ASSOCIATION MINING OF FOLK MUSIC GENRES AND TOPONYMS

Kerstin Neubarth[1,2] **Izaro Goienetxea**[3] **Colin G. Johnson**[2] **Darrell Conklin**[3,4]

[1]Canterbury Christ Church University, Canterbury, United Kingdom
[2]School of Computing, University of Kent, Canterbury, United Kingdom
[3]Department of Computer Science and Artificial Intelligence,
University of the Basque Country UPV/EHU, San Sebastián, Spain
[4]IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

ABSTRACT

This paper demonstrates how association rule mining can be applied to discover relations between two ontologies of folk music: a genre and a region ontology. Genre–region associations have been widely studied in folk music research but have been neglected in music information retrieval. We present a method of association rule mining with constraints consisting of rule templates and rule evaluation measures to identify different, musicologically motivated, categories of genre–region associations. The method is applied to a corpus of 1902 Basque folk tunes, and several interesting rules and rule sets are discovered.

1. INTRODUCTION

In recent years music information retrieval (MIR) research has increasingly turned towards folk and ethnic music and its contexts [6, 19], and perspectives for collaboration between MIR and ethnomusicology have been outlined [22, 23]. While musicologists consider interactions between folk music genres, their geographical distribution and musical characteristics [12, 18], MIR research has mainly focused on geographically organised folk music corpora [10, 11, 21]. A recent study on Cretan folk music extracts distinctive melodic interval patterns for genre and for area classes, but does not link genres and areas, although the idea of conjunctive genre–area classes is mentioned [5]. In this paper we analyse relations between genres and regions in a collection of Basque folk music, through association rule mining.

Association rule mining has been developed extensively in the wider field of knowledge discovery and data mining, but has seen only limited attention in MIR (e.g. [4]). Work on cultural heritage, not restricted to but also covering music, has used association rule mining to populate a heritage ontology with new relations between concepts: annotations of heritage objects were mined and discovered associations proposed to a domain expert, who categorised them as subclass or associative relations [13]. Our research goes beyond these studies by applying association rule mining to suggest different, labelled, categories of associations.

2. TASK DESCRIPTION

The folk music collection Cancionero Vasco was originally compiled by the musicologist Padre Donostia for a musical heritage competition in 1912 [18, article 'Basque Music']. It has been digitised and curated by Fundación Euskomedia in collaboration with musicologists at Fundación Eresbil, under the auspices of the Basque Studies Society. [1] The digitised collection contains 1902 Basque folk songs and dances. The examples are annotated with genre information and the location (toponym) where they were collected. The annotation vocabulary is defined in two ontologies: a geographical ontology of provinces, municipalities and towns or villages, and an ontology of hierarchically organised genres. The aim of applying association rule mining to this collection is to discover patterns of genres and regions co-occurring in the annotations, which suggest that certain genres are particularly associated with certain regions and vice versa.

A common challenge in applications of association rule mining is to manage potentially large numbers of discovered association rules and identify those rules which are interesting to a user [14]. Genre–region relations feature prominently in folk music research, and the musicological interest goes beyond simple mappings: In a qualitative analysis we examined surveys of traditional music in 25 European countries [18], extracted statements linking genres and regions, grouped them according to similar meaning and for each group suggest a shared interpretation to facilitate translation into association rules (Table 1). As these association categories are based on recurring observations in musicological reference articles, they are considered relevant to users of folk music collections. The data mining task then is not only to discover associations between the two ontologies of genres and regions, but to distinguish associations of the categories listed in Table 1.

3. DATA AND METHODS

This section describes the data representation and the data mining method for discovering genre–region associations in the Cancionero Vasco. In order to identify association

[1] http://www.euskomedia.org/cancionero

Category	Example observation	Interpretation
Present	"survivals of calendar ritual and wedding music are found in the Opole area"	genre present in region
Absent	"the [two-part] form does not exist at all in Kosovo and Metohija"	genre not present in region
Local	"localized dances include the corridinho (Algarve)"	genre present in exactly one region
Mainly	"krakowiak dances [...] are found mainly in Małopolska"	genre over-represented in region with respect to other regions
Dominant	"the polka is the most popular dance in these regions"	genre over-represented in region with respect to other genres
Typical	"the dance-song seguidillas is typical of New Castile"	genre over-represented in region with respect to other regions and genres
Hardly	"the klarino style is hardly found at all on the islands"	genre under-represented in region with respect to other regions
Rare	"Genres [...] such as the epic are quite rare in central European repertories, whereas genres [...] such as the ballad are quite common."	genre under-represented in region with respect to other genres

Table 1. Categories of genre–region associations, with example quotations from the New Grove [18].

rules of the different categories summarised in Table 1, these categories are translated into association constraints consisting of rule templates and rule evaluation measures.

