

# MIR USER STUDIES THROUGH THE LENS OF RELEVANCE: PROMOTING THE IMPACT OF MIR USER RESEARCH

David M. Weigl,<sup>1,2</sup> Joan C. Bartlett,<sup>1</sup> Daniel Steele,<sup>1</sup> Catherine Guastavino<sup>1</sup>

<sup>1</sup>School of Information Studies, McGill University, Canada

<sup>2</sup>Oxford e-Research Centre, University of Oxford, UK

david.weigl@oerc.ox.ac.uk

## ABSTRACT

Although there have been repeated calls in the MIR literature for a greater emphasis on the (potential) users of music information systems—complementing valuable and thriving research on MIR algorithms—the impact of formal investigations of user information needs and information behaviour on MIR system design has been limited [5] [7] [8]. Challenges of generalization, a lack of systematic synthesis of results, and the disconnect between system/evaluation task designers and user studies researchers have been proposed as potential reasons for this situation [5]. To address these issues, we present an analytical interface and query mechanism operating over a large and extensible corpus of empirical findings identified during a systematic analysis of 159 research articles presenting MIR user studies, coded according to a conceptual framework for *relevance*, a central notion at the heart of information searching and information retrieval [3].

## 1. MOTIVATION: RELEVANCE IN MIR

The definition and operationalisation of experiential similarity and relevance measures—a key research priority for Music Information Retrieval (MIR) [1]—stands to benefit greatly from user-focussed research efforts. Discussions on the nature of relevance and its place at the heart of Information Retrieval (IR) have been ongoing for decades, and have been traced to the very beginnings of organized academic research into textual IR [6]. While some initial investigations have probed this notion in the music information domain, finding significant overlap with textual IR in terms of the criteria users apply in their relevance judgements [2] [4], formal consideration of relevance in MIR has remained scarce. To address this, we have adapted an established conceptual framework for relevance in textual IR—the stratified model of relevance interactions [6] [9]—to clarify the variety of relevance considerations for MIR.

## 2. CORPUS AND CODING ACTIVITY

We have conducted a systematic analysis of a set of 159 articles reporting on user studies in the MIR literature, compiled by Lee and Cunningham in previous research [5] presenting a predominantly bibliometric investigation. In contrast, our research is situated at the content level, codifying empirical findings reported within each article according to our conceptual framework for relevance. At present, the resulting corpus contains 866 discrete findings derived from 176 studies contained within the corpus (publications can report on more than one study.) We have developed tools to enable real-time updates as new studies are added.

The coding activity was undertaken by four researchers employing a collaborative coding tool. The tool allows coders to assign one or more studies to each article in the set, capturing descriptions of the given study’s purpose, the sample size and frame, and the methodology employed. Findings are assigned to each study by providing a short textual description, and by choosing descriptors from a set of drop-down menus corresponding to each stratum (conceptual category) of the relevance model. The initial set of descriptors was determined based on a preliminary application of the model to the music domain [9], and on subsequent discussions during the initial phase of the coding process. In cases where none of the available descriptors suitably convey the intended meaning, new descriptors can be minted by specifying a stratum of the model and providing a corresponding new label and explication. Provenance information describing the researcher responsible, as well as the specific article, study, and finding that prompted the addition, is captured automatically by the tool. Newly added descriptors become immediately available to all researchers in order to promote inter-coder consistency.

## 3. PROVIDING ACCESS

The coded corpus of findings is publicly accessible through an analytical interface<sup>1</sup> presenting two interaction modes: a co-occurrence matrix (Figure 1) providing access to a two-dimensional projection of the classification space defined by the relevance model, and a query mechanism filtering the corpus of findings according to individual stratum descriptors. The former interaction mode provides an overview of the differing extent to which the classification



© David M. Weigl,<sup>1,2</sup> Joan C. Bartlett,<sup>1</sup> Daniel Steele,<sup>1</sup> Catherine Guastavino<sup>1</sup>. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** David M. Weigl,<sup>1,2</sup> Joan C. Bartlett,<sup>1</sup> Daniel Steele,<sup>1</sup> Catherine Guastavino<sup>1</sup>. “MIR user studies through the lens of relevance: Promoting the impact of MIR user research”, Extended abstracts for the Late-Breaking Demo Session of the 17th International Society for Music Information Retrieval Conference, 2016.

