

# SentiMusic – Mash-Up media extins cu analiza sentimentelor

Anca Marginean  
Universitatea Tehnica din Cluj-Napoca  
[anca.marginean@cs.utcluj.ro](mailto:anca.marginean@cs.utcluj.ro)

Cristina Nan  
Universitatea Tehnica din Cluj-Napoca  
[cristinutzanan@yahoo.com](mailto:cristinutzanan@yahoo.com)

**Abstract** — Analiza sentimentelor este o componenta importanta in Web2.0. Desi interesul principal vine din directia marketingului in vederea identificarii tendintelor sociale, consideram ca multe alte domenii nu trebuie neglijate. Astfel, am aplicat tehnici specifice analizei sentimentelor pe versurile pieselor muzicale in cadrul unei aplicatii de tip mashup. Aceasta aplicatie prezinta functiile standard de randare a muzicii si de colectare a datelor despre artisti si cantece, precum si informatii despre sentimentele transmise prin versuri. Premisa de start a fost ca versurile cantece ascultate de catre o persoana ar putea da informatii despre starea emotionala a persoanei.

**Index Terms** — analiza sentimentelor, servicii Rest, Web Api.

## I. INTRODUCTION

Dezvoltarea aplicatiilor mobile folosind servicii web a cunoscut un avant puternic in ultimii ani. Aplicatiile mobile din domeniul muzical reprezinta unul dintre primele trei motive pentru care oamenii cumpara dispozitive mobile.

Dar dincolo de componenta de divertisment, aplicatiile muzicale pot deveni sursa pentru construirea unui profil utilizator, dincolo de preferintele musicale. Detectarea sentimentelor exprimate in cantece prin versuri ar putea oferi beneficii domeniilor precum psihologie, ingrijirea copiilor sau media. Tipul de muzica ascultat spune adeseori mai multe despre starea emotionala a utilizatorului decat o conversatie directa. Cand vine vorba de relatia parinti-copii, o mai buna intelegere a starii emotionale a copiilor i-ar putea ajuta pe parinti sa aiba o atitudine mai potrivita fata de acestia. Mai mult, o filtrare a cantece se poate realiza pe baza mesajelor malitioase transmise. In ce priveste recomandarea cantece, aceasta poate fi ajustata in functie de starea emotionala curenta a utilizatorului. De exemplu, pot fi recomandate cantece similare in sentiment cu cele anterioare, sau dimpotriva, cantece care sa contribuie in mod pozitiv la starea ascultatorului.

In acest context, s-a construit o aplicatie de tip mashup. Aceste aplicatii sunt aplicatii hibride care utilizeaza mai multe surse. Aplicatia are la baza mai multe servicii REST. Conform site-ului ProgrammableWeb, unul dintre site-urile care colecteaza servicii REST sau WEB APIs, in prezent exista in jur de 11 457 API-uri, dintre care 276 incadrate in categoria muzica. LastFM<sup>1</sup>, MusicBrainz<sup>2</sup>, AlchemyAPI<sup>3</sup> si ChartLyrics<sup>4</sup> sunt principalele API-uri utilizate in aplicatie.

Lucrarea e structurata in 4 sectiuni. Principalele concepte sunt descrise in sectiunea II, iar sectiunea III introduce arhitectura sistemului propus. De asemenea introduce si rezultatele obtinute si o comparatie cu alte sisteme asemanatoare. Concluziile sunt prezentate in sectiunea IV.

## II. CONTEXT

### A. Analiza sentimentelor cu AlchemyAPI

Atitudinea unui vorbitor sau a unui scriitor fata de o tema este scopul principal al analizei sentimentelor. Analiza automata a sentimentelor implica tehnici ale Procesarii Limbajului Natural, precum extragerea entitatilor cu nume sau a caracteristicilor unui obiect.

Conform lucrurilor [1], [2], [3], o opinie este o cvintupla  $(o_j, f_{jk}, so_{ijkl}, h_i, t_l)$  formata din:

- $o_j$  obiectul studiat
- $f_{jk}$  o caracteristica a  $o_j$
- $h_i$  detinatorul opiniei
- $t_l$  momentul exprimarii opiniei
- $so_{ijkl}$  valoare sentimentului detinatorului opiniei  $h_i$  asupra caracteristicii  $f_{jk}$  la momentul  $t_l$ .