3.1 Ontologies

The annotations mined in this study are standardised and structured in two ontologies: a genre and a geographical ontology, formalised in description logic (DL) [3, 9].

The ontology of folk music genres consists of two parts: a set of statements $G \sqsubseteq G'$ defines the genres and their subsumption relations, e.g.

$$\textsf{work songs} \sqsubseteq \textsf{life-cycle songs}$$

$$\textsf{life-cycle songs} \sqsubseteq \textsf{genre}$$

state that work songs are subsumed by, i.e. more specific than, life-cycle songs and that life-cycle songs are a genre. The second part of the ontology is a set of assertions $G(e)$ where G is a genre concept and e a folk tune example: this part formalises the genre annotations of the examples in the Cancionero Vasco, using 31 genres. Examples are asserted with their most specific known genre annotation, which is not necessarily the lowest level of the subsumption hierarchy. The genre counts used in the association rule mining (Section 3.2) can be derived from the ontology by querying for the examples instantiating a genre; here the inference capabilities of DL allow to count examples not only for the directly asserted genre but also for more general, subsuming genres. Out of the 1902 examples in the corpus 341 examples are without a genre annotation.

The geographical ontology covers Euskal Herria, the Basque speaking areas in North-East Spain and South-West France. The ontology is organised into the three levels of provinces (7 toponyms), municipalities (681 toponyms) and towns or villages (2280 toponyms). Locations are formalised as instances in DL, assigned to levels, e.g.

$$\textsf{province(Lapurdi)} \quad \text{and} \quad \textsf{municipality(Azkaine)}.$$

In DL the geographical relationships are defined in terms of spatial containment roles, e.g.

$$\textsf{contains(Lapurdi, Azkaine)}$$

asserts that Lapurdi (a province) contains Azkaine (a municipality). Each folk tune example e is asserted with the region R in which it was collected: $\textsf{collected}(R, e)$. As with the genre ontology, assertions can be made at any level of the hierarchy and higher-level counts are inferred based on the transitivity of the containment relation. Out of the 1902 folk tunes in the Cancionero Vasco 272 tunes are without a toponym annotation.

The original annotation terms for genres are Spanish or Basque. For the presentation in this paper we give English translations for genres. As toponyms we use the Basque, rather than Spanish or French, names.

3.2 Association Rule Mining

Association rules are rules of the form $a \rightarrow b$ with an antecedent item set a and a consequent item set b ($a \cap b = \emptyset$) [20]. A rule $a \rightarrow b$ with confidence c states that $c\%$ of the data records containing items a also contain items b. Rule templates [14] define the form of a rule and specify which items can occur in the antecedent and consequent. In this study we mine for rules with one item in the antecedent and one item in the consequent. Here an item can be a genre (denoted G in the rule templates), a region (denoted R) or the complement of a genre or region (denoted $\overline{G}$ and $\overline{R}$ respectively).

	a	$\overline{a}$	
b	n_{ab}	$n_b - n_{ab}$	n_b
$\overline{b}$	$n_{a\overline{b}} = n_a - n_{ab}$	$n_{\overline{b}} - n_{a\overline{b}}$	$n_{\overline{b}} = n - n_b$
	n_a	$n_{\overline{a}} = n - n_a$	n

Figure 1. Contingency table for a rule $a \rightarrow b$.

The rule templates determine how the genre and region of a candidate association are mapped onto a contingency table from which a rule evaluation measure can be calculated [15]. Figure 1 presents a 2×2 contingency table for an association rule $a \rightarrow b$, where a is the antecedent item, b the consequent item, n_a the number of folk tunes annotated with a, n_b the number of folk tunes annotated with b, n_{ab} the number of folk tunes annotated with both a and b, and n the total number of folk tunes in the corpus. The notation $\overline{a}$ denotes the complement of item a. For example, with the rule template $G \rightarrow R$, n_a refers to the genre count and n_b refers to the region count; n_{ab} is the number of tunes instantiating both the genre and the region.

