

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Jure Demšar

**Razvoj spletnega portala za zbiranje in  
ocenjevanje kratkih finančnih sporočil**

DIPLOMSKO DELO  
NA UNIVERZITETNEM ŠTUDIJU

Ljubljana, 2010

Št. naloge: 01713/2010

Datum: 05.10.2010



Univerza v Ljubljani, Fakulteta za računalništvo in informatiko izdaja naslednjo nalogo:

Kandidat: **JURE DEMŠAR**

Naslov: **RAZVOJ SPLETNEGA PORTALA ZA ZBIRANJE IN OCENJEVANJE  
KRATKIH FINANČNIH SPOROČIL**  
**THE DEVELOPMENT OF WEB PORTAL FOR ASSEMBLY AND  
RANKING OF SHORT FINANCE MESSAGES**

Vrsta naloge: Diplomsko delo univerzitetnega študija

Tematika naloge:

V diplomii predlagajte in razvijte spletno rešitev namenjeno pregledovanju aktualnih finančnih informacij na spletu. Portal naj finančne informacije pridobi iz sistema Twitter. Razvijte algoritme za iskanje po tovrstnih novicah in za njihovo rangiranje oziroma ocenjevanje informativnosti. Uporabnost sistema s stališča ocene možne koristnosti prikazanih informacij ovrednotite s praktičnim preiskusom.

Mentor:

prof. dr. Blaž Zupan



Dekan:

prof. dr. Nikolaj Zimic

Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

*Besedilo je oblikovano z urejevalnikom besedil  $\LaTeX$ .*

UNIVERZA V LJUBLJANI  
FAKULTETA ZA RAČUNALNIŠTVO IN INFORMATIKO

Jure Demšar

**Razvoj spletnega portala za zbiranje in  
ocenjevanje kratkih finančnih sporočil**

DIPLOMSKO DELO  
NA UNIVERZITETNEM ŠTUDIJU

Mentor: prof. dr. Blaž Zupan

Ljubljana, 2010

Rezultati diplomskega dela so intelektualna lastnina Fakultete za računalništvo in informatiko Univerze v Ljubljani. Za objavljanje ali izkoriščanje rezultatov diplomskega dela je potrebno pisno soglasje Fakultete za računalništvo in informatiko ter mentorja.

*Besedilo je oblikovano z urejevalnikom besedil  $\LaTeX$ .*

Namesto te strani **vstavite** original izdane teme diplomskega dela s podpisom mentorja in dekana ter žigom fakultete, ki ga diplomant dvigne v študentskem referatu, preden odda izdelek v vezavo!



# IZJAVA O AVTORSTVU

diplomskega dela

Spodaj podpisani Jure Demšar,

z vpisno številko 63040025,

sem avtor diplomskega dela z naslovom:

Razvoj spletnega portala za zbiranje in ocenjevanje kratkih finančnih sporočil

S svojim podpisom zagotavljam, da:

- sem diplomsko delo izdelal/-a samostojno pod mentorstvom prof. dr. Blaža Zupana,
- so elektronska oblika diplomskega dela, naslov (slov., angl.), povzetek (slov., angl.) ter ključne besede (slov., angl.) identični s tiskano obliko diplomskega dela,
- soglašam z javno objavo elektronske oblike diplomskega dela v zbirki "Dela FRI".

V Ljubljani, dne 3.12.2010

Podpis avtorja:



# Zahvala

Za pomoč pri izdelavi diplomske naloge se zahvaljujem mentorju prof. dr. Blažu Zupanu in članom Google skupine "Twitter Development Talk".



# Kazalo

<b>Povzetek</b>	<b>2</b>
<b>Abstract</b>	<b>3</b>
<b>1 Uvod</b>	<b>4</b>
<b>2 Problemska domena</b>	<b>6</b>
2.1 Socialno omrežje Twitter . . . . .	6
2.2 Vrste sporočil na Twitterju . . . . .	7
2.3 Osnovne funkcionalnosti omrežja Twitter . . . . .	7
2.4 Finančne informacije na omrežju Twitter . . . . .	8
2.5 Spletni portal StockTwits . . . . .	9
<b>3 Portal TwittRank</b>	<b>11</b>
3.1 Orodja in tehnike . . . . .	11
3.2 Primeri uporabe . . . . .	13
3.3 Baza podatkov . . . . .	16
3.4 Zbiranje finančnih sporočil . . . . .	20
3.5 Osnovne operacije nad sporočili . . . . .	21
3.6 Prijava na portal . . . . .	23
3.7 Pošiljanje novih sporočil . . . . .	25
3.8 Ocenjevanje sporočil . . . . .	28
3.9 Profil . . . . .	30
3.10 Spremljanje prometa . . . . .	32
3.11 Oblikovanje . . . . .	33
<b>4 Testiranje</b>	<b>35</b>
<b>5 Sklepne ugotovitve</b>	<b>39</b>



# Povzetek

Cilj diplomskega dela je bil razvoj rešitve, ki nam omogoča enostavno pregledovanje aktualnih finančnih informacij na internetu. V ta namen smo razvili spletni portal, ki pridobiva informacije iz socialnega omrežja Twitter. Za lažje pregledovanje informacij smo razvili algoritme za iskanje in filtriranje novic. Implementirali smo tudi tehnike za ocenjevanje novic, ki nam omogočajo ločevanje informativnih novic od teh, ki ne vsebujejo zelenih informacij. Poleg teh dveh glavnih funkcionalnosti razviti portal omogoča tudi komentiranje že obstoječih sporočil, objavljanje novih in sledenje zanimivim pošiljateljem oziroma izbranim delnicam. Sistem smo preverili na primeru enomesečnega valutnega trgovanja. Z doseženimi rezultati smo bili zadovoljni. Naš portal se je izkazal za uporabno in dobro rešitev omenjenih problemov.

## **Ključne besede:**

Finančne informacije, Twitter, načrtovanje programske opreme, spletni portal

# Abstract

In our thesis we focused on developing a solution that allows us to browse actual financial information on the internet easily. We have developed a web portal, which gathers financial information from the social network Twitter. We have also implemented algorithms for searching and filtering the news, and for ranking those messages. The ranking system helps us to distinguish between news that include useful information and those who do not. Besides those two main functionalities our portal also has mechanics for commenting existing messages, posting new messages and follow interesting posters or stocks. We tested the usefulness of our system at the problem of monthly foreign exchange trading. We were satisfied with our results. The portal proved as a good and useful solution to the aforementioned problems.

## **Key words:**

Financial informations, Twitter, software planning, web portal

# Poglavje 1

## Uvod

Finančne naložbe postajajo vse bolj popularne. Če bi stvar zelo posplošili, nam dobre finančne naložbe prinašajo materialna sredstva brez vsakih fizičnih naporov. Cilj vsakega, ki se z njimi ukvarja, je čim večji zaslužek. Ena bolj popularnih in varnih naložb je vezava denarja na banki. Problem pri vezavi je razmeroma majhna donosnost, le okoli 5% letno. Zato so se kmalu pojavila podjetja, ki z vašim denarjem upravljajo namesto vas, in obljublajo donosnost okoli 15% letno, toda pri vseh poslovanjih vzamejo visoko provizijo.

Zaradi majhne donosnosti pri vezavi v bankah in visoke provizije finančno naložbenih podjetij so ljudje začeli s svojim denarjem upravljati kar sami, brez provizij. Problem pri tem je razmeroma veliko tveganje, če nismo zelo dobro informirani o mnogih faktorjih, ki vplivajo na gibanje vrednosti delnic in valut. Z rastjo popularnosti finančnih naložb na internetu raste tudi število informacij, ki nam pomagajo pametno in dobro upravljati z našimi prihranki. Na internetu danes lahko najdemo ogromno finančnih informacij. Iskanje po teh je težavno in dolgotrajno, saj so raztresene po mnogih lokacijah. Naslednji problem je ocenjevanje njihove kvalitete, saj je ta navadno vnaprej določena s slovesom njihovega avtorja. Tako smatramo, da so dobre tiste informacije, ki jih je napisal znan avtor, kar pa ne drži vedno. Posledica tega je tudi to, da se je novim dobrim avtorjem težko prebiti v ospredje. Poleg tega pa so finančne informacije zelo časovno občutljive - aktualne so le kratek čas, nato pa zastarajo in postanejo neuporabne, zato je omrežje Twitter idealen sistem za posredovanje le-teh.

Naša želja je tako narediti spletni portal, ki učinkovito reši te probleme. Razviti portal iz Twitterja zbira finančna sporočila in ponuja učinkovite mehanizme za urejanje, filtriranje in iskanje po večih različnih kriterijih. Uporabniki imajo tudi možnost ocenjevati posamezna sporočila, ocene (angl. *Rank*)

se beležijo v bazi podatkov, tako da ima vsak uporabnik svojo oceno (angl. *UserRank*), ki nam pove, kako dobra sporočila je pošiljal v zadnjem času. S pomočjo ocene uporabnikovega vpliva na socialnem omrežju (angl. *KloutScore*) nato izračunamo skupno oceno informacije (angl. *TweetRank*). Ta pa uporabnikom, ki nimajo znanja na tem področju, pomaga ločiti dobre informacije od slabih. S pomočjo dobrih informacij se nato uporabniki lahko odločajo za boljše in pametnejše naložbe, kot so se v preteklosti.

# Poglavje 2

## Problemska domena

### 2.1 Socialno omrežje Twitter

Socialno omrežje je struktura zgrajena iz posameznikov ali organizacij, kjer ponostavljeno, posamezniki oziroma organizacije predstavljajo vozlišča v grafu, ta pa so med seboj povezana z ozirom na prijateljstva, sorodstva, skupne interese, razmerja, trgovanja, in podobnimi.



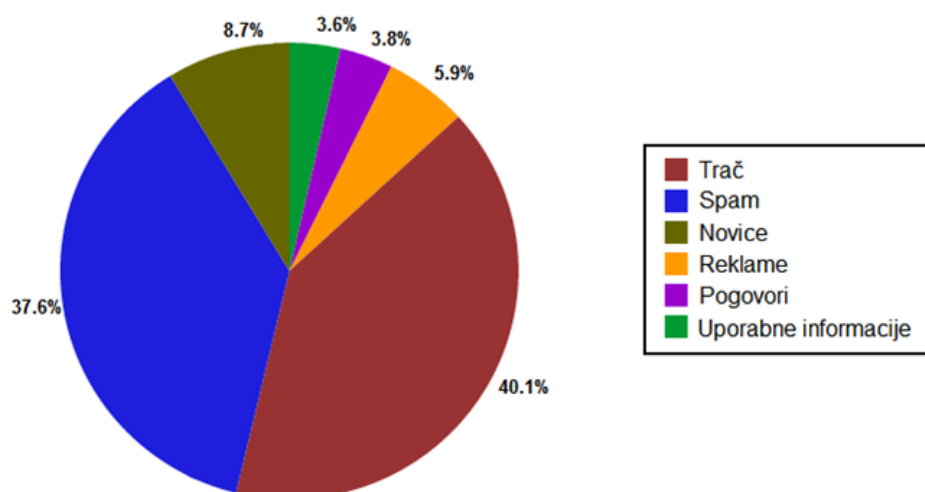
Slika 2.1: Spletni portal Twitter.

