THE MILLION MUSICAL TWEETS DATASET: WHAT CAN WE LEARN FROM MICROBLOGS

David Hauger
Johannes Kepler University Linz
Austria
david.hauger@jku.at

Markus Schedl
Johannes Kepler University Linz
Austria
markus.schedl@jku.at

Andrej Košir
University of Ljubljana
Slovenia
andrej.kosir@ldos.fe.uni-lj.si

Marko Tkalčič
Johannes Kepler University Linz
Austria
marko.tkalcic@jku.at

ABSTRACT

Microblogs and Social Media applications are continuously growing in spread and importance. Users of Twitter, the currently most popular platform for microblogging, create more than a billion posts (called tweets) every week. Among all the different types of information being shared, some people post their music listening behavior, which is why Twitter became interesting for the Music Information Retrieval (MIR) community. Depending on the device and personal settings, some users provide geographic coordinates for their microposts.

Having continuously crawled and analyzed tweets for more than 500 days (17 months) we can now present the "Million Musical Tweet Dataset" (MMTD) – the biggest publicly available source of microblog-based music listening histories that includes geographic, temporal, and other contextual information. These extended information makes the MMTD outstanding from other datasets providing music listening histories.

We introduce the dataset, give basic statistics about its composition, and show how this dataset allows to detect new contextual music listening patterns by performing a comprehensive statistical investigation with respect to correlation between music taste and day of the week, hour of day, and country.

1. INTRODUCTION

Microblogs and social media have continuously been growing in importance over the past years — for end users, but also for industry and academia. Compared to other sources of information, they show high actuality and benefit from a large number of users. Twitter [1], for instance, the currently largest platform for microblogging, already has about 500 million users as of October 2012, according to Twitter CEO Dick Costolo [2] (the last official numbers date back to 200 million users in April 2011 [3]).

Microblogs have already been proven successful in a number of different contexts (see Section 2), but up to now they have hardly been exploited within the field of Music Information Retrieval (MIR). As there are no strict rules or specified formats for the up to 140 characters a Twitter post (tweet) consists of, Twitter is a relatively noisy source of information. However, thanks to hashtags and plugins for music players that automatically post music listening events, Twitter is a valuable source of information for MIR, when trying to incorporate or explicitly evaluate the user context. Compared to other sources like Last.fm [4], Twitter also provides information on geographic positions if available (for instance, if posted from a GPS-enabled device).

In this paper, we present a novel dataset of information derived from microblogs (tweets), describing the music listening habits of users. It is composed of preprocessed tweets, being consistently mapped to artists and songs from MusicBrainz [5]. As far as we are aware of, this dataset is currently the largest source of information on geospatial music listening events publicly available.

The remainder of the paper is structured as follows. First we give an overview of existing datasets and their applications in Section 2. In Section 3 we provide information about the acquisition of the dataset and some basic statistics. In Section 4 we conduct a statistical analysis of the correlation between music taste (measured via genre distribution) and temporal as well as geographical properties. In Section 5 we briefly present some ideas on how the dataset can be exploited, focusing on music visualization and contextual clustering. In Section 6 we summarize the work and outline possibilities of further exploiting the dataset.

2. RELATED WORK

Microblogging services like Twitter are continuously growing in importance. They have already been exploited

[1] http://www.twitter.com

[2] http://www.telegraph.co.uk/technology/twitter/9945505/Twitter-in-numbers.html

[3] http://huffingtonpost.com/2011/04/28/twitter-number-of-users_n_855177.html

[4] http://www.last.fm

[5] http://musicbrainz.org

for research in different areas, for instance, detection of breaking news [16], trends [12] [23], earthquakes [17], or health issues [14], even exploring spatio-temporal dynamics [10].

Nevertheless, within the MIR community, microblogs are still a relatively new source of information. As recent work stresses the importance of adding contextual information to music recommendation [1] [22], we believe that microblogs in general, and the proposed dataset in particular, provide valuable information for MIR. `Twitter` has already been used as a source for music similarity estimation [20], music recommendation [24], and for identifying cultural listening patterns [19].