The evaluation measures most commonly used in association rule mining are support (frequency of the co-occurrence) and confidence (conditional probability of the co-occurrence given the antecedent). These two measures, however, are not sufficient to distinguish all association categories of Table 1, e.g. *Typical* against *Mainly* and *Dominant*. We thus considered further existing measures and their properties (e.g. [8, 16]). Measures for each category were selected in two steps. First, we defined the requirements for each category, given its interpretation (Table 1), based on established measure properties:

- Asymmetric vs symmetric measures: For all categories except *Present* and *Typical* the measure should be asymmetric, i.e. distinguish between $a \rightarrow b$ and $b \rightarrow a$.
- Increasing function with number of examples: The measures for *Mainly*, *Dominant*, *Hardly* and *Rare* should increase with n_{ab} for fixed n_a, while the measure for *Typical* should increase with n_{ab} for both n_a and n_b fixed.
- Decreasing function with number of counter-examples: The measures for *Mainly*, *Dominant*, *Hardly* and *Rare* should decrease with increasing $n_{a\overline{b}}$, and thus n_a, for fixed n_{ab}, while the measure for *Typical* should decrease with both increasing $n_{a\overline{b}}$ and $n_{\overline{a}b}$, and thus both n_a and n_b, for fixed n_{ab}.
- Sensitivity vs. insensitivity to sample size: As the measures are used to capture the relationship between a and b, they should be insensitive to changes in $n_{\overline{ab}}$ and thus to changes in n for fixed n_{ab}, n_a and n_b.

Second, we determined measures that match the requirements. Where more than one measure meets the category criteria, a measure is preferred in which variations of the measure value can easily be related to values in the contingency table [16]. Table 2 lists the resulting constraints for each association category.

Category	Template	Measure
Present	$G - R$	support
Absent	$G \rightarrow \overline{R}$	confidence ($c = 1$)
Local	$G \rightarrow R$	confidence ($c = 1$)
Mainly	$G \rightarrow R$	confidence
Dominant	$R \rightarrow G$	confidence
Typical	$G - R$	Jaccard
Hardly	$G \rightarrow \overline{R}$	Sebag-Schoenauer
Rare	$R \rightarrow \overline{G}$	Sebag-Schoenauer

Table 2. Constraints for the association categories.

Measure	Definition
support	$s = n_{ab}$
confidence	$c = n_{ab}/n_a$
Jaccard	$J = n_{ab}/(n_a + n_b - n_{ab})$
Sebag-Schoenauer	$S = n_{ab}/n_{a\overline{b}}$

Table 3. Definitions of the rule evaluation measures.

The definitions of the measures are given in Table 3. To ensure invariance with changes in n, absolute rather than relative support is applied for *Present*. For the categories describing under-representation (*Hardly* and *Rare*), the measure of Sebag-Schoenauer was found to discriminate better than confidence: confidence accepts most infrequent pairs, while Sebag-Schoenauer accepts pairs that are less frequent than comparison pairs.

During the mining, each candidate pair of a genre and a region is evaluated against the category constraints, i.e. the genre, region and pair counts are mapped onto n_a, n_b and n_{ab} according to the template, and the measure value is calculated. Pairs are tested for all categories, and associations can be assigned to more than one category (e.g. *Present* and *Mainly*).

4. RESULTS

Table 4 lists selected highly ranked rules for all categories. The p-values in Table 4 are calculated according to Fisher's exact test with left tail for *Absent*, *Hardly* and *Rare* and right tail for all other categories [7]. The p-value measures the probability of finding at most (left tail) or at least (right tail) the number of co-occurrences given by the pair count, under the conditions of the genre and region counts. To account for multiple comparisons, i.e. testing multiple hypotheses on the same data, results are marked for both significance level α and the Bonferroni-corrected significance level β; it should be noted, though, that the Bonferroni correction is highly conservative. Rules above the significance level are not necessarily rejected as uninteresting (see also [16]), rather the p-values provide additional information for interpreting discovered rules, with respect to the distribution of genres and regions in the total corpus.

