Recognizing Musical Entities in User-generated Content

Lorenzo Porcaro[1], Horacio Saggion[2]

[1] Universitat Pompeu Fabra,
Music Technology Group, Barcelona,
Spain

[2] Universitat Pompeu Fabra,
TALN Natural Language Processing Group, Barcelona,
Spain

{lorenzo.porcaro, horacio.saggion}@upf.edu

Abstract. Recognizing Musical Entities is important for Music Information Retrieval (MIR) since it can improve the performance of several tasks such as music recommendation, genre classification or artist similarity. However, most entity recognition systems in the music domain have concentrated on formal texts (e.g. artists' biographies, encyclopedic articles, etc.), ignoring rich and noisy user-generated content. In this work, we present a novel method to recognize musical entities in Twitter content generated by users following a classical music radio channel. Our approach takes advantage of both formal radio schedule and users' tweets to improve entity recognition. We instantiate several machine learning algorithms to perform entity recognition combining task-specific and corpus-based features. We also show how to improve recognition results by jointly considering formal and user-generated content.

Keywords. Named entity recognition, music information retrieval, user-generated content.

1 Introduction

The increasing use of social media and microblogging services has broken new ground in the field of Information Extraction (IE) from user-generated content (UGC). Understanding the information contained in users' content has become one of the main goals for many applications, due to the uniqueness and the variety of this data [4]. However, the highly informal and noisy status of these sources makes it difficult to apply techniques proposed by the NLP community for dealing with formal and structured content [21].

In this work, we analyze a set of tweets related to a specific classical music radio channel, *BBC Radio 3*[1], interested in detecting two types of musical named entities, *Contributor* (person related to a musical work) and *Musical Work* (musical composition or recording).

The method proposed makes use of the information extracted from the radio schedule for creating links between users' tweets and tracks broadcasted. Thanks to this linking, we aim to detect when users refer to entities included into the schedule. Apart from that, we consider a series of linguistic features, partly taken from the NLP literature and partly specifically designed for this task, for building statistical models able to recognize the musical entities. To that aim, we perform several experiments with a supervised learning model, Support Vector Machine (SVM), and a recurrent neural network architecture, a bidirectional LSTM with a CRF layer (biLSTM-CRF).

The contributions of this work are summarized as follows:

— A method to recognize musical entities from user-generated content which combines contextual information (i.e. radio schedule)

---

[1] www.twitter.com/BBCRadio3

with Machine Learning models for improving the accuracy while recognizing the entities.

— The release of language resources such as an user-generated and bot-generated Twitter corpora manually annotated, usable for both MIR and NLP researches, and domain specific word embeddings.

The paper is structured as follows. In Section 2, we present a review of the previous works related to Named Entity Recognition, focusing on its application on UGC and MIR. Afterwards, in Section 3 it is presented the methodology of this work, describing the dataset and the method proposed. In Section 4, the results obtained are shown. Finally, in Section 5 conclusions are discussed.

2 Related Work

Named Entity Recognition (NER), or alternatively Named Entity Recognition and Classification (NERC), is the task of detecting entities in an input text and to assign them to a specific class. It starts to be defined in the early '80, and over the years several approaches have been proposed [11]. Early systems were based on handcrafted rule-based algorithms, while afterward thanks to advancements in Machine Learning techniques, probabilistic models started to be integrated into NER systems.

In particular, new developments in neural architectures have become an important resource for this task. Their main advantages are that they do not need language-specific knowledge resources [6], and they are robust to the noisy and short nature of social media messages [7]. Indeed, according to a performance analysis of several Named Entity Recognition and Linking systems presented in [1], it has been found that poor capitalization is one of the main issues when dealing with microblog content. Apart from that, typographic errors and the ubiquitous occurrence of out-of-vocabulary (OOV) words also cause drops in NER recall and precision, together with shortenings and slang, particularly pronounced in tweets.

Table 1. Examples of user-generated tweets

1	No Schoenberg or Webern?? Beethoven is there but not his pno sonata op. 101??
2	Heard some of Opera 'Oberon' today... Weber... Only a little....
3	Cavalleria Rusticana...hm..from a Competition that very nearly didn't get entered!