Unfortunately, most existing datasets of `Twitter` posts, such as "The Edinburgh Twitter Corpus" [15], are not suited for geospatial MIR tasks as only very few tweets are related to music and less than 3% of the tweets have geolocalized information available. Also the frequently used dataset of the TREC 2011 and 2012 Microblog tracks [6] [13] is not suited. Although it contains approximately 16 million tweets, this dataset is not tailored to music-related activities, i.e. the amount of music-related posts is marginal.

Moreover, tweets have no specific format ($\leq$ 140 characters of unstructured text) and therefore need preprocessing to be assigned to a specific artist and/or song. Nevertheless, `Twitter` is one of the few sources with geolocalized listening information being available. Therefore, this source could be valuable for previous work like [2], where `Last.fm` was used as a source for listening histories. The most recent similar, but much smaller dataset for music-related microblogs is the "MusicMicro" dataset [18]. Besides being smaller, the "MusicMicro" dataset does not offer genre information nor other musical metadata (except for artist and song name).

As for datasets targeted at the music domain, the "Million Song Dataset" [3] and the "Yahoo! Music Dataset" [5] are quite popular. However, they do either not provide listening information ("Yahoo! Music Dataset") or only taste profiles without geo infomation ("Million Song Dataset"). The "Million Musical Tweet Dataset" (MMTD) presented here, in contrast, provides geolocalized listening information and links plain text tweets to respective artists and tracks, which allows for combination with content-based [4] and other contextual features [7]. The MMTD currently provides the biggest publicly available dataset for geolocalized music listening behavior. It can be downloaded from `http://www.cp.jku.at/datasets/MMTD/`.

The two pieces of information that make the MMTD unique are *temporal information* and *geographic information*. The former was shown to be extremely useful in the recommendation systems domain. Koren et al. [9] developed a matrix factorization-based method for modeling temporal dynamics in order to improve the rating prediction in the movies domain. Similarly, Koenigstein et al. [8], used this method to improve the rating prediction on the "Yahoo! Music Dataset" in the 2011 KDD Cup. Information on countries are valuable for culture-specific MIR approaches [21].

[6] `http://trec.nist.gov/data/tweets`

rank	country code	number of users	rank	country code	number of tweets
1	US	70,204	1	US	227,432
2	ID	30,605	2	DE	153,163
3	BR	24,985	3	BR	145,049
4	MY	14,771	4	GB	130,951
5	FR	13,890	5	ID	94,245
6	GB	9,006	6	FR	65,525
7	RU	5,234	7	MY	50,648
8	NL	5,223	8	CA	27,370
9	MX	4,538	9	RU	23,542
10	TR	2,878	10	MX	18,717
11	ES	2,847	11	NL	18,320
12	SG	2,422	12	TR	14,479
13	PH	2,385	13	ES	11,811
14	CA	2,344	14	SG	7,637
15	IT	1,521	15	IT	6,412
16	JP	1,501	16	AR	6,319
17	ZA	1,297	17	PH	5,723
18	DE	1,133	18	JP	5,529
19	UA	1,062	19	UA	4,940
20	AR	874	20	ZA	3,733

Table 1. Top-20 countries by number of users and tweets.

3. DATA ACQUISITION AND BASIC STATISTICS

In this section, we provide the background of the data acquisition and processing that led to the presented dataset. We further give some basic dataset statistics.

3.1 Data Acquisition

Between September 2011 and April 2013 we crawled the `Twitter` Streaming API [7], which provides a random subset of 1% of all tweets. We retained only tweets with geographic information attached (less than 3% of all tweets) and including potentially music-related hashtags that have already been proven successful [6], e.g. `#nowplaying`, `#np`, `#itunes`, `#musicmonday` and `#thisismyjam`. We employed the pattern-based approach described in [6] to map the content of the tweets to artists and tracks. We used the `MusicBrainz` database for indexing, which covers $\approx$ 30% of all tweets including the desired hashtags. Of course this method creates some bias in terms of cultural music listening patterns as the dataset is restricted to `Twitter` users posting musical information and there might be other conventions for the usage of hashtags in non-western countries.