AlchemyAPI este un framework specializat pe Analiza Textelor, permitand Extragerea Entitatilor cu Nume, a cuvintelor cheie, precum si Analiza Sentimentului. Este o abordare hibrida [11], combinand tehnicile statistice cu cele lingvistice si de machine learning, mai precis retele neuronale. Unul dintre marile sale avantaje este viteza cu care poate procesa date de dimensiuni mari.

Sistemul este intr-o continua schimbare, invatand din toate datele pe care le proceseaza. Setul de antrenare momentan are mai mult de 200 bilioane de cuvinte, de 285 de ori mai mult decat Wikipedia. Aici intra mai multe limbi precum si Entitati cu Nume extrase din documentele analizate care vin din diverse domenii de pe toate continentele.

In ce priveste Analiza sentimentelor, AlchemyAPI o permite la diverse nivele, incluzand nivelul document, entitate, cuvand cheie, sau specificat de utilizator. Acest nivel detaliat de analiza a sentimentului permite extragerea unor informatii valoroase din textul analizat.

### B. Servicii REST

Flexibilitatea in utilizare a serviciilor a facut ca acestea sa cunoasca o foarte larga raspandire. Un serviciu web este

<sup>1</sup> <http://www.last.fm/home>

<sup>2</sup> <http://musicbrainz.org/>

<sup>3</sup> <http://www.alchemyapi.com/>

<sup>4</sup> <http://www.chartlyrics.com/>

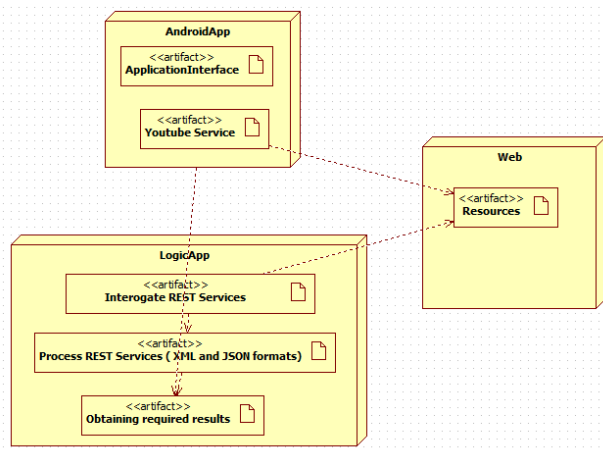


Fig. 1 Arhitectura sistemului

un sistem software construit sa suporte interactiunea masina-masina intr-o retea. Serviciul REST foloseste mesaje XML standardizate si este independent de o anumita platforma sau sistem de operare.

Serviciile web au doua caracteristici importante. In primul rand, sunt reutilizabile. Cele mai intalnite exemple in acest sens sunt serviciile de conversie monetara sau serviciile meteo. In al doilea rand, sunt interoperabile, facilitand integrarea aplicatiilor construite pe platforme diferite.

Din perspectiva arhitecturii lor, serviciile sunt impartite in doua mari categorii [10]: RPC (Remote Procedure Call) si REST (Representational State Transfer). De interes pentru noi sunt serviciile REST. Acestea sunt orientate pe resurse disponibile pe Internet, cu un URI asociat. Ofera operatii fara stare si manipuleaza reprezentari XML ale resurselor.

### III. DESCRIEREA SISTEMULUI

Arhitectura sistemului poate fi vizualizata in figura 1. Sistemul este compus din doua module. Modulul AndroidApp contine interafata aplicatiei si prezentarea cantece folosind serviciul Youtube. LogicApp foloseste serviciile REST, realizand apelurile si prelucrand rezultatele obtinute in formate XML sau JSON. Rezultatul prelucrării va fi utilizat de catre modulul AndroidApp.

Principalele functii ale sistemului sunt:

- 1) cautarea unui artist sau a unui cantece,
- 2) vizualizarea informatiilor despre un artist si cantece (precum top al albumelor, cantece, anul nasterii, artisti similari, poze),
- 3) ascultarea unui cantece,
- 4) vizualizarea versurilor unui cantece,
- 5) detectia sentimentului dintr-un cantece.

Informatiile despre artisti au fost extrase din colectia MusicBrainz. Aceasta este o enciclopedie care colecteaza informatii despre muzica si ofera functii de acces. Informatiile extrase sunt (i) numele artistului (ii) data si locul nasterii, (iii) artisti similari, (iv) concerte viitoare. Informatii de acelasi tip sunt extrase si din colectia LastFM pe baza serviciilor acesteia. ChartLyrics este serviciul folosit pentru extragerea versurilor. Alchemy API ofera servicii de analiza a textului, dintre care s-a folosit cu

precaderea cea de analiza a sentimentului. Pentru randarea si vizualizarea cantece s-a folosit Youtube API.

