vom Ausfall des diesjährigen Literatur-Nobelpreises erklärt der Chief der Seniors, wie man mit einfachen Mitteln sich – ohne viel Ahnung zu haben – in der Welt zurecht findet. Man muss sich nur ein paar passende Signale zulegen. Die Seniors bezweifelt wie immer, dass die Welt so einfach ist.



Free trade can be wonderful if you have the power of nuclear weapons.

(@Deepdrumpf)

Highland Analytica oder der Twitter-Bot

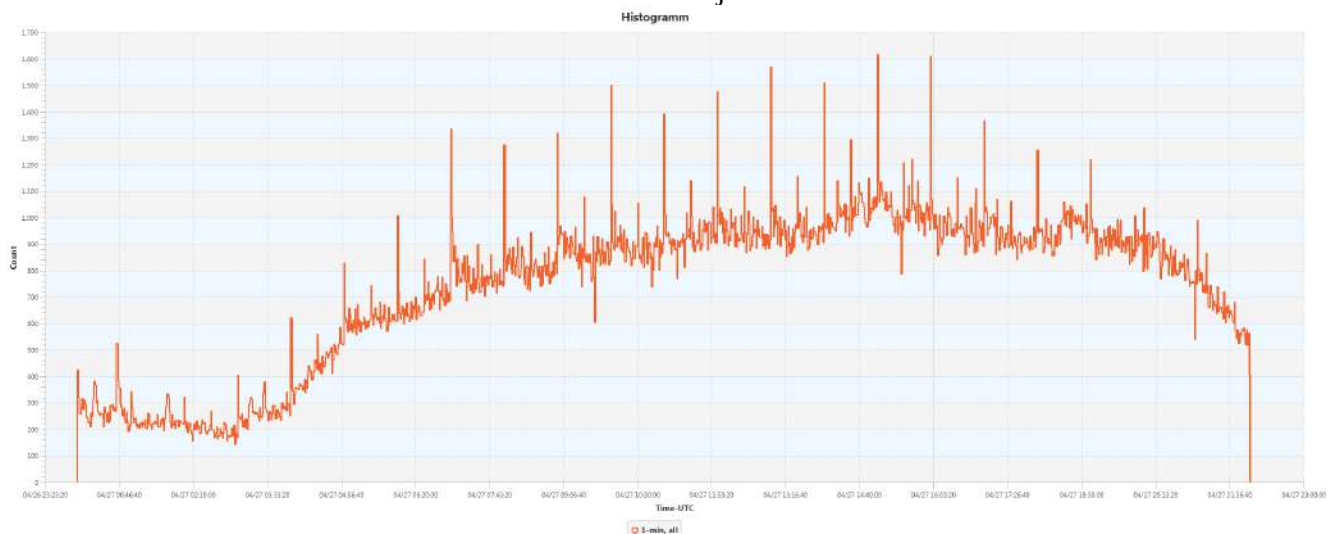
Vanellus:



Die Berichte über Cambridge Analytica haben mich inspiriert, selbst einen Twitter-Bot namens **Vanellus** vulgo **der Kiebitz** zu schreiben. Er hält bisher aber entgegen seinen Namen bis auf Testtweets das Maul. Vanellus sammelt alle deutschsprachigen Tweets (rund 1 Million/Tag) und analysiert diese nach einer Reihe von Kriterien.

Um einen Bot zu schreiben, muss man sich zunächst als Benutzer bei Twitter anmelden. In einem zweiten Schritt füllt man ein Formular für Bots aus und erhält von Twitter Zugangs-Schlüssel für das Application Programming Interface (API). Mit diesem API kann man Tweets senden und aus der Twitter-Datenbank Tweets herunter laden. Tatsächlich schlägt man sich nicht direkt mit den Feinheiten dieses APIs herum. Alle wichtigen Programmiersprachen haben eigene Bibliotheken, die den Umgang damit sehr erleichtern. Ich verwende für Vanellus die Programmiersprache Java, weil meine Trading Software *CashBot* in Java läuft und ich einige Teile von CashBot direkt übernehmen konnte. Java gehört zu den wichtigsten Programmiersprachen und hat mit twitter4J eine ausgereifte Twitter-Library. Ich bin nicht der Erste und Einzige, der auf diese Idee gekommen ist.

Vanellus hat sich nach kurzer Zeit bereits geklont. Es gibt inzwischen A- und B-Vanella, sowie C- und D-Vanellus. Die weibliche Form ist auf die Benutzerin Amalia_vEnz angemeldet. Vermutlich hätte Twitter keine Einwände gehabt, wenn ich die Viererbande nur auf meinem Account registriert hätte. Jeder zusätzliche Benutzer, jeder zusätzliche Tweet, sind ein Plus in der Twitter-Bilanz. Ob in einem Rechenzentrum ein paar Rechner mehr oder weniger stehen, ist auch schon wurscht. Das Klonen war notwendig, weil mit dem freien Zugang die Datenrate auf 12 Abfragen pro Minute (mit jeweils maximal 100 Tweets) beschränkt ist. Wenn man diese Grenze verletzt bekommt man eine Zeitsperre von 15 Minuten. Wenn man etwas Geld hinlegt, kann man größere Mengen herunter saugen. Mit einem freien Zugang kann man nicht – wie geplant - alle deutschsprachigen Tweets erfassen. Es teilen sich nun 3 Bots die Arbeit auf (der 4. dient zu Testzwecken). Sie fragen Tag und Nacht im Abstand von 2,2 Sekunden die Twitter-Datenbank nach dem Muster „lang:de“ ab. Der Twitter-Server schickt alle Tweets zurück die seit der letzten Abfrage vor 2,2 Sekunden angefallen sind. Twitter kann diesen einfachen Trick leicht erkennen. Bisher hatte ich jedoch keine Probleme.



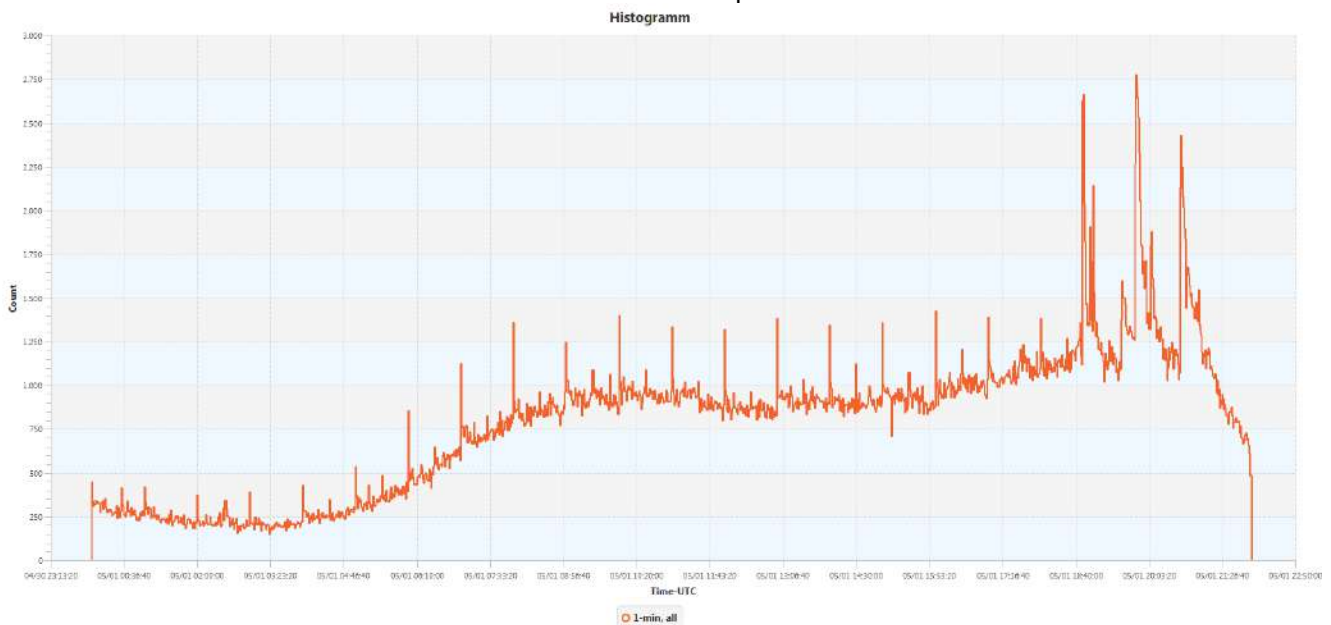
Grafik 1: Histogramm, Tweets/min am Freitag 27. April 2018

Grafik 1 zeigt den typischen täglichen Verlauf der Tweets/min. Die regelmäßigen Spitzen finden zu jeder vollen Stunde statt. Das ist die Aktivität von Nachrichtenportalen, von Bots die Turmuhr spielen und vor allem von Wetterstationen, die zu jeder vollen Stunde ihre Messdaten durchgeben. Teilweise sind es auch die Reaktionen von Usern auf diese Tweets. Der Twitter-Server liefert zwischen 0:00 und 2:00 MEST keine Daten. Twitter nützt wahrscheinlich after midnight zum Server-Service.

Am 1. Mai gab es am Abend heftige Ausschläge. Das sind die Tore im Champions-League-Schlager Real-Madrid gegen die Bayern. Man kann durch die Anzahl der eintrudelnden Tweets bereits feststellen, dass irgendwas Wichtiges passiert ist. Man weiß allerdings nicht, wer die Tore geschossen hat. Ich verwende diese Methode auch bei meiner Trading-Software. Es läuft rund um die Uhr der S&P-500-Aktienindex mit. Wenn er plötzlich nach unten sackt, ist auf der Welt was Größeres passiert. Ich schaue anschließend am Netz nach, was es ist. Meist tauchen die Meldungen aber erst nach einiger Zeit auf. Die großen Trader sitzen näher an der Quelle und reagieren vor der Herde.

Ähnliche Effekte gab es beim Pokal-Schlager zwischen Eintracht-Frankfurt und den Bayern oder dem CL-Finale zwischen Real und Liverpool.

Anmerkung: Twitter vergibt jeden Tweet einen Zeitstempel. Dieser ist in UCT (Greenwich). Vanellus verwendet ebenfalls UCT und rechnet im Moment nicht um. Die Mitteleuropäische Sommerzeit ist UCT+2h.



Grafik 2: Histogramm, Tweets/min am Dienstag 1. Mai 2018

Als Fan von FC Erzgebirge Aue hat mich das Abstiegsderby gegen Darmstadt98 mehr interessiert. Man kann durch „hineinzoomen“ auch den Verlauf dieses Spieles und die Folgen sehr schön verfolgen. Vanellus hat zu diesem Zweck eine flexible Abfragesprache implementiert.

```
histogramm{ query: "<i>#D98Aue <or> <i>#Erzgebirge Aue <or> <i>#AUE <or> <w>VAR", stepSz: 1, since: "2018-05-13", until: "2018-05-14 14:00"}; Go < >
```

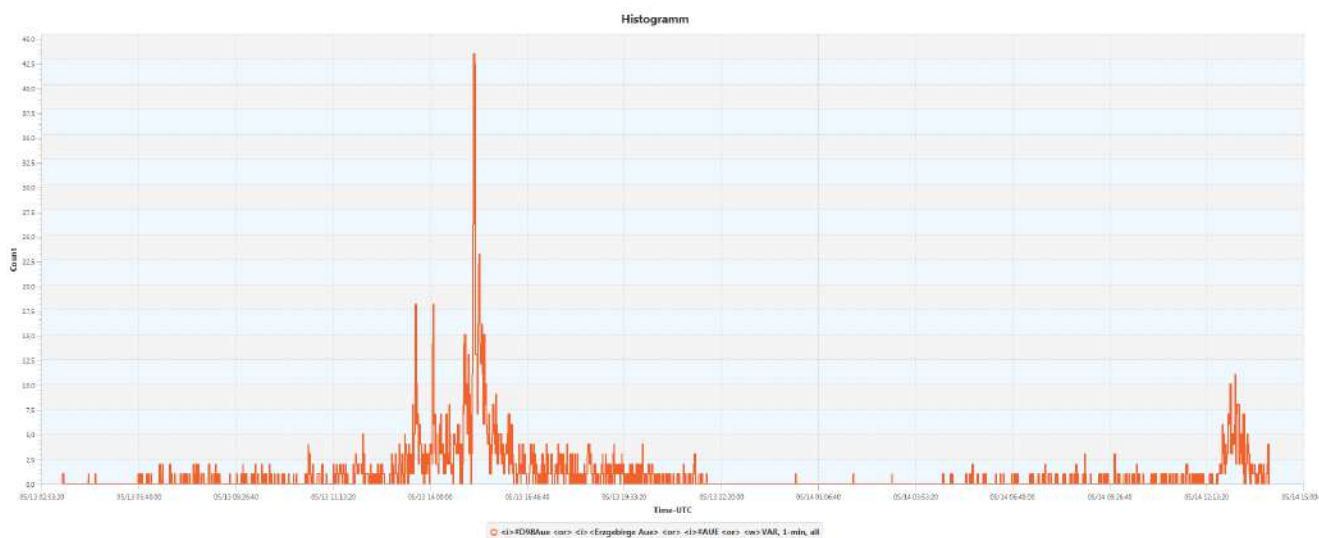


Man kann bei Twitter nicht davon ausgehen, dass sich die Benutzer an die Regeln der Groß- und Kleinschreibung halten. Mit dem vorgestellten <i> wird auch „#aue“ oder „#aueforever“ gefunden. Der Begriff „VAR“ muss wegen des vorgestellten „<w>“ jedoch exakt so als Wort geschrieben werden. Ansonsten hat man zu viele Treffer die mit dem Video-Assistent-Referee nix zu tun haben. Der VAR spielte in diesem

Spiel eine wichtige Rolle, da Aue bereits in der 4. Minute ein klares Tor aberkannt wurde (siehe Bild). Aue gehörte zu jenen Teams, die sich zu Beginn der Meisterschaft gegen die Einführung dieses Systems ausgesprochen haben und heimste deswegen einige Häme ein.

| | | | |
|---------|----------------|-----------------|---|
| lang:de | 05/13 15:17:43 | 9zehn100_8und70 | Alles was mir zu Aue einfällt: Das kann ja wohl nicht VAR sein! #D98AUE #SkyBuli |
| lang:de | 05/13 15:25:15 | iMia_San_FCB | Eine Schande, was da in Darmstadt passiert ist. Das ist doch der Beweis, dass man VAR oder wenigstens die Torlinientechnik brauch!! #D98AUE |
| lang:de | 05/13 15:36:35 | RC_KH | Liebe Gegner von VAR oder Torlinientechnologie. Schöne Grüße aus Aue. #Liga2 |
| lang:de | 05/13 15:37:27 | SveMaHe | RT @RC_KH: Liebe Gegner von VAR oder Torlinientechnologie. Schöne Grüße aus Aue. #Liga2 |

In Grafik 3 sieht man links das übliche „Vorgeplänkel“ zum Spiel. Die erste Spitze ist die Aberkennung des Tores. Die zweite Spitze ein nicht gegebener – klarer – Elfer für Aue, kurz vor Schluss hat der Schiri Aue einen weiteren Elfer vorenthalten. Die hohe Spitze markiert das Spielende und die aufgeregte Diskussion über das Spiel. Zu diesem Zeitpunkt steigen – bei jedem Spiel – die Nachrichten-Medien ein, die Twitter-User reagieren darauf. Die Aufregung legt sich aber schnell. Aue legte am Montag Protest gegen das Spiel ein und beschwerte sich über den Schiri. Das ist die Spitze rechts. Es endete wie bei Aschenputtel. Der von Karlsruhe wegen einer Verletzung nach Aue abgegebene Sören Bertram schießt mit 3 Toren seinen Ex-Klub in der Relegation vom Platz. Die „Macht aus dem Schacht“ bleibt in der 2. Liga.



Grafik 3: Histogramm, Erzgebirge Aue gegen Darmstadt, So./Mo. 13./14. Mai

Am Dienstag, 15. Mai fiel das O2-Handynetz für längere Zeit aus. Mit der obenstehenden Abfrage

retrieve(query: "<x>O2 &and> (<i>ausfall <or> down <or> <i><Kein Empfang>)", display: 1);

Go < >

| Query | Time-UTC | User | Text |
|---------|----------------|---------------|---|
| lang:de | 05/15 10:47:09 | Alien_timi | Spinnt bei euch auch O2 ... Kein Empfang mehr seit 1 std... |
| lang:de | 05/15 10:57:22 | der_sugasoph | Das E-Plus & O2 Handynetz sind grad down. ☐ |
| lang:de | 05/15 11:50:45 | diss_kurs | Netzausfall E-Plus o2 Nerv ja |
| lang:de | 05/15 12:53:44 | DerLarsDE | Aldi Talk und das O2 Netz ist vor allem in NRW und Bayern komplett down. Hoffentlich wird es schnell behoben. https://t.co/1jpb62IvwN |
| lang:de | 05/15 12:57:03 | FOCUS_TopNews | Störung bei E-Plus, O2 und Aldi Talk - Kunden klagen über flächendeckenden Netzausfall https://t.co/Y5TVGNub8n |

erhält man die darunter stehenden Tweets. „<x>O2“ bedeutet, die Groß- und Kleinschreibung ist egal, aber es muss ein eigenes Wort sein. Ansonsten erhält man auch Echo2018 als Treffer. Zusätzlich muss es noch um den Zusammenbruch des Netzes gehen und z.B. nicht um günstige O2-Tarife. Bei „ausfall“ soll auch „Ausfall“, „Netzausfall“ oder „Totalausfall“ erkannt werden. Eine Feinheit dieses Ereignisses ist: Die Betroffenen können ihre Unpässlichkeit nicht direkt tweeten. Trotzdem ist der User *Alien_timi*

FOCUS um mehr als 2 Stunden voraus. Man kann sich wie bei Aue das Histogramm anschauen und eine Reihe von anderen Auswertungen machen. Eine gute handverlesene Datenbank und eine mächtige Abfragesprache ist jedoch die notwendige Basis um aus der Masse der Tweets Information heraus filtern zu können. Umgekehrt muss man die vollständigen Daten haben um auch speziellen Ereignisse erfassen zu können. Bei 1 Million Tweets pro Tag wird das auf Dauer auf einem normalen PC eine Herausforderung.

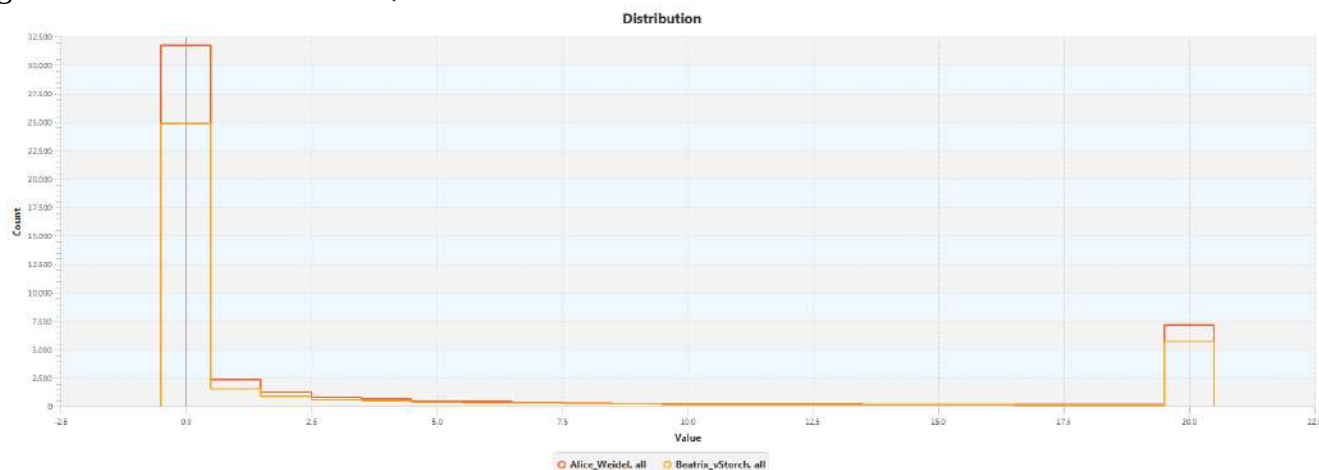
Spieglein, Spieglein an der Wand, wer hat die meisten Follower im Land?

Am 29. April hat die AfD Politikerin Alice Weidel einen Tweet der Pollytix-Strategic-Research weiter geleitet. Sie ist vor ihrer Parteifreundin Beatrix v. Storch die Follower-Königin im AfD-Land.

lang.de 04/29 14:28:16 Alice_Weidel RT @pollytix_gmbh: Die #MdB's der @AfDImBundestag mit den meisten #FollowerIn bei #Twitter sind @Alice_Weidel @Beatrix_vStorch @PetrBystro...

Um das fest zu stellen braucht man keinen Strategic-Research, man braucht nur auf den Twitter-Accounts der Politiker nachschauen. Das sollte ein Feriapraktikant oder AzuBi in einer Stunde problemlos schaffen. Die ungekrönte AfD-Twitterkönigin ist die ehemalige CDU Politikerin Erika Steinbach. Aber Steinbach ist keine Abgeordnete und so hat Pollytix schon richtig gezählt.

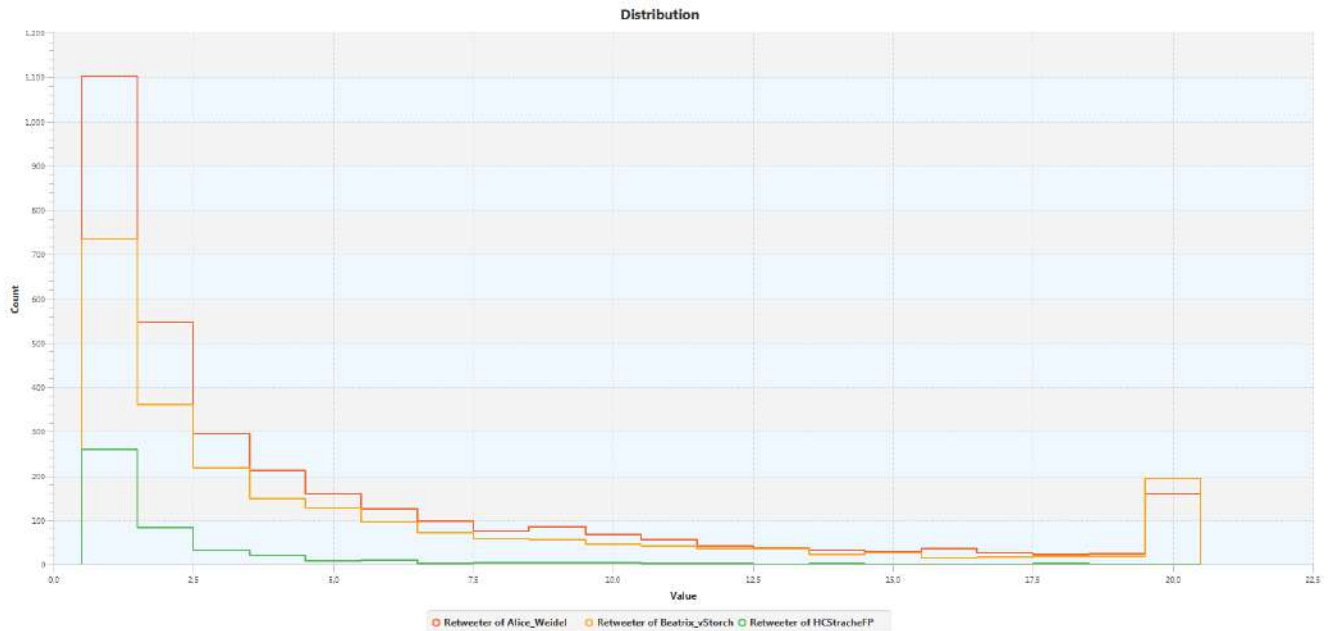
Interessanter ist schon die Frage: Gibt es diese Follower überhaupt und leisten sie Gefolgschaft? Man kann von Twitter die Benutzer-Kennung der Follower herunter laden. Diese Kennung wird bei jedem Tweet mit geschickt. Vanellus kann so dem Treiben der Follower folgen. Für die folgenden Ergebnisse habe ich die Daten von 4 Wochen, vom Fr. 27. April bis Do. 24. Mai, verwendet. In Summe sind das gut 27 Millionen Tweets von 3,4 Millionen verschiedenen Usern.



Grafik 4: Verteilung der (in-)aktiven Follower von Alice Weidel (rot) und Beatrix v. Storch (gelb)

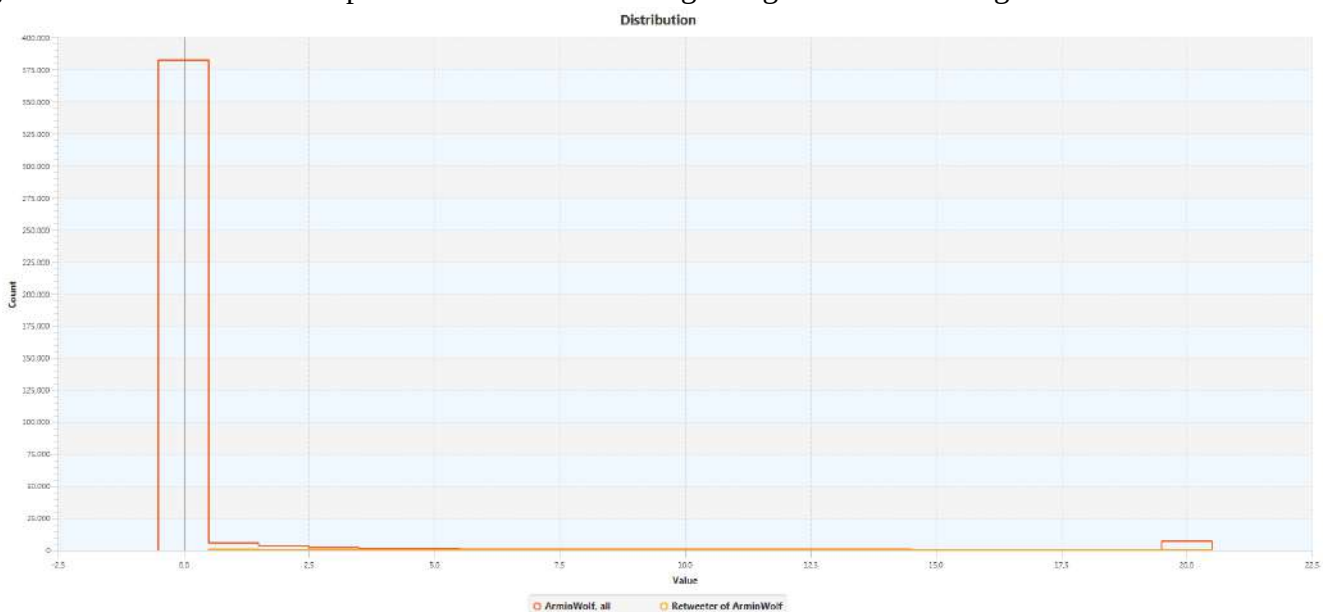
Alice Weidel hat gut 47.000, ihre Parteifreundin Beatrix v. Storch 36.000 Follower. Grafik 4 zeigt, wie viele Tweets diese Follower in den letzten 4 Wochen versandt haben. Bei 32.000 dürfte das Interesse für Twitter nicht sehr ausgeprägt sein. Sie haben innerhalb dieses Zeitraum nichts gezwitschert. Das heißt nicht unbedingt, dass sie die Tweets von Alice Weidel gänzlich ignorieren. Aber es dürfte dieser mit Abstand größten Gruppe Twitter nicht so wichtig sein. 2.500 haben ein Mal getweetet. Ganz rechts sieht man die Gruppe der „Power-User“ die mehr als 20 Tweets versandt haben. Diese Statistik erfasst jeden Tweet. Wenn jemand bei Real gegen Bayern „Tooor“ getweetet hat, dann ist er bereits ein aktiver User. Das Verhältnis von aktiven zu inaktiven Usern ist bei Beatrix v. Storch ähnlich.

Aussagekräftiger ist wohl die Anzahl der Follower die zumindest eine Meldung retweeten. Wie man in Grafik 5 sieht, sind es bei Alice Weidel 1.100 User die das einmal machen. Die Verteilung sinkt mit höheren Anzahl von Retweets steil ab. Es gibt aber auch hier wieder eine Gruppe von 200 Power-Usern. Einige davon sind jedoch Bots die so ziemlich alles weiterleiten was ihnen in die Fänge kommt. Die Verteilung von Beatrix v. Storch ist wieder aliquot zur Follower-Zahl.



Grafik 5: Retweeter-Verteilung, Alice Weidel (rot), Beatrix v. Storch (gelb), H.C. Strache (grün)

Die grüne Linie zeigt die Verteilung der Retweets für H.C. Strache. Strache hat mit 49.000 Followern etwas mehr als Alice Weidel. Er ist jedoch im Twitter-Universum eine vernachlässigbare Größe geworden. Man muss sich um seinen Gesundheitszustand ernsthaft Sorgen machen: Soviel Kreide wie H.C. seit dem Eintritt der FPÖ in die Regierung frisst, kann nicht gesund sein. Um ja nichts Falsches zu sagen, retweetet er primär den Regierungssprecher. Darüber hinaus meldet er, dass er mit Norbert Hofer Kränze für die Opfer des Faschismus niedergelegt hat, er ist von Arik Brauers Rede zur Befreiung vom Faschismus tief berührt und sendet in verschiedensten Varianten Israel Glückwünsche zum 70. Jahrestag der Staatsgründung. Das reißt seine Fans nicht vom Stockerl. Es zeigt sehr schön das generelle Dilemma eines Populisten. Einmal in der Regierung schaut die Welt ganz anders aus.



Grafik 6: Follower (rot) und Retweeter-Verteilung (gelb) von ArminWolf (ORF).

Der offizielle Österreichische Twitter-König ist der ORF-Journalist Armin Wolf. Er hat 400.000 Follower. Twitter erlaubt – bei einem freien Zugang – das Herunterladen von 75.000 Follower-

Kennungen innerhalb von 15 Minuten. Vanellus lädt den ersten Schwung herunter, schläft 15 Minuten, lädt den zweiten Schwung, schläft 15 Minuten ... Diese Aufgabe wird nebenläufig erledigt. Der allgemeine Betrieb wird davon nicht beeinträchtigt.

Wie man in Grafik 6 sieht, sind von den 400.000 Wolf-Followern 380.000 Twitter-Zombies. Die Anzahl der Retweeter (gelb) ist in der Grafik kaum auszumachen. Die Twitter Gemeinde interessiert inzwischen Armin Wolf genauso wenig wie H.C. Strache. Möglicher Weise gibt es einen direkten Zusammenhang. Armin Wolf hat sich als wackerer Kämpfer gegen den pöhszen H.C. positioniert. Seit dieser auf dem Kreide-Trip ist, fehlt ihm die Angriffsfläche. Möglicher Weise läuft sich auf einem so flüchtigen und schnelllebigen Medium wie Twitter auch alles schnell tot.

| User | Level | Retweets |
|-----------------|-------|----------|
| Beatrix_vStorch | 0 | 1 |
| AfDimBundestag | 1 | 2 |
| JoachimPaul_AfD | 1 | 1 |
| BlauesWunder18 | 1 | 1 |
| RolandTichy | 1 | 2 |
| BlondJedi | 1 | 1 |
| krk979 | 1 | 2 |
| Dieter_Stein | 1 | 1 |
| MarcFelixSerrao | 1 | 1 |
| UbbelohdeBerlin | 1 | 1 |
| Hartes_Geld | 1 | 1 |
| AfD | 1 | 6 |
| AfD_SE | 1 | 1 |
| FamUnt | 1 | 1 |
| LSteinwandter | 1 | 2 |
| WMuhsal | 1 | 1 |
| ronaldglaeser | 1 | 2 |
| Alice_Weidel | 1 | 2 |
| Joerg_Meuthen | 1 | 1 |
| StBrandner | 1 | 1 |
| PetrBystronAfD | 1 | 1 |
| DrDavidBerger | 1 | 1 |
| SteinbachErika | 1 | 1 |
| Martin_Hess_AfD | 1 | 1 |
| JoanaCotar | 1 | 1 |
| AfDBerlin | 1 | 2 |
| maxotte_says | 1 | 1 |
| AfDKompakt | 1 | 4 |
| UdoHemmelgarn | 1 | 1 |
| GtzFrmming | 1 | 2 |

Man kann sich mit Vanellus mehrere Ebenen von Retweet-Beziehungen ansehen. Level-0 ist der ausgewählte User. Level-1 sind die User die dieser retweeted, Level-2 die Benutzer, die von Level-1 Usern retweeted werden Man erhält so einen Retweeter-Baum. Dieser wird jedoch sehr schnell unübersichtlich und groß.

| User | Level | Retweets |
|-----------------|-------|----------|
| Alice_Weidel | 0 | 0 |
| AfDimBundestag | 1 | 7 |
| AfD | 1 | 4 |
| GtzFrmming | 1 | 1 |
| pollytix_gmbh | 1 | 1 |
| Beatrix_vStorch | 1 | 1 |
| Mewinator89 | 1 | 1 |

Die Screenshots zeigen links den Retweet-Baum (bis Ebene-1) von Beatrix v. Storch und rechts von Alice Weidel. Weidel geht mit Retweets – von Parteifreunden - sehr sparsam um. In größerer Anzahl retweeted sie nur die Personen neutralen Partei- und Fraktions-Tweets.

Beatrix v. Storch ist kollegialer. Sie schreckt aber auch nicht davor zurück sich selbst zu retweeten.

Bei H.C. Strache hat der Baum nur einen Stamm. Er retweeted wie bereits erwähnt ausschließlich und ausführlich „RegSprecher_AT“.

Twitter-Zwillinge und Bot-Erkennung:

Speziell seit dem US-Präsidentenwahlkampf 2016 gibt es eine intensive Diskussion um den Einfluss von Social-Media-Bots. Es hat auch das amerikanische Verteidigungsministerium einen Wettbewerb zur Bot-Erkennung ausgeschrieben. Es gibt verschiedene Ansätze zur Bot-Erkennung. Der klassische Versuch ist inhaltlicher Natur. Man versucht durch Spracherkennung und Sprachanalyse den Bots auf die Spur zu kommen. Der andere ist rein statistisch. Man untersucht das Zeitverhalten und auch die Tatsache, dass Bots in der Regel im Rudel auftreten. Man kann beide Ansätze verbinden. Bei Vanellus dominiert der statistische Ansatz.

Text-Zwillinge:

Unter einem Text-Zwilling versteht Vanellus zwei User die ungefähr die selbe Anzahl von Tweets senden und die eine hohe Anzahl von Tweets gemeinsam haben. Die Feststellung von Zwillingen ist für einen alten Hacker eine nette Herausforderung. Bei 3,4 Millionen Usern gibt es 5,8 Milliarden User-Paare. Wenn man jeden User mit jeden anderen direkt vergleicht, wird man nie fertig. Durch eine Reihe von Tricks geht die Berechnung bei Vanellus ziemlich flott von statten.

Auto-Zwillinge:

Zu meiner Überraschung erzeugte die Methode zur Erkennung von Zwillingen auch sogenannte Auto-Zwillinge. Ein Bot ist der Zwilling von sich selbst. Das passiert, wenn ein Bot dieselbe Nachricht

mehrmals schickt. Ein Beispiel dafür ist „*phil_de_bot*“. Dieser Bot verschickt alle 30 Minuten einen Tweet mit einem zufällig ausgewählten philosophischen Zitat. Der Zitatenschatz ist nicht sehr groß und so wiederholt „*phil_de_bot*“ regelmäßig Zitate. Es gibt Bots, die Kirchturm-Glocke spielen und jede volle Stunde die Zeit durchgeben. Auch diese Meldungen wiederholen sich periodisch. Ein anderes Beispiel sind Radiosender, die den jeweiligen Liedtitel tweeten. Aktuelle Hits werden häufig gespielt und es wiederholt sich dementsprechend der Tweet. Die Auto-Zwillinge sind mir passiert. Ich habe den Bug (Fehler) zum Feature gemacht. Man kann die Erkennung auch ausschalten bzw. nur nach Auto-Zwillingen suchen. Ich fand die Erkennung der Auto-Zwillinge am Ende durchaus interessant und habe dieses ungeplante Feature im Code gelassen. Es gibt Bots, die sowohl Auto- als auch echte Zwillinge sind. Sie wiederholen eine Botschaft mehrmals und tun dies im Rudel.

Echte Zwillinge:

Bei echten Zwillingen handelt es sich um zwei verschiedene User. Wobei diese in der Regel ganze Rudel bilden.

| | | | |
|-----|----------|----------------|---|
| 148 | Hashtags | #Homöopathie | Ein Beispiel dafür ist der Homöopathie-Rudel. |
| 12 | Hashtags | #MachAuchDuMit | Der User „ <i>sand_iris</i> “ bildet einen Zwilling mit |
| 10 | Hashtags | #machauchdumit | „ <i>Ann_Walters_</i> “, „ <i>Barbara_Hofer_</i> “, „ <i>H_Schneider_</i> “ |
| 8 | Hashtags | #homeopathy | Diese bilden mit weiteren Usern wieder Zwillinge. |
| 7 | Hashtags | #Homeopathy | Ziel dieser Gang ist die Propagierung der Homöopathie. |
| 6 | Hashtags | #Globuli | Der Screenshot zeigt die häufigsten Hashtags des Bots |
| 4 | Hashtags | #Placebo | „ <i>sand_iris</i> “ Es hat eine gewisse Ironie, wenn Anhänger |
| | | | der sanften Medizin ein umfangreiches Bot-Netzwerk |
| | | | zur Propagierung ihrer Ansichten aufbauen. |

Die mit Abstand größte Bot-Gang ist mit dem User „*Arikatym31*“ verbunden (man könnte ein beliebig anderes Rudelmitglied als Repräsentanten auswählen). Diese Gang hat einige Hundert hoch aktiver Mitglieder. Wobei sich – und auch das ist für Bot-Gangs typisch – die Bots häufig umbenennen.

| | | | |
|---------|----------------|------------|---|
| lang:de | 04/27 00:13:32 | Arikatym31 | RT RTLisaFitzpatrick #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 00:28:22 | Arikatym31 | RT Scotttdrives #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 00:46:04 | Arikatym31 | RT KristinePThomas #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 00:46:09 | Arikatym31 | RT glmh101 #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 00:46:12 | Arikatym31 | RT mindwanderer #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 13:38:22 | Arikatym31 | RT RTgbrodermann #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 13:51:23 | Arikatym31 | RT sofiaberen #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 14:08:24 | Arikatym31 | RT RT DizzyDortch #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 14:08:30 | Arikatym31 | RT RTDizzyDortch #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |
| lang:de | 04/27 14:23:11 | Arikatym31 | RT RT JasonMehrtens #MTVBRPETNUGGET #MTVBRSHADETAYLORKATY #PremiosMTVMiaw #MTVLAINSTAGLCAMILAC |

Man erkennt sie trotzdem leicht, weil die Twitter-Kennung gleich bleibt. Der Screenshot zeigt die ersten 10 Tweets von *Arikatym31*. Alle Tweets dieser Gang enthalten nur Hashtags. Offensichtliches Ziel ist das Hashtag-„Pumpen“. Der von dieser Gang „gepumpte“ Hashtag „*#PremiosMTVMiaw*“ hat innerhalb des Deutschen Netzes die Poleposition. Ich kenne mich in der Welt der Social Media zu wenig aus, um beurteilen zu können, welchen Zweck es hat, mit „*#PremiosMTVMiaw*“ die Poleposition zu erringen. Angeblich nimmt Twitter sozial auffällige Bots bzw. Bot-Gangs aus dem Rennen. Bei dieser sehr leicht erkennbaren Gang ist dies nicht der Fall. Bei „*phil_de_bot*“ oder den Bots der Radiosender die den aktuellen Titel ansagen, gibt es für Twitter keinen Grund, dagegen vorzugehen. Die Existenz eines API und ausgereifter Programmibliotheken beweist hinlänglich, dass Bots ein akzeptierter und wichtiger Teil des Twitter Kommunikationsnetzes sind. Der Eingangs zitierte Bot [DeepDrumpf](#) hat es sogar zu einiger Medienberühmtheit gebracht. Dieser am M.I.T. entwickelte Bot baut aus den Tweets des Users „*RealDonaldTrump*“ nach einer relativ einfachen Methode neue

Texte zusammen. RealDonaldTrump eignet sich sehr gut als Vorlage. Seine Sprache ist einfach gestrickt, der Inhalt ist sprunghaft. Man bekommt das auch mit minimaler Bot-Intelligenz hin.

Fan oder Gegner, die Sentiment-Analyse:

Ich habe mich primär auf die AfD und ihre markantesten Repräsentanten konzentriert, weil diese auf Twitter sehr präsent sind und es innerhalb der User zwei klar erkennbare Lager gibt. Es ist – für einen Menschen – sehr leicht erkennbar, ob jemand die AfD mag bzw. sie verabscheut. Teilweise deklarieren die User dies bereits in ihrem User-Namen. Z.B. heißen die beiden aktivsten Gegner „KeineAFD2017“ und „ContraAfD1“. AfD Politiker hängen – wenn sie nicht so bekannt wie die beiden Damen sind – häufig ein „AfD“ an ihren Namen an. Ortsgruppen nennen sich z.B. „AfD_Frankenthal“.

Für einen Bot ist die Unterscheidung hingegen gar nicht so trivial. Die automatische Klassifikation von Benutzern ist unter dem Titel „Sentiment Analyse“ ein sehr heißes Forschungsgebiet. Man kann diese Klassifizierung auch für zielgerichtete Werbung einsetzen. Ein verwandtes Gebiet sind „Recommender Systems“. Es schlägt einem z.B. Amazon auf Grund des bisherigen Verhaltens neue Produkte vor. Der Netflix-Recommender empfiehlt Filme. Netflix hat einen mit 1 Million \$ dotierten Wettbewerb zur Verbesserung dieses Systems ausgeschrieben. Diese Systeme werden von den Benutzern in der Regel als nützliches Feature und nicht als pöhsze Verführer wahr genommen. Die Techniken von Cambridge Analytica bewegten sich nach meinen Informationen innerhalb dieses allgemein bekannten Spektrums. Die Firma hat sich primär durch Angeberei ihres Chefs einen schlechten Ruf eingehandelt. Das war ihr Glück und Ende. Es war dieser Ruf und nicht die besonders fiesen Methoden die große Kunden zur Beendigung der Verträge bewegen hat.

| User | Matches | Total-Tweets | Ratio% |
|-----------------|---------|--------------|--------|
| mrstone0856 | 1735 | 1843 | 94 |
| kardasiapat | 1094 | 1491 | 73 |
| DerDude72445273 | 1 | 1382 | 0 |
| KeineAFD2017 | 2 | 1365 | 0 |
| artepmobil | 900 | 1139 | 79 |
| ContraAfD1 | 3 | 1138 | 0 |
| SchafDas | 764 | 1096 | 69 |
| torsten_luer | 963 | 1046 | 92 |
| Tinimaus1110 | 866 | 1018 | 85 |
| GertWalterWolf1 | 780 | 991 | 78 |
| mginp | 1 | 965 | 0 |
| APVogt | 394 | 951 | 41 |
| m_aus_s | 797 | 936 | 85 |
| LC180666 | 668 | 928 | 71 |
| e_pitzky | 791 | 911 | 86 |
| ruthlissy | 809 | 902 | 89 |
| wmoebius | 7 | 901 | 0 |
| Chat_Atkins | 0 | 881 | 0 |
| Renft1964 | 688 | 854 | 80 |
| genauhinsehen | 719 | 819 | 87 |
| schnaggi | 1 | 775 | 0 |

Die Retweet-Ratio:

Eine naheliegende Klassifizierung ist, sich die Retweets eines Benutzers anzuschauen. Wenn jemand einen AfD-Politiker Retweeted, dann kann man dies als Zustimmung werten. Der Screenshot zeigt die Ratio der AfD-Retweets innerhalb der Poweruser. Z.B. sind 69% der Tweets des Users „SchafDas“ AfD Retweets. Es ist auch eine klare Unterscheidung in Freund und Feind erkennbar. Es kommt zwar vor, dass deklarierte Gegner wie „KeineAFD2017“ AfD-Quellen retweeten. In diesem Fall will der Gegner die Ansichten eines AfD-Politikers an den Pranger stellen. Man muss für eine derartige Untersuchung die Tweets inhaltlich filtern. Es können politisch Zerstrittene gemeinsam den Bayern die Daumen halten. Es mussten im Tweet der Begriff „AfD“ bzw. die bekanntesten Personen oder eindeutig zuordenbare Schlagwörter wie „#MerkelMussWeg“ vorkommen.

Anmerkungen: Man könnte eine ähnliche Untersuchung über die Pro- und Contra Bayern Fans machen. Es böte sich auch Rasenball Leipzig an.

Zunächst habe ich per Hand eine Liste der bekanntesten AfD-Politiker erstellt. Das war erstens mühsam und zweitens

sind mir u.A. AfD-Ortsgruppen durch die Lappen gerutscht. Diese retweeten primär ihre lokalen Repräsentanten. Ich bin dann auf die glorreiche Idee gekommen, die Retweets von eindeutig erkennbaren AfD-Powerusern – konkret von „mrstone0856“ und „torsten_luer“ als Vorlage zu nehmen. Vanellus erstellt aus den Retweets dieser beiden unermüdlichen hardcore-Fans die Liste der AfD-Retweet Kandidaten. Damit hatte ich das Problem der Lokalhäuptlinge und der AfD-Leithammel

im Netz gelöst. Allerdings retweeteten diese Poweruser auch Meldungen von neutralen Quellen wie SPIEGELONLINE oder NZZ. Ich habe zunächst wieder per Hand eine Liste dieser neutralen Benutzer erstellt, die nicht zur AfD Retweet-Ratio gezählt werden sollen. Das hatte einen ähnlichen Effekt wie der ursprüngliche Ansatz. Man findet immer wieder Medien oder User die man nicht zählen sollte. Bis ich auf die ebenfalls naheliegende Idee gekommen bin: Wenn jemand von den aktiven Gegnern häufig retweetet wird, dann kann man diesen im AfD Sinn ignorieren. Konkret verwende ich zu diesem Zweck die User „Chat_Atkins“ und „DerDude72445273“.

Diese Klassifizierung funktioniert auch bei Usern, die weniger aktiv sind, recht gut. Allerdings hatte die AfD_Frankenthal eine AfD-Ratio von 0%. Es handelt sich eindeutig um eine sehr linientreue Lokalorganisation. Nur hält sich – der vermutlich Ältere – Twitter Verantwortliche nicht an die informellen Twitter-Standards. Entsprechend diesem Standard schreibt man „RT @AfD Kompakt“ und nicht wie er „AfD Kompakt:“. Vanellus erkennt dieses Format nicht als Retweet. Der Versuch auch derartige Tweets als Retweet zu klassifizieren, erzeugte bei den Gegnern wieder zu viele Fehlalarme. Ich bin schlussendlich zum Entschluss gekommen: Auf diese Art und Weise ist die AfD_Frankenthal nicht zu erfassen.

| | | | |
|---------|----------------|-----------------|---|
| lang:de | 04/27 03:31:27 | AfD_Frankenthal | AfD Kompakt: Bundesregierung völlig kenntnislos bei Ein- und Ausreisen https://t.co/wjcnZh0esj https://t.co/42tEL8Tsub |
| lang:de | 04/27 03:46:27 | AfD_Frankenthal | AfD Kompakt: Eklat im Potsdamer Stadtschloss https://t.co/p3vEj1Zq2R https://t.co/lBb4rjHbow |
| lang:de | 04/27 04:01:26 | AfD_Frankenthal | AfD Kompakt: Neuerliche Überprüfung von NO2-Grenzwerte herbeiführen https://t.co/sTmKkKj9i9 https://t.co/LZYGtoNae5 |
| lang:de | 04/27 04:16:26 | AfD_Frankenthal | AfD Kompakt: Stoppt Genozid an Christen im Nahen Osten https://t.co/RkoWmFSZvo https://t.co/X4Fn1Mwqqi |
| lang:de | 04/27 04:31:26 | AfD_Frankenthal | AfD Kompakt: Deutschland und Europa können Weltklima nicht retten https://t.co/gN9XBdZDAK https://t.co/5nX9cQcF9A |
| lang:de | 04/27 10:58:05 | AfD_Frankenthal | AfD Kompakt: Muslime werden antisemitische Passagen im Koran nicht freiwillig schwärzen https://t.co/1oCDtcYHUK https://t.co/VWXrdRNyOw |
| lang:de | 04/27 16:33:03 | AfD_Frankenthal | Presseschau: Wohlfahrtsverbände tabuisieren Behinderungen durch Inzest https://t.co/RShgnzsmzB |

Man kann die Methode umdrehen und mit vertauschen Leithammeln auch eine Anti-AfD Ratio berechnen. Das funktioniert nicht so gut. Die AfD-Fans bilden – schon alleine durch die Existenz von AfD-Parteistrukturen – einen relativ geschlossenen Block. Die Gegner sind - was das Retweeten angeht - diffuser.

Die Hashtag-Ratio:

| User | Matches | Total-Tweets | Ratio% |
|-----------------|---------|--------------|--------|
| mrstone0856 | 729 | 1843 | 39 |
| kardasiapat | 430 | 1491 | 28 |
| DerDude72445273 | 49 | 1382 | 3 |
| KeineAFD2017 | 39 | 1365 | 2 |
| artepmobil | 318 | 1139 | 27 |
| ContraAfD1 | 32 | 1138 | 2 |
| SchafDas | 266 | 1096 | 24 |
| torsten_luer | 399 | 1046 | 38 |
| Tinimaus1110 | 242 | 1018 | 23 |
| GertWalterWolf1 | 261 | 991 | 26 |
| mginpl | 40 | 965 | 4 |

Man kann dieselbe Idee auch für die Twitter-Hashtags verwenden. Von den Retweets war ich schon gewieft genug um die Arbeit den Powerusern zu überlassen. Bei den Hashtags wäre das eine Sisyphus Arbeit, da immer wieder neue hinzukommen und alte nicht mehr verwendet werden. Diese Arbeit erledigen die Poweruser. Bei Hashtags ist die Angabe von „don‘t care“ besonders wichtig. Es dominiert z.B. auf beiden Seiten der Hashtag „#afd“. Mit diesem und anderen Hashtags kann man die Gruppen nicht unterscheiden. Es wird jedoch kein Gegner „#merkelmussweg“ verwenden und umgekehrt kein AfD-Fan „#noafd“. Man sucht nach der Menge aller Hashtags der AfD-Fans die von der Gegenseite

nicht verwendet werden. Diese Methode ist nicht so trennscharf wie die Retweet-Ratio. Allerdings bekommt man auch sinnvolle Werte für Benutzer, die von der ersten Methode nicht erfasst werden. Sie funktioniert auch mit umgekehrten Rollen recht gut. Die AfD-Gegner haben keine klar definierten Leithammeln, sie haben jedoch ebenfalls eine relativ geschlossene Agenda. Beide Methoden ergeben zusammen schon eine relativ präzise Gruppeneinteilung. Nur ein kleines Dorf, die AfD_Frankenthal, leistete noch Widerstand. Es greift auch die Hashtag-Methode voll daneben, weil der dafür Verantwortliche sich auch nicht an Hashtag-Regeln hält. Sie kommen in seinen Tweets nicht vor.

Die Schindel-Ratio:

Der Shingling (Dt. Schindel) Algorithmus ist eine 1993 erfundene Methode zur allgemeinen

Klassifizierung von Dokumenten. Man möchte z.B. wissen, ob ein Artikel in die Kategorie Sport, Wetter, Politik ... fällt. Man gibt Muster-Berichte vor. Ein neuer Artikel wird entsprechend der größten Übereinstimmung klassifiziert.

Die erste Schindel besteht aus den Zeichen 1..11 des Textes. Die zweite Schindel von 2..12, die dritte von 3..13. Die Textfragmente überlappen sich wie die Schindeln eines Daches. Ein üblicher Wert ist eine Schindellänge von 9. Für diese Analyse hat der Wert 11 etwas besser funktioniert. Man kann z.B. auch vier aufeinanderfolgende Worte als Schindeln definieren. Man geht genauso wie bei den Hashtags vor. Man bildet die Schindeln der AfD-Power-User und zieht von diesen die gleichlautenden Schindeln der AfD-Gegner ab. Danach durchsucht man die Tweets aller User nach übereinstimmenden Schindeln. In einem Tweet können mehrere passende Schindeln vorkommen. Diese werden extra gezählt. Es kann die Ratio wesentlich größer als 100% werden bzw. es haben auch Gegner relativ hohe Schindel-Ratios. Die Frage ist nur, ob sich die Werte signifikant unterscheiden.

| | | | |
|-----------------|-----|----|------|
| AfD_Frankenthal | 893 | 72 | 1240 |
|-----------------|-----|----|------|

Diese Methode pickt auch die Frankenthaler in das richtige Töpfchen. Mit einem Schindel Wert von 1240 gehören sie eindeutig ins AfD Lager. Es gibt jedoch einen schwierigen User namens AfD_Hesse. Nach den Tweets zu schließen, bezeichnet „Hesse“ keine Ortsgruppe. Es ist möglicher Weise eine Referenz an den Dichter Hermann Hesse. AfD_Hesse hat bei allen Methoden nur eine mittelmäßig ausgeprägte AfD-Ratio. Er gehört aber – auf Grund der Ratio – auch nicht ins Lager der Gegner. Er ist bekennender AfDler, engagiert sich – auf Twitter – allerdings auch stark für den Tierschutz und retweeted regelmäßig PETA. Auf einen Tweet, wie das denn zusammen passt, antwortet er mit:

| | | | |
|---------|----------------|-----------|--|
| lang:de | 04/28 13:49:45 | AfD_Hesse | @Andrino11elf @LarsSteinke @neomagazin Duerfen AfD-Waehler keine Tierfreunde sein? |
|---------|----------------|-----------|--|

Insofern war die nicht so eindeutige Klassifizierung nicht ganz daneben.

Es ist kaum möglich, die Klassifizierung von Abertausenden Usern per Hand zu überprüfen. Meine Stichproben ergaben eine recht gute Quote von richtigen Einschätzungen des Bots. Dem Mann (es könnte auch eine Frau sein) aus Frankenthal bin ich dankbar, dass er mich nach dem Motto – es wäre doch gelacht, wenn ich ihn nicht richtig hin bekomme – zu zusätzlichen Verfahren angeregt hat.

Potpourri:

Vanellus unterstützt noch eine Reihe von anderen Analyse-Methode. Man kann z.B. eine Liste jener User erstellen, die am ausdauerndsten twittern. Damit erkennt man Bots. Man kann sich die Peak-Tweet-Rate innerhalb eines Zeitraumes ansehen. Mit den oben bereits beschriebenen Methoden kann man u.A. untersuchen, wie lange der Tweet eines Promi ein Thema ist. In der Regel sind es ein paar Stunden. Man kann sich anschauen, welche Tweets bei den Usern ankommen und welche nicht. Man kann bei Powerusern ein Tagesprofil erstellen. Auf Grund des Histogramms erkennt man auf einen Blick, dass Alice Weidel am 1. Mai entgegen ihren sonstigen Gewohnheiten schon ganz zeitig in der Früh ihr Twitter-Büchserl geschultert hat. Sie dürfte gegen Gewerkschafter eine ähnliche Abneigung haben wie der [Da Wildschütz](#) vor dem Jaga. Man kann sich auch anschauen, welche Hashtags gerade in sind. Wobei man nach Themen bzw. Schlagwörtern selektieren kann.

Ich habe für diese Analyse die Tweets auf dem Rechner der Lektorin herunter geladen und diese auf meinem mit mehr RAM ausgestatteten Arbeits-PC analysiert (ansonsten ist der neuere PC der Lektorin aber besser ausgestattet). Die Analyse kann aber auch in Echtzeit durchgeführt werden. Es ist für die Analyse-Routinen egal, ob die Daten von der Platte geladen werden oder vom Twitter-Server kommen. Sie haben kein Mascherl. Allerdings verändern die aktuellen Daten laufend das Ergebnis. Das ist für eine derartige Analyse lästig. Vanellus ist so gebaut, dass er unmittelbar auf äußeres Geschehen reagieren kann. Er tut es allerdings noch nicht.

Aussichten:

Die aktuelle Version von Vanellus ist V0.31. Der Vogel ist gerade erst dem Netz entfliegen. Ich habe vor, die Software weiter zu entwickeln. Es gibt sehr viele nette Probleme, die das Herz eines alten Hackers erfreuen und ihm auch wieder das Gefühl von „forever young“ geben. Es wirkt sich jedoch auf mein Sozialverhalten negativ aus. Ich kehre in solchen Phasen der Welt der Rücken zu und fühle mich im Elfenbeinturm sehr wohl. Anmerkung der Lektorin: Der Chef für Alles ist immer so.

Es gibt bereits Interesse – aber noch keine fixen Vereinbarungen – sowohl von akademischer als auch kommerzieller Seite.

Ursprünglich sollte Vanellus auf meinem PC mit 16GB Speicher rund 50 Millionen Tweets auf einmal analysieren können. Tatsächlich ist bei 30 Millionen Schluss. Die Ursache sind die letztklassigen Java-Standardroutinen zur Behandlung von Text/Dokumenten (Für Details siehe den technischen Anhang). Im Grunde genügen auch 30 Millionen Tweets. Aber ein derartiger Mist tut meiner Hacker Seele weh. Mein Vater war ein leidenschaftlicher Energiesparer. Als er meine Schwester in Amerika besucht hat, hat er ständig unter der in seinen Augen unglaublichen Energieverschwendung der Amis gelitten. Mir geht es mit der Java-Text-Bibliothek so.

Man kann sich noch eine Reihe von anderen Analyse Methoden ausdenken. Z.B. kann man analog zur Page-Rank Methode von Google eine User-Rank einbauen. Wer wird wie oft retweetet und wie wichtig sind diejenigen, die die Botschaft weiter geben. Vanellus könnte auch das Maul aufmachen, aber ich weiß nicht recht, was er sagen soll. Ein wichtiger und großer Schritt wäre die Einbeziehung von anderen sozialen Kanälen und insbesondere von Facebook.

Generell habe ich etwas das Problem von [Der Wilde auf seiner Maschin](#). „*I hab zwar ka Ahnung wo I hin fahr, aber dafür bin I schneller durt*“. Sollte es zur Zusammenarbeit mit einem kommerziellen Interessenten kommen, dann wäre es dessen Aufgabe dieses Problem zu lösen und ich könnte mich auf das konzentrieren was ich kann: Bitschnitzen.

„Die Software wird schneller langsamer, als die Hardware schneller wird“.
(N.Wirth).

Technischer Anhang: Das Elend der Java-String-Class:

Anmerkung: Nur für Programmierer geeignet und gedacht.

Vanellus hält alle Daten in einer selbst gebauten In-Memory-Datenbank. Die hereinkommenden Tweets werden in dieser Datenbank abgespeichert. Es wird von einem Tweet die Twitter-UserId, der Zeitstempel, die Tweet-Id, der User-Name, der Tweet-Text und die Abfrage auf Grund dessen er selektiert wurde, gespeichert. Es wird ein Tweet in mehrere Tabellen aufgeteilt. Es wird für den User-Namen und den Tweet-Text eine Hashzahl erzeugt. Sowohl der User-Name als auch der Tweet-Text werden in eine Hashtabelle mit ihrem CRC32-Key eingetragen. Damit werden User-Namen und auch Tweet-Texte die sich wiederholen nur 1x gespeichert. Unmittelbar wird ein Tweet in einer Tabelle mit dem Record

<Message-Id>, <Time>, <Query-Hash>, <User-Hash>, <Text-Hash>

gespeichert. Wobei jedes Element eine Long ist. Die Query, der User-Namen und der Tweet-Text wird bei Bedarf aus der jeweiligen Hashtabelle gelesen.

Ich habe mir dafür eigene spezialisierte Collections gebaut, da bei den Java-Collections eine long nicht als 8-Byte Zahl sondern als Long-Objekt mit mindestens 24-Byte overhead gespeichert wird. Sich eigene Collections zu bauen, widerspricht den üblichen Regeln. Meiner Meinung nach stammen diese Regeln von Leuten, die noch probiert haben die vorhandenen Ressourcen effektiv zu nutzen („down to the metal“) bzw. es ist für Leute gedacht, die das nie tun werden. Davon abgesehen findet ein alter Hacker nur das Rad gut, das er selbst erfunden hat.

Nach meinen Berechnungen sollten – wenn man der JVM 12 GByte Speicher zuweist - 50 Millionen

Tweets relativ locker in den Hauptspeicher passen. Beim Plattenimage der Datenbank geht es sich auch sehr schön aus. Zu meiner Überraschung waren es nur 30 Millionen. Das Rätsels Lösung: Java verwendet die UTF-16 Darstellung von Strings. Einst dachte man, dass man damit das Problem zusätzlicher Sprachen und Zeichensätze auf einfache Weise lösen kann. Man denkt sich als Software Entwickler nix und vertraut, dass die Hardware-Ingenieure das Problem schon lösen werden. Tatsächlich hat sich wieder einmal das Wirtsche- gegenüber dem Moorschen Gesetz durchgesetzt. UTF-16 war auch nur eine sehr kurzfristige Lösung. Man ist schnell drauf gekommen, dass für asiatische Sprachen auch 2 Bytes nicht ausreichend sind. UTF-16 ist die schlechteste aller Welten. Sie ist für die lateinischen Sprachen ineffizient und löst nicht das Problem von verschiedenen langen Codes. Diese einstige Fehlentscheidung treibt noch Jahrzehnte später in vielen Programmiersystemen (neben Java u.A. Qt oder Windows) ihr Unwesen. Eine wesentlich elegantere und effizientere Methode ist das von Ken Thompson und Robert Pike 1992 vorgeschlagene UTF-8 Format. Die Java Textwriter Klasse speichert Strings in diesem Format auf die Festplatte ab. Sie werden beim Einlesen jedoch wieder zu UTF-16 aufgeblasen. Nach meinen Tests benötigt UTF-8 für die deutschsprachigen Tweets nur 51% des Speicherplatzes von UTF-16.

Es gibt zwar mehrere Change Requests in Richtung UTF-8, diese wurden bisher von einer Java-Version auf die nächste verschoben. Eine derartige Umstellung ist kein triviales Problem.

Richtig bitter wird es, wenn man sich den Code der String-Klasse anschaut. IndexOf() ist so programmiert, wie sie jeder Anfänger ohne viel Nachdenken programmieren würde. Effizientes String-Matching ist eines der besten erforschten Gebiete der Informatik. Diese Forschungen sind bei den Autoren der Klasse offensichtlich spurlos vorüber gegangen. Eine effizientere Suchmethode erzeugt auch keinerlei Kompatibilitätsprobleme. Ein besonders trauriges Kapitel ist der String-Hashcode. Auch das ist ein sehr gut erforschtes Gebiet. Die Umstellung auf eine brauchbare Hashfunktion würde jedoch so manchen Code brechen. In Vanellus verwende ich bereits bisher die wesentlich bessere CRC32-Methode als Hashwert. Obwohl diese Zahl wie der Name schon sagt 32-Bit lang ist, gibt Java einen long-Wert zurück, weil das im Checksum-Interface so vorgesehen ist. Es rächt sich hier die fehlende Unterstützung von unsigned int durch den Sprachstandard. Ich habe es bei Long belassen. Das ist nicht das eigentliche Problem und möglicher Weise kommt man einmal drauf, dass CRC32 zu viele Kollisionen erzeugt. Dann kann man relativ problemlos auf einen 40- oder 64-Bit Hashmethode umstellen.

Fest steht, dass ich nicht mit so einem Murks weiter arbeiten will. Zu Beginn eines Projektes kann man noch derartige Fehler korrigieren. Man entkommt den Java-Strings nicht gänzlich, da z.B. die Anzeige von Tweets im GUI Strings als Input verlangt. Diese Daten sind jedoch sowieso streng von der Datenbank getrennt und man zeigt auch nicht 50 Millionen Tweets an.

Eine Alternative ist die Verwendung einer Library wie Lucene. Lucene verwendet – wie wohl jede big-data Library die etwas auf sich hält - UTF-8. Die andere Möglichkeit ist eine eigene UTF-8 String-Klasse zu schreiben. Theoretisch sollte auch ein 1-Byte Kode wie ISO-8859-15 für Deutschsprachige Tweets genügen. Man bekommt jedoch bei Emojis Probleme.

Bisher habe ich mich hauptsächlich mit Bit-Schnitzen und Numerischen Operationen beschäftigt.

Strings habe ich nur zum Beschriften von Tabellen verwendet. Die Welt der Big-String-Data ist neu – und spannend – für mich. Ich hielt es nicht für möglich, dass eine derartig zentrale Klasse derartig viel Schrott enthält und habe mir über diesen Punkt – abgesehen von der Hashfunktion - anfangs keinerlei Gedanken gemacht.

Nachtrag: Ich habe nun mit dem Bau einer eigenen UTF-8 Klasse begonnen und habe bei der Konvertierung von Groß- in Kleinschreibung Fehler beim türkischen İ ( 0130) gefunden. Die Tweets sind auf Deutsch, enthalten jedoch türkische Namen. Bei der Suche am Netz stellte sich heraus, der Bug liegt nicht auf meiner sondern auf Java Seite. Eine vollständige Konvertierung in allen Alphabeten ist eine ziemliche Herausforderung. Allerdings ist Türkisch keine exotische Sprache. Das sollte eine Standard-Library schon können.



Mirko Wenig
@BaerenGazette



Über Nacht war [#Tinder](#) down, weil [#Facebook](#) nach den jüngsten Datenskandalen den Zugang von Apps beschränkt hatte. Viele Singles waren so verzweifelt, dass sie sogar in ihrem eigenen Bett schliefen.

6:33 AM - Apr 5, 2018

♡ 6 👤 See Mirko Wenig's other Tweets



König Zuckerbergs Glück ohne Ende:

Mitte März tauchten in den Medien Berichte über den ruchbaren Missbrauch von Facebook-Daten durch Cambridge Analytica auf. Am 29. April war Twitter dran. Obwohl ich mich nun intensiv mit diesen Fragen beschäftige, konnte ich nicht eruieren, was Pöhszes und vor allem Illegales geschehen ist. Mein Eindruck ist: Es ist das übliche Geschnatter. Man sagt auch Medien-Agenda dazu. Facebook hat im Moment einen Marktwert von 535 Milliarden \$. Man müsste den größten Flughersteller Boeing (211 Milliarden) und den größten Erdölkonzern Exxon (333 Milliarden) fusionieren um mit Facebook gleich zu ziehen. Facebook sitzt nicht auf schwarzem sondern auf virtuellem Gold: Die Userdaten. Mit diesen Daten kann man das Grundproblem der modernen Konsumwelt lösen. Die Produkte sowie die Gewohnheiten, Lebensstile der Konsumenten werden immer differenzierter. Es gab einst beim Greissler für jede Produktgruppe 1-2 Angebote. Im Supermarkt sind es ein paar Dutzend, im Online-Handel von Amazon ein paar Tausend. Es gab einst in einem gut sortierten Plattengeschäft ein paar Hundert, maximal ein paar Tausend Platten. Heute sind es auf Spotify ein paar Millionen und es kann niemand mehr das Filmangebot auf Netflix überblicken. Es ist zur Herausbildung einer sogenannten [Long-Tail Ökonomie](#) gekommen. Diese stellt die Werbeindustrie vor Probleme. Man kann für die Unzahl der Nischenprodukte - die in Summe den Hauptmarkt ausmachen - nicht mehr allgemein werben. Eine klassische Lösung war die Gründung von Spezial-Zeitschriften. Die Auflage dieser Zeitschriften ist klein, aber man kann die Werbung wesentlich besser fokussieren.

Facebook war mit seinem Konzept zur richtigen Zeit am richtigen Ort. Es ist die Lösung für die Probleme der Long-Tail Ökonomie. Wobei die Link-Struktur, die Freundschaftslisten, eine zentrale Rolle spielen. Selbst wenn man über einen Benutzer relativ wenig weiß, kann man ihn mit Hilfe der Informationen seines „Freundeskreises“, seiner Likes, einordnen. Der Fachausdruck dafür ist „Collaborative Filtering“.

Vanellus verwendet keine direkte persönliche Information der Twitter-User. Trotzdem erfährt man mit einfachen Mitteln sehr viel über ihre Ansichten und Gewohnheiten, man kann sie Gruppen zuordnen ... Facebook hat - wie Twitter - eine Programmierschnittstelle (API) und es gibt für die wichtigsten Programmiersprachen Bibliotheken zur einfachen Nutzung dieses API. Es ist dokumentiert, welche Daten man mit diesem API bekommt. Das hängt auch davon ab, wie viel man für den Zugang zahlt. Bei einem freien Twitter-Zugang erhält man z.B. keine exakten geographischen Daten der User. Die Daten sind vorhanden und man kann nachlesen, wie viel diese Zusatzinformation kostet. Beim Twitter-„Skandal“ hat Cambridge-Analytica die gesammelten US-Tweets der letzten 5 Monate gekauft. Sie hatten damit jenen Datenbestand, denn ich (abgesehen von der Lokalität) durch kontinuierliches Herunterladen aufgebaut habe bzw. aufbaue. Es ist kein Ruhmesblatt, dass sie diese Daten nicht eh schon hatten. Twitter verweist in seinen Geschäftsbedingungen ausdrücklich darauf hin: „*Unsere Server erlaubt den Zugriff auf Daten der letzten 6 Tage. Wenn sie ältere Daten brauchen, müssen sie*

diese nachkaufen“. Daraus geht eindeutig hervor, dass Twitter die Daten gegen Bares gerne liefert. Ähnliches gilt für Facebook.

Möglicher Weise hat Cambridge-Analytica die vorhandenen Informationen sehr geschickt genützt (ich bezweifle selbst dies). Ich habe aber noch nie gelesen, durch welchen schmutzigen Trick die Firma an Daten heran gekommen ist, an denen man nicht herankommen sollte und konnte.

Wie auch immer. Als der „Skandal“ Mitte März ruchbar wurde, verlor Facebook an der Börse massiv an Wert. Manche orakelten schon von König Zuckerbergs Glück und Ende. Wie man in der Grafik sieht, ist jedoch auch der Aktienindex S&P-500 stark gefallen. Nicht so stark wie Facebook, aber das ist die normale Relation zwischen einer einzelnen Aktie und einem Index. Mark Zuckerberg ist am 11. und 12. April vor einem Ausschuss des Kongresses „gegrillt“ worden. Ich hatte eher den Eindruck: Die von ihm gesponserten Senatoren ermöglichten einen Gratis-Werbeauftritt. Soweit ich das verfolgt habe, waren die Fragen der Politiker von erschütternder Schlichtheit und Ahnungslosigkeit. „Ob er sein prinzipielles Datenmodell ändern wolle?“. Die analoge Frage für den Boeing Chef wäre „Wollen Sie weiterhin Flugzeuge bauen?“. Ich stellte mir mehrmals die Frage „Sind die so ahnungslos und deppert oder stellen sie sich nur so?“. Mark Zuckerberg hat sich geschickt ein bisserl Asche aufs Haupt gestreut. Er denkt natürlich nicht in Traum daran, sein grundlegendes Geschäftsmodell zu ändern. Er hat bereits vor dem Hearing im Senat guten Willen demonstriert und die Regeln für das API etwas verschärfen lassen. Man hatte nicht bedacht, dass man dadurch den Dating-Dienst Tinder lahm legt. Wenn man sich ein Date ausmacht, will man vorab ein paar Informationen über den Nachtabschnitts-Partner haben. Die Tinder Gemeinde reagierte nicht mit „Wie arg, was die bisher über mich gesammelt haben“ sondern verstört, dass nix mehr geht.

Der FB Kurs hat sich – so wie jener des S&P-500 – inzwischen erholt. Die Nachrichten von Ableben waren etwas verfrüht.



Kurs von Facebook (dunkel-blau) und dem S&P-500 in den letzten 6 Monaten

| Spanischer Originaltext | Deutsche Übersetzung |
|---|--|
| Yo soy un hombre sincero, de donde crece la palma, y antes de morirme quiero echar mis versos del alma. | Ich bin ein aufrichtiger Mensch von da, wo die Palme wächst, und bevor ich sterbe, möchte ich mir meine Verse von der Seele singen. |
| Mi verso es de un verde claro y de un carmín encendido. Mi verso es un ciervo herido que busca en el monte amparo. | Mein Vers ist von hellem Grün und von entflammtem Rot Mein Vers ist ein verwundeter Hirsch der im Gebirge Zuflucht sucht |
| Cultivo una rosa blanca, En julio como en enero, Para el amigo sincero Que me da su mano franca. | Ich ziehe eine weiße Rose heran, im Juli wie im Januar, für den ehrlichen Freund der mir seine offene Hand reicht. |
| Con los pobres de la tierra quiero yo mi suerte echar. El arroyo de la sierra me complace más que el mar. | Mit den Ärmsten der Erde will ich mein Los teilen. Der Bach im Gebirge erfreut mich mehr als das Meer. |

Guantanamera : Text Jose Marti, Musik: Jose Fernandez Diaz.
Außerhalb von Kuba wurde das Lied von Pete Seeger bekannt gemacht.



The Nobel Prize is run by a self-perpetuating committee. They vote for themselves and get the world's publishing industry to jump to their tune.
(Doris Lessing, über das Literatur-Nobelpreis-Komitee).

*Oh Christ. I couldn't care less. ...
I can't say I'm overwhelmed with surprise.
I'm 88 years old and they can't give the Nobel to someone who's dead, so I think they were probably thinking they'd probably better give it to me now before I've popped off.*
(Doris Lessing, nachdem sie den Literatur-Nobelpreis 2007 erhalten hat).

Hört, sieht und riecht die Signale:

Chief, hast schon gehört, heuer gibt es keinen Literatur-Nobel.

Ja, Seniora, es wackelt auch 2019. Ich habe mir die Frage gestellt, ob es wie bei 6 aus 45 dann einen Mehrfach Jackpot gibt oder ob sich die Nobel-Stiftung das Preisgeld ein Zeitl erspart.

Ha, ha Chief. Der Mann, der nur die Zahlen liebt. Ich glaub, du bist noch sauer, weil voriges Jahr der Barde gewonnen hat und willst dich nur über das Komitee lustig machen.

Seniora, bei der Entscheidung für den Barden waren sie schon auf der richtigen Spur, nur hat sie unterwegs der Mut verlassen.

Hmm, Chief, das versteh ich jetzt nicht. Aber wie ich dich kenne, hat es irgendwas mit Zahlen zu tun. Seniora, du liegst nicht ganz falsch. Ich täte einen Korpus von Liedtexten auswählen. Jeder kann daraus sein eigenes Lied machen und einschicken. Es muss nur jedes Wort im Korpus vorkommen. Bei der Preisverleihung wird eine Zahl gezogen. Die ist der Startwert für einen Zufallszahlen Generator und ein Programm generiert auf Grund dessen einen eigenen Liedtext. Wer mit seiner Einsendung am nächsten zu diesem Text ist, dem gehört der Nobel-Jackpot.

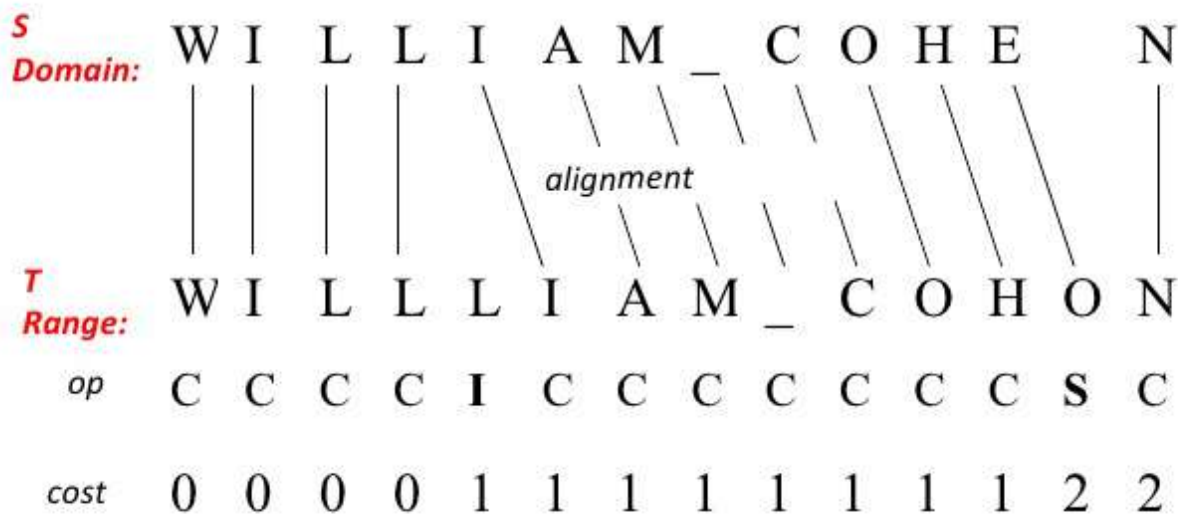
Chief, und wer soll entscheiden, welche von Millionen von Einsendungen am nächsten ist? Das geht doch gar nicht.

Seniora, es ist der Computer schon erfunden worden. Google stellt für eine Millisekunde ein Tausendstel seiner Rechenpower zur Verfügung und der Nobelpreisträger steht fest.

Chief, wie soll ein Computer beurteilen, ob ein Text ähnlich ist?

Levenshtein distance - example

- distance("William Cohen", "William Cohon")



Seniora, das ist einfach. Wenn du dich bei der Google-Suche vertippst findet der Algo trotzdem ähnliche Treffer. Es muss nur innerhalb einer gewissen [Levenshtein-Distanz](#) sein. Das hat der Wladimir Lewenstein schon lange vor Google im Jahr 1965 erfunden. Mathematiker drehen der Welt gerne den Rücken zu, spielen sich mit was und sind deswegen ihrer Zeit voraus.

Könnte mich der Chief an seiner Mathematiker-Weisheit teilhaben lassen?

Gerne, Seniora. Nehmen wir an, es hat wer „*Er steht im Tor*“ eingesandt. Der Nobel-Computer hat aber „*Es steht ein Tier*“ gedichtet. Man kann „*Tor*“ in „*Tier*“ mit 2 elementaren Editieroperationen umwandeln. Aus dem „*o*“ wird ein „*i*“ und zusätzlich fügt man noch das „*e*“ ein. Bei „*Er*“ muss man das „*r*“ in „*s*“ umwandeln, aus „*im*“ das „*m*“ am Ende in ein „*n*“ und am Anfang noch ein „*e*“ einfügen. Seniora, wennst mit gezählt hast, kennst jetzt die Levenshtein-Distanz.

Hmm, Chief, das war ein bisserl schnell. 2 von „*Tor*“ in „*Tier*“, 1 von „*Er*“ nach „*Es*“ und 2 von „*im*“ nach „*ein*“. Ich würde 5 sagen.

Super, Seniors.

Chief, und wie stellt man sicher, dass das Komitee beim Gedicht nicht schummelt?

Ganz einfach, Seniors, indem der Gedicht-Algorithmus bekannt ist. Zusammen mit der Zufallszahl kann jeder auf der Welt nachrechnen, ob sie getrickst haben.

Chief, dann weiß man doch schon im vornherein, wie das beste Gedicht lauten wird.

Seniors, du solltest deinem Geliebten ernst nehmen. Er macht zwar komische Vorschläge, aber er denkt sich was dabei. Die Zufallszahl, mit der alles beginnt, wird so wie bei 6 aus 45 erst bei der Preisverleihung gezogen. Wenn man die nicht kennt, weiß man nicht, was herauskommt.

Chief, ich habe zwar einen Koch und Gärtner, aber von einem Geliebten weiß ich jetzt nix. Glaubst, dass da gute Texte herauskommen?

Seniors, mit den Texten vom Barden werden sie mithalten können und singen kann er wie der Troubadix. Der [DeepDrumpf](#) Bot macht aus den Tweets vom RealDonaldTrump eigene Texte. Die sind nicht blöder als das Original.

Chief, ganz hast du den vorjährigen Nobel doch nicht überwunden.

Seniors, es hat mich nur gestört, dass das Komitee halbe Sachen macht. Man hätte dem Barden den Preis schon viel früher geben sollen.

Chief, es hat sich auch die Doris Lessing darüber aufgeregt, dass man ihr den Preis erst beim Probeliegen im Holzpyjama verliehen hat. Aber warum ist ausgerechnet dir ein Preis für den jungen Barden ein Anliegen?

Seniors, so ein Preis ist doch ein Warnsignal. So wie die gelbe Warnfarbe einer Wespe oder die Augenzeichnung einer sich aufrichtenden Kobra. Da weiß man gleich. Achtung, Gefahr, Hände weg. Wahrscheinlich hätte ich mir damals gedacht „*Hmm, so gut und aufmüpfig und überhaupt kann der Typ doch gar nicht sein, wenn er den Nobel bekommt*“.

Chief, in solchen Fällen zitierst du immer den Papa „*Und wenn alle vom Traunstein hüpfen, hüpf ich noch lange nicht hinterher*“. Jetzt beschwerst du dich, weil du in deiner Jugend ein glühender Fan vom Barden warst und dich das Komitee nicht vor ihm gewarnt hat.

Nein, Seniors, ich beschwer mich, dass das Komitee kein Schild mit der Aufschrift „*Hinunter-hüpfen, eine interessante Erfahrung mit umwerfenden Finale*“ aufgestellt hat. Wenn man das Komitee kennt, kann man das Signal schon interpretieren. Kannst dich noch erinnern wie wir

Hiiiiiiiiiffffe, Chief, jetzt kommst wieder mit dem Geschichtl von der Korsika-Durchquerung daher. Ja Chief, es war genial von dir, dass du jeden Abend auf der Suche nach einem guten Platzl für das Zelt nach dem Schild „Campieren Verboten“ Ausschau gehalten hast. Aber du willst damit nur von deinen



Jugendsünden ablenken. Es hat dich Pete Seeger sehr deutlich gewarnt. Er hat beim Auftritt des Barden den [Stecker gezogen](#).

Seniors, ich gebe zu, die Geschichte war mir bekannt, aber ich habe das vom Pete ziemlich puristisch gefunden. Man kann auch auf der E-Gitarre und E-Harp gute Musik machen.

Chief, ich bin zwar von der braven Beatles-Fraktion und nicht wie du so ein aufmüpfiger Fan des Barden und der Stones. Aber eines weiß ich schon: Darum ist

es den Pete doch gar nicht gegangen.

Ja Seniors, nach seinen eigenen Worten war die Anlage sehr schlecht eingestellt, man konnte den Gesang nicht verstehen. Er hat zum Tontechniker gesagt „*Stell das besser ein*“. Der meinte „*Die wollen das so*“. Da ist er zornig geworden und hat der Stecker gezogen.

Chief, der Barde war trotzdem beleidigt, er hat das Newport-Folk-Festival für die nächsten 37 Jahre gemieden. Das hätte dir doch zu denken geben müssen.



27/05
2018

gez
A.D.

Seniora, du stierst in meinen Wunden. Ich gebe zu, man hätte schon damals wissen können, dass der Barde ein ziemliches Arschloch und der Pete der *hombre sincero* ist. Wie die Revoluzzer-Zeit vorbei war, ist er auf der frommen Welle geritten, und wie man sich im Patriotismus überschlagen hat, war er der Super-Patriot. Normalerweise kann man die Nobel-Preisrede auf der Seite des Komitees nachlesen. Er war in seiner Sierigkeit der erste Preisträger, der sich das Copyright behalten hat und die Rede stattdessen selber auf Amazon verscherbelt. Es gibt ja genug Deppen, die aus ihren Jugend-Torheiten nix lernen und den Barden noch immer anbeten.

Chief, ich dachte, er hat überhaupt keine Rede gehalten.

Seniora, er war nicht bei der Preisverleihung. Das ist schon öfters vorgekommen, weil die Preisträger zu alt waren oder so wie die Jelinek keine Leute aushalten. Die Jelinek hat stattdessen eine artige Videobotschaft geschickt. Das war okay. Aber den Text einer Rede muss jeder abliefern. Das haben sie ihm klargemacht: Keine Rede, kein Cash. Da hat er im letzten Moment doch was geliefert, aber sie dürfen es nicht veröffentlichen.

Chief, ich war nie so ein Fan vom Barden, darum ist das mir wurscht. Aber ein bisschen hat deine Theorie schon was auf sich. Ich habe früher was vom Robert Menasse gehalten. Seit er mit Preisen eingedeckt wird, finde ich ihn nur mehr peinlich.

Seniora, du hast nicht nur was von ihm gehalten. Du bist wochenlang mit stolzgeschwellter Brust durch die Gegend gelaufen, weil er seinen Roman „*Don Juan de la Mancha oder die Erziehung der Lust*“ im Hochland Bildungswerk das erste Mal der Öffentlichkeit vorgestellt hat. Was ist schon die Leipziger Buchmesse gegenüber deinem Bildungswerk. Und wie er uns in seiner Dichterklausur zum Abendessen eingeladen hat, bist du überhaupt hingeschmolzen.

Chief, du bist jetzt nur neidig, weil ich die Kochkünste des Dichters gelobt habe. Er kann auch kochen und vielleicht ist er sogar ein guter Liebhaber. Er kann darüber zumindest besser schreiben wie du.

Seniora, er hat lang und breit über seine missratene Dillsauce lamentiert.

Chief, der RoM ist halt ein selbstkritischer Mensch. Du hättest von einem Vier Hauben Meisterwerk geschwärmt.

Seniora, sein Jammer über die Dillsauce war aber nichts gegen seinen Hader über die Welt, die sein Genie nicht anerkennt.

Ja, Chief. Wahrscheinlich war das sein Unglück und Ende als kritischer Intellektueller. Er ist beim Ringen nach Anerkennung ein Minnesänger der herrschenden und realen EU geworden. Deine Signaltheorie hat aber ein Problem. Es hat in früheren Zeiten diese Preise noch nicht gegeben.

Seniora, in früheren Zeiten gab es noch weit bessere Signale. Für die Deutsche Geistesgeschichte ist das sehr einfach. Wer von der Obrigkeit drangsaliert wurde, ist interessant. Der Rest sind alte, eitle



Menasses. Der Fichte wurde in Jena wegen angeblicher Atheistischer Umtriebe von der Uni gemobbt. Der [Feuerbach](#) hat den größten Teil seines Lebens in der 46 Häuser Gemeinde Bruckberg verbracht. Sie haben extra eine Polizeistation in diesem Kaff errichtet, damit man ihn und seinen Freundeskreis unter ständiger Beobachtung halten kann. [Hoffmann v. Fallersleben](#) hat die Preußische Regierung die Staatsbürgerschaft entzogen. Er irrte als Staatenloser durch die Lande. [Börne](#) und [Heine](#) wurden überhaupt ins Exil vertrieben.

Stopp, Chief, du brauchst mir jetzt nicht die ganze Deutsche Geistesgeschichte aufzählen. Es passt zumindest für den Jose Marti.

Ja, Seniora. Die innigsten Oden auf die Hoamad entstehen in der Ferne. Der Jose Marti hat den Text von Guantanamo im New Yorker Exil geschrieben. Es ist eine Ironie der Geschichte, dass ein vertriebener Staatenloser den Text der Deutschen Nationalhymne verfasst hat.

Chief, und wie passt das zum Herrn Geheimrat? Von dem kann man nicht behaupten, dass er von der Obrigkeit drangsaliert wurde. Falsifiziert

das nicht deine Theorie?

Seniora, es hat jedes Signal Fehler. Es gibt Insekten die schauen absichtlich wie eine Wespe aus, haben aber gar keinen Stachel. Die Fehlerrate muss nur so gering sein, dass das Signal trotzdem brauchbar ist. Außerdem kann man Fehlerkennungs-Kodes einbauen. Beim IBAN sind die ersten beiden Zahlen nach dem „AT“ eine Checksumme. Dieser Wert muss gemäß einer Formel mit den übrigen Zahlen übereinstimmen. Wenn sich jemand beim IBAN vertut, wird das erkannt. Die Überweisung wird nicht durchgeführt. Gute Fehlercodes können – falls die Anzahl der Fehler nicht zu groß ist – sogar das Ergebnis korrigieren.

Chief, man sagt was, was nicht zu deinen Hirngespinnsten passt und du kommst einem mit der Mathematik und sonst noch allem Möglichen daher. Du solltest wissen, dass ich auf diesen Trick schon lange nicht mehr hereinfalle.

Seniora, da musst du dir schon ein besseres Beispiel wie den Geheimrat einfallen lassen, um meine „nur ein von der Obrigkeit Gepiesackter ist ein guter Denker und Dichter“ Theorie zu widerlegen. Der Geheimrat war ein aufgeblasener Wichtigtuier, der sich noch weit mehr überschätzt hat, wie der Menasse.

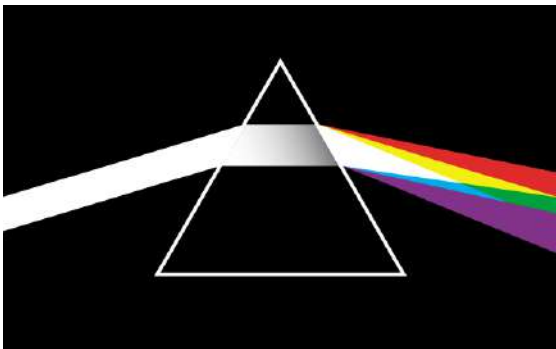
Chief, wenn sich jetzt wer aufbläst, dann bist du es.

Seniora, der Menasse ist zwar der festen Meinung, dass ihm und nicht der Jelinek der Nobel gebührt hat, er weit über den Thomas Mann steht und nur Büchner und der Geheimrat ebenbürtig sind. Aber zumindest kränkt er sich nicht, dass man ihm noch nicht den Physik-Nobel verliehen hat.

Chief, und was hat der Physik-Nobel mit dem Geheimrat zu tun?

Seniora, der Herr Geheimrat hat seine Farbentheorie höher bewertet als sein Dichterisches Werk. Immerhin hat er mit dieser Theorie einen Titanen wie Newton widerlegt. Tatsächlich ist das Ganze nur peinlich. Oder wie es der Wittgenstein ausgedrückt hat: „*Goethe's theory of the constitution of colours of the spectrum has not proved to be an unsatisfactory theory, rather it really isn't a theory at all. Nothing can be predicted by means of it*“. Der Herr Geheimrat hat nicht begriffen, was eine wissenschaftliche Theorie ist, ja er hat nicht einmal begriffen, was ein wissenschaftliches Experiment ist.

Chief, ich weiß, du magst den Geheimrat noch weniger als den Barden. Aber vielleicht könntest du das einer wissenschaftlich ungebildeten Person in verständlichen Worten erklären.



Seniora, Newton hat weißes Licht durch einen schmalen Spalt und anschließend durch ein Glasprisma fallen lassen. Er hat einen schönen Regenbogen bekommen und damit gezeigt, dass weißes Licht aus verschiedenen Spektren besteht. Der Regenbogen kommt zustande, weil sich die Farben im Glas mit unterschiedlicher Geschwindigkeit ausbreiten. Und was macht dieser eingebildete Trottel: Er hält das Prisma gegen eine weiße Wand, sieht kein Spektrum und erklärt: Ich, das GRÖGAZ, habe Newton und seine Farbentheorie widerlegt.

Chief, und was ist daran so trottelhaft?

Seniora, erstens muss man ein wissenschaftliches Experiment genauso wiederholen, wie es ursprünglich gemacht wurde. Wenn bei einem anderen Versuchsaufbau was anderes herauskommt, ist das überhaupt keine Widerlegung des ursprünglichen Werkes. Und zweitens hätte er – wenn er den Newton nur ein bisserl verstanden hätte – schon vorher wissen müssen, dass bei seinem Experiment kein Regenbogen herauskommt.

Chief, was ist der Unterschied?

Seniora, der Unterschied ist, dass das Licht im Newtonschen Experiment aus einer Richtung kommt, beim Herrn Geheimrat hingegen aus vielen Richtungen. Es heben sich bei seinem Experiment die Effekte gegenseitig auf. Die verschiedenen Regenbogen mischen sich wieder zu weißem Licht

zusammen. Nur am Fensterkreuz hat er leichte Regenbogen Effekte gesehen, weil das Fensterkreuz das Licht etwas gebündelt hat.

Chief, ich will mit dir nicht über Physik streiten. Selbst wenn sich der Geheimrat in dieser Frage vergaloppiert hat, dann kann er trotzdem ein guter Literat gewesen sein.

Seniora, ein berühmtes Zitat aus dem Faust ist:



*Habe nun, ach! Philosophie,
Juristerei und Medizin,
Und leider auch Theologie
Durchaus studiert, mit heißem Bemühn.
Da steh ich nun, ich armer Tor!
Und bin so klug als wie zuvor;*

Das ist doch ein totaler Blödsinn. Wenn man etwas ernsthaft betreibt, egal ob es Kochen, Tischlern, Musizieren, Bogenschießen oder von mir aus auch Theologie ist, wird man gescheiter und ein reiferer Mensch. Aber man muss dafür mindestens 10.000 Stunden hart hackeln. Der Geheimrat kannte diese Erfahrung nicht, dazu war er zu faul. Talent alleine ist zu wenig.

Knüppel aus dem Sack

[1] Von allen Wünschen in der Welt
Nur Einer mir anjetzt gefällt,
Nur: Knüppel aus dem Sack!
Und gäbe Gott mir Wunschesmacht,
Ich dächte nur bei Tag und Nacht,
Nur: Knüppel aus dem Sack!

Dann brauchst' ich weder Gut noch Gold,
Ich machte mir die Welt schon hold
Mit: Knüppel aus dem Sack!
Ich wär' ein Sieger, wär' ein Held,
Der erst' und beste Mann der Welt
Mit: Knüppel aus dem Sack![1]

Ich schaffte Freiheit, Recht und Ruh
Und frohes Leben noch dazu
Beim: Knüppel aus dem Sack!
Und wollt' ich selbst recht lustig sein,
So ließ' ich tanzen Groß und Klein
Beim: Knüppel aus dem Sack!

O Märchen, würdest du doch wahr
Nur Einen einz'gen Tag im Jahr,
O Knüppel aus dem Sack!
Ich gäbe drum, ich weiß nicht was,
Und schlüge drein ohn' Unterlaß:
Frisch: Knüppel aus dem Sack
Aufs Lumpenpack!
Aufs Hundepack!

Chief, du pickst dir eine Stelle heraus und hackst darauf los. Seniora, die Stelle ist sprichwörtlich geworden. Du kannst aber auch die Bürgschaft hernehmen. Die Vorstellung, der Hitler hätte den Graf Stauffenberg wieder laufen lassen und anschließend zu seinem engsten Freundeskreis gehört, ist doch jenseits von dumm und dümmer. Generationen von Schülern sind mit diesem Schmus gequält worden. Warum stand nicht [Knüppel aus dem Sack](#) vom Hoffmann v. Fallersleben am Unterrichtsplan? Fack ju Gemeinrad.

Chief, darf ich dich in aller Freundschaft auf ein kleines Detail aufmerksam machen. Die Bürgschaft ist vom Schiller. Seniora, ist das der mit „Loch in Erde, Bronze rin. Glocke fertig, nix bim bim“? Der Meister hat den Klöppel vergessen. Chief, du hast von der Literatur noch weniger Ahnung wie der Geheimrat von der Physik, velwechsert Schirrel und Geheimrat. Trotzdem führst du dich als großer Literaturkritiker auf. Seniora, wenn man von etwas keine Ahnung hat braucht man Signale. Ich vertrau der Staatsmacht, dem Zensor und Ludwig Börne. Der Börne hat – vom Pariser Exil aus – den Geheimrat schon damals sehr feinsinnig vom Podest geholt. Laut dem Börne passen auch im Erbkönig die Motive nicht zueinander. Super Chief, der Erbkönig ist wirklich vom Geheimrat. Es gibt übrigens eine geniale [Lesung vom Oskar Werner](#).

Seniora, der Oskar Werner kann aus dem Wiener Telefon lesen und man glaubt, es ist große Kunst.

Chief, und wie erklärst du die Bewunderung für den Geheimrat. Ganz deppert waren die auch nicht.

Seniora, das hat nix mit deppert zu tun. Der gebildete Mensch verwendet den Großteil seiner Intelligenz dazu um die eigene Existenz schön zu färben. Mit dem Rest führt er Phöszes im Schilde.



Wenn man den Fürstenknecht zum Gott erhebt, weiht man den eigenen Opportunismus. Mark Twain schildert in seinem [Bericht über Bayreuth](#) mit beißender Ironie, wie die Deutschen Bürger vor der Fürstenloge buckeln. Ein Ami versteht nicht, warum jemand, der in seinen Leben noch nie was geleistet hat, gesellschaftliche Anerkennung genießt. Chief, du ziehst so über den Geheimrat und seinen Faust her, warst aber ganz empört, wie wir mit der Petra in Weimar eine Faustaufführung gesehen haben und der Regisseur das Stück verhunzt hat. Seniora, was soll man zu einer Aufführung sagen, bei der Mephisto unvermittelt „Einmal um die ganze Welt“ von Karel Gott anstimmt? Chief, du hättest so wie der Ullrich Magenschmerzen bekommen können. Seniora, ich bin mit der Petra nicht verheiratet. Sie hat extra wegen uns die Karten in Weimar ergattert, da kann ich doch nicht sagen, lass mich mit diesem depperten Stück in Ruhe. Außerdem mag Ullrich den Geheimrat noch weniger wie ich. Er hat aber ein anderes Signal. Der Geheimrat stand unterm Erich hoch in Kurs und alles, was unterm Erich hoch im Kurs war, bereitet ihm Magenschmerzen.

Chief, und wie passen die Signale vom Ullrich mit deinen zusammen? Seniora, ob und was ein Signal ist und wie man es interpretiert, hängt auch vom Empfänger ab. Ein Hund riecht weit besser wie wir. Trotzdem oder gerade deswegen findet er den Geruch von Scheiße oder verschwitzten Schuhen hoch interessant. Das ist für ihn Big Data.



Chief, du lenkst schon wieder ab. Seniora, es gibt eine hohe Korrelation zwischen meinen und Ullrichs Signalen. Der Erich gab sich zwar revolutionär, tatsächlich war es eine spießig-öde Biedermänner-Diktatur. Die haben den Herrn Geheimrat bei jeder passenden und unpassenden Gelegenheit [zitiert](#).

Chief, so leidend warst du bei der Aufführung gar nicht. Du hast mit der Petra kichernd geschwätzt.

Ja, Seniora, die Petra hat gemeint, dass in der Liste der Studien die Mathematik fehlt. Das war eine sehr feinsinnige Beobachtung. Zumindest von der Mathematik hat der Herr Universaldilettant die Finger gelassen. Er hat erkannt, dass er kein Gauß ist. Nur muss man dann auch die Physik meiden. Physik ist Mathematik plus Experiment.

Chief, ich fand euer Verhalten eher peinlich.

Seniora, es war dir peinlich, dass uns der von uns sitzende Biederbürger mit Blicken ermordet hat. Ein Stück vom Geheimrat ist ein sakraler Akt, selbst wenn auf der Bühne Karel Gott gesungen wird. Da schwätzt man nicht und macht keine Anmerkungen zum Text.

Chief, du warst nachher noch ziemlich provokant. Er hat – wie es sich gehört – die Aufführung beklatscht, und du fragst ihn am Ausgang „Sagen Sie, hat Ihnen das wirklich gefallen?“.

Seniora, er hat leicht säuerlich geantwortet „Man muss alles vergessen, was man über das Stück weiß. Dann war es sehr gut“.

Chief, ich hab mitgehört. Da warst du ziemlich Schmädstad.

Ja, Seniora, dazu ist mir nichts eingefallen. Ich weiß bis heute nicht, ob die Antwort sehr klug oder sehr deppert war.

Siehst Chief, manchmal sind die Signale der Welt doch nicht so leicht zu deuten.

Seniora, mir ist die Welt des Biederbürgers, der Oberschul- und Hofräte vollkommen fremd. Ich versteh diese Leute nicht, ich habe keine Nase dafür.

Chief, vielleicht würdest du diese Welt durchaus interessant finden, wenn du einen feineren Geruchssinn hättest.

Ja, Seniors, das hat was auf sich. Es könnte sich wie bei der Scheiße verhalten. Für uns riecht sie nur ekelig, für einen Hund enthält es ein Bukett von Duftnoten.

Chief, du wirst alt und mild. So nett habe ich dich noch nie über Biederbürger sprechen gehört.

Ja, Seniors, vor allem alt.

Chief, ich weiß, es ist dir peinlich. Aber du wirst auch mild. Du suchst verzweifelt nach Signalen um erklären zu können, warum das alles ein Schas ist. Früher hast du so markige Sprüche „wenn ich einmal merke, dass ich ein Biederbürger geworden bin, geb ich mir die Kugel“ von dir gegeben. Das hab ich schon lang nicht mehr gehört. Weil du einer geworden bist.

Seniors, ich taumle nach diesem rechten Uppercut in den Seilen.



Chief, du jammerst, dass deine Göttinnen unter der Erde sind und die jungen, chicen Dirndl nicht singen können. Statt das Abendland, lässt du den Blues untergehen. Die Verklärung der Vergangenheit ist ein typisches Merkmal des Biederbürgers.

Seniors, die Göttinnen sind unter der Erde, aber es gibt interessante Nymphen. Schau, ich zeigt dir die [Sofie Reed](#).

Chief, von der Figur her könnt sie beim Wagner auftreten. Aber gut, sie ist wenigstens kein chices Mädel.

Chief, kennst du sie von deinem Harp-Lehrer?

Nein, Seniors, von einem [Werbevideo](#).

Hmm Chief, ich überleg gerade, für was die Frau werben könnte?

Seniors, das ist jetzt einfach. Es ist ein Werbevideo für Lee Oskar Harps. Da darf man schon ein bisserl eine wilde Henn sein.

Chief, dann hat es doch mit dem Stephan zu tun.

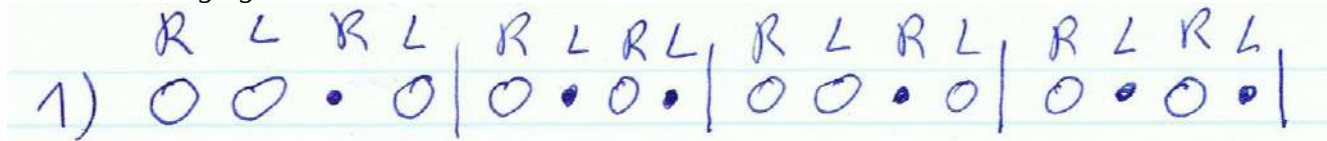
Ja, Seniors, er hat gesagt, ich soll mir Blues-Harp Nummern anhören. Nachdem ich eine Lee Oskar spiele, habe ich mir die Sophie angeschaut. Man kann davon ausgehen, dass der Oskar nur jemanden präsentiert, der harpen kann.

Chief, hast mit Stephan darüber geredet?

Ja, Seniors, sie ist ein ungeschliffener Harp-Diamant.

Chief, du spielst nicht eine Lee Oskar. Das ist schon fast so arg wie bei den Büchern. Du spielst den Tennessee Waltz in jeder Dur und zum Drüberstreuen in A-Moll und quälst mich mit der Frage „Was hat dir jetzt am besten gefallen?“. Ich darf dann ja nicht sagen, es ist immer das gleiche Gedudel.

Seniors, in A-Moll ist es nicht das gleiche Gedudel. Der Waltz klingt in A-Moll nach „Maria durch den Dornwald ging“.



Chief, du lenkst schon wieder ab. Da hast für deine Trommel-Rehleins normale Noten in diese Kreiserl und Punkterl übertragen, damit sie sich leichter tun. Durch die Hintertür hast Noten lesen gelernt. Auch wenn du es nicht hören willst: Du bist alt, mild und biederbürgerlich geworden.

Ja, Seniors, denn Liebe ist, wenn er dank ihrer weisen Ratschläge mild und gelassen wird.

Ja, Chief, denn wahre Liebe ist, wenn er dank ihres Einflusses im Theater nur depperte Fragen stellt, mit der Petra ein bisserl schwätzt und nicht gleich grob herum pöbelt und ein Handgemenge anzettelt.

Verwendete Literatur:

Christoph Weiß (Hg.): Ludwig Börnes Goethe-Kritik.