To enable experimenting with the MMTD on a semantic level, we used a web service provided by `Mapquest` [8] to map geographic coordinates to cities, countries, and other geographic entities. To get comparable local time (which is not directly provided by `Twitter`) we used `GeoNames` [9] for retrieving the time zones for the geographic coordinates. The top 20 countries in terms of number of users, respectively number of tweets, are listed in Table 1.

In addition to this contextual information, we added genre information by querying `Last.fm` for tags on the

[7] `https://dev.twitter.com/docs/streaming-apis`
[8] `http://www.mapquest.com`
[9] `http://geonames.org`

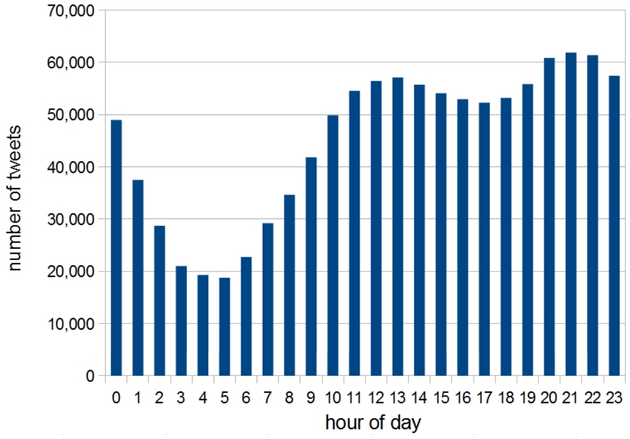

Figure 1. Number of tweets per hour of the day.

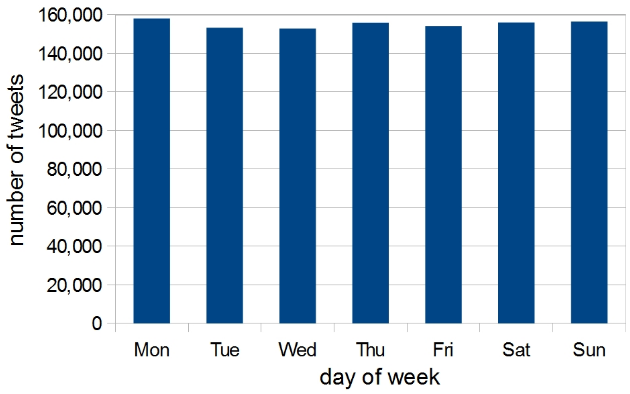

Figure 2. Number of tweets per day of the week.

artist and song level, filtering them by the 20 most popular genres from `Allmusic`[10]. This resulted in a multi-genre feature vector for each tweet.

3.2 Basic Statistics

The "Million Musical Tweets Dataset" (MMTD) is based on 1,086,808 tweets referring to 133,968 unique tracks ($mean$ = 8.11 tweets/track; σ = 44.94; $median$ = 1) by 25,060 different artists ($mean$ = 43.37 tweets/artist; σ = 327.03; $median$ = 3). The tweets were created by 215,375 users from 202 different countries. On average we have 1,078 users per country (σ = 5,848; $median$ = 29.5) and 5,381 tweets per country (σ = 25,032.45; $median$ = 74.5), each user creating on average 5.08 tweets (σ = 268.19; $median$ = 1).

Analyzing the temporal distribution of tweets shows that twitterers are less active during night, as expected (see Figure 1). However, there is no significant difference between the days of the week (see Figure 2).

Using the pre-filtered `Last.fm` tags for multiple genre assignment, we obtained 276,697 tweets per genre (σ = 256,662). The distribution of genres is unbalanced as can be seen in Figure 3. For future work, different genre classifications could be investigated, for instance, using ontologies or synonyms for the provided tags.