Music Information Retrieval (MIR) is an interdisciplinary field which borrows tools of several disciplines, such as signal processing, musicology, machine learning, psychology and many others, for extracting knowledge from musical objects (be them audio, texts, etc.) [10]. In the last decade, several MIR tasks have benefited from NLP, such as sound and music recommendation [15], automatic summary of song review [23], artist similarity [22] and genre classification [12].

In the field of IE, a first approach for detecting musical named entities from raw text, based on Hidden Markov Models, has been proposed in [26]. In [13], the authors combine state-of-the-art Entity Linking (EL) systems to tackle the problem of detecting musical entities from raw texts. The method proposed relies on the *argumentum ad populum* intuition, so if two or more different EL systems perform the same prediction in linking a named entity mention, the more likely this prediction is to be correct. In detail, the off-the-shelf systems used are: DBpedia Spotlight [8], TagMe [2], Babelfy [9]. Moreover, a first Musical Entity Linking, MEL[2] has been presented in [14] which combines different state-of-the-art NLP libraries and SimpleBrainz, an RDF knowledge base created from MusicBrainz[3] after a simplification process.

Furthermore, *Twitter* has also been at the center of many studies done by the MIR community. As example, for building a music recommender system [24] analyzes tweets containing keywords like *nowplaying* or *listeningto*. In [22], a similar dataset it is used for discovering cultural listening patterns.

---

[2]http://mel.mtg.upf.edu
[3]https://musicbrainz.org

Publicly available *Twitter* corpora built for MIR investigations have been created, among others the *Million Musical Tweets* dataset[4] [5] and the *#nowplaying* dataset[5] [25].

3 Methodology

We propose a hybrid method which recognizes musical entities in UGC using both contextual and linguistic information. We focus on detecting two types of entities:

— Contributor: person who is related to a musical work (composer, performer, conductor, etc).

— Musical Work: musical composition or recording (symphony, concerto, overture, etc).

As case study, we have chosen to analyze tweets extracted from the channel of a classical music radio, *BBC Radio 3*. The choice to focus on classical music has been mostly motivated by the particular discrepancy between the informal language used in the social platform and the formal nomenclature of contributors and musical works. Indeed, users when referring to a musician or to a classical piece in a tweet, rarely use the full name of the person or of the work, as shown in Table 2.

We extract information from the radio schedule for recreating the musical context to analyze user-generated tweets, detecting when they are referring to a specific work or contributor recently played. We manage to associate to every track broadcasted a list of entities, thanks to the tweets automatically posted by the *BBC Radio3 Music Bot*[6], where it is described the track actually on air in the radio. In Table 3, examples of bot-generated tweets are shown.

Afterwards, we detect the entities on the user-generated content by means of two methods: on one side, we use the entities extracted from the radio schedule for generating candidates entities in the user-generated tweets, thanks to a matching algorithm based on time proximity and string similarity. On the other side, we create a statistical model capable of detecting entities directly from

Table 2. Example of entities annotated (*top*) and corresponding formal forms (*down*), from the user-generated tweet (1) in Table 1

Informal form
Schoenberg
Webern
Beethoven
pno sonata op. 101
Formal form
Arnold Franz Walter Schoenberg
Anton Friedrich Wilhelm Webern
Ludwig Van Beethoven
Piano Sonata No. 28 in A major, Op. 101

Table 3. Examples of bot-generated tweets

1	Now Playing Joaquín Rodrigo, Goran Listes - 3 Piezas españolas for guitar #joaquínrodrigo,#goranlistes
2	Now Playing Robert Schumann, Luka Mitev - Phantasiestücke, Op 73 #robertschumann,#lukamitev
3	Now Playing Pyotr Ilyich Tchaikovsky, MusicAeterna - Symphony No.6 in B minor #pyotrilyichtchaikovsky, #musicaeterna

the UGC, aimed to model the informal language of the raw texts. In Figure 1, an overview of the system proposed is presented.

3.1 Dataset

In May 2018, we crawled Twitter using the Python library *Tweepy*[7], creating two datasets on which *Contributor* and *Musical Work* entities have been manually annotated, using Inside-outside-beginning tags [19].

The first set contains user-generated tweets related to the *BBC Radio 3* channel. It represents the source of user-generated content on which we aim to predict the named entities. We create it filtering the messages containing hashtags related to *BBC Radio 3*, such as *#BBCRadio3* or *#BBCR3*. We obtain a set of 2,225 unique user-generated tweets.