Genre (count)	Region (count)	Pair count	Template	Measure	p-value
Present					
life-cycle songs (477)	Nafarroa (897)	259	$G - R$	$s = 259$	0.00089^{++}
Artaxuriketak (38)	Nafarroa (897)	30	$G - R$	$s = 30$	4.3e-05**
Absent					
dances (495)	Hazparne (23)	0	$G \rightarrow \overline{R}$	$c = 1$	0.00093^{++}
sacred songs (301)	Araba (27)	0	$G \rightarrow \overline{R}$	$c = 1$	0.00922*
Local					
smugglers' songs (1)	Lapurdi (383)	1	$G \rightarrow R$	$c = 1$	0.02321^{+}
smugglers' songs (1)	Azkaine (61)	1	$G \rightarrow R$	$c = 1$	0.03207^{+}
Mainly					
Artaxuriketak (38)	Nafarroa (897)	30	$G \rightarrow R$	$c = 0.79$	4.3e-05**
moral songs (11)	Nafarroa (897)	8	$G \rightarrow R$	$c = 0.73$	0.04470^{+}
Dominant					
dances (495)	Araba (27)	24	$R \rightarrow G$	$c = 0.89$	7.8e-12**
dances (495)	Eugi (15)	13	$R \rightarrow G$	$c = 0.87$	1.4e-06**
Typical					
Carlism songs (1)	Biriatu (3)	1	$G - R$	$J = 0.33$	0.00158*
Hardly					
dances (495)	Atharratze (23)	1	$G \rightarrow \overline{R}$	$S = 494$	0.00857
life-cycle songs (477)	Bizkaia (21)	2	$G \rightarrow \overline{R}$	$S = 237.5$	0.07232
Rare					
Artaxuriketak (38)	Lapurdi (383)	1	$R \rightarrow \overline{G}$	$S = 382$	0.01112^{+}
moral songs (11)	Nafarroa (897)	8	$R \rightarrow \overline{G}$	$S = 111.13$	0.98935

* significance level $\alpha = 0.01$, ** with Bonferroni correction $\beta = 0.0002$
$^{+}$ significance level $\alpha = 0.05$, $^{++}$ with Bonferroni correction $\beta = 0.001$

Table 4. Examples of discovered association rules.

In all cases of *Local*, genres are represented by only one example in the corpus and thus are necessarily identified as local. The occurrence of the smugglers' song in Lapurdi could be linked to the site of Lapurdi, stretching from the coast inland across the Pyrenees between Spain and France: in fact, the example in the Cancionero Vasco was collected more specifically in the municipality of Azkaine, in the Larrun area known to have been used by smugglers; in nearby Sara the "smugglers' race", celebrated in August, has become part of 20th-century folklore [17].

5. DISCUSSION

The interest of this study lies in discovering genre–region associations in the Cancionero Vasco which are potentially interesting to users who browse or analyse the collection. The association categories defined in this paper provide additional semantics for rules as compared to traditional association rule mining, and can help organise and understand the folk music corpus. Multiple labels can further specify an association (Example 1) and even capture different aspects of the same genre–region pair (Example 2):

Example 1: Not only are the corn-harvesting songs Artaxuriketak present in Nafarroa, the Artaxuriketak examples in the Cancionero Vasco are mainly from Nafarroa.

Example 2: Within the Cancionero Vasco moral songs are rare in Nafarroa, i.e. under-represented with respect to other genres in Nafarroa (8 out of 897 instances), but are over-represented in Nafarroa with respect to their occurrence in other regions, i.e. they occur mainly in Nafarroa (8 out of 11 instances).

The p-value is a symmetric measure, i.e. it does not distinguish between e.g. rule $G \rightarrow R$ and rule $R \rightarrow G$. Using different rule templates with confidence or Sebag-Schoenauer as evaluation measure, on the other hand, al-

Region (count)	Pair count	Category	Measure	p-value
Araba (27)	0	*Absent*	$c = 1$	0.57771
Bizkaia (21)	0	*Absent*	$c = 1$	0.65307
Gipuzkoa (175)	4	*Present*	$s = 4$	0.46864
Lapurdi (383)	1	*Present*	$s = 1$	0.9964
		Rare	$S = 382$	0.01112
Nafarroa-Beherea (47)	1	*Present*	$s = 1$	0.61722
Nafarroa (897)	30	*Present*	$s = 30$	4.3e-05
		Mainly	$c = 0.79$	4.3e-05
Zuberoa (80)	2	*Present*	$s = 2$	0.48504

Table 5. Group of rules for the genre Artaxuriketak (genre count 38).

lows one to identify the asymmetric cases of over-representation (*Mainly* vs. *Dominant*) or under-representation (*Hardly* vs. *Rare*), which correspond to different observations in folk music surveys.