<sup>1</sup> Available at <http://relevance.linkedmusic.org>

space outlined by the model has been explored in previous research; users can drill down to access a detailed list of corresponding findings by clicking on a given cell of the diagram. We anticipate that this mode will be of use to MIR user researchers, aiding both in synthesis of the literature by demonstrating areas of converging findings across different studies, and in hypothesis generation by identifying gaps in our present knowledge. The latter interaction mode is aimed primarily at MIR system designers and evaluators wishing to rapidly gain an overview of MIR user research findings pertaining to their specific area of interest. For instance, developers addressing a mood classification task are provided with simple and straightforward access to summary descriptions of the 29 findings on mood management, 24 findings on emotion connotation, and 17 findings on music emotion elicitation identified in the collection of user studies, each presented in context of their respective generating study, and accompanied with Google Scholar queries to the corresponding article for each finding.

#### 4. PROMOTING THE IMPACT OF MIR USER RESEARCH

Lee and Cunningham [5] outlined a number of challenges limiting the impact of MIR user research. These include challenges of generalization, whereby the common reliance on convenience sampling and limitations of sample size require triangulation between different studies in order to identify findings that generalize across multiple groups of users; a lack of systematic synthesis of research results, complicated by the highly diffuse nature of their dissemination, the reviewed articles having appeared in 83 distinct publication venues across a wide variety of research fields; and, perhaps as a consequence of the preceding challenges, a disconnect between system/evaluation task designers and user studies researchers. Lee and Cunningham publicly share their list of articles with full citation information, in the spirit of collaborative and transparent research synthesis, as a first step toward addressing these issues. We gratefully adopt this collection, and reciprocate by making our corpus of coded studies and findings available, accessible and queryable through a simple user interface. Further, we invite other MIR user studies researchers to help grow the corpus in order to keep up with the fast-changing nature of our field, initially by contacting the corresponding author. We hope that the outcomes of this process will be useful to future MIR user research, providing a reference of current knowledge, a tool for hypothesis generation, and a lens to promote the synergic impact of these studies in terms of implications for MIR system design.

#### 5. ACKNOWLEDGEMENTS

This work forms part of doctoral research undertaken at the School of Information Studies, McGill University. Support in part by a McGill University William Dawson Award to C. Guastavino, and by the EPSRC *Fusing Semantic and Audio Technologies for Intelligent Music Production and Consumption (FAST IMPACT)* project (EP/L019981/1).

Cultural	3.2	0.1	0.2	2.1	2.4	1.4	2.2	1.8	7.6
Social	2.9	0.3	0.3	4.3	4.6	1	3.8	13.7	1.8
Situational	6	2.2	0.9	6.5	9.7	4.6	24.6	3.8	2.2
Affective	3.6	0.9	0	3.1	5.5	12.4	4.6	1	1.4
Cognitive	13	6.6	0.5	17	42.7	5.5	9.7	4.6	2.4
Interface	20.7	8.1	1.3	49.7	17	3.1	6.5	4.3	2.1
Engineering	1.3	0	2.9	1.3	0.5	0	0.9	0.3	0.2
Processing	6.1	15.2	0	8.1	6.6	0.9	2.2	0.3	0.1
Content	40.6	6.1	1.3	20.7	13	3.6	6	2.9	3.2

**Figure 1.** Distribution of findings according to the strata outlined in the relevance model, expressed as percentages of the total number of findings ( $N=866$ ).

#### 6. REFERENCES

- [1] J. S. Downie. Music information retrieval. *Annual review of information science and technology*, 37(1):295–340, 2003.
- [2] C. Inskip, A. MacFarlane, and P. Rafferty. Creative professional users’ musical relevance criteria. *Journal of Information Science*, 36(4):517–529, June 2010.
- [3] B. J. Jansen and S. Y. Rieh. The 17 theoretical constructs of information searching and information retrieval. *Journal of the American Society for Information Science and Technology*, 61(8):1517–1534, 2010.
- [4] A. Laplante. User’s relevance criteria in music retrieval in everyday life: An exploratory study. In *Proc. ISMIR 2010*, pages 601–606, 2010.
- [5] Jin Ha Lee and Sally Jo Cunningham. The impact (or non-impact) of user studies in music information retrieval. In *Proc. ISMIR 2012*, pages 391–396, 2012.
- [6] Tefko Saracevic. Relevance: A review of the literature and a framework for thinking on the notion in information science. Part II: nature and manifestations of relevance. *Journal of the American Society for Information Science and Technology*, 58(13):1915–1933, 2007.
- [7] Markus Schedl, Arthur Flexer, and Julián Urbano. The neglected user in music information retrieval research. *Journal for Intelligent Information Systems*, 41(3):523–539, 2013.
- [8] D. M. Weigl and C. Guastavino. User studies in the music information retrieval literature. In *Proc. ISMIR 2011*, pages 335–340, 2011.
- [9] D. M. Weigl and C. Guastavino. Applying the stratified model of relevance interactions to music information retrieval. *ASIST 2013*, 50(1):1–4, 2013.