---

[4]http://www.cp.jku.at/datasets/MMTD
[5]https://zenodo.org/record/2594483
[6]https://twitter.com/BBCR3MusicBot

[7]http://www.tweepy.org

Table 4. Example of musical named entities annotated

Beethoven	is	there	but
B-CONTR	O	O	O
not	his	pno	sonata
O	O	B-WORK	I-WORK

The second set consists of the messages automatically generated by the *BBC Radio 3 Music Bot*. This set contains 5,093 automatically generated tweets, thanks to which we have recreated the schedule.

In Table 5, the amount of tokens and entities annotated are reported for the two datasets. For evaluation purposes, both sets are split in a training part (80%) and two test sets (10% each one) randomly chosen. Within the user-generated corpora, entities annotated are only about 5% of the whole amount of tokens. In the case of the automatically generated tweets, the percentage is significantly greater and entities represent about the 50%.
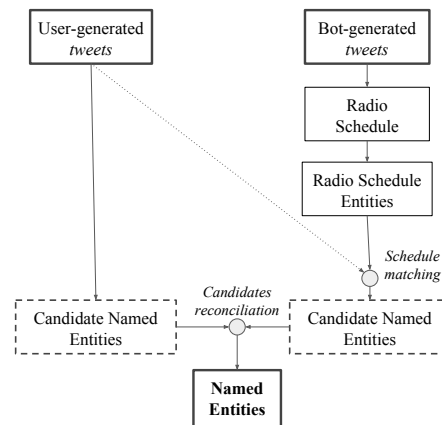
3.2 NER System

According to the literature reviewed, state-of-the-art NER systems proposed by the NLP community are not tailored to detect musical entities in user-generated content. Consequently, our first objective has been to understand how to adapt existing systems for achieving significant results in this task.

In the following sections, we describe separately the features, the word embeddings and the models considered. All the resources used are publicy available[8].

3.2.1 Features' Description

We define a set of features for characterizing the text at the token level. We mix standard linguistic features, such as Part-Of-Speech (POS) and chunk tag, together with several gazetteers specifically built for classical music, and a series of features representing tokens' left and right context.

---

[8] https://github.com/MTG/music-ner

Fig. 1. Overview of the NER system proposed

For extracting the POS and the chunk tag we use the Python library *twitter_nlp*[9], presented in [21].

In total, we define 26 features for describing each token: 1) POS tag; 2) Chunk tag; 3) Position of the token within the text, normalized between 0 and 1; 4) If the token starts with a capital letter; 5) If the token is a digit. Gazetteers: 6) Contributor first names; 7) Contributor last names; 8) Contributor types ("soprano", "violinist", etc.); 9) Classical work types ("symphony", "overture", etc.); 10) Musical instruments; 11) Opus forms ("op", "opus"); 12) Work number forms ("no", "number"); 13) Work keys ("C", "D", "E", "F" , "G" , "A", "B", "flat", "sharp"); 14) Work Modes ("major", "minor", "m"). Finally, we complete the tokens' description including as token's features the surface form, the POS and the chunk tag of the previous and the following two tokens (12 features).

3.2.2 Word Embedding

We consider two sets of GloVe word embeddings [16] for training the neural architecture, one pre-trained with 2B of tweets, publicy downloadable[10], one trained with a corpora of 300K tweets collected during the 2014-2017 *BBC Proms Festivals* and disjoint from the data used in our experiments.

---

[9] https://github.com/aritter/twitter_nlp
[10] https://github.com/stanfordnlp/GloVe

Table 5. Tokens' distributions within the two datasets: user-generated tweets (*top*) and bot-generated tweets (*bottom*)

	Training	TestA	TestB
Contributor	1.069 (3,12%)	119 (2,96%)	127 (2,97%)
Musical Work	964 (2,81%)	118 (2,93%)	163 (3,81%)
Total tokens	34.247	4.016	4.275
Contributor	15.162 (27,50%)	1.852 (22,93%)	1.879 (27,30%)
Musical Work	12.904 (23,40%)	1.625 (23,56%)	1.689 (24,48%)
Total tokens	55.122	6.897	6.881

3.2.3 Models

The first model considered for this task has been the John Platt's sequential minimal optimization algorithm for training a support vector classifier [17], implemented in WEKA [3]. Indeed, in [18] results shown that SVM outperforms other machine learning models, such as Decision Trees and Naive Bayes, obtaining the best accuracy when detecting named entities from the user-generated tweets.

