

Digital Musicology and MIR: Papers, Projects and Challenges

Frans Wiering

Department of Information and Computing
Sciences, Utrecht University
f.wiering@uu.nl

Emmanouil Benetos

Music Informatics Research Group, Department
of Computer Science, City University London
emmanouil.benetos.1@city.ac.uk

ABSTRACT

In this paper we report on the ISMIR 2013 Demo and Late Breaking Session entitled Digital Musicology and MIR. Five papers were discussed as examples of interesting MIR contributions to musicology. Two important projects, *Transforming Musicology* and *CompMusic*, were briefly presented. Finally, this paper reports the first results of a questionnaire about challenges from Digital Musicology for MIR research. The most important outcomes are that lack of suitable musical data is still an important obstacle and that there is a great demand for tools and methods that make integrated access and analysis of symbolic and audio data possible.

1. INTRODUCTION

Each year many exciting new technologies are presented at ISMIR. Quite a few of these have a potential to generate new insights in music, whether from a psychological, cultural or musicological viewpoint: in other words, to contribute to the growing field of Digital Musicology. In this Demo and Late Breaking Session around 25 participants discussed a number of papers that the audience found particularly relevant to musicological research (section 2). After this, a few projects were presented (section 3). It was also an aim of this session to discuss what the current technical needs of the various music research communities might be, and how these needs could be translated into MIR research challenges. Although a few challenges were mentioned in passing, there was no time to discuss these in sufficient detail. Therefore we created a brief online questionnaire about these questions. In section 4, we present the first outcomes.

This session was co-organised by the International Musicological Society's Study Group on Digital Musicology (<http://www.projects.science.uu.nl/music/StudyGroup/>), chaired by Frans Wiering.

2. PAPERS

In the first part of the session we discussed papers that might be shortlisted for the Award for Best Digital Mu-

sicology Paper at ISMIR 2013, had such an award existed. The following papers were nominated:

- Anja Volk and W. Bas de Haas, 'A Corpus-Based Study on Ragtime Syncopation' [6], nominated by Andre Holzapfel. Anja explained that she plans to expand the corpus and explore it more deeply. Andre proposed an idea to explore syncopation in Turkish makam music.
- Vignesh Ishwar *et al.*, 'Motif Spotting in an Alapana in Carnatic Music' [3], nominated by Tim Crawford. Frans Wiering remarked that the paper took the melodic concept of a raga very seriously: Matthias Mauch agreed. Xavier Serra said that a collection of vocal Carnatic music will be made available online (along with pitch annotations) through the CompMusic project.
- Tom Collins *et al.*, 'SIARCT-CFP: Improving Precision and the Discovery of Inexact Musical Patterns in Point-Set Representations' [1] was nominated by Laurent Pugin, who considers the method important for MIR and digital musicology. Tom Collins wants to apply it to different types of music, and appreciates being sent examples that do not work with the current approach. The system cannot yet be used by musicologists, and making it usable is an open challenge. Frans stated that you need a computer scientist to sit next to a musicologist in order to make a system usable. Jordan Smith asked if the US Society for Music Theory could propose grand challenges for MIR. Frans Wiering confirms that these would be very welcome.
- Nanzhu Jiang and Meinard Müller, 'Automated Methods for Analyzing Music Recordings in Sonata Form' [4]. Frans Wiering, who proposed it, is not sure if the method is already usable for musicological research, but it seems to have great potential. Laurent Pugin agrees and says that it might also be adapted in the future to highlight deviations from the classic sonata form. Frans says that such deviations from the norm are important in the context of digital humanities in general.
- Nicolas Gonzalez Thomas *et al.*, 'A Methodology for the Comparison of Melodic Generation Models Using Meta-Melo' [2] was proposed by Maarten Grachten, as a useful contribution for musicologists. Tim Crawford was a bit puzzled by it; Tillman Weyde commented on the usefulness of the method for

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

© 2013 International Society for Music Information Retrieval

automating some of the processes in evaluation of generative music, and was interested in what we could learn about the process. Maarten Grachten was concerned if the problem could be left to solely to machines.