[10] `http://www.allmusic.com`

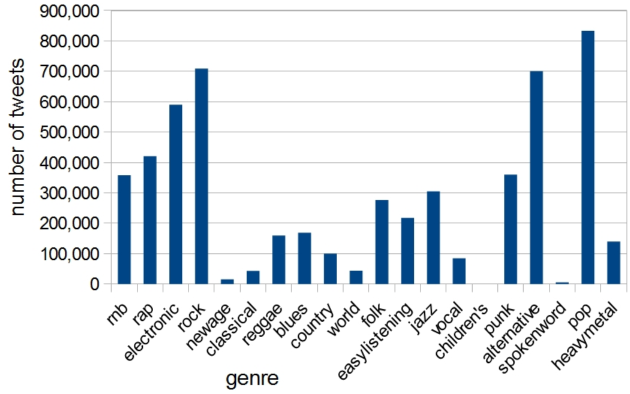

Figure 3. Number of tweets per genre.

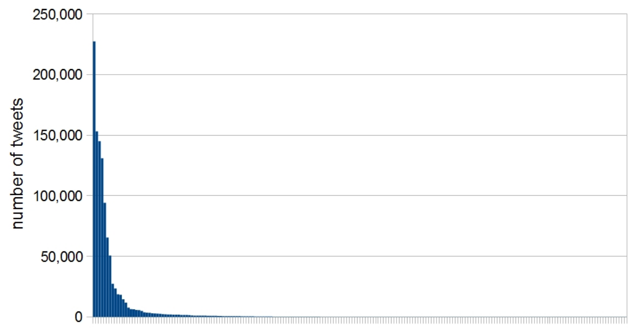

Figure 4. Number of tweets per country.

4. SPATIO-TEMPORAL STATISTICAL ASPECTS

In order to show the utility of the presented dataset we performed two experiments: (i) a geographical analysis of the listening preferences reflected in the dataset and (ii) a temporal analysis. For that purpose, we aggregated the dataset on a geographical and on a temporal basis. The tweets were first grouped by country. Then for each country the tweets were grouped by the day of the week. Finally, for each country and day of the week, the tweets were grouped by the hour of the day (e.g., all tweets from 2:00 to 2:59 were grouped together). When grouping, we summed the genre vectors of all the tweets belonging to a group, which yielded a summed vector, or histogram, of the genre distribution in each group. An excerpt of the aggregated data is shown in Table 2.

The distribution of tweets among countries, on the daily basis, and on the hour-of-the-day basis are shown, respectively, in Figures 4, 2, and 1.

4.1 Geographical analysis of the dataset

In this experiment we addressed the research question The research question addressed in this experiment is whether there are any differences between musical tastes among different countries that are reflected in the dataset. When we use the term *musical taste* we refer to how often specific genres have been played in the observed country (or other cluster in the next subsection).

To answer this question we first conducted the Kruskal-Wallis ANOVA (Analysis of Variance) (see [11] for de-

