A Dataset of Multimedia Material About Classical Music: PHENICX-SMM

Markus Schedl David Hauger
Johannes Kepler University
Department of Computational Perception
Linz, Austria
Email: {markus.schedl, david.hauger}@jku.at

Marko Tkalčič
Free University of Bolzano
Faculty of Computer Science
Bozen–Bolzano, Italy
Email: marko.tkalcic@unibz.it

Mark Melenhorst Cynthia C.S. Liem
Multimedia Computing Group
Delft University of Technology
the Netherlands
Email: {M.S.Melenhorst, C.C.S.Liem}@tudelft.nl

***Abstract*—We present a freely available dataset of multimedia material that can be used to build enriched browsing and retrieval systems for music. It is one result of the EU-FP7 funded project "Performances as Highly Enriched aNd Interactive Concert eXperiences" (PHENICX) that aims at enhancing the listener experience when enjoying classical music. The presented PHENICX-SMM dataset includes in total more than 50,000 multimedia items (text, image, audio) about composers, performers, pieces, and instruments.**

In addition to presenting the dataset, we detail one possible use case, that of building a personalized music information system that suggests certain types and quantities of multimedia material, based on personality traits and musical experience of its users. We evaluate the system via a user study and show that people generally prefer the personalized results over non-personalized.

I. Introduction and Context

Classical music is a great asset and important part of our cultural heritage. Nowadays, however, fewer and fewer people attend respective orchestra concerts. There is strong evidence that in particular the younger generation is reluctant to attend concerts, due to various reasons, not least because of a lack of appealing technological pervasion, e.g., via smartphone or tablet apps supporting concerts by offering additional (multimedia) material or extensive activity on social media [8].[1]

In the EU-FP7 funded project "Performances as Highly Enriched aNd Interactive Concert eXperiences" (PHENICX)[2] this work is connected to, a central aim is to make classical music accessible to new audiences [7]. To this end, several strategies have been elaborated. One of them is to create informative, appealing, and easy-to-use systems to visualize and interact with multimodal information in order to support and enhance the listening experience. Taking into account the different individual preferences towards particular pieces of such information (e.g., biographies, tags, images, or audio samples) enables the creation of personalized systems to experience various facets of classical music.

In order to provide such supporting multimedia material, it obviously needs to be collected first. Within the PHENICX project, we did so by elaborating several strategies to harvest this data from a variety of web sources. Here, we present the resulting PHENICX-SMM dataset, where SMM refers to "supporting multimedia material". Among other tasks, the dataset can be used to create digital program notes for classical concert performances [10], which support concert goers with multimedia material tailored to their information or entertainment need, in real time during the concert. The presented dataset focuses on classical music items, e.g., performers and pieces, but its usage is in principle not restricted to this kind of music, e.g., material about instruments is included, too.

While there exists a wealth of other music-related datasets, such as the Million Song Dataset[3] [1], the Yahoo! Music dataset [2], or the MagnaTagATune[4] dataset [6], just to name very few of the more popular ones, they typically provide different kinds of computational or social features (e.g., audio-based descriptors, collaborative tags, ratings, or similarity information). To the best of our knowledge, the PHENICX-SMM dataset is the first to offer a selection of multimedia items about various kinds of music items.

The remainder is organized in the following way. In Section II, we detail the data acquisition procedure, the composition of the dataset, and its availability. We further present and discuss basic statistics of the dataset. In Section III, we showcase the use of PHENICX-SMM in a personalized system to access various multimedia material of different types (text, image, audio) for different music items (composer, performer, artist, instrument, piece). We round off the paper with a conclusion in Section IV.

II. Dataset

A. Acquisition

In order to acquire the dataset, we first created lists of seed items (pieces, artists, instruments, etc.). To this end, we exploited different music-related data sources, e.g., Wikipedia lists,[5] Classical Net,[6] Last.fm,[7] and Freebase.[8] The lists based

[1] This also became evident in personal discussions between the authors and members and leaders of the Royal Concertgebouw Orchestra (RCO) Amsterdam, with whom we closely collaborate.

[2] http://phenicx.upf.edu

[3] http://labrosa.ee.columbia.edu/millionsong

[4] http://mi.soi.city.ac.uk/blog/codeapps/the-magnatagatune-dataset

[5] https://en.wikipedia.org/wiki/List_of_Classical-era_composers

[6] http://classical.net

[7] http://www.last.fm

[8] http://www.freebase.com

978-1-4673-8695-1/16/$31.00 © 2016 IEEE

on Wikipedia and Classical Net were directly taken from the respective websites and parsed accordingly. For Last.fm, we used