Socialno omrežje Twitter je spletni portal, ki uporabnikom omogoča ob-

javljanje in pregledovanje sporočil z največjo dolžino 140 znakov (slika 2.1). Twitter je bil predstavljen javnosti leta 2006 in je zaradi njegove enostavnosti kmalu postal popularen za pisanje kratkih blog sporočil. Iz dneva v dan ga je uporabljajo več in več uporabnikov in tako še danes velja za eno najhitreje rastočih omrežij, registriranih ima kar 100 milijonov uporabnikov. Uvrščen je med 15 najbolj popularnih strani na spletu.

## 2.2 Vrste sporočil na Twitterju

Twitter se večinoma uporablja za posredovanje sporočil zabavne narave (slika 2.2). Na dan se na njem objavi 50 milijonov sporočil, le peščica izmed njih pa je finančnih, zato jih je brez posebnega iskalnega mehanizma izredno težko najti.



Slika 2.2: Vrste informacij na Twitterju.

## 2.3 Osnovne funkcionalnosti omrežja Twitter

Twitter nam omogoča naslednje operacije [2]:

- Objavljanje sporočil – objavljamo lahko sporočila do dolžine 140 znakov. V večini primerov so sporočila javna, imamo pa tudi nekaj osnovnih nastavitev za zasebnost.

- Brisanje sporočil – svoja sporočila lahko brišemo.
- Sledenje – lahko sledimo nekemu uporabniku, tako vidimo na našem profilu vsako objavo tega uporabnika.
- Odgovor – odgovor na sporočilo nekoga drugega.
- Posredovanje – še enkrat objavimo sporočilo nekoga drugega, da ga vidijo še vsi tisti, ki nam sledijo.
- Sezname – skupine uporabnikov z enakim zanimanjem se povezujejo v sezname uporabnikov, ki jim nato mi lahko sledimo.

## 2.4 Finančne informacije na omrežju Twitter

Med vsemi milijoni sporočil se najde tudi peščica takih, ki vsebujejo finančne informacije (slika 2.3). Uporabniki Twitterja, ki jih zanimajo te informacije, so se zmenili, da bodo ta sporočila označevali, tako da jih lahko ločimo od ostalih. Dogovor je, da vsako sporočilo finančne narave vsebuje znak \$, temu pa nato sledijo črke, ki bolj natančno določajo tip sporočila oziroma določajo institucijo o kateri sporočilo poroča. Na primer \$GOOG nam pove, da bo v besedilu sporočila informacija o podjetju Google, \$\$ pomeni splošno informacijo, recimo \$EURUSD pa govori o razmerju med evrom in ameriškim dolarjem.



Slika 2.3: Finančna sporočila na Twitterju.

S pomočjo teh pravil lahko s Twitter vmesnikom API med vsemi sporočili poiščemo tiste, ki vsebujejo finančne informacije za predmet te naloge.

## 2.5 Spletni portal StockTwits

Spletni portal StockTwits (<http://stocktwits.com/>) je najbrž najbolj znan portal, ki iz Twitterja zbira finančne informacije. Uporablja ga preko 75 tisoč uporabnikov. Britanska revija Times ga je za letošnje leto uvrstila med 50 najbolj uporabnih spletnih strani (<http://www.time.com>).

The screenshot shows the StockTwits website interface. At the top left is the StockTwits logo with the tagline 'Real Investors. Real Ideas. Real Time.' To the right is a badge from TIME MAGAZINE naming it a '2010 50 BEST WEBSITES WINNER'. Below the header, there's a navigation bar with 'Join StockTwits', 'Share Ideas & Learn From Professional Investors', and a 'Sign In' button for existing members. The main content area is divided into several sections: a 'Join StockTwits' sidebar with benefits like 'Join a Professional Community of Thousands' and a 'Sign up now' button; a search bar for 'Symbol or User'; a 'Hourly Top Trending Tickers' section with line charts for symbols like \$ARNA, \$STUDY, \$XLF, \$SPY, \$QQQQ, \$WFC, \$DIA, \$KBE, \$USDJPY, and \$GBPNZD; and a 'Watch StockTwitsTV LIVE' section. The central part of the page features a stream of tweets from users, with 'TraderFlorida' being a prominent user. The footer contains copyright information for 2010 StockTwits and a disclaimer stating that the site is not a financial advisor.

Slika 2.4: Spletni portal StockTwits.

StockTwits iz Twitterja zbira finančne informacije (slika 2.4), te pa lahko uporabniki prebirajo z uporabo brskalnika. Omogoča sledenje določenim uporabnikom (zainteresirane uporabnike obvesti, ko nek drug uporabnik objavi nekaj novega), iskanje in objavljanje novih sporočil. Zelo dobra funkcionalnost je, da si lahko zgradimo svoj seznam delnic, ki nas zanimajo, in nato spremljamo samo sporočila, ki se tičejo teh delnic.

Glavne pomanjkljivosti portala StockTwits so, da še vedno ne vemo, katera informacija je dobra in katera ne, prav tako nimamo komentiranja posameznih novic. Lahko jih komentiramo samo tako, da objavimo novo sporočilo.

# Poglavje 3

## Portal TwittRank

### 3.1 Orodja in tehnike

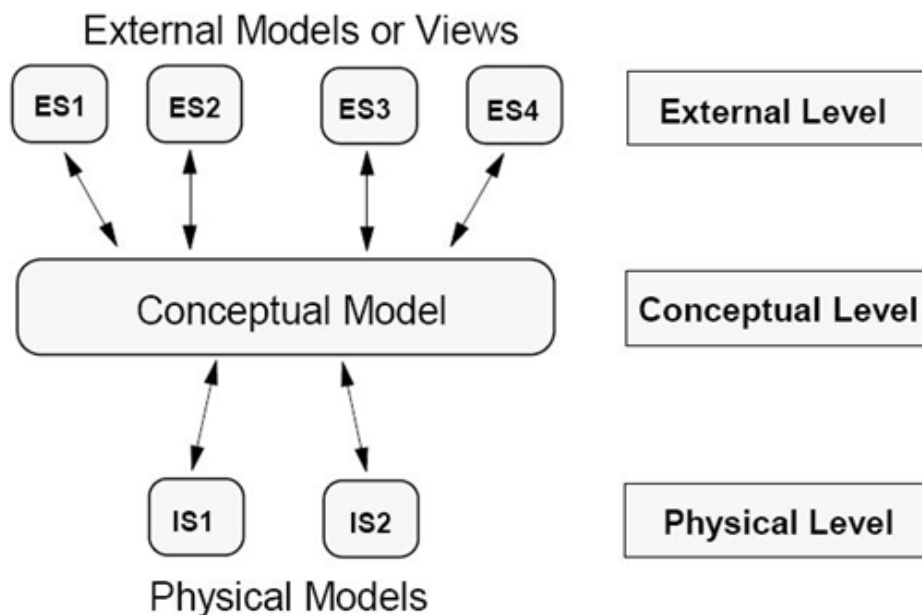
Pri načrtovanju in implementiranju portala smo uporabili več različnih orodij in tehnik. Za izgradnjo modela primerov uporabe in modelov podatkovne baze smo uporabili orodje Power Designer. To je orodje podjetja Sybase, ki velja za enega najbolj priljubljenih programov za načrtovanje programske opreme. Teče na operacijskem sistemu Microsoft Windows. Poleg samega načrtovanja nam ponuja tudi mnoge funkcije za samodejno izgradnjo programske kode in skript SQL [7].

Za postavitev baze smo uporabili Microsoft SQL Server, ki podpira relacijske modele baz. Osnovna jezika, ki se na tej platformi uporabljata, sta T-SQL in ANSI-SQL.

Do baze podatkov dostopamo s tri nivojsko arhitekturo (ANSI-SPARC arhitektura) [6]. Namen te je, da ločimo uporabniški pogled na bazo in njeno fizično implementacijo (slika 3.1). S tem:

- dovoljuje uporabniško prilagojene poglede na podatke - vsak lahko sam prilagodi pogled na podatke, toda podatki so še vedno skupni za vse uporabnike,
- uporabniki ne vidijo točno, kako je baza fizično shranjena,
- administrator lahko popravlja strukturo baze, ne da bi s tem spremenil poglede uporabnikov.

Pri izgradnji vmesnika za dostop do baze sem uporabil brezplačno orodje OR.NET (<http://www.ornetmapper.com/>), ki nam iz že postavljene baze na



Slika 3.1: ANSI-SPARC tri nivojska arhitektura.

Microsoft SQL strežniku samo zgradi kodo, ki preko konceptualnega nivoja skrbi za dostope do baze. V tej kodi imamo vse potrebne funkcije za upravljanje podatkov (dodajanje, spreminjanje, brisanje, ...) (slika 3.2).

Za pisanje kode programa smo uporabili skupino orodij Microsoft Visual Studio 2010. Portal je napisan v jeziku ASP.NET C#, uporablja pa tudi nekaj gradnikov AJAX.

Microsoft Visual Studio [5] je integrirano okolje (IDE) za razvijanje programske opreme. Uporablja se za razvoj spletnih strani, grafičnih programov, spletnih aplikacij, Windows strežnih programov ... Paket Microsoft Visual Studio 2010 vsebuje vsa potrebna orodja (urejevalnik kode, razhroščevalnik,...) za razvoj naprednih spletnih portalov, zato za naš portal nismo uporabljali nobenih drugih programov. Poleg tega ima sprogramiranih že ogromno gradnikov, ki jih lahko uporabimo v naših programih (okna za vnos besedila, gumbi, seznami,...).

C# programski jezik je moderen, objektno orientiran jezik, ki ga je razvil Microsoft v okviru njihovega projekta .NET. Kasneje je bil tudi priznan kot ISO (ISO/IEC 23270) in Ecma (ECMA-334) standard. Ecma je evropsko ([http://en.wikipedia.org/wiki/Ecma\\_International](http://en.wikipedia.org/wiki/Ecma_International)) združenje proizvajalcev računalnikov. Zadnja verzija je bila izdana 12. aprila 2010 [8]. Za C#

```
// if already contains do not insert
User u = new UserAssembler().ReadSingleByKey(username);

if (u == null)
    u = new User(username, avatar, kloutscore, 0, user_ID, 0);
else if (u.Avatar != avatar)
    u.Avatar = avatar;
else if (u.KloutScore != kloutscore && kloutscore != 0)
    u.KloutScore = kloutscore;

u.MessageCount++;
new UserAssembler().MakePersistent(u, true);
```

Slika 3.2: Primer uporabe funkcij OR.NET za dostop do podatkovne baze.

smo se odločili, ker je močno podoben Javi, ki smo jo na fakulteti obravnavali ves čas študija. Pri izgradnji spletnih portalov Visual Studio uporablja Microsoftovo ogrodje ASP.NET, ki nam omogoča izgradnjo dinamičnih spletnih strani, ki jih lahko pišemo v kateremkoli jeziku platforme .NET (C#, Visual Basic,...). Uporabili smo tudi nekaj gradnikov AJAX. AJAX (asinhroni JavaScript) je skupina povezanih tehnik za razvijanje spletnih aplikacij, ki se zaganjajo na klientovi strani. Uporabili smo ga predvsem zato, ker podpira asinhrono komunikacijo s strežnikom, ki teče nevidno v ozadju, zato uporabniki lahko nemoteno pregledujejo stran.