#### A. Comparatie cu alte sisteme asemanatoare

TABLE I. COMPARATIE CU ALTE SISTEME

SentiMusic	Deezer	Shazam	AudioGlow	Func
X	X	X	X	ascultare
X	X	X	X	vizualizare
X	-	-	-	sentiment
X	X	-	-	versuri
X	X	X	-	informatii
-	X	-	-	radio
X	X	-	-	artisti similari

Exista multe aplicatii muzicale pentru telefoane mobile<sup>5</sup>. MusicBrainz si LastFm, cele doua surse importante de informatii ale aplicatiei, pun la dispozitie si aplicatii pentru platforma Android. Intre alte aplicatii asemanatoare amintim: Shazam, Deezer, 8Tracks, N7Player, Pandora, AudioGlow.

In tabelul I este prezentata comparatia aplicatiei noastre cu trei dintre cele mai utilizate aplicatii. Din cunostintele noastre, nu exista nici o alta aplicatie muzicala care sa prezinte si detectia sentimentului unei piese pe baza versurilor acesteia.

In ce priveste analiza sentimentului in afara domeniului muzical, o aplicatie mobila care detecteaza starea emotionala a utilizatorului este dezvoltata de compania Samsung<sup>6</sup>.

#### B. Experimente

##### 1) Analiza unei liste de piese

Aplicatia foloseste doua metode oferite de AlchemyAPI: (i) una care extrage sentimentul dintr-un string si (ii) una care extrage sentimentul dintr-un URL. Primeia i se furnizeaza versurile, in timp ce celei de-a doua URL-ul unde se afla versurile unei piese. Cele doua metode masoara in principiu acelasi lucru, insa rezultatele sunt usor diferite, datorita in principal dublării titlului piesei in pagina versurilor.

Un prim experiment a vizat masurarea timpului mediu de analiza a unei piese. Pentru aceasta s-au analizat mai multe piese. In testarea automata a aplicatiei pe mai multe cantece, s-a tinut cont de constrangerea data de utilizarea gratuita a serviciilor, si anume limitarea serviciului ChartLyrics Api a apelurilor la unul o data la 20 de secunde. Testul s-a realizat pe 45 de cantece, timpul total fiind putin peste 15 minute, ceea ce inseamna ca intarzierea data de analiza sentimentului este nesemnificativa.

Un al doilea experiment a vizat comportamentul modulului de analiza a sentimentului pe o lista de 20 de cantece considerate de catre utilizatori fericite<sup>7</sup>, respectiv triste<sup>8</sup>. Rezultatele obtinute sunt descrise in tabelul II.

<sup>5</sup><http://www.techradar.com/news/phone-and-communications/mobile-phones/10-best-android-music-players-1114284>

<sup>6</sup><http://www.technologyreview.com/news/426560/a-smart-phone-that-knows-youre-angry/>

<sup>7</sup><http://www.billboard.com/articles/list/5915801/top-20-happy-songs-of-alltime>

<sup>8</sup><http://www.buzzfeed.com/awesomer/songs-you-cant-make-it-throughwithout-crying>

TABLE II. EXEMPLE DIN TESTAREA PE SETUL DE 20+20 PIESE

Numele cantecului	Sentiment din String
<b>Cantece vesele</b>	<b>Valoare asteptata: pozitiva</b>
Sha La La (Make Me Happy)	Positiv = 0.13
Don't Worry, Be Happy	Positiv = 0.03
Happy Together	Positiv = 0.31
Happy Birthday, Sweet Sixteen	Positiv = 0.01
<b>If It Makes You Happy</b>	<b>Negativ = -0.11</b>
Shiny Happy People	Positiv = 0.67
If You Wanna Be Happy	Positiv = 0.07
Hotel Happiness	Positiv = 0.15
<b>Cantece triste</b>	<b>Valoare asteptata: negativa</b>
Hurt	Negativ = -0.29
Mad World	Negativ = -0.15
<b>I See A Darkness</b>	<b>Positiv = 0.03</b>
To Build a Home	Negativ = -0.18
<b>Siren Song</b>	<b>Positiv = 0.03</b>
Roads	Negativ = -0.59
About Today	Negativ = -0.19
Dagger	Negativ = -0.30

AlchemyAPI returneaza o valoare pozitiva pentru un sentiment pozitiv, respectiv una negativa pentru un sentiment negativ de tristete. Din lista de 20 de cantece, 14 au obtinut valori pozitive, 3 au obtinut negative, si 3 valoare neutra. Din lista de 20 de cantece triste, 12 au fost identificate negative si 8 pozitive. Pentru cele identificate in alta categorie decat cea reala, valorile sunt totusi apropiate de 0.