country	day of week	time of day	rnb	rap	el	ro	na	cl	re	bl	co	wo	fo	el	ja	vo	ch	pu	al	sw	pop	hm
Brazil	4 (Fri)	00	199	343	544	752	10	52	121	167	78	62	290	157	273	44	0	464	707	1	778	262
Brazil	4 (Fri)	05	29	30	49	70	2	8	13	19	5	6	22	17	28	6	0	39	65	0	69	17
Brazil	4 (Fri)	12	312	476	700	920	4	74	190	186	103	58	331	250	313	59	0	535	880	1	950	296
Brazil	4 (Fri)	17	260	424	647	873	6	60	150	164	99	44	299	201	315	45	0	533	847	3	880	310
Brazil	4 (Fri)	21	332	578	862	1104	6	60	194	220	149	62	385	247	382	63	0	684	1073	0	1164	426
France	6 (Sun)	00	213	277	194	208	4	10	90	50	17	30	69	56	90	21	0	80	196	6	260	24
France	6 (Sun)	05	22	29	22	18	1	0	8	4	0	1	4	3	9	2	0	6	17	0	24	0
France	6 (Sun)	12	422	529	440	417	4	12	173	105	27	51	112	116	203	40	0	173	403	3	528	41
France	6 (Sun)	17	280	325	276	280	1	13	108	75	26	32	72	89	113	28	0	118	302	5	366	20
France	6 (Sun)	21	265	331	283	283	1	20	113	72	30	22	91	83	129	34	0	106	268	2	348	22
Indonesia	0 (Mon)	00	128	145	253	352	7	19	60	88	72	18	173	155	154	41	1	187	322	2	395	76
Indonesia	0 (Mon)	05	27	34	42	59	2	5	13	12	13	1	26	23	27	6	0	32	57	0	67	15
Indonesia	0 (Mon)	12	206	223	359	509	3	29	58	85	81	17	185	211	175	47	0	226	466	0	580	104
Indonesia	0 (Mon)	17	245	295	428	619	16	35	84	108	84	17	206	255	201	55	0	296	569	1	706	110
Indonesia	0 (Mon)	21	273	316	511	722	25	68	81	141	137	31	302	360	287	85	0	360	680	2	843	144
Malaysia	2 (Wed)	00	133	169	245	306	8	23	40	58	47	15	140	98	122	32	0	163	295	2	358	59
Malaysia	2 (Wed)	05	9	9	18	21	1	1	2	3	1	2	8	10	6	4	0	13	20	0	29	5
Malaysia	2 (Wed)	12	75	84	137	173	6	8	26	29	27	5	62	53	63	14	0	92	158	0	209	37
Malaysia	2 (Wed)	17	167	178	269	327	9	12	53	45	53	20	133	100	111	38	0	157	302	2	392	72
Malaysia	2 (Wed)	21	173	180	271	355	4	18	52	51	68	22	133	150	123	39	0	157	332	2	432	69
United States	1 (Tue)	00	992	1062	865	923	12	47	392	218	145	59	313	267	466	122	0	502	994	13	1348	143
United States	1 (Tue)	05	248	266	219	251	2	9	100	53	40	14	80	79	125	29	0	132	259	4	346	42
United States	1 (Tue)	12	632	757	667	762	15	30	283	183	124	45	266	198	347	75	0	415	785	18	973	144
United States	1 (Tue)	17	696	833	678	712	7	35	270	145	107	35	239	210	341	69	0	359	792	14	991	104
United States	1 (Tue)	21	1125	1311	1076	1168	15	52	459	297	180	57	405	361	543	126	0	617	1241	17	1649	153

Table 2. Some random examples for aggregated data on a per-country, per-day-of-week, and per-hour-of-day basis from the Top-5 countries in terms of numbers of users. The order of genres corresponds to that in Figure 3.

tails) on the whole dataset grouping the data by genre and comparing each genre separately. The analysis showed that all p-values were $p < 0.001$, meaning that there are significant differences among all variables according to the geographical location of tweets.

After the ANOVA showed significant differences, we proceeded with a pair-wise comparison between the countries. Since there are 202 countries, this would mean roughly 20,000 comparisons which makes it very hard, if not impossible, to interpret. As such an analysis would be hard to perform, we opted to choose a subset of 20 countries to perform the pairwise comparison. We chose the 20 countries with the biggest number of tweets (see Table 1).

In the pairwise comparison, we compared the histograms of all the 20 genres, among 20 selected countries using the Chi-square goodness of fit test (see [11] for details). The test result shows a p-value $p < 0.001$ for all pairs. Hence, all countries are significantly different from each other (compare to previous experiments on cultural listening patterns in [19]).

4.2 Temporal analysis of the dataset

We compared the distribution of genre preferences among different days of the week using the Kruskal-Wallis ANOVA to see whether listening habits are different on weekdays and on weekends. However, the test did reveal no significant differences among days of the week. Even when we narrowed the selection of countries down to five culturally distinct countries (Saudi Arabia, Malaysia, Germany, United Kingdom and Brazil), the differences were not significant. This means that the listening habits of the users present in this dataset do not differ between days. Two possible explanations for this finding are: (i) although someone might listen to the own favorites on week-ends and to artists reflecting the "common taste" among colleagues on working days, only music that is really liked is also posted via `Twitter` (if it is not automatically tweeted), or (ii) the aggregation already reflects this "common taste" for any day of the week.