Example 3: Artaxuriketak occur mainly in Nafarroa as compared to other regions (79% of the Artaxuriketak instances in the Cancionero Vasco are from Nafarroa). On the other hand, dances are dominant in Araba with respect to other genres in the same region (89% of examples from Araba are dances).

The association categories can facilitate the analysis of groups of discovered rules. Folk music collections are often organised according to genres or regions [12], and folk music surveys may review genres against an underlying geographical classification [18]. For example, a summary of folk-dance in Finland states: "Although most polska melodies have been collected in Pohjanmaa, the dance was known throughout the country except in the far north and Karelia." [18]

Example 4: The geographical distribution of Artaxuriketak at the level of provinces can be described in a similar way (Table 5): While most Artaxuriketak in the corpus were collected in Nafarroa, the genre is also known in the other provinces except – within the Cancionero Vasco – in Araba and Bizkaia. Given the large number of instances for Nafarroa (897 instances, 47% of the corpus) it is not surprising that most Artaxuriketak were collected in Nafarroa, but the high proportion (nearly 80% of Artaxuriketak) is statistically significant. It is interesting to note that traditionally farmers in Nafarroa have cultivated corn (maize) while the chief crops in Araba have been cereals as well as fruit, wine and olives [2]. For Lapurdi, surveys in the 1980s reported that nearly half of the area was dedicated to wine (45%), followed by other crops (40%), woods (10%) and urban or uncultivated areas (5%) [1].

Other folk music surveys follow a mainly geographical organisation. For example, the New Grove article on traditional music in Croatia is structured according to regions; the following statement is taken from the section on Western Croatia: "This region is characterised by the tanac and balun dances. [...] Other dances, like the polka and the valcer, are also performed." [18]

Example 5: Folk music in the province of Araba, according to the Cancionero Vasco, is dominated by dances (89% of the Araba examples are dances, p-value $= 7.8e$-12). Related rules, not shown in Table 4, indicate that of the other genres only life-cycle songs were also collected in Araba (pair count 3, *Hardly*, $S = 158$, p-value $= 0.06411$), and more particularly within life-cycle songs the subgroup of work songs (*Present*, $s = 3$, p-value $= 0.02580$).

6. CONCLUSIONS

In this paper we have shown how association rule mining can be used to discover relevant relations between genres and geographical locations of folk tunes, more specifically how association rule mining with rule templates and evaluation measures can be used to identify different, musicologically motivated, categories of genre–region associations. The method was applied to a collection of Basque folk music, and example associations discussed in the context of folk music research.

This research represents an original contribution to MIR both in terms of retrieval task (discovering associations between folk music genres and their geographical distribution) and method (systematically combining rule templates and evaluation measures to distinguish different association categories). As a case study of interdisciplinary collaboration between MIR and musicology it demonstrates how musicology can inform the task definition, method design and discussion of results; the examples illustrate how MIR can both support musicological observations and stimulate further analysis.

Our work can be extended in several ways. The mining results can support more traditional information retrieval of music, like browsing and searching of music collections guided by association categories or by groups of association rules. Classification using association rules could be explored to suggest genre and toponym annotations for unlabelled tunes. The method could also be adapted to search for associations between annotations and music content classes. In addition the approach could be applied to data in other MIR areas such as user tagging, folksonomies and music recommendation.

7. ACKNOWLEDGEMENTS

We thank Fundación Euskomedia and Fundación Eresbil for providing the Cancionero Vasco for study. This research was partly supported by a grant *Análisis Computacional de la Música Folclórica Vasca* (2011-2012) from the Diputación Foral de Gipuzkoa, Spain.

8. REFERENCES

[1] Enciclopedia Auñamendi. Fondo Bernardo Estorés Lasa. http://www.euskomedia.org/aunamendi.

[2] Encyclopaedia Britannica Online. academic edition, 2012. http://original.search.eb.com/.

[3] F. Baader, D. Calvanese, D. McGuiness, D. Nardi, and P. F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, Implementation and Applications*. Cambridge University Press, Cambridge, 2003.