However, recent advances in Deep Learning techniques have shown that the NER task can benefit from the use of neural architectures, such as biLSTM-networks [6, 7]. We use the implementation[11] proposed in [20] for conducting three different experiments. In the first, we train the model using only the word embeddings as feature. In the second, together with the word embeddings we use the POS and chunk tag. In the third, all the features previously defined are included, in addition to the word embeddings. For every experiment, we use both the pre-trained embeddings and the ones that we created with our *Twitter* corpora. In section 4, results obtained from the several experiments are reported.

3.3 Schedule Matching

The bot-generated tweets present a predefined structure and a formal language, which facilitates the entities detection. In this dataset, our goal is to assign to each track played on the radio, represented by a tweet, a list of entities extracted from the tweet raw text. For achieving that, we experiment with the algorithms and features

---
[11] https://github.com/UKPLab/ emnlp2017-bilstm-cnn-crf

presented previously, obtaining an high level of accuracy, as presented in section 4. The hypothesis considered is that when a radio listener posts a tweet, it is possible that she is referring to a track which has been played a relatively short time before. In this cases, we want to show that knowing the radio schedule can help improving the results when detecting entities.

Once assigned a list of entities to each track, we perform two types of matching. Firstly, within the tracks we identify the ones which have been played in a fixed range of time (*t*) before and after the generation of the user's tweet. Using the resulting tracks, we create a list of candidates entities on which performing string similarity. The score of the matching based on string similarity is computed as the ratio of the number of tokens in common between an entity and the input tweet, and the total number of token of the entity.

In order to exclude trivial matches, tokens within a list of stop words are not considered while performing string matching. The final score is a weighted combination of the string matching score and the time proximity of the track, aimed to enhance matches from tracks played closer to the time when the user is posting the tweet.

The performance of the algorithm depends, apart from the time proximity threshold *t*, also on other two thresholds related to the string matching, one for the *Musical Work* (*w*) and one for the *Contributor* (*c*) entities. It has been necessary for avoiding to include candidate entities matched against the schedule with a low score, often source of false positives or negatives. Consequently, as last step *Contributor* and *Musical Work* candidates entities with respectively a string matching score

lower than *c* and *w*, are filtered out. In Figure 2, an example of *Musical Work* entity recognized in an user-generated tweet using the schedule information is presented.

3.3.1 Candidates Reconciliation

The entities recognized from the schedule matching are joined with the ones obtained directly from the statistical models. In the joined results, the criteria is to give priority to the entities recognized from the machine learning techniques. If they do not return any entities, the entities predicted by the schedule matching are considered. Our strategy is justified by the poorer results obtained by the NER based only on the schedule matching, compared to the other models used in the experiments, to be presented in the next section.

4 Results

The performances of the NER experiments are reported separately for three different parts of the system proposed.

Table 6 presents the comparison of the various methods while performing NER on the bot-generated corpora and the user-generated corpora. Results shown that, in the first case, in the training set the F1 score is always greater than 97%, with a maximum of 99.65%. With both test sets performances decrease, varying between 94-97%. In the case of UGC, comparing the F1 score we can observe how performances significantly decrease. It can be considered a natural consequence of the complex nature of the users' informal language in comparison to the structured message created by the bot.

In Table 7, results of the schedule matching are reported. We can observe how the quality of the linking performed by the algorithm is correlated to the choice of the three thresholds. Indeed, the *Precision* score increase when the time threshold decrease, admitting less candidates as entities during the matching, and when the string similarity thresholds increase, accepting only candidates with an higher degree of similarity. The behaviour of the *Recall* score is inverted.

Finally, we test the impact of using the schedule matching together with a biLSTM-CRF network. In this experiment, we consider the network trained using all the features proposed, and the embeddings not pre-trained. Table 8 reports the results obtained. We can observe how generally the system benefits from the use of the schedule information. Especially in the testing part, where the neural network recognizes with less accuracy, the explicit information contained in the schedule can be exploited for identifying the entities at which users are referring while listening to the radio and posting the tweets.