Spletne strani narejene v ASP.NET se grafično oblikujejo s pomočjo datotek CSS. Za vizualno oblikovanje gradnikov smo uporabili Adobe Photoshop.

## 3.2 Primeri uporabe

Do našega portala lahko dostopata dva tipa uporabnikov, neregistrirani in registrirani. Za prijavo v sistem uporabimo kar naše Twitter uporabniško ime in geslo. Neregistrirani uporabniki (gosti), lahko uporabljajo naslednje funkcije portala:

- Prijava – prijava v sistem z našim Twitter uporabniškim imenom in geslom. S prijavo preidemo iz gosta v prijavljenega uporabnika.
- Ogled – na osnovni strani imamo pregledno zapisanih zadnjih 10 objav, s pomočjo pod strani pa lahko dostopamo tudi do starejših.

- Iskanje – iskanje po objavljenih sporočilih (iskanje avtorja, delnice, iskanje po vsebini).
- Filtriranje po borzah – če nas zanimajo samo delnice neke določene borzne hiše, lahko s klikom na ustrezen gumb omogočimo prikaz samo te skupine delnic.
- Urejanje – objavljena sporočila lahko uredimo po skupni oceni, oceni Klout, uporabniškem imenu, oceni sporočila in času. Vse rubrike lahko urejamo padajoče in naraščajoče.
- Pregled profilov – gosti lahko pregledujejo profile ostalih uporabnikov, kjer vidijo, katere pošiljatelje sporočil kdo spremlja, kateri spremljajo njih in katere delnice spremlja.

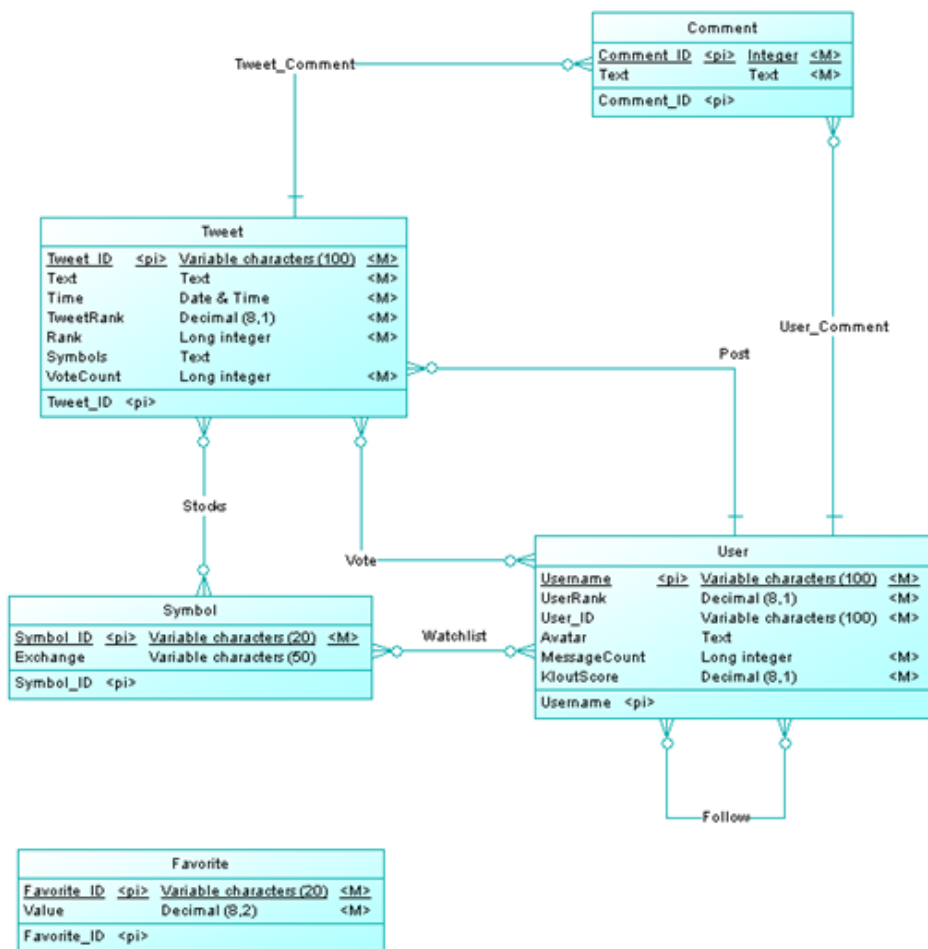
Registrirani uporabniki pa imajo poleg zgornjih funkcionalnosti na voljo še:

- Odjava – z odjavo iz sistema prijavljen uporabnik postane gost.
- Ocenjevanje – vsako sporočilo lahko ocenimo kot dobro (+) ali slabo (-), za vsako sporočilo ima uporabnik na voljo samo en glas.
- Brisanje – svoja sporočila lahko pobrišemo.
- Pošiljanje – prijavljeni uporabniki lahko objavljajo nova sporočila.
- Uvoz prijateljev – na svoji strani profila lahko uvozimo prijatelje (tiste, ki jim sledimo, oziroma sledijo nam) iz Twitterja.
- Komentiranje – posameznim sporočilom lahko dodajamo kratke komentarje.
- Urejanje seznama delnic – na seznam delnic lahko dodajamo delnice, ki nas zanimajo. Iz njega pa lahko pobrišemo tiste, ki jih ne potrebujemo več.
- Urejanje profila – na tej strani lahko odstranimo tiste uporabnike, ki jim nočemo več slediti.
- Sledenje – tistim uporabnikom, ki so nam všeč lahko začnemo slediti.
- Filtriranje po seznamih – prijavljeni uporabniki imajo na voljo še naslednje štiri mehanizme za filtriranje sporočil:

- prikaz naših sporočil,
- prikaz sporočil avtorjev, ki jim sledimo mi,
- prikaz sporočil avtorjev, ki sledijo nam,
- prikaz sporočil, ki se tičejo tistih delnic, ki smo jih dodali na naš seznam zanimivih.

### 3.3 Baza podatkov

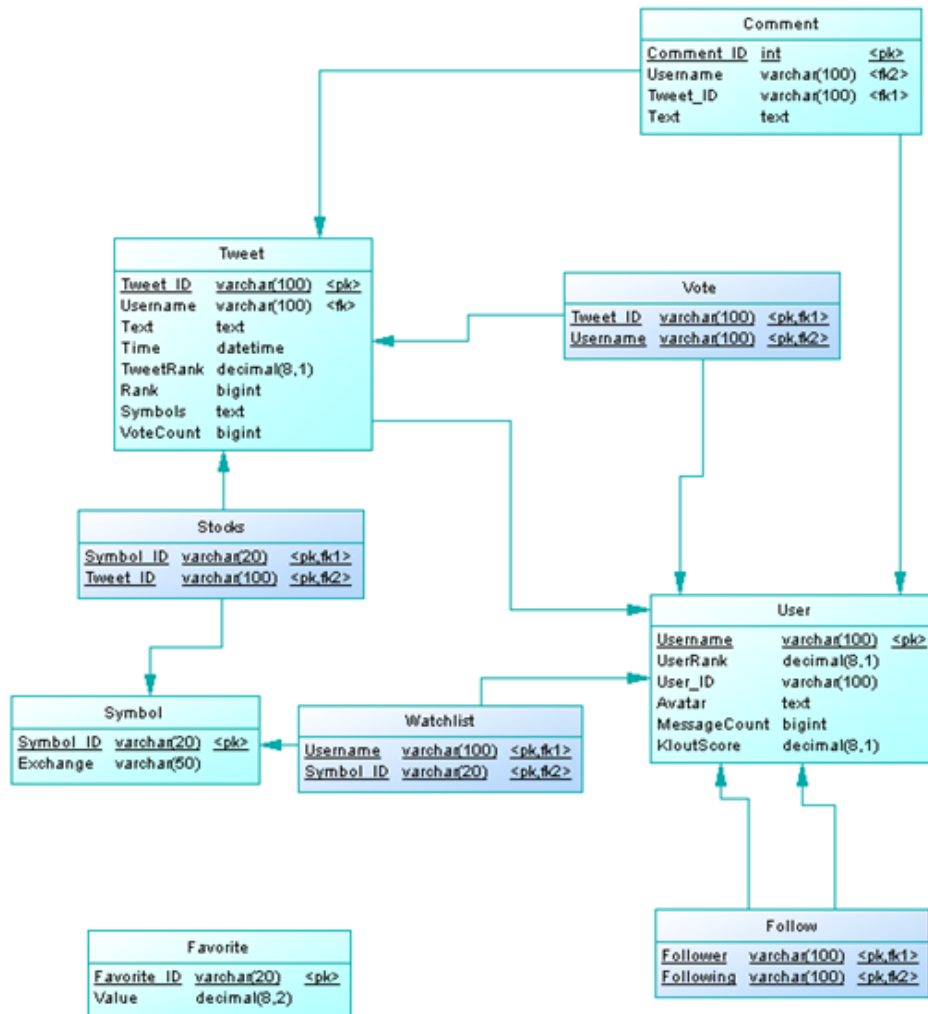
Konceptualni model (slika 3.3) smo zgradili v programu PowerDesigner. Vsebuje 5 relacij in 7 povezav (razmerij) med njimi.



Slika 3.3: Konceptualni model baze podatkov.

Orodje PowerDesigner nam nato iz konceptualnega modela zgradi primeren fizični model. Za natančnejši opis baze bomo uporabili fizični model, ker se nam zdi bolj primeren in bolj nazorno kaže strukturo podatkovne baze.

V fizičnem modelu (slika 3.4) imamo 9 tabel, prav toliko jih bo imela tudi naša baza, ko bo narejena na strežniku. Te tabele so:



Slika 3.4: Fizični model podatkovne baze.

- Tweet – tu so shranjena sporočila. Poleg identifikacijske številke, ki je hkrati primarni ključ tabele, imamo zapisanega tudi avtorja sporočila (tuj ključ - povezava s tabelo User), besedilo, čas in datum, skupno oceno sporočila, oceno iz glasovanja uporabnikov, simbole (delnice, ki nastopajo v sporočilu) in število glasov.
- User – o uporabniku hranimo samo osnovne podatke, uporabniško ime, njegovo oceno, identifikacijsko številko, osebno sliko, število sporočil in njegovo oceno Klout (<http://klout.com/kscore>), ki nam pove kako vpliven je uporabnik na Twitterju.
- Symbol – tu so shranjeni vsi simboli, ki označujejo različne delnice, poleg samega simbola shranimo tudi, na kateri borzi se s tem simbolom trguje, kar potrebujemo za filtriranje po borzah.
- Vote – vsebuje podatke o glasovanju uporabnikov. Vsakič, ko uporabnik neko sporočilo oceni s pozitivnim oziroma negativnim glasom, se v tej tabeli naredi zapis, ki mu onemogoča nadaljnje glasovanje nad tem sporočilom.
- Comment – tu so shranjeni komentarji, ki jih uporabniki prispevajo nad sporočili.
- Follow – v tabeli Follow je zapisano, kdo komu sledi.
- Stocks – tu imamo zapisano, katere delnice nastopajo v posameznih sporočilih. To nam olajša iskanje posameznih delnic po sporočilih.
- Watchlist – tu ima vsak uporabnik shranjene tiste delnice, ki ga najbolj zanimajo, tudi ta tabela služi samo za filtriranje.
- Favorite – v tabeli favorite je shranjenih 5 simbolov, ki so bili popularni v zadnjih dveh urah. Teh 5 delnic in njihove vrednosti prikazujemo na uvodni strani.