Based on the distribution of the listening habits (tweets grouped by the hour of the day), as depicted in Figure 1, we clustered the tweets in the following hour-of-day groups, using the minimum at $5am$, and the local maxima at $noon$ and $10pm$ as borders:

- from 05:00 to 11:59 (morning)
- from 12:00 to 21:59 (afternoon/evening)
- from 22:00 to 04:59 (night)

Performing the cross-group test for all countries with the Kruskal-Wallis ANOVA showed that all p-values are $p < 0.05$, which means that there are significant differences among all the hour-of-day groups, e.g. people listen to different music on Monday mornings vs. Saturday nights. Repeating the test for the 5 culturally different countries, however, not all the p-values are $p < 0.05$, which we report in Table 3. This means that, for these countries, having $p > 0.05$, some genres do not show significantly different playing frequency among the various hour-of-day groups.

5. USING THE MILLION MUSICAL TWEETS DATASET FOR VISUALIZATION

Having discussed statistics of our dataset, the current section briefly points out an example of how the information within the MMTD may be used for clustering and visualizion. As multi-genre-vectors are not suited to map tweets to a certain color, we decided to use the approach presented by Hauger and Schedl [6] using non-negative matrix factorization of `Last.fm` genre tags and latent factors to assign tweets to a color.

We aggregated the tweets of our dataset by their geographic coordinates and displayed them as circles on a map, where the size of the circle represents the number of tweets for this area and the color represents the most

	rnb	rap	el	ro	na	cl	re	bl	co	wo	fo	el	ja	vo	ch	pu	al	sw	pop	hm
Chi-square	15.86	25.74	23.61	23.59	2.09	5.41	11.61	6.82	8.05	4.58	13.22	15.74	10.50	2.53	1.99	21.61	24.44	1.92	23.99	16.00
df	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
Asymp. Sig.	0	0	0	0	0.352	0.067	0.003	0.033	0.018	0.101	0.001	0	0.005	0.282	0.371	0	0	0.384	0	0

Table 3. Kruskal-Wallis ANOVA results for the differences among the hour-of-day variable grouped by genre for the 5 culturally different countries.

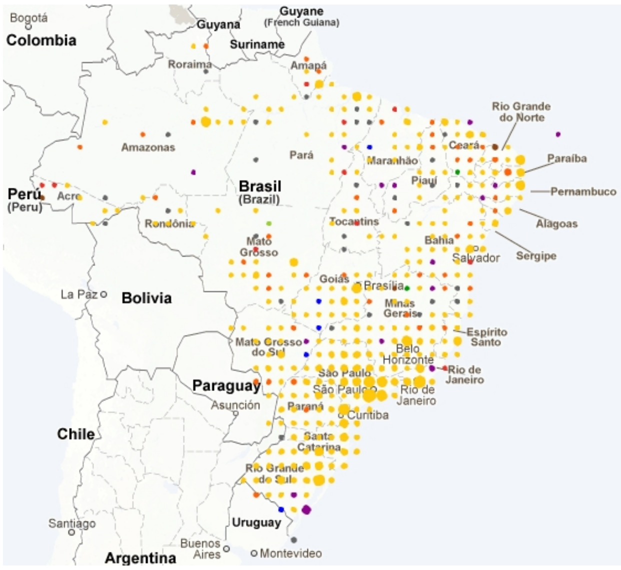

Figure 5. Visualizing genres / latent factors for Brazil.

popular genre or latent factor within this set of tweets. According to our findings that there are significant differences in the genre distribution for countries, we hence visually show differences between countries.

Figure 5 shows that in Brazil the genre cluster for “Rock” (yellow) is by far the most popular one. Comparing it to France, the European country with the largest number of `Twitter` users, we see that French twitterers show a strong preference for the group representing the “Rap” and “Hip-Hop” cluster (violet).