### 2) Analiza detaliata pe un cantec:

Un alt experiment a vizat analiza sentimentului pe fragmente din cantec, precum si analiza sentimentului din entitatile identificate in versuri. S-au ales doua cantece si s-au aplicat 4 metode: (i) extragerea sentimentului dintr-un string care contine toate versurile, (ii) extragerea sentimentului din strofele cantecului luate separat, (iii) extragerea sentimentului din entitatile identificate in toate versurile, (iv) analiza in URL. S-a ales un cantec considerat foarte vesel impreuna cu un cantec considerat foarte trist.

Conform datelor partial prezentate in tabelul III, sentimental dominant al piesei este in concordanta cu sentimental asociat intregii piese. O evolutie a sentimentului pe fragmentele cantecului, fara refren, poate fi vazuta in figura 2. Aici se observa ca maximum de sentiment pozitiv (0,83) este atins in prima strofa, urmand ca in strofa a doua sa scada puternic (0,43), dar refrenul contine valori pozitive. Este de asteptat ca refrenul sa fie mai apropiat de valoarea globala a sentimentului decat strofele luata in parte, unde sentimentul fluctueaza.

Similar cantecului vesel, se poate observa in tabelul IV ca refrenul este o componenta puternica a mesajului intregii piese triste.

### 3) Comparatii cu alte unele de analiza a sentimentului

Un ultim experiment a urmarit compararea rezultatelor oferite de diferite unele de analiza a sentimentului aplicate din nou pe o piesa vesela si una trista. S-au considerat

versiunile demo pt Lexalytics<sup>9</sup>, Textalytics<sup>10</sup> si Vivekin<sup>11</sup>.

TABLE III. CANTEC FERICIT: R.E.M-"SHINY HAPPY PEOPLE"

Text analizat	Sentiment	Scor
Toate versurile	pozitiv	0.67
<b>Analiza detaliata</b>		
1. "Shiny happy people laughing"	pozitiv	0.83
2. "Meet me in the crowd People, people Throw your love around Love me, love me Take it into town Happy, happy Put it in the ground where the flowers grow Gold and silver shine."	Pozitiv	0.59
3. "Shiny happy people holding hands Shiny happy people holding hands Shiny happy people laughing"	Pozitiv	0.77
4. strofa 2	Pozitiv	0.45
5. refren1	Pozitiv	0.77
6. "Woah, here we go"	Pozitiv	0.65
7. refren2	Pozitiv	0.75
<b>URLGetTextSentiment</b>	Pozitiv	0.42
<b>URLGetRankedNamedEntities</b> Entity Type- Music Group Text- R.E.M.	Pozitiv	0.66

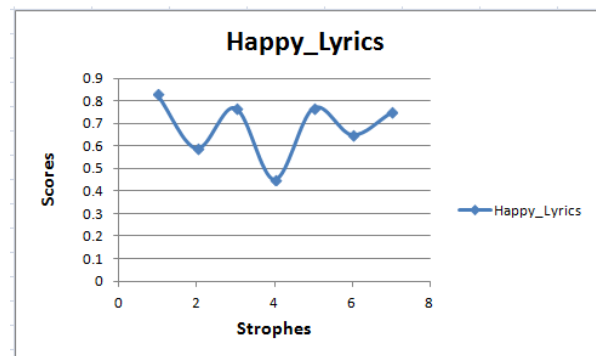


Fig. 2 Evolutia sentimentului in cantecul fericit ales

TABLE IV. CANTEC TRIST: R.E.M. "TRY NOT TO BREATHE"

Text analizat	Sentiment	Score
Toate versurile	Negativ	-0.45
<b>Analiza detaliata</b>		
1. "I will try not to breathe I can hold my head still with my hands at my knees These are the eyes of the old Shiver and fold"	Negativ	-0.53
2. "I need something to fly over my grave again I need something to breathe"	Negativ	-0.69
3. strofa	Negativ	-0.12
4. strofa	Negativ	-0.51
5. refren	Negativ	-0.57
6. "I will try not to worry you I have seen things you will never see Leave it to memory me Dont dare me to breathe"	Negativ	-0.09
<b>URLGetTextSentiment</b>	Negativ	-0.39
<b>URLGetRankedNamedEntities</b>	Negativ	-0.39