Iz fizičnega modela, nam orodje PowerDesigner nato generira ustrezno skripto za kreiranje podatkovne baze. V našem primeru gre za skripto namenjeno strežniku Microsoft SQL Server 2008 (slika 3.5). S pomočjo te nato brez večjih težav v orodju Microsoft SQL Server Management 2008 postavimo podatkovno bazo.

```
if exists (select 1
           from sysobjects
           where id = object_id('Watchlist')
                 and type = 'U')
    drop table Watchlist
go

/*=====*/
/* Table: Comment */
/*=====*/

create table Comment (
    Comment_ID          int          identity,
    Username            varchar(100) not null,
    Tweet_ID           varchar(100) not null,
    Text               text         not null,
    constraint PK_COMMENT primary key nonclustered (Comment_ID)
)
go

/*=====*/
/* Index: Tweet_Comment_FK */
/*=====*/

create index Tweet_Comment_FK on Comment (
    Tweet_ID ASC
)
go
```

Slika 3.5: Del skripte v povpraševalnem jeziku SQL.

### 3.4 Zbiranje finančnih sporočil

Potrebno je bilo sprogramirati rešitev, ki bo iz vseh Twitter sporočil izbiral samo tista s finančno vsebino in jih sproti zapisoval v bazo. To smo realizirali s Twitter API vmesnikom ([http://dev.twitter.com/pages/streaming\\_api](http://dev.twitter.com/pages/streaming_api)) za zbiranje sporočil, s katerim lahko med vsemi objavami na Twitterju z nastavljanjem določenih filtrov dobimo samo tiste, ki jih potrebujemo. To tehniko uporablja že portal StockTwits, ki vsa sporočila objavlja v datoteki XML (slika 3.6), ki je javno dostopna vsem uporabnikom. Naš program vsako minuto preveri, če je v tej datoteki prišlo do kakšnih novih objav, če pa bi se pri StockTwits odločili, da njihovi podatki ne bodo več javni, bi z uporabo Twitter vmesnika API to datoteko XML brez problema začeli izdelovati sami.

```

<?xml version="1.0" encoding="UTF-8" ?>
- <stream>
  <title>All StockTwits Updates</title>
  <page>1</page>
  <total>30</total>
  - <tweets type="array" size="30">
    - <tweet>
      <tweet_text>China buys another 4.4 million bushels of US soybeans this AM; Trade chatter that more business coming $ZS_F
      #farm</tweet_text>
      <tweet_datetime>Mon Oct 18 13:26:23 +0000 2010</tweet_datetime>
      <twitter_username>ArlanFF101</twitter_username>
      <avatar_url>http://a3.twimg.com/profile_images/340374039/suderman150a_normal.jpg</avatar_url>
      <status_id>27733973230</status_id>
      <twitter_user_id>26336221</twitter_user_id>
      <source><a href="http://www.tweetdeck.com" rel="nofollow">TweetDeck</a></source>
    - <stocks type="array" size="1">
      - <stock>
        <ticker>ZS_F</ticker>
      </stock>
    </stocks>
  </tweet>
  - <tweet>
    <tweet_text>Factory production decreased 0.2% reflecting declines inconsumer durable goods, like appliances and furniture -
    #consumercannonball $$</tweet_text>
    <tweet_datetime>Mon Oct 18 13:26:23 +0000 2010</tweet_datetime>
    <twitter_username>Hedgeye</twitter_username>
    <avatar_url>http://a3.twimg.com/profile_images/1138704847/Screen_shot_2010-10-
    06_at_12.01.52_PM_normal.png</avatar_url>
    <status_id>27733964337</status_id>
    <twitter_user_id>18719020</twitter_user_id>
    <source><a href="http://www.tweetdeck.com" rel="nofollow">TweetDeck</a></source>
    <stocks type="array" size="0" />
  </tweet>

```

Slika 3.6: Del datoteke XML z novimi objavami.

Ker je najbolje, da se odkrivanje znanj iz podatkov dogaja nevidno v ozadju, smo ta del portala sprogramirali kot strežni program, ki teče na Microsoft Windows operacijskih sistemih nevidno v ozadju in ne potrebuje nobenih posegov s strani uporabnika. Nastavljen je tako, da se zažene vsakič, ko se na strežniku zažene operacijski sistem Windows. S tem imamo minimalno izgub v primeru nenačrtovanih izpadov strežnika (težave z elektriko, ...). Vsako minuto naš strežni program pregleda datoteko XML in v bazo podatkov zapiše nove objave ter shrani uporabnika, če le-tega še ni v bazi. Iz datoteke XML dobimo naslednje podatke o sporočilu:

- avtorja sporočila,
- besedilo sporočila,
- čas in datum,
- simbole (delnice, ki nastopajo v sporočilu),
- skupno oceno sporočila, oceno sporočila in število glasov pa nastavimo na 0.

Naš strežni program smo porabili še za nekaj ostalih funkcij, ki so programirane z večnitno tehnologijo, da ne motijo ena druge. Tako v strežnem programu vsako uro poiščemo pet delnic, ki so bile popularne v tem intervalu. Te so izpisane na začetni strani, v desnem zgornjem kotu (slika 3.7). Če nas zanimajo bolj podrobne stvari o delnici, nas klik nanjo usmeri na njeno stran na finančnem Yahooju.

<a href="#">CIB</a> - 66,06
<a href="#">EURUSD</a> - 1,38
<a href="#">GOLD</a> - 96,62
<a href="#">GOOG</a> - 624,82
<a href="#">C</a> - 4,30

Slika 3.7: Izpis popularnih delnic.

Ker so finančne informacije časovno občutljive in postanejo hitro zastarele, je nesmiselno imeti v bazi shranjena stara sporočila. Odločili so se, da bomo enkrat dnevno pobrisali vsa sporočila, starejša od enega meseca. Tudi to se zgodi nevidno v ozadju s pomočjo našega strežnega programa.

Poleg zgornjih treh funkcij v strežnem programu tudi ocenjujemo uporabnike, več o tem bomo povedali v poglavju 3.8 Ocenjevanje sporočil.

## 3.5 Osnovne operacije nad sporočili

Na osnovni strani spletnega portala TwittRank (slika 3.8) imamo prikazana sporočila, ki so urejena po času padajoče, torej bomo najnovejše sporočilo našli na vrhu seznama. Zaradi občutljivosti finančnih sporočil na čas (hitro zastarajo) imamo trenutno prikazana samo tista, ki so prispela v zadnjih 24 urah. S tem pospešimo urejanje, iskanje in filtriranje. Sporočila lahko uredimo glede na čas, skupno oceno, oceno Klout, uporabniško ime in oceno sporočila. Po vseh kategorijah lahko uredimo padajoče in naraščajoče. Pod vsakim sporočilom

Slika 3.8: Začetna stran portala.

imamo tudi zapisane simbole, ki v njem nastopajo, klik nas odpelje na stran te delnice na finančnem Yahooju.

Poleg urejanja imamo na voljo še iskanje po sporočilih. Z vpisom zelenega iskalnega gesla in pritiskom na gumb za iskanje, nam portal prikaže samo sporočila, ki vsebujejo izbrano geslo (v besedilu, uporabniškem imenu ali simbolih).

Implementacija filtriranja po borzah je bila bolj kompleksna. Na spletni strani (<http://www.eoddata.com/symbols.aspx>) so zapisane vse pomembnejše svetovne borze in vse oznake delnic, ki na njih kotirajo (slika 3.9). Tako smo napisali program, ki iz te strani prebere simbole in jih zapiše v datoteko XML pod ustrezno borzo. Ker pa podjetja ves čas nastajajo in propadajo, bomo morali vsake toliko časa program ponovno zagnati in posodobiti datoteko XML.

Ko imamo simbole enkrat v datoteki XML (slika 3.10), se samo sprehodimo po vrsticah in jih zapišemo v bazo podatkov. Tako imamo sedaj v bazi zapisane vse simbole in borze, katerim pripadajo. Odločili smo se, da bo ta vrsta filtriranja podpirala samo šest največjih svetovnih borz, že s to omejitvijo imamo v bazi več kot 17.000 simbolov. Ker ima tudi vsako sporočilo točno

List of Symbols for New York Stock Exchange [NYSE]											DOWNLOAD														
New York Stock Exchange																									
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Code	Name													High	Low	Close	Volume	Change							
A	Agilent Technologies Inc.													34.74	34.08	34.48	2,781,600	0.25	↑	0.73					
AA	Alcoa Inc.													13.29	12.77	13.13	40,965,900	0.00	▬	0.00					
AAI	AirTran Holdings Inc.													7.410	7.350	7.390	5,589,600	0.020	↑	0.27					
AAN	Aaron's Inc.													17.88	17.41	17.77	445,300	0.10	↑	0.57					
AAN.A	Aaron's Inc.													17.70	17.52	17.62	6,900	0.06	↑	0.34					
AAP	Advance Auto Parts Inc													61.15	60.14	60.96	844,400	1.17	↑	1.96					
AAR	AMR Corporation													23.50	23.06	23.25	14,800	0.17	↓	0.73					
AAV	Advantage Oil & Gas Ltd													6.530	6.420	6.490	338,300	0.010	↓	0.15					
AB	Alliance Capital Management Holding L.P.													26.60	25.88	26.25	653,000	0.55	↓	2.05					
ABA	Alabama Power Company													26.23	26.07	26.21	48,300	0.03	↓	0.11					
ABB	ABB Ltd													22.61	22.21	22.31	1,606,200	0.23	↓	1.02					
ABC	AmerisourceBergen Corporation (Holding													32.60	32.18	32.29	2,340,400	0.02	↓	0.06					
ABD	Acco Brands Corporation													6.120	5.880	6.040	259,700	0.070	↑	1.17					
ABG	Asbury Automotive Group Inc													14.94	14.51	14.58	278,000	0.03	↓	0.21					
ABK	Ambac Financial Group Inc.													1.120	1.000	1.050	59,060,000	0.100	↑	10.53					
ABK-Z	Ambac Financial Group Inc.													10.030	9.500	9.740	40,300	0.440	↑	4.73					
ABM	ABM Industries Incorporated													22.93	22.07	22.08	372,800	0.52	↓	2.30					

Slika 3.9: Simboli za borzo NYSE - New York Stock Exchange.

