

A Shared Vocabulary for Audio Features

Alo Allik, György Fazekas, Simon Dixon, and Mark Sandler

Queen Mary University of London

{alo.allik,gyorgy.fazekas,simon.dixon,mark.sandler}@eecs.qmul.ac.uk

Abstract. The aim of the Shared Open Vocabulary for Audio Research and Retrieval project is to foster greater agreement on the representation of content-based audio features within music research communities. The Audio Feature Ontology has been developed for this purpose as part of a library of modular ontologies in order to increase interoperability, reproducibility and sustainability in music information retrieval workflows. The ontology provides a descriptive framework for expressing different conceptualisations of the audio features domain and allows for publishing content-derived information about audio recordings.

Introduction. Due to profusion of digital audio content there is a growing demand for vocabularies that facilitate interoperability between music related data sources. This includes content-based audio feature data, which can be utilised for various commercial and research purposes such as automatic genre classification, query by humming services and many others. Researchers in audio and music information retrieval increasingly use common sets of features to characterise audio material, and large data sets are released for commercial and scientific use. The development of data sets and research tools however is not governed by shared open vocabularies and common data structures. Current formats for describing content-based audio features, including standard as well as generic structured data formats are limited in their extensibility, modularity and interoperability. This limits their ability to support reproducible research, sustainable research tools, and the creation of shared data sets.

The Audio Feature Ontology provides a model for describing acoustical and musicological data and allows for publishing content-derived information about audio recordings. It was initially created within the Music Ontology framework[1]. The Music Ontology has been adopted by researchers and user communities, as well as the industry, including the BBC and its music website (<http://bbc.co.uk/music>). The audio feature ontology subsumes concepts from other ontologies in this framework, including the Event and Timeline ontologies. With regards to different conceptualisations of feature representations, the Audio Feature ontology deals with data density and temporal characteristics[1]. However, its vocabulary is incomplete with regards to user needs within the audio research communities as a number of popular features extracted in research and commercial software are not supported.

Updated Audio Feature Ontology. The aim of this work is to foster greater agreement on the representation of audio features within research communities and to extend the Audio Feature Ontology. It may not be feasible however that

a single ontology may completely represent the different conceptualisations of the domain that exist in the research communities. For example, in some contexts, audio features are categorised according to musicological concepts, such as pitch, rhythm and timbre, while in others the computational workflow used in calculating the features determines the taxonomic hierarchy. The main scope of the present ontology is to provide a framework for communication, feature representation, and describe the association of features and audio signals.

Approach. In order to gain a better understanding of the domain and user needs, a catalogue of audio features was first compiled based on a thorough review of relevant literature, existing feature extraction tools and vocabularies, and research workflows. The catalogue was created in linked data format listing feature objects and their attributes and serves as the foundation for a hybrid ontology engineering process, combining manual and semi-automatic approaches[2]. This catalogue facilitates conceptual scaling of feature attributes to generate concept lattices for Formal Concept Analysis (FCA)[3]. Concept lattices facilitate the extraction of conceptual hierarchies and thereby can provide a foundation for shared vocabularies. Subsets of popular features computed in feature extraction tools can be queried from the linked data graph to help define the scope and domain boundaries of the ontology. The catalogue identifies approximately 400 different features and significantly increases the scope of the ontology.

Conclusion. A shared vocabulary for audio features has the potential to increase research interoperability, reproducibility and sustainability in music information retrieval research communities. Leveraging open linked data formats, it provides a flat vocabulary to facilitate task and tool specific ontology development and serves as a descriptive framework for audio feature extraction. Future work includes developing specific ontologies based on the existing tools and vocabularies, and updating Sonic Annotator, a command line application for automatically analysing large audio collections. SAWA, an open access demonstrator will also be updated to adhere to the revised ontology. Finally, Sonic Visualiser, a desktop application for visualising audio features will be updated to handle RDF files adhering to the updated ontology.

Acknowledgements. This research was funded by JISC (<http://jisc.ac.uk>) through the Shared Open Vocabulary for Audio Research and Retrieval project.

References

1. Fazekas, G., Raimond, Y., Jakobson, K., Sandler, M.: An overview of semantic web activities in the OMRAS2 project. *Journal of New Music Research (JNMR)* 39(4) (2010)
2. Cimiano, P., Hotho, A., Stumme, G., Tane, J.: Conceptual knowledge processing with formal concept analysis and ontologies. In: Eklund, P. (ed.) *ICFCA 2004. LNCS (LNAI)*, vol. 2961, pp. 189–207. Springer, Heidelberg (2004)
3. Wille, R.: Concept lattices and conceptual knowledge systems. *Computers and Mathematics with Applications* 23, 493–515 (1992)