<sup>9</sup> <http://www.lexalytics.com/>

<sup>10</sup> <http://textalytics.com/>

<sup>11</sup> <http://sentiment.vivekn.com/>

TABLE V. COMPARATIE A UNELTELOR PE UN CANTEC TRIST

Sentiment/ Unealta	String	Grupul1	Grupul2	Grupul6
Alchemy	-0.45 Negativ	-0.53	-0.69	-0.09
Lexalytics	+0.52 Pozitiv	-0.49	0 neutru	+0.52
Textalytics	Pozitiv	-	-	Pozitiv
Vivekin	Negativ Grad de incredere 98.93	Negativ Grad de incredere 76.90	Negativ Grad de incredere 96.15	Negativ Grad de incredere 65.00

TABLE VI. COMPARATIE A UNELTELOR PE UN CANTEC VESEL

Sentiment/ Unealta	String	Grupul1	Grupul2	Grupul3
Alchemy	+0.67 Pozitiv	+0.83	+0.59	+0.77
Lexalytics	+0.53 Pozitiv	+0.49	+0.55	+0.49
Textalytics	Pozitiv	Pozitiv	Pozitiv	Pozitiv
Vivekin	Pozitiv Grad de incredere 87.12	Nega Grad de incredere 64.98	Pozitiv Grad de incredere 72.29	Negativ Grad de incredere 64.98

Rezultatele obtinute sunt prezentate in tabelul V, respectiv VI. Toate uneltele folosite reprezinta sentimentul de fericire cu valori pozitive, respectiv pe cel de tristete cu valori negative, mai putin Vivekin care returneaza grade de incredere pentru cele doua sentimente.

#### IV. CONCLUZII

S-a realizat o aplicatie mash-up pentru ascultarea muzicii si gasirea informatiilor despre artisti si cantece. O functionalitate importanta o reprezinta analiza sentimentului din versurile cantece folosind serviciile AlchemyAPI. S-a realizat o suita de teste vizand comportamentul pe o lista de piese, pe fragmente din cantece sau analize comparative cu alte unelte de analiza a sentimentului. Rezultatele obtinute sunt promitatoare si justifica, in opinia noastra, explorarea utilizarii analizei de sentiment si in industria media.

#### ACKNOWLEDGMENTS

Parte din cercetare s-a derulat in cadrul proiectului "Argumentare structurata pentru suport decizional cu constrangeri normative", programul PN-II Cooperari Bilaterale Romania-Republica Moldova, 2013-2014, UEFSCDI, respectiv GREEN-VANETS, proiect intern UTCN, 2013-2014.

#### REFERINTE

- [1] Bing Liu. "Sentiment Analysis and Opinion Mining (Synthesis Lectures on Human Language Technologies). Morgan & Claypool Publishers, 2012
- [2] Y. Yang, H.Chen "Machine recognition of music emotion: A review". ACM Transactions on Intelligent Systems and Technology, vol. 3, no. 3., 2012
- [3] Bing Liu, "Sentiment Analysis and Subjectivity. A Chapter in Handbook of Natural Language Processing", 2nd Edition, 2010.
- [4] Barthelet M., Fazekas G., and Sandler M. "Multidisciplinary perspectives on music emotion recognition: Implications for content and context-based models". In Proc. CMMR, pp. 492-507, 2012
- [5] H. Kopka and P. W. Daly, Advances in Knowledge-Based and Intelligent Information and Engineering Systems, Part1, 3rd ed. Harlow, England: Addison-Wesley, 1999.
- [6] Lee,Wei-Meng. "Android Application Development Cookbook". Available: <http://itebooks.info/book/1428/>
- [7] Steve Graham. "Servicii WEB cu Java. XML, SOAP, WSDL si UDDI"
- [8] Florian Mircea Boian. "Servicii WEB. Modele, Platforme, Aplicatii"
- [9] Leonard Richardson, Sam Ruby. "RESTful Web Services Web services for the real world"
- [10] Sabin Buraga, Lenuta Alboaie. 2006 "Servicii Web. Concepte de baza si implementari"
- [11] WhitePaper AlchemyAPI. Sentiment Analysis with AlchemyAPI: A Hybrid Approach, 2014