In conclusion, Mathias Mauch feels like we are talking about musicology as if it is separate from us. He suggests giving training to musicologists to use command line tools, so that they have a better understanding of how powerful they are. Frans Wiering is in two minds about training as a necessary precondition for tool use and advocates the design of usable, intuitive tools. Finally, Frans said that the large number of relevant papers shows a good impact from digital musicology in the MIR community this year.

3. PROJECTS

Tim Crawford speaks about exciting new project called *Transforming Musicology* (UK AHRC Large Grant). The project includes the following activities:

1. Bring together three communities of interest, to further the aims of musicology and provide a new focus in serving musicology, and also to serve the much larger community of people who engage in music. Thousands of music experts exist in the world outside academia and their expertise is highly valuable.
2. Ambitious, heuristic research. The research is designed to make a valuable contribution to musicology, even when some things might not work. We need to accept failure! Musicologists typically do not report failure, and the project will try to change that.
3. Scope: 16th century lute and vocal music; 19th century music specialising on Wagner leitmotifs; and music and social media, specifically on how people share music and what they do with it. In addition, a series of mini-projects is envisaged that will put together musicologists with technologists who will help them to explore possibilities on particular datasets, using musicological motivations as driving force.

This project can lead to larger proposals that will drive computational musicology, and to a changed attitude to technology in the next generation of scholars. Lastly, the project underlines the importance of some kind of semantic framework, where we do not only record our results, but we record the data that we are using and the provenance of it, record the processes that were applied to that, as well as the parameter settings and the results (in a broader sense). The project aims to capture all these in a framework, which is very ambitious. The aim is thereby to record the entire scholarly process, so that our work can be reusable by others, and also testable. Tim commented that ‘This sounds suspiciously like science, but is that so frightening?’

Frans Wiering is international co-investigator in *Transforming Musicology*. He says that this project can lead to

a whole international network for computational musicology, and asks Tim about potential collaborations. Tim talks about high-profile events, e.g. at the American Musicological Society, for getting maximum impact. He is however more interested in what happens along the way in the project. The above-mentioned mini-projects will also be very helpful; a call for these was issued in December 2013 (see <http://www.transforming-musicology.org/>). If this interest in participation is reported back to AHRC, this will create traction. Several projects reported at ISMIR 2013 might be suitable as mini-projects.

Matthias Mauch comments that the project description might be a bit opaque. Matthias has recently his first experience on a humanities workshop, where there seems to be no notion on if something is true or not. Matthias asks if Tim intends to transform musicology as being more scientific. Tim responds that he indeed wants to change it, but not necessarily towards being more scientific. He explains that the notion of proof in humanistic research is completely different from scientific proof. For example, an elegant argument might be more convincing to a humanist than any amount of statistics. There is no objective truth, although there are many interesting ideas and influences. But the humanities world is changing and is faced with severe challenges, e.g. Big Data. You have to perform some kind of quantitative analysis on Big Data; otherwise the whole concept loses its point. At least in the UK, funding bodies are steering things towards verifiable arguments.

Next Xavier Serra talked about CompMusic (<http://compmusic.upf.edu/>). The project is in the middle of its duration (5 years). The main effort has been defining scope and compiling the corpus from 5 music traditions, focusing on melodic and rhythmic analysis. The data will be published in a web repository, and will be useful for many other things. For example, for Turkish makam analysis, the score corpus can be very useful. From an MIR perspective, these datasets are more complex than a collection of audio recordings, since they contain also contextual info. Andre Holzapfel notes that much research from CompMusic was already presented at ISMIR, especially in Carnatic music. Andre managed to obtain funding for a Marie Curie fellowship, showing that the CompMusic project made funding for ethnomusicological research more viable, at least in the EU. Tim says that big projects can make such funding priorities possible. Xavier finally says that big projects now encourage open data and source software creation.