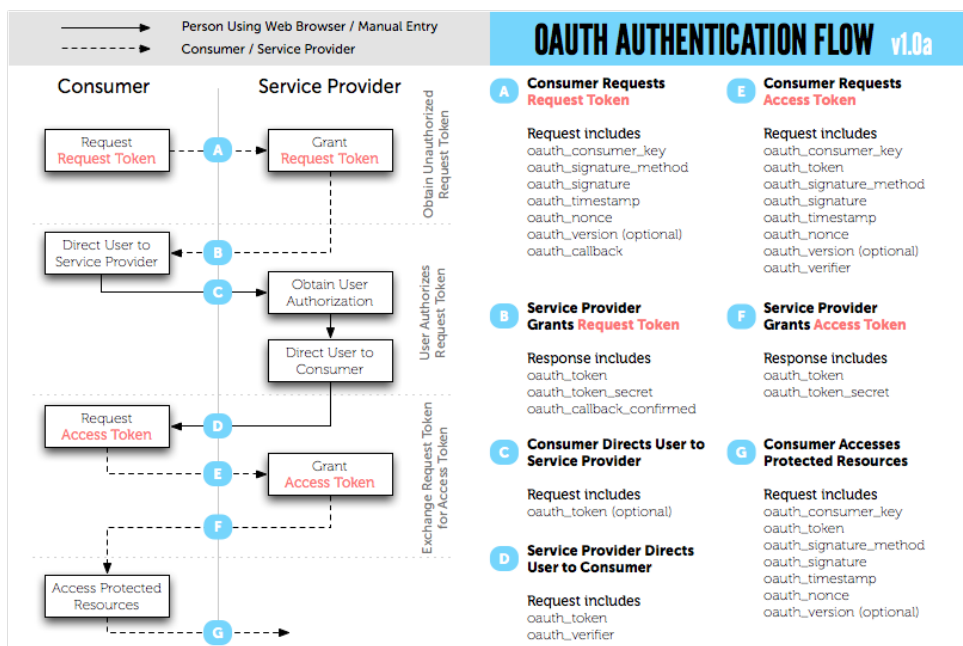
zapisano, katere delnice vsebuje, je bilo implementiranje filtriranja dokaj enostavno, samo pogledati moramo, če se katera od delnic, ki nastopa v sporočilu, tiče izbrane borze. Filtriranje je zelo uporabna funkcija, saj navadno določena oseba trguje le na nekaj borzah.

```
<symbols>
  <exchange name="FOREX">
    <symbol>AEDAUD</symbol>
    <symbol>AEDCAD</symbol>
    <symbol>AEDCHF</symbol>
    <symbol>AEDDKK</symbol>
    <symbol>AEDEUR</symbol>
    <symbol>AEDGBP</symbol>
    <symbol>AEDINR</symbol>
    <symbol>AEDJPY</symbol>
    <symbol>AEDNZD</symbol>
    <symbol>AEDPKR</symbol>
    <symbol>AEDUSD</symbol>
```

Slika 3.10: Delček datoteke XML z zapisi borz in njihovih simbolov.

## 3.6 Prijava na portal

Za prijavo v naš sistem moraš biti registriran na portalu Twitter. S tem smo dosegli, da lahko nemoteno pošiljamo sporočila hkrati na naš portal ter na Twitter. Tak način realizacije prijave pa nam omogoča tudi uvoz naših prijateljev. Uvozimo lahko tako tiste, ki sledijo nam, kot tudi tiste, ki jim sledimo mi. Samo prijavo smo realizirali s pomočjo Twitter API vmesnika, ki za avtentikacijo uporablja standard OAuth.

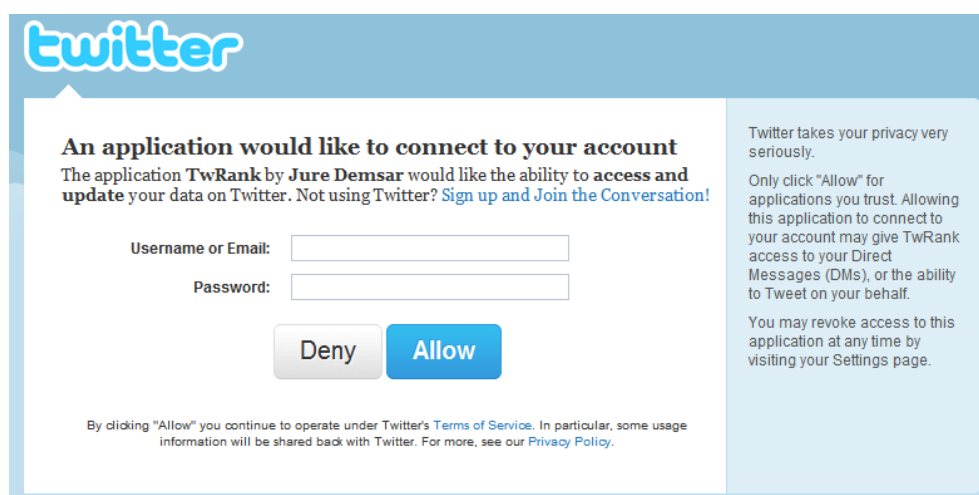


Slika 3.11: Potek avtorizacije z OAuth standardom.

Pred OAuthom je Twitter uporabljal kar pošiljanje uporabniških imen in gesel, kar pa seveda ni najbolj varno. Zato so se pri Twitterju odločili, da preidejo na OAuth, ki je odprt standard za avtorizacijo. Uporabnikom dovoljuje, da delijo svoje privatne vire (slike, filme, seznam kontaktov, ...), shranjene na neki strani, ne da bi izdal podatke o svojem računu (uporabniško ime, geslo). Navadno se uporablja, ko do te strani dostopamo z neke druge strani. Namesto uporabniških podatkov se pošiljajo žetoni (slika 3.11), vsak pa pove, za koliko časa se bo odprl dostop in za katere vire se bo odprl. To dovoljuje, da s pomočjo našega programa pridobivamo zaupne podatke, na neki strani. V našem primeru tako lahko iz našega portala opravljamo praktično vse funkcije, ki jih ponuja Twitter. Sredi leta 2010 je tako Twitter ukinil uporabo starega načina prijavljanja in danes se lahko prijavljamo samo preko OAuth (<http://oauth.net/core/1.0/#anchor9>).

Za uspešno uporabljanje Twitter avtentikacije je potrebno svojo aplikacijo registrirati na portalu Twitter. Ko uspešno zaključimo potek registracije, dobimo svoje skrivne žetone, ki jih uporabimo v naši kodi. V veliko pomoč pri avtentikaciji nam je bila brezplačna knjižnica OAuth za C#, ki vsebuje vse osnovne procedure in funkcije (<http://devblog.yedda.com/>). Potrebni so bili samo še določeni popravki in prilagoditve, da smo proces prijave in odjave pri-

lagodili potrebam našega portala. S prijavo v naš sistem smo tako prijavljeni tudi na Twitter (slika 3.12). Vse pomembne informacije o prijavi se zapišejo v podatke o seji, tako da lahko tudi na podstraneh v vsakem trenutku izvemo, ali je uporabnik prijavljen ali ne. Če je uporabnik prijavljen, o njem hranimo tudi vse potrebne podatke za uporabljanje funkcij Twitter vmesnika API, zato v podatke o seji shranimo vse OAuth žetone in uporabniško ime.



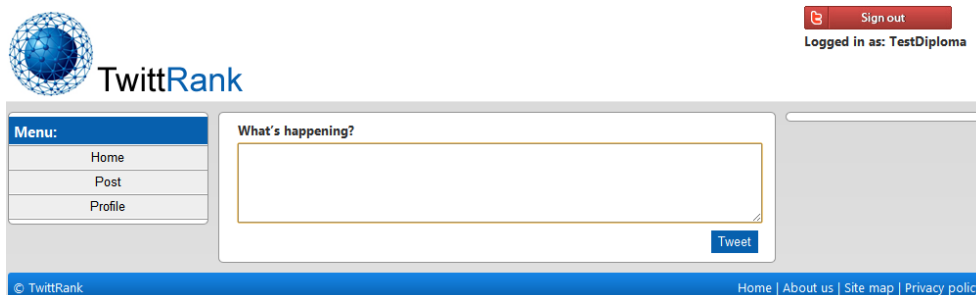
Slika 3.12: Prijava preko portala Twitter.

S klikom na gumb za odjavo se lahko prijavljeni uporabniki odjavijo iz našega portala, hkrati pa nam v novem oknu odpre še stran za odjavo iz Twitterja.

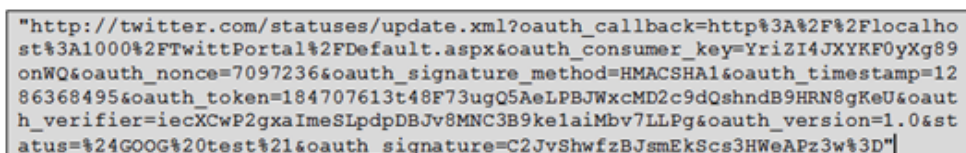
## 3.7 Pošiljanje novih sporočil

Če smo prijavljeni na portal, imamo omogočen dostop do strani za pošiljanje novih sporočil (slika 3.13). V bistvu sporočilo najprej objavimo na Twitterju, nato pa ga prenesemo še v našo bazo podatkov. Tudi to se naredi s Twitter API vmesnikom (<http://dev.twitter.com/doc/post/statuses/update>), s pomočjo OAuth knjižnice.

Za vsako interakcijo s Twitterjem moramo uporabljati OAuth varnostne žetone, ki smo jih dobili ob prijavi. Proces pošiljanju predstavlja kreiranje določenega URL naslova, ki je sestavljen iz žetonov in besedila, ki ga hočemo poslati. Če hočemo poslati sporočilo “\$GOOG test” dobimo spodnji zahtevek (slika 3.14).

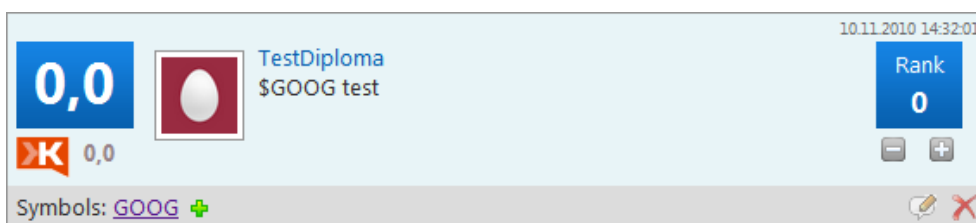


Slika 3.13: Stran za pošiljanje sporočil.



Slika 3.14: Generiran zahtevek za pošiljanje preprostega sporočila.

Na tem sestavljenem URL naslovu se ustvari dokument XML z vsemi informacijami o sporočilu (datum, pošiljatelj, vsebina, ...). Preostane nam samo še, da iz datoteke XML pobremo podatke, ki jih potrebujemo in jih zapišemo v bazo. S tem preverimo tudi vsebino sporočila in v njem poiščemo oznake delnic (\$GOOG - oznaka za Google, \$AAPL - oznaka za Apple, ...), tako dobimo delnice, ki nastopajo v besedilu novice. S pomočjo tega podatka pa lahko kasneje filtriramo sporočila.