5 Conclusion and Future Work

We have presented in this work a novel method for detecting musical entities from user-generated content, using a combination of linguistic and domain features with statistical models and extracting contextual information from a radio schedule. We analyzed tweets related to a classical music radio station, integrating its schedule to connect users' messages to tracks broadcasted. We focus on the recognition of two kinds of entities related to the music field, *Contributor* and *Musical Work*.

According to the results obtained, we have seen a pronounced difference between the system performances when dealing with the *Contributor* instead of the *Musical Work* entities. Indeed, the former type of entity has been shown to be more easily detected in comparison to the latter, and we identify several reasons behind this fact. Firstly, *Contributor* entities are less prone to be shorten or modified, but due to their length *Musical Work* entities often represent only a part of the complete title of a musical piece.

Furthermore, *Musical Work* titles are typically composed by more tokens, including common words which can be easily misclassified. The low performances obtained in the case of *Musical Work* entities can be a consequences of these observations. On the other hand, when referring to a *Contributor* users often use only the surname, but in most of the cases it is enough for the system to recognizing the entities.
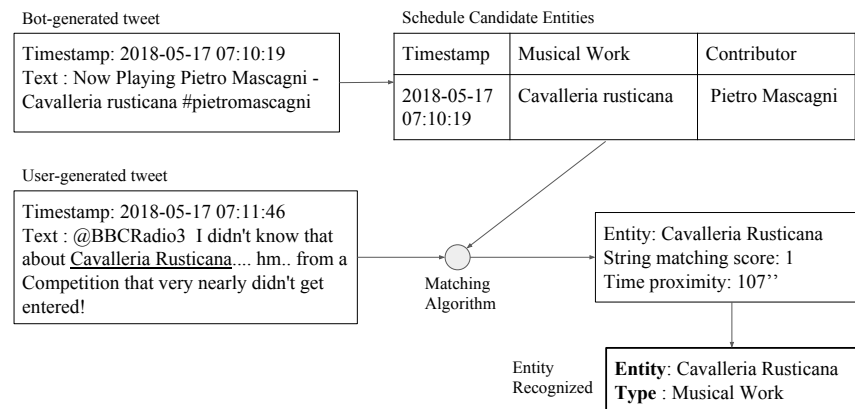
Fig. 2. Example of the workflow for recognizing entities in UGC using the information from the radio schedule

Table 6. F1 score for *Contributor* (C) and *Musical Work* (MW) entities recognized from user-generated tweets (*top*) and top-generated tweets (*bottom*)

Model	Features	GloVe vectors	Training		TestA		TestB	
			C	*MW*	*C*	*MW*	*C*	*MW*
SVM	all	–	95.44	80.80	64.91	**33.48**	61.02	36.21
biLSTM-CRF	–	trained	79.09	51.51	60.00	26.66	67.02	31.48
		pre-trained	85.51	69.28	70.00	33.33	**71.26**	32.08
biLSTM-CRF	POS+chunk	trained	79.37	50.90	61.23	28.98	62.03	**40.00**
		pre-trained	73.51	37.28	**71.62**	25.00	63.74	25.53
biLSTM-CRF	all	trained	97.42	**88.92**	66.22	28.17	69.11	36.36
		pre-trained	**98.46**	87.35	68.79	23.68	70.41	29.51
SVM	all	–	99.12	97.70	97.74	94.32	97.88	95.42
biLSTM-CRF	–	trained	98.95	97.07	98.06	92.99	98.33	95.59
		pre-trained	99.34	94.94	97.88	91.40	98.27	92.35
biLSTM-CRF	POS+chunk	trained	**99.94**	98.28	97.99	**94.68**	98.03	**95.97**
		pre-trained	99.69	97.23	98.12	93.30	98.49	93.61
biLSTM-CRF	all	trained	99.80	98.22	97.70	91.99	98.36	94.48
		pre-trained	99.90	**99.40**	**98.24**	90.46	**98.78**	94.23

From the experiments we have seen that generally the biLSTM-CRF architecture outperforms the SVM model. The benefit of using the whole set of features is evident in the training part, but while testing the inclusion of the features not always leads to better results.