Finally, Frans Wiering mentions that he was involved in writing a report for the European Science Foundation (ESF) entitled *Musicology (Re-)Mapped* [5]. Among other things, this report points out the cultural importance of music research in general, the great potential for interdisciplinary music research and the importance of creating a strong technical infrastructure for musicology. This report

can be cited in support of grant applications related to digital musicology.

4. CHALLENGES

Since there was not enough time during the session to discuss MIR research challenges from Digital Musicology, Frans Wiering created an online questionnaire. It was announced early December 2013 via various mailing lists, including the Music-IR list and the list of the IMS Study Group on Digital Musicology. By 20 December, 30 responses were collected; the questionnaire will remain open until further notice at <http://bit.ly/1kFRkGB>.

profession	number
musicologist	15
computational/digital musicologist	7
computational/digital musicologist and MIR researcher	3
MIR researcher	2
Music librarian	1
Music librarian, composer	1
Composer, information engineering	1

Table 1. Overview of respondents.

Fifteen respondents describe themselves as musicologists; a further 10 as computational or digital musicologists. Five participants come from outside musicology but have another professional interest in music (see Table 1). Not many participants answered the (optional) question about the most useful computer tool or digital resource in musicology. Only DIAMM (<http://www.diamm.ac.uk/>), music21 (<http://web.mit.edu/music21/>) and RILM (<http://www.rilm.org/>) were put forward multiple times. Interestingly, these come from three different, important areas: digitised sources, automatic music processing and online scholarly literature. The questionnaire had three main questions:

1. What, in your opinion, is the most important open problem (technical or otherwise) in Digital Musicology?
2. What are other important open problems in Digital Musicology?
3. Specifically, what ‘Grand Challenge’ should Digital Musicology propose to the MIR community?

4.1 Responses

The responses were labelled using an informal taxonomy developed while studying the responses. Table 2 gives an overview of the numbers of labels assigned. Note that some of the responses to question 1 contained either multiple suggestions or suggestions that fell under two labels. The labels were grouped under four categories.

Institutional and organisational issues. (1) Access to online resources is still an important issue. Digital musicology is critically dependent on researchers being able to

access the resources they need, but there is no common infrastructure to guarantee that. (2) The problem is particularly felt by researchers without academic affiliation experiencing ‘institutional orphanhood’ and by extension, most likely also by the lay experts targeted by *Transforming Musicology*. (3) Lack of understanding of intellectual property (IP) issues is seen as part of the problem: musicologists still seem to habitually sign their IP rights away to publishers rather than to put their output online. (4) Finally, lack of funding may threaten the sustainability of already existing resources and tools.

Methodology. Methodological issues concern (5) the mutual understanding of disciplinary values and practices, in particular the understanding of musicological research goals by computer scientists. (6) Such understanding must next be translated into computational methods that can deal with the research questions in a satisfactory manner. (7) Software created for this purpose must be usable and have high-quality interfaces and where necessary proper training must be provided.

Resources. (8) Researchers feel a strong need for further data creation both in classical and popular music, and for high-quality Optical Music Recognition (OMR) software to support this endeavour. Two related issues are (9) quality and (10) standardisation. Once data has been created it may be (11) difficult to find it or (12) difficult to track its use by other researchers.

Processing. (13) There is a general need for more tools for various high-level tasks such as annotation, search and analysis. (14) There seems to be a widely experienced need to be able to handle both audio and scores jointly, within a single application.

4.2 Initial findings

The most important open problem in Digital Musicology is clearly (8) data creation: 20% of the observations relate to this issue. OMR is often cited as a possible solution but we need ‘more sophisticated tools in OMR, allowing for working with typical musical notation, and facilitating the automatic encoding of large amounts of musical data’. Some other problems that are closely related to data creation, notably (1) access to online resources and (10) standardisation, also receive high scores.

The second important problem is (7) interfaces, usability and training. Existing software is found difficult to handle, user interfaces are ugly or uninviting, training for scholars is lacking. One respondent suggests that ‘a database of terms and practices’ would already be helpful. Closely related to this are other often-mentioned methodological issues that concern (5) interdisciplinary understanding and (6) matching tools and research questions.