6. DISCUSSION AND FUTURE WORK

In this paper, we presented the “Million Musical Tweets Dataset” (MMTD). Its unique property of including time and geo-location data allows the research community to follow novel research avenues.

The results of the statistical tests showed that there are significant differences in musical tastes (expressed through multi-genre vectors) among different clusters (both on the geographical basis and on the temporal basis). These differences could be exploited for developing adaptive systems/services on the geo-temporal basis (e.g., contextual filtering of music based on country and/or time of day). We also presented one way of visualizing differences in geospatial music listening patterns.

The proposed dataset is publicly available and may be used, for instance, for contextual music recommendation or similarity estimation. As for the latter, the MMTD could be used to build hybrid similarity functions including audio features, contextual music features (tags, playlist co-

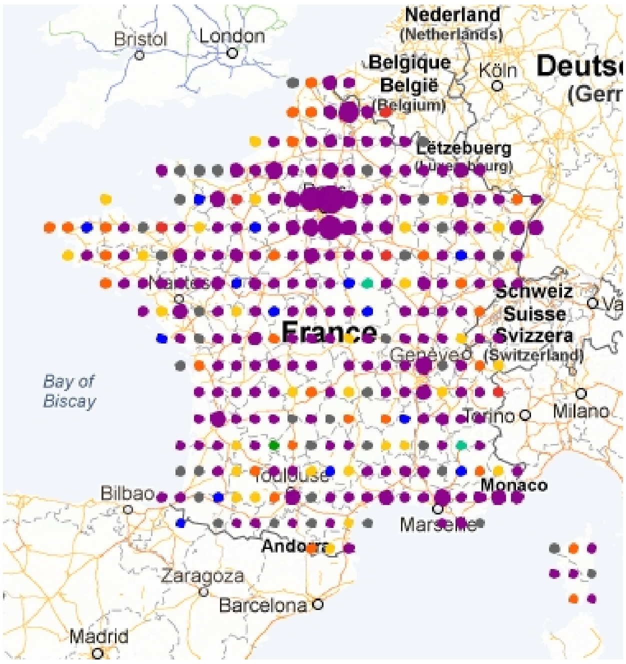

Figure 6. Visualizing genres / latent factors for France.

occurrences), and user context information.

Future work will put emphasis on user-related examination of the dataset, looking into other user-centric properties like age, gender, or twitting activity, as well as on the search for “cultural clusters” of different countries and/or cities. In this vein, it would also be interesting to examine whether there are differences between urban and rural areas.

7. ACKNOWLEDGMENTS

This research is supported by the Austrian Science Funds (FWF): P22856, P25655, and by the European Union FP7 programme through the PHENICX project (grant agreement no. 601166).

8. REFERENCES

[1] Linas Baltrunas, Bernd Ludwig, Stefan Peer, and Francesco Ricci. Context-Aware Places of Interest Recommendations and Explanations. *Joint Proceedings of the Workshop on Decision Making and Recommendation Acceptance Issues in Recommender Systems (DEMRA) and the 2nd Workshop on User Models for Motivational Systems: The affective and the rational routes to persuasion (UMMS)*, 2011.

[2] D. Baur, F. Seiffert, M. Sedlmair, and S. Boring. The streams of our lives: Visualizing listening histories in

context. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1119–1128, 2010.

[3] Thierry Bertin-Mahieux, Daniel P.W. Ellis, Brian Whitman, and Paul Lamere. The million song dataset. In *Proceedings of the 12th International Conference on Music Information Retrieval (ISMIR)*, 2011.

[4] Michael A. Casey, Remco Veltkamp, Masataka Goto, Marc Leman, Christophe Rhodes, and Malcolm Slaney. Content-Based Music Information Retrieval: Current Directions and Future Challenges. *Proceedings of the IEEE*, 96:668–696, April 2008.

[5] Gideon Dror and Noam Koenigstein and Yehuda Koren and Markus Weimer. The Yahoo! Music Dataset and KDD-Cup'11. In *Proceedings of KDDCup 2011*, 2011.