In addition, some of the features designed in our experiments are tailored to the case of classical music, hence they might not be representative if applied to other fields. We do not exclude that our method can be adapted for detecting other kinds of entity, but it might be needed to redefine the features according to the case considered.

Similarly, it has not been found a particular advantage of using the pre-trained embeddings instead of the one trained with our corpora.

Table 7. *Precision* (P), *Recall* (R) and *F1* score for *Contributor* (C) and *Musical Work* (MW) of the schedule matching algorithm. *w* indicates the *Musical Work* string similarity threshold, *c* indicates the *Contributor* string similarity threshold and *t* indicates the time proximity threshold in seconds

			t=800			*t*=1000			*t*=1200		
w	*c*		P	R	F1	P	R	F1	P	R	F1
0.33	0.33	*C*	72.49	16.49	26.87	69.86	17.57	28.08	68.66	**17.93**	**28.43**
		MW	26.42	4.78	8.10	26.05	5.29	8.79	23.66	5.29	8.65
0.33	0.5	*C*	**76.77**	14.32	24.14	74.10	15.64	25.83	73.89	16.00	26.30
		MW	27.1	4.95	8.37	26.67	5.46	**9.06**	24.24	**5.46**	8.91
0.5	0.5	*C*	**76.77**	14.32	24.14	74.71	15.64	25.87	73.89	16.00	26.30
		MW	**30.43**	4.78	8.26	30.30	5.12	8.76	27.52	5.12	8.63

Table 8. *Precision* (P), *Recall* (R) and *F1* score for *Contributor* (C) and *Musical Work* (MW) entities recognized from user-generated tweets using the biLSTM-CRF network together with the schedule matching. The thresholds used for the matching are *t*=1200, *w*=0.5, c=0.5

		Training			TestA			TestB		
		P	R	F1	P	R	F1	P	R	F1
biLSTM-CRF	*C*	**98.22**	96.64	**97.42**	69.01	63.64	66.22	**67.35**	70.97	**69.11**
	MW	**91.54**	86.44	**88.92**	43.48	20.83	28.17	**45.83**	30.14	36.36
biLSTM-CRF +	*C*	95.92	**97.81**	96.86	**74.19**	**71.88**	**73.02**	63.29	**74.63**	68.49
Sch. Matcher	*MW*	87.33	**87.03**	87.18	38.46	**22.73**	**28.57**	42.55	**32.26**	**36.70**

Furthermore, we verified the statistical significance of our experiment by using Wilcoxon Rank-Sum Test, concluding that there have been not significant difference between the various model considered while testing.

The information extracted from the schedule also presents several limitations. In fact, the hypothesis that a tweet is referring to a track broadcasted is not always verified. Even if it is common that radio listeners do comments about tracks played, or give suggestion to the radio host about what they would like to listen, it is also true that they might refer to a *Contributor* or *Musical Work* unrelated to the radio schedule.

Acknowledgements

References

1. **Derczynski, L., Maynard, D., Rizzo, G., Van Erp, M., Gorrell, G., Troncy, R., Petrak, J., & Bontcheva, K. (2015).** Analysis of named entity recognition and linking for tweets. *Information Processing and Management*, Vol. 51, No. 2, pp. 32–49.

2. **Ferragina, P. & Scaiella, U. (2012).** Fast and accurate annotation of short texts with Wikipedia pages. *IEEE Software*, Vol. 29, No. 1, pp. 70–75.

3. **Frank, E., Hall, M. A., & Witten, I. H. (2016).** *The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*. Morgan Kaufmann, Fourth Edition.

4. **Habib, M. B. & Keulen, M. V. (2014).** Information Extraction for Social Media. *Workshop on Semantic Web and Information Extraction*, July, pp. 9–16.

5. **Hauger, D., Schedl, M., Košir, A., & Tkalcic, M. (2013).** The million musical tweets dataset: What can we learn from microblogs. *Proceedings of the 14th International Society for Music Information Retrieval Conference*, pp. 189–194.

6. **Lample, G., Ballesteros, M., Subramanian, S., Kawakami, K., & Dyer, C. (2016).** Neural architectures for named entity recognition. *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, July, pp. 260–270.