Amongst the MIR challenges, two seem to stand out: (8) data creation and (1) joint handling of scores and recordings. The remarks concerning data creation focus on tool support for symbolic data. Funding for creating high-quality symbolic data has never been very generous; the

Category	Open problems			MIR challenges
	main	other	total	
Institutional / organisational issues				
1. access to online resources	5	1	6	3
2. institutional orphanhood	0	2	2	0
3. awareness of digital publication and IP	1	1	2	0
4. sustainability and funding	2	1	3	0
Methodology				
5. musicological goals and collaboration	4	1	5	1
6. relating tools and research questions	1	6	7	1
7. interfaces, usability and training	4	6	10	2
Resources				
8. data creation	7	8	15	7
9. quality of resources		3	3	1
10. music encoding standards	5	3	8	1
11. findability	1	1	2	0
12. usage and circulation	0	1	1	0
Processing				
13. automatic analysis tools	3	2	5	4
14. joint handling of scores and recordings	4	2	6	7
Total	37	38	75	27

Table 2. Overview of the responses, showing how often each label was assigned. Column ‘main’ corresponds to question 1; column ‘other’ to question 2; column ‘total’ gives the sum of the previous two. Column ‘MIR challenges’ corresponds to question 3.

key issue seems to be to create an infrastructure for collective data creation. Professional attitude is critical, too. One of the world’s leading musicologists included in his response a passionate plea to ‘musicologists who work on critical editions to stop “giving away” their work to publishers,’ otherwise ‘digital musicology for the next 100 years will be stuck with inferior (public-domain) editions.’

Finally, ‘joint handling of scores and recordings in terms of search, annotation and information extraction’ emerged as an issue surprisingly often. This might signal a methodological change in musicology caused by the availability of audio ‘Big Data’. At the same time, it is a genuine research challenge for MIR to create a ‘musical scene-analyser’ that connects the two domains that have been artificially kept separate in MIR research.

5. CONCLUSION

From the preliminary results of the questionnaire, two main challenges for MIR stick out:

1. support for data creation;
2. joint handling of scores and recordings in terms of search, annotation and information extraction.

Both require significant development of computational methods as well as understanding of fundamental musicological processes and values. They can only be successfully resolved if sufficient attention is given to interface design and usability. One could claim that these issues are under-researched in MIR and deserve much stronger

research effort, whether in the context of Digital Musicology or in general.

6. REFERENCES

- [1] T. Collins, A. Arzt, S. Flossmann, G. Widmer: ‘SIARCT-CFP: Improving Precision and the Discovery of Inexact Musical Patterns in Point-Set Representations.’ *Proceedings ISMIR 2013*, pp. 549-554, 2013.
- [2] N. Gonzalez Thomas, P. Pasquier, A. Eigenfeldt, J.B. Maxwell: ‘A Methodology for the Comparison of Melodic Generation Models Using Meta-Melo.’ *Proceedings ISMIR 2013*, pp. 561-566, 2013.
- [3] V. Ishwar, S. Dutta, A. Bellur, H. Murthy: ‘Motif Spotting in an Alapana in Carnatic Music.’ *Proceedings ISMIR 2013*, pp. 499-504, 2013.
- [4] N. Jiang and M. Müller: ‘Automated Methods for Analyzing Music Recordings in Sonata Form.’ *Proceedings ISMIR 2013*, pp. 595-600, 2013.
- [5] Standing Committee for the Humanities. *Musicology (Re-) Mapped*. Ed. E. Dahlig-Turek, S. Klotz, R. Parncutt, F. Wiering. European Science Foundation. 2012. ISBN 978-2-918428-84-8. http://www.esf.org/fileadmin/Public_documents/Publications/musicology.pdf
- [6] A.Volk, W. B. de Haas: ‘A Corpus-Based Study on Ragtime Syncopation.’ *Proceedings ISMIR 2013*, pp. 163-168, 2013.