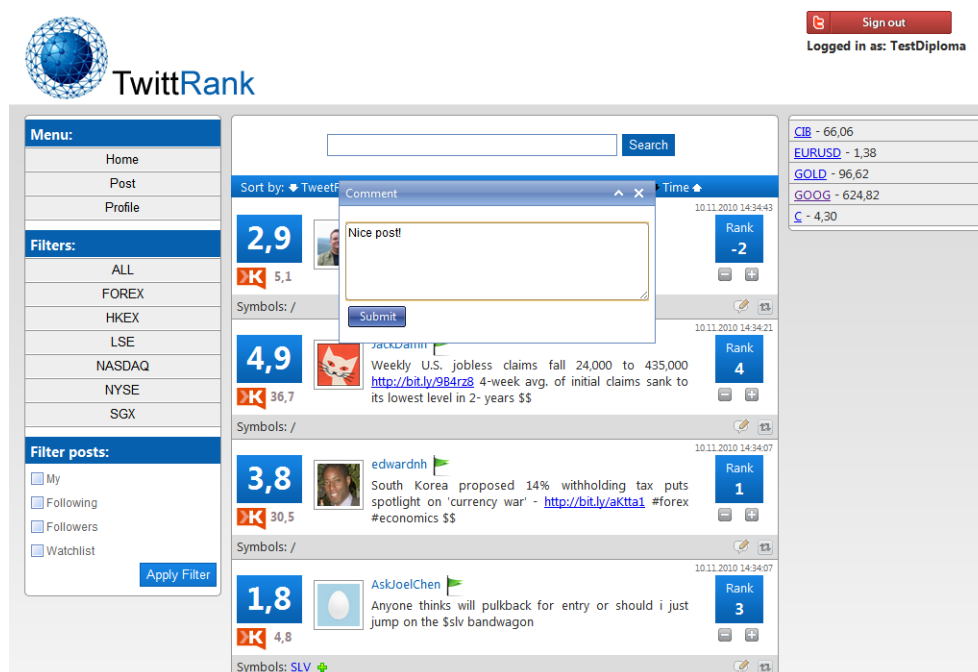


Slika 3.15: Prikaz našega novega sporočila.

Ko je sporočilo poslano, ga vidimo na osnovni strani (slika 3.15). Če smo avtor sporočila, imamo poleg običajnih možnosti na voljo tudi gumb za njegov izbris. Ta izbriše sporočilo tako na našem portalu kot na Twitterju. Zgodba o brisanju je dokaj podobna pošiljanju, le da namesto vsebine sporočila povemo identifikacijsko številko sporočila, ki ga hočemo odstraniti. S pomočjo te

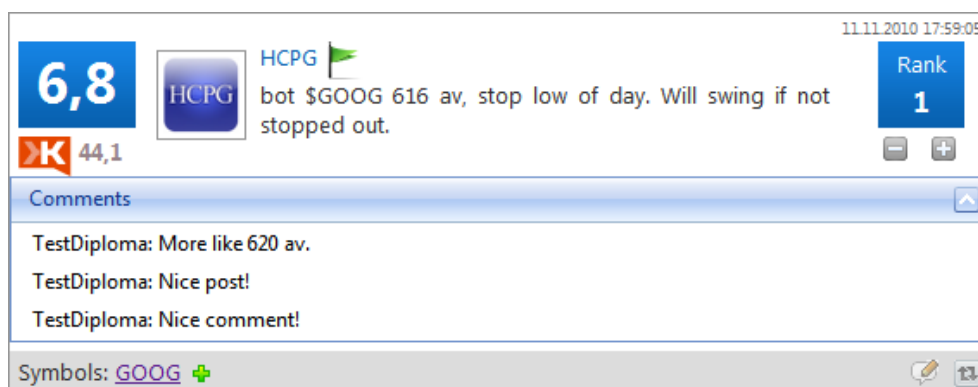
identifikacijske številke in OAuth žetonov kreiramo podoben URL naslov kot pri pošiljanju. Preusmeritev na ustvarjeni naslov pa poskrbi za izbris zelenega sporočila.

Če smo prijavljeni imamo na voljo tudi možnost komentiranja že obstoječih sporočil (slika 3.16). Glede na to, da so sporočila dolga največ 140 znakov, se nam je zdelo nesmiselno, da bi v komentarjih lahko objavljali neka dolgo-vezenja, zato smo dolžino komentarjev omejili na 70 znakov. Ob kliku na gumb za komentiranje, se nam odpre novo okno, ki je precej podobno strani za pošiljanje novih sporočil. Na voljo imamo samo vnosno polje in gumb za potrditev komentarja.



Slika 3.16: Komentiranje sporočila.

Ob kliku na gumb za objavo, se ta zabeleži v bazo, na sporočilu pa se pokaže spustni seznam, ki vsebuje vse komentarje, ki se nanj nanašajo (slika 3.17). Ko seznam odpremo, se nam dinamično poveča velikost sporočila, tako da lahko nemoteno pregledujemo tudi ostala sporočila. Vsaka kontrola, ki pripada sporočilu, ima na svojih atributih zapisano identifikacijsko številko sporočila, tako da pri vsakem pritisku na gumb točno vemo, na katero sporočilo se morajo navezovati spremembe.



Slika 3.17: Prikaz komentarjev.

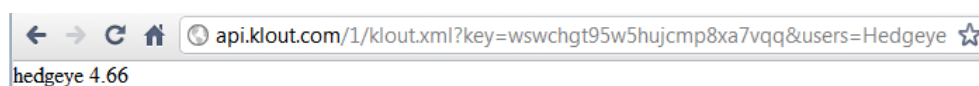
### 3.8 Ocenjevanje sporočil

Glavni namen našega portala, poleg preglednega prebiranja informacij, je njihovo ocenjevanje. Prva ideja je bila, da bi ocenjevali vsebino novic, toda kmalu smo ugotovili, da bi bilo to najbrž preveč kompleksno in verjetno tudi neuporabno. Novice, ki jih dobimo, se nanašajo na zelo različna področja in so tudi vsebinsko različno strukturirane. Novice o Forexu so tako povsem drugačne kot recimo novice o trgovanju z običajnimi delnicami. Poleg tega ogromno sporočil vsebuje samo oznake delnic, na katere se nanaša, in povezavo na neko spletno stran, ki nato te delnice natančneje opisuje. Če bi hoteli dobro ocenjevati vsebino, bi bilo potrebno pregledovati tudi te strani, saj nas včasih povezava pelje na stran, kjer spet najdemo samo povezave do določenih informacij. Zaradi teh razlogov smo idejo o preučevanju vsebine sporočil kmalu opustili. Odločili smo se, da bomo stvar rahlo poenostavili in našo skupno oceno informacije izračunali iz treh podocen, ki bodo odvisne od pošiljatelja sporočila ter uporabnikov, ki prebirajo novice na strani.

Prvi del skupne ocene je ocena, ki jo sporočilu namenijo uporabniki. Vsak uporabnik ima možnost posamezno sporočilo oceniti s pozitivnim oziroma negativnim glasom. Če uporabnik meni, da je sporočilo dobro, bo s klikom na gumb “+” dal novici pozitiven glas, s tem se bo ocena sporočila povečala za 1. Če pa uporabnik misli, da je novica slaba, lahko s klikom na gumb “-” odda negativen glas. Glasovanje je anonimno, da ne pride do zamer med avtorji. Ko uporabnik odda oceno, se mu nadaljnje glasovanje za to sporočilo zaklene.

Drugi del ocene je ocena Klout [3], ki je merilo za vpliv nekega uporabnika na Twitter omrežje. Je število na intervalu od 1 do 100, višje številke pomenijo večjo sfero vpliva. Uporablja več kot 35 spremenljivk, s katerimi

meri doseg, verjetnost akcije in mrežno oceno uporabnika. Doseg je odvisen od velikosti množice, ki spremlja in odgovarja na njegova sporočila. Verjetnost akcije ocenjuje, kako pogosto bodo sporočila uporabnika generirale akcije (odgovore, posredovanja, komentarjem, ...). Mrežna ocena kaže, kakšen vpliv ima uporabnik na celotno množico ostalih uporabnikov Twitterja. Vse te 3 ocene se izračunajo iz več kot 35 spremenljivk, način izračuna je skriven in naj bi bil plod zapletenih metod strojnega učenja. Iz teh treh se nato izračuna končna ocena Klout, ki jo uporabimo pri izračunu naše skupne ocene. Dobimo jo s pomočjo Klout API vmesnika, katerega uporaba je zelo enostavna. Najprej smo našo aplikacijo registrirali na strani (<http://www.klout.com>), kjer so nam dodelili ključ, s katerim nato s preprostim obiskom določenega URL naslova dobimo oceno. URL naslov je sestavljen iz našega ključa in uporabniškega imena, za katerega hočemo dobiti oceno. Naša aplikacija ima ključ "wswchgt95w5hujcmp8xa7vqq", ki ga lepo vidimo v spodnjem primeru (slika 3.18), kaže nam URL naslov ter rezultat poizvedbe, če hočemo dobiti oceno Klout za uporabnika Hedgeye.



Slika 3.18: Rezultat zahtevka za oceno Klout.

Tretja je ocena uporabnikov, ki pošiljajo finančne informacije. Na višino le-te vplivajo posamezne ocene uporabnikovih sporočil (ocene, ki jih sporočilom dodelijo ostali uporabniki). Torej je odvisna od kvalitete sporočil, ki jih nekdo objavlja. Ker sporočila starejša od enega meseca brišemo iz naše baze podatkov, nam ta del končne ocene pove, kako dobra sporočila je avtor pošiljal v zadnjem mesecu. Za izračun ocene smo uporabili neke vrste uravnoteženo vsoto. Spodaj imate primer izračuna za uporabnika TestDiploma (slika 3.19), če je v zadnjem mesecu objavil 10 sporočil z ocenami: 3, 5, 2, 2, 7, 7, 7, 9, 10, 2. Shranimo jih v seznam, ki ga najprej uredimo po velikosti od najmanjšega do največjega. Nato pa preštejemo, koliko je v seznamu ocen z enako ali višjo vrednostjo. Zapomnimo si manjšega od para: trenutna ocena, število večjih ali enakih. Za našo oceno vzamemo največjega izmed teh zapomnjenih števil. Dobra stran takega ocenjevanja je, da avtor počasi napreduje oziroma nazaduje po lestvici, saj ima ena majhna odklonska ocena majhen pomen pri končni. Nazadnje dobljeno oceno še ustrezno umestimo na interval od 0 do 10.