[6] David Hauger and Markus Schedl. Exploring Geospatial Music Listening Patterns in Microblog Data. In *Proceedings of the 10th International Workshop on Adaptive Multimedia Retrieval (AMR)*, Copenhagen, Denmark, October 2012.

[7] Peter Knees and Markus Schedl. A Survey of Music Similarity and Recommendation from Music Context Data. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP)*, 2013.

[8] Noam Koenigstein, Gideon Dror, and Yehuda Koren. Yahoo! music recommendations. *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys)*, page 165, 2011.

[9] Yehuda Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89, April 2010.

[10] Chung-Hong Lee, Hsin-Chang Yang, Tzan-Feng Chien, and Wei-Shiang Wen. A Novel Approach for Event Detection by Mining Spatio-temporal Information on Microblogs. In *International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 254–259, July 2011.

[11] E L Lehman, Joseph P. Romano, and Erich L. Lehmann. *Testing Statistical Hypotheses*. Springer Texts in Statistics. Springer New York, New York, NY, 2005.

[12] Michael Mathioudakis and Nick Koudas. Twittermonitor: trend detection over the twitter stream. In *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*, pages 1155–1158, New York, NY, USA, 2010. ACM.

[13] Richard McCreadie, Ian Soboroff, Jimmy Lin, Craig Macdonald, Iadh Ounis, and Dean McCullough. On Building a Reusable Twitter Corpus. In *Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, Portland, OR, USA, August 12-–16 2012.

[14] Michael J. Paul and Mark Dredze. You Are What You Tweet : Analyzing Twitter for Public Health. *Artificial Intelligence*, pages 265–272, 2011.

[15] Saša Petrović, Miles Osborne, and Victor Lavrenko. The edinburgh twitter corpus. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Linguistics in a World of Social Media*, pages 25–26, Los Angeles, California, USA, June 2010. Association for Computational Linguistics.

[16] S. Phuvipadawat and T. Murata. Breaking news detection and tracking in twitter. In *Web Intelligence and Intelligent Agent Technology (WI-IAT), 2010 IEEE/WIC/ACM International Conference on*, volume 3, pages 120–123, 2010.

[17] Takeshi Sakaki, Makoto Okazaki, and Yutaka Matsuo. Earthquake Shakes Twitter Users: Real-Time Event Detection by Social Sensors. In *Proceedings of the 19th International Conference on World Wide Web (WWW)*, May 2010.

[18] Markus Schedl. Leveraging Microblogs for Spatiotemporal Music Information Retrieval. *Lecture Notes In Computer Science*, 7814:796–799, 2013.

[19] Markus Schedl and David Hauger. Mining Microblogs to Infer Music Artist Similarity and Cultural Listening Patterns. In *Proceedings of the 21st International World Wide Web Conference (WWW): 4th International Workshop on Advances in Music Information Research: "The Web of Music" (AdMIRe)*, Lyon, France, 2012.

[20] Markus Schedl, David Hauger, and Julián Urbano. Harvesting microblogs for contextual music similarity estimation - a co-occurrence-based framework. *Multimedia Systems*, 2013.

[21] Xavier Serra. Data Gathering for a Culture Specific Approach in MIR. In *Proceedings of the 21st International World Wide Web Conference (WWW): 4th International Workshop on Advances in Music Information Research (AdMIRe)*, Lyon, France, April 17 2012.

[22] Bo Shao. *User-centric music information retrieval*. PhD thesis, Florida International University, Miami, FL, USA, 2011. AAI3472062.

[23] Beaux Sharifi, Mark-Anthony Hutton, and Jugal Kalita. Summarizing Microblogs Automatically. In *Proceedings of NAACL HLT*, June 2010.

[24] Eva Zangerle, Wolfgang Gassler, and Günther Specht. Exploiting twitter's collective knowledge for music recommendations. In *Proceedings of the 21st International World Wide Web Conference (WWW): Making Sense of Microposts (#MSM2012)*, pages 14–17, 2012.