[4] D. Conklin. Melodic analysis with segment classes. *Machine Learning*, 65(2-3):349–360, 2006.

[5] D. Conklin and C. Anagnostopoulou. Comparative pattern analysis of Cretan folk songs. *Journal of New Music Research*, 40(2):119–125, 2011.

[6] O. Cornelis, M. Lesaffre, D. Moelants, and M. Leman. Access to ethnic music: Advances and perspectives in content-based music information retrieval. *Signal Processing*, 90:1008–1031, 2010.

[7] S. Falcon and R. Gentleman. Hypergeometric testing used for gene set enrichment analysis. In F. Hahne, W. Huber, R. Gentleman, and S. Falcon, editors, *Bioconductor Case Studies*, pages 207–220. Springer, 2008.

[8] L. Geng and H. J. Hamilton. Interestingness measures for data mining: a survey. *ACM Computing Surveys*, 38(3):1–32, 2006.

[9] I. Goienetxea, J. Arrieta, I. Bagüés, A. Cuesta, P. Leiñena, and D. Conklin. Ontologies for representation of folk song metadata. Technical Report EHU-KZAA-TR-2012-01, Department of Computer Science and Artificial Intelligence, University of the Basque Country UPV/EHU, 2012. http://hdl.handle.net/10810/8053.

[10] R. Hillewaere, B. Manderick, and D. Conklin. Global feature versus event models for folk song classification. In *Proceedings of the 10th International Society for Music Information Retrieval Conference (ISMIR 2009)*, pages 729–733, Kobe, Japan, 2009.

[11] Z. Juhász. A systematic comparison of different European folk music traditions using self-organizing maps. *Journal of New Music Research*, 35(2):95–112, 2006.

[12] I. J. Katz. The traditional folk music of Spain: explorations and perspectives. *Yearbook of the International Folk Music Council*, 6:64–85, 1974.

[13] T. Kauppinen, H. Kuittinen, K. Seppäla, J. Tuominen, and E. Hyvönen. Extending an ontology by analyzing annotation co-occurences in a semantic cultural heritage portal. In *Proceedings of the ASWC 2008 Workshop on Collective Intelligence, 3rd Asian Semantic Web Conference (ASCW 2008)*, pages 1–6, Bangkok, Thailand, 2008.

[14] M. Klemettinen, H. Mannila, P. Ronkainen, H. Toivonen, and A. I. Verkamo. Finding interesting rules from large sets of discovered association rules. In *Proceedings of the 3rd International Conference on Information and Knowledge Management*, pages 401–407, Gaithersburg, Maryland, 1994.

[15] N. Lavrač, P. Flach, and B. Zupan. Rule evaluation measures: a unifying view. In *Proceedings of the 9th International Workshop on Inductive Logic Programming (ILP-99)*, pages 174–185, Bled, Slovenia, 1999.

[16] P. Lenca, P. Meyer, B. Vaillant, and S. Lallich. On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid. *European Journal for Operational Research*, 184(2):610–626, 2008.

[17] J. A. Perales Díaz. Fronteras y contrabando en el Pirineo Occidental. *Zainak. Cuadernos de Antropología-Etnografía*, 17:127–136, 1998.

[18] S. Sadie, editor. *New Grove Dictionary of Music and Musicians*. Macmillan, London, 2001.

[19] X. Serra. A multicultural approach in music information research. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR 2011)*, pages 151–156, Miami, Florida, USA, 2011.

[20] R. Srikant, Q. Vu, and R. Agrawal. Mining association rules with item constraints. In *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining (KDD-97)*, pages 67–73, Newport Beach, CA, USA, 1997.

[21] J. Taminau, R. Hillewaere, S. Meganck, D. Conklin, A. Nowé, and B. Manderick. Descriptive subgroup mining of folk music. In *2nd International Workshop on Machine Learning and Music (MML 2009)*, Bled, Slovenia, 2009.

[22] G. Tzanetakis, A. Kapur, W. Schloss, and M. Wright. Computational ethnomusicology. *Journal of Interdisciplinary Music Studies*, 1(2):1–24, 2007.

[23] P. van Kranenburg, J. Garbers, A. Volk, F. Wiering, L. Grijp, and R. C. Veltkamp. Collaborative perspectives for folk song research and music information retrieval: The indispensable role of computational musicology. *Journal of Interdisciplinary Music Studies*, 4(1):17–43, 2010.