7. **Lin, B. Y., Xu, F. F., Luo, Z., & Zhu, K. Q. (2017).** Multi-channel bilstm-crf model for emerging named entity recognition in social media. *Proceedings of the 3rd Workshop on Noisy User-generated Text*, pp. 160–165.

8. **Mendes, P. N., Jakob, M., García-Silva, A., & Bizer, C. (2011).** DBpedia Spotlight: Shedding light on the Web of documents. *Proceedings of the 7th International Conference on Semantic Systems (I-Semantics)*, Vol. 95, pp. 1–8.

9. **Moro, A., Raganato, A., & Navigli, R. (2014).** Entity linking meets word sense disambiguation: a unified approach. *Transactions of the Association for Computational Linguistics (TACL)*, Vol. 2, No. 0, pp. 231–244.

10. **Müller, M. (2015).** *Fundamentals of Music Processing: Audio, Analysis, Algorithms, Applications*. Springer Publishing Company, Incorporated, 1st edition.

11. **Nadeau, D. (2007).** A survey of named entity recognition and classification. *Lingvisticae Investigationes. International Journal of Linguistics and Language Resources*, Vol. 30, No. 1, pp. 3–26.

12. **Oramas, S., Espinosa-Anke, L., Lawlor, A., Serra, X., Saggion, H., Music Technology Group, & Universitat Pompeu Fabra (2016).** Exploring customer reviews for music genre classification and evolutionary studies. *Proc. 17th International Society for Music Information Retrieval Conference*, pp. 150–156.

13. **Oramas, S., Espinosa-Anke, L., Sordo, M., Saggion, H., & Serra, X. (2016).** ELMD: An automatically generated entity linking gold standard dataset in the music domain. *Language Resources and Evaluation Conference - LREC*, pp. 3312–3317.

14. **Oramas, S., Ferraro, A., Correya, A., & Serra, X. (2017).** Mel: a music entity linking system. *Proceedings of the 18th International Society for Music Information Retrieval Conference*, Suzhou, China.

15. **Oramas, S., Ostuni, V. C., Di Noia, T., Serra, X., & Di Sciascio, E. (2015).** Sound and music recommendation with knowledge graphs. *ACM Transactions on Intelligent Systems and Technology*, Vol. 9, No. 4.

16. **Pennington, J., Socher, R., & Manning, C. (2014).** Glove: Global vectors for word representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 1532–1543.

17. **Platt, J. (1998).** Fast training of support vector machines using sequential minimal optimization. *Advances in Kernel Methods - Support Vector Learning*, MIT Press, pp. 41–65.

18. **Porcaro, L. (2018).** Information extraction from user-generated content in the classical music domain. *Master Thesis. Universitat Pompeu Fabra*.

19. **Ramshaw, L. & Marcus, M. (1995).** Text chunking using transformation-based learning. *Third Workshop on Very Large Corpora*, pp. 82–94.

20. **Reimers, N. & Gurevych, I. (2017).** Reporting score distributions makes a difference: Performance study of LSTM-networks for sequence tagging. *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pp. 338–348.

21. **Ritter, A., Clark, S., & Etzioni, O. (2011).** Named entity recognition in tweets: an experimental study. *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, pp. 1524–1534.

22. **Schedl, M. & Hauger, D. (2012).** Mining microblogs to infer music artist similarity and cultural listening patterns. *International Conference Companion on World Wide Web (WWW)*, pp. 877.

23. **Tata, S. & Di Eugenio, B. (2010).** Generating fine-grained reviews of songs from album reviews. *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*, July, pp. 1376–1385.

24. **Zangerle, E., Gassler, W., & Specht, G. (2012).** Exploiting twitter's collective knowledge for music recommendation. *2nd workshop on Making Sense of Microposts #MSM2012*, February, pp. 14–17.

25. **Zangerle, E., Pichl, M., Gassler, W., & Specht, G. (2014).** #nowplaying music dataset: Extracting listening behavior from twitter. *Proceedings of*

the First International Workshop on Internet-Scale Multimedia Management - WISMM '14*, pp. 21–26.

26. **Zhang, X., Liu, Z., Qiu, H., & Fu, Y. (2009).** A hybrid approach for chinese named entity recognition in music domain. *8th IEEE International Symposium on Dependable, Autonomic and Secure Computing, DASC 2009*, pp. 677–681.