Skupno oceno izračunamo kar kot povprečje vseh treh podocen. Oceno Klout in oceno sporočila moramo najprej še umestiti na interval od 0 do 10, tako je tudi končna skupna ocena umeščena na interval od 0 do 10.

```
TestDiploma - 3, 5, 2, 2, 7, 7, 7, 9, 10, 2
TestDiploma[2] = 10 => min = 2
TestDiploma[3] = 7 => min = 3
TestDiploma[5] = 6 => min = 5
TestDiploma[7] = 5 => min = 5
TestDiploma[9] = 2 => min = 2
TestDiploma[10] = 1 => min = 1
```

Slika 3.19: Primer izračuna ocene uporabnika.

### 3.9 Profil

S klikom na povezavo do profila, se nam odpre stran z našim profilom (slika 3.20). Če pa kliknemo na uporabniško ime v seznamu sporočil, se nam odpre tuj profil. Razlika je v tem, da na tujem profilu lahko podatke samo gledamo, na našem pa jih lahko tudi urejamo. Poleg tega imamo na tujem profilu tudi gumb, s katerim lahko začnemo izbranemu uporabniku slediti.

Na vrhu profila imamo o vsakem uporabniku napisane osnovne podatke:

- uporabniško ime,
- oceno uporabnika,
- oceno Klout uporabnika,
- število sporočil, ki jih je prispeval posamezni uporabnik na naš portal.

Na profilu imamo tri sezname, na prvem so napisani uporabniki, ki jim sledimo mi, na drugem uporabniki, ki sledijo nam, tretji pa služi za zapis delnic, ki se nam zdijo zanimive.

Če nekemu uporabniku sledimo, lahko s preprostim klikom na osnovni strani takoj pridemo do vseh sporočil, ki jih je objavil. Na voljo imamo tudi možnost za uvoz sledenj iz Twitterja. Ta nam iz našega Twitter profila na naš portal skopira vse uporabnike, ki jim sledimo na portalu Twitter, ampak samo tiste, ki so aktivni tudi na našem portalu. Za aktivne smatramo uporabnike, ki imajo svoj zapis v naši bazi uporabnikov. S klikom na gumb za odstranitev sledenja lahko sledenje prekinemo. To storimo takrat, ko nek uporabnik začne pošiljati slaba sporočila, ali pa ko postane nedejaven.

Drug seznam predstavlja uporabnike, ki sledijo nam. Tudi te lahko uvozimo s Twitterja, zgodba o uvozu je enaka kot pri prvem seznamu. Teh iz seznama ne moremo odstraniti, odstranijo se lahko sami, tako da nam nehalo slediti.

The screenshot shows the TwittRank user profile for 'TestDiploma'. The page includes a navigation menu, user information, and three lists: Following, Followers, and Watchlist.

**Menu:**

- Home
- Post
- Profile

**User Profile:** TestDiploma, 30 messages, UserRank: 4,5, 0,0

**Following:**

Username	Remove
BlitzFX	X
EquityClock	X
ForexHeadlines	X
IndianaGrainCo	X
infoarbitrage	X
innovinance	X
JosefHiggins	X
MartinT25	X
theback9	X
tigertrader	X

**Followers:**

Username
DiplomaTest
Lumpizaver

**Watchlist:**

Symbol	Exchange	Remove
AUDUSD	FOREX	X
EURGBP	FOREX	X
EURUSD	FOREX	X
GBPEUR	FOREX	X
LVS	NYSE	X
USDAUD	FOREX	X
USDEUR	FOREX	X

© TwittRank | Home | About us | Site map | Privacy policy

Slika 3.20: Profil uporabnika.

Na tretjem pa imamo zbrane delnice, ki se nam zdijo zanimive. Zberemo tiste, s katerimi aktivno trgujemo ali pa načrtujemo, da bomo v prihodnosti mogoče začeli z njimi trgovati. Dodajanje je preprosto, samo vpišemo ime delnice v vnosno polje in jo dodamo v seznam s klikom na gumb. Delnica se doda samo, če vnos v polje dejansko ustreza neki delnici v naši bazi podatkov. Večino pomembnih delnic smo v bazo podatkov uvozili v eni od zgodnjih faz razvoja portala. Če kakšno na novo objavljeno sporočilo vsebuje delnice, ki jih še nimamo v bazi, hkrati z vnosom sporočila v bazo zapišemo še te. Tudi pri

seznamu delnic imamo gumb za njeno odstranitev.

S pomočjo teh treh seznamov imamo v primeru, da smo na stran prijavljeni, dodatne možnosti filtriranja (slika 3.21). Imamo štiri dodatne možnosti filtriranja, ki jih lahko med sabo poljubno kombiniramo. Prikažemo lahko:

- svoja sporočila,
- sporočila uporabnikov, ki jim sledimo,
- sporočila uporabnikov, ki sledijo nam,
- sporočila, ki se tičejo delnic na našem seznamu zanimivih delnic.

The screenshot shows the TwittRank website interface. At the top left is the TwittRank logo. In the top right corner, there is a 'Sign out' button and a notification 'Logged in as: TestDiploma'. On the left side, there is a 'Menu' with options: Home, Post, Profile. Below the menu is a 'Filters' section with a list of stock exchanges: ALL, FOREX, HKEK, LSE, NASDAQ, NYSE, SGX. A red box highlights the 'Filter posts:' section, which contains four checkboxes: 'My' (unchecked), 'Following' (checked), 'Followers' (unchecked), and 'Watchlist' (checked). Below these checkboxes is an 'Apply Filter' button. The main content area shows a list of tweets. Each tweet entry includes a rank in a blue box, a user profile picture, the user's name, the tweet text, a date and time stamp, and a 'Rank' button. Below each tweet, there are symbols and a plus sign. The tweets shown are:
 

- Rank 3,3 by ForexBatman: 'Again: support level to watch in \$EURUSD 1,3697, a break here could open the door to 1.33 as we head into year end! #forex'. Symbols: EURUSD. Rank button: Rank 3.
- Rank 5,2 by piptee: 'tempted to buy the Seurusd off a daily bounce.'. Symbols: EURUSD. Rank button: Rank -1.
- Rank 5,9 by Street\_Insider: 'Las Vegas Sands \$LVS 1.8% lower following downgrade to Neutral at UBS which told clients to cash in their chips'. Symbols: LVS. Rank button: Rank 0.
- Rank 2,9 by ForexBatman: 'Good morning guys! Again a choopy London session for \$EURUSD #forex'. Symbols: EURUSD. Rank button: Rank 2.

 On the right side of the interface, there is a table of stock prices:
 

CIB	- 66,06
EURUSD	- 1,38
GOLD	- 96,62
GOOG	- 624,82
C	- 4,30

Slika 3.21: Dodatne možnosti filtriranja sporočil.

### 3.10 Spremljanje prometa

Google analitika (<http://www.google.com/analytics/>) je brezplačna storitev, ki jo ponuja Google. Ustvari zelo podrobno statistiko glede obiskovalcev spletne strani. Glavna tarča produkta so oglaševalci in ne administratorji, za katere

je bila sprva namenjena analiza obiska spletnih strani. Google analitika je najbolj priljubljena služba za spletno statistiko, uporablja jo kar 57% med 10.000 najbolj popularnimi spletnimi stranmi, oziroma 49,95% med prvim milijonom popularnih spletnih strani ([http://en.wikipedia.org/wiki/Google\\_Analytics](http://en.wikipedia.org/wiki/Google_Analytics)). Implementacija je zelo enostavna. Svojo stran registriramo na strani Google analitike, kjer nam samodejno generira kodo, ki jo prilepimo v datoteko naše spletne strani, nato lahko preko enostavnega vmesnika pregledujemo vse podatke o obiskovalcih naše strani.

## 3.11 Oblikovanje

Spletne strani narejene v ASP.NET se grafično oblikujejo s pomočjo datotek CSS (slika 3.22). Datoteke CSS so predloge, ki določajo izgled spletnih strani. Z njihovo pomočjo določimo pisavo, velikost črk ter vizualno predstavitev spletne strani. HTML naj bi predstavljal semantično strukturo in smiselno hierarhijo dokumenta, CSS pa predstavitevno vlogo [4].

```
/* PRIMARY LAYOUT ELEMENTS
-----*/

.page
{
    width: 1000px;
    background-color: #fff;
    margin: 20px auto 0px auto;
    border: 1px solid #496077;
}

.header
{
    position: relative;
    margin: 0px;
    padding: 0px;
    background: #4b6c9e;
    width: 100%;
}
```

Slika 3.22: Primer oblikovanja z uporabo datoteke CSS.

Za vizualno oblikovanje gradnikov (gumbov, seznamov, ...) smo uporabili Adobe Photoshop, ki je profesionalni računalniški program za obdelavo

fotografij in drugih grafik. Razvili so ga pri podjetju Adobe Systems, ki je trenutno eden izmed vodilnih izdelovalcev grafičnih programov na svetu ([http://sl.wikipedia.org/wiki/Adobe\\_Photoshop](http://sl.wikipedia.org/wiki/Adobe_Photoshop)). Poleg oblikovanja gradnikov z risanjem v Photoshopu, lahko gradnike oblikujemo tudi s pomočjo preoblek in tem.

# Poglavje 4

## Testiranje

Uporabnost našega portala smo preizkusili z eno mesečnim trgovanjem na Forexu, ki je svetovni trg denarnih valut, na njem se trguje 5 dni v tednu, od ponedeljka do petka, 24 ur na dan. Unikaten je zaradi:

- velike količine trgovanja, kar povzroča visoko likvidnost,
- geografske porazdelitve,
- trgovanja 24 ur na dan,
- majhne marže,
- vzvodov – z vzvodom 200:1 je 1.000 € dejansko vrednih 200.000 €.

Mnogo finančnikov je mnenja, da je prav zaradi zgoraj naštetih razlogov Forex najboljši način trgovanja. Vsak dan na trgu za valutno trgovanje obrne 3.98 trilijona ameriških dolarjev, ta številka pa se še povečuje [1]. Ta trg je večji od delniškega trga in trga s standardiziranimi terminskimi pogodbami skupaj. Pri tej višini prometa nenehno prihaja do gibanj in s tem možnosti za dobiček, tudi ko drugi trgi stagnirajo. Dodatna dobra lastnost je, da ob pravilnih napovedih vedno lahko ustvarjamo dobiček, medtem ko je pri delnicah drugače. V zadnjih letih (v času recesije) je bilo s trgovanjem delnic zelo težko kaj zaslužiti. Pri trgovanju na Forexu vlagatelj vedno trguje s kombinacijo dveh valut (valutni par ali križ), pri čemer eno valuto kupi (dolga pozicija) in drugo proda (kratka pozicija). To pomeni, da vlagatelj špekulira glede pričakovane rasti vrednosti ene od valut v primerjavi z drugo. Trgovanje s kritjem (z vzvodi) vlagateljem omogoča nakup in prodajo sredstev z večjo vrednostjo od kapitala na njihovem računu. Trgovanje Forex se navadno izvaja na računih gibljivega

kritja, praksa panoge pa je trgovanje z razmeroma nizkimi zneski kritja, saj znaša nihanje valutnega menjalnega tečaja na dan manj kot en odstotek ali dva. Trgovanje s kritjem vsebuje tudi določeno stopnjo tveganja. Ker se ohranja pozicija, ki presega dejansko vrednost računa, bi lahko v primeru neželenih gibanj na trgu vlagatelj doživel znatne izgube. Zaradi tega trgovanje s kritjem zahteva pozorno spremljanje uporabe kritja, to je zneska zavarovanja za ohranjanje pozicij kritja. Če uporaba kritja preseže znesek zavarovanja, ki je na voljo za trgovanje s kritjem, je treba pozicije zapreti, zmanjšati ali pa je treba za kritje pozicij zagotoviti dodatna sredstva.

Ob registraciji na Forex strani (<http://www.forex.com/>) dobimo na naš račun preizkusnih 50.000 €. Za naš test smo dobili štiri uporabnike, ki so bili pripravljeni en mesec ocenjevati in spremljati novice na portalu TwittRank in nato pridobljeno znanje uporabiti v praksi, na poizkusnem računu Forex. Uporabniki pred trgovanjem niso imeli veliko izkušenj z valutnim trgovanjem. Za bolj merodajne rezultate testa bi bilo seveda bolje, če bi bilo obdobje trgovanja daljše, toda naš brezplačni račun je na žalost odprt samo en mesec. V mesecu, ko so naši kandidati trgovali, je tudi prišlo do lepe rasti pri nam najbolj domačem paru EUR/USD (slika 4.1), tako da je večina kandidatov prav zaradi tega para dosegla zelo lepe rezultate. Kandidati so si konec vsakega dneva zapisali trenutno stanje na računu.

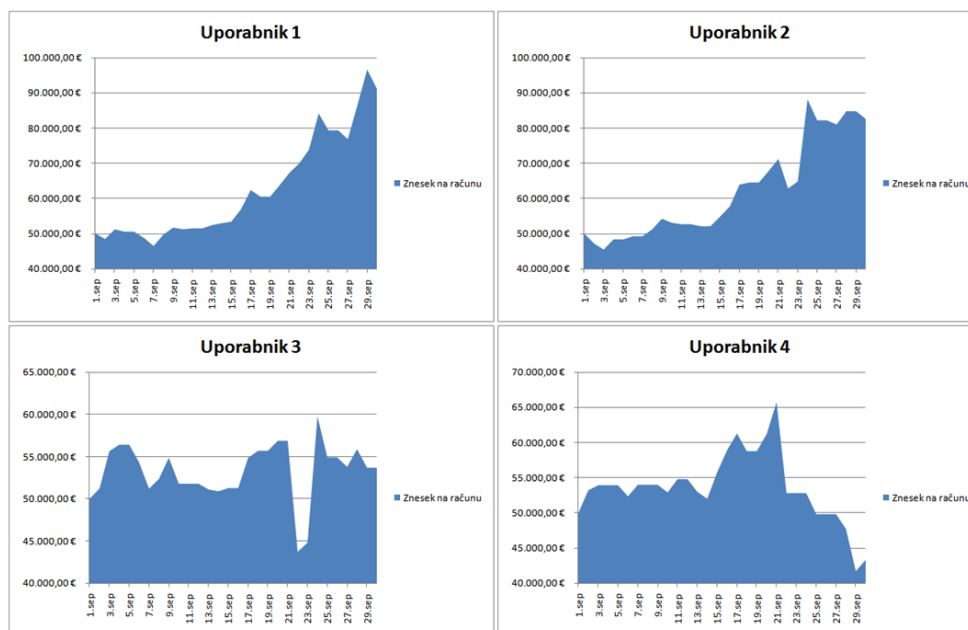


Slika 4.1: Gibanje para EUR/USD za september 2010.

Ko so našim 4 uporabnikom zaprli račune, smo vse podatke zbrali in jih analizirali, rezultati so bili presenetljivo dobri. Vendar kot smo že omenili, je en mesec dokaj kratko obdobje, tako da trenutno še ne moremo podati dejan-

skega obsega uporabnosti portala, vsekakor pa so začetni odzivi in rezultati spodbudni. Poleg tega smo imeli v mesecu testiranja lepo pomoč s strani para EUR/USD.

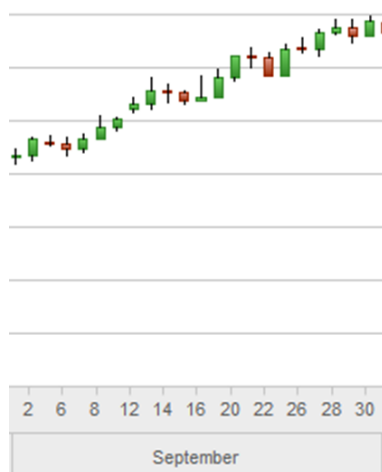
Po enomesečnem trgovanju sta 2 uporabnika zaključila poslovanje s kar lepim plusom, en uporabnik je bil približno tam, kjer je začel, en uporabnik pa je bil v minusu. V grafih tudi opazimo podobno gibanje stanja na računu. Vsi uporabniki so se na začetku gibali okoli ničle, v srednjem delu meseca se pri vseh opazi rast, v zadnjem delu pa je prišlo do razlik pri rezultatih. Če grafe primerjamo z grafom para EUR/USD, vidimo, da so si zelo podobni. Tudi pri podrobnejši analizi trgovanja in rezultatov smo opazili, da je večina dobička prišla s pomočjo tega para. Do različnih končnih rezultatov je prišlo zaradi razlik v vložnem času in različnih odločitev pri izbiri parov trgovanja. Prva dva uporabnika sta poleg EUR/USD trgovala še z AUD/USD (slika 4.3). Preostala dva uporabnika pa sta malo bolj eksperimentirala, kar ni ravno pametno početje glede na to, da nista imela pred preizkusom nobenih pravih izkušenj z valutnim trgovanjem. Rezultate uporabnikov vidimo na spodnjih grafih (slika 4.2).



Slika 4.2: Rezultati enomesečnega FOREX trgovanja.

Skupno so naši testni osebki zaslužili 70.530 €, kar zneso 17.632 € na osebo. Torej je šlo kar za 35% mesečno rast portfelja. Povsem nerealno je pričakovati,

da bi tak trend zaslužka lahko obdržali na daljše obdobje. Imeli smo obilo sreče z izbiro meseca, saj so v septembru praktično vsi resni Forex posredniki zaslužili lepe vsote denarja. Za mnenje smo vprašali nekaj profesionalnih Forex posrednikov na spletnem forumu BabyPips (<http://forums.babypips.com/>). Vsi so pohvalili našo idejo in rezultate, vendar so rekli, da bomo kot novinci na tem področju lahko zelo srečni, če bomo imeli letno rast portfelja nad 25%.



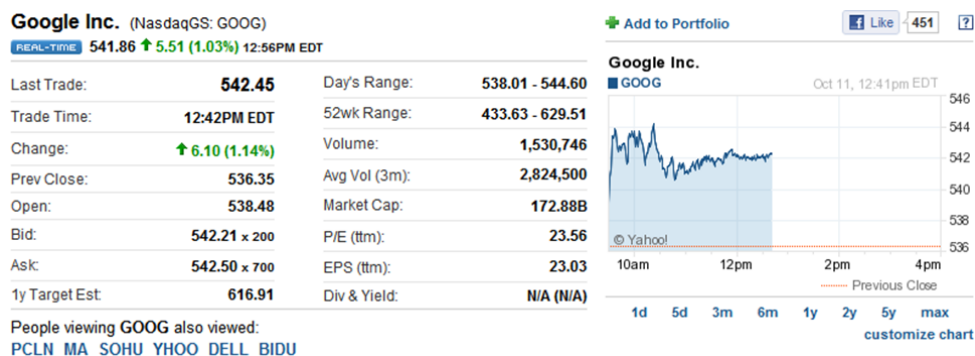
Slika 4.3: Gibanje para AUD/USD za september 2010.

## Poglavje 5

# Sklepne ugotovitve

V portal smo vgradili vse funkcionalnosti, ki smo si jih zamislili pred začetkom razvoja. Med samim razvojem pa se je pojavilo še kar nekaj zanimivih idej. Te bi bile lahko temelji nadaljnjim izboljšavam projekta. Trenutno pa jih zaradi pomanjkanja virov oziroma zaradi prevelike kompleksnosti nismo realizirali:

- Ocenjevanje vsebine sporočil. Kot smo že omenili, smo med razvojem to opustili zaradi prevelike kompleksnosti, vsekakor pa bi bila lahko to neka unikatna in zelo dobra lastnost našega portala.
- Boljši algoritem za skupno oceno sporočila. Trenutno se skupna ocena izračuna kar kot povprečje treh podocen, šele čas bo pokazal, kako dobra je ta metoda. Problem pri finančnih informacijah je, da njihovo kvaliteto lahko ocenimo šele, ko je novica že zastarela, zato bi bilo potrebno informacije oceniti ponovno, ko že vemo, ali je bila napoved dobra ali slaba. Na ta način bi s strojnimi učenjem najbrž lahko dobili boljšo metodo od povprečja treh ocen.
- Portfelj. Vsi večji finančni portali nam ponujajo možnost vnašanja portfelja. Na teh portalih lahko tudi spremljajo vrednost vnešenega portfelja. Da bi dodali to funkcionalnost naši strani, potrebujemo realno časovno spremljanje borz, kar pa je zelo drago. Že za realno časovne podatke o slovenski borzi (ki je ena manjših na svetu) je potrebno plačati 800 € na mesec.
- Grafi delnic. Tudi za prikaz grafov potrebujemo realno časovne podatke, poleg teh pa tudi zgodovino delnice. Trenutno nas klik na oznako delnice popelje na njen graf na finančnem Yahooju (slika 5.1).



Slika 5.1: Podatki o delnici GOOG na finančnem Yahooju.

Naš portal se je tudi brez zgornjih funkcij izkazal za uporabnega. Že samo možnost enostavnega pregledovanja večine finančnih sporočil iz Twitterja uporabnikom prihrani mnogo časa. Tudi mehanizem za ocenjevanje je zelo uporabna stvar, vendar ga še nismo dobro preizkusili, saj je portal uporabljalo le nekaj oseb. Ocenjevanje pa bo prišlo do izraza šele, ko bo število uporabnikov naraslo.

# Literatura

- [1] Bank for International Settlements, *Triennial Central Bank Survey*, 2007.
- [2] T. Collins, *The Little Book of Twitter: Get Tweetwise!*, London: Micheal O'Mara Books Limited, 2009.
- [3] D. Evans, *Social Media Marketing: The Next Generation of Business Engagement*, Indiana: Wiley Publishing Inc., 2010.
- [4] S. Jenkins, *Web Design All-in-one For Dummies*, Indiana: Wiley Publishing Inc., 2009.
- [5] J. Mayo, *Microsoft Visual Studio 2010: A Beginner's Guide*, Columbus: McGraw-Hill Osborne Media, 2010.
- [6] D. M. Prigmore *An Introduction to Databases with Web Applications*, Essex: Prentice Hall, 2007.
- [7] S. Sumathi *Fundamentals of Relational Database Management Systems*, Warsaw: Polish Academy of Sciences, 2007.
- [8] A. Troelsen *Pro C# 2010 and the .NET 4 Platform, Fifth Edition*, Atlanta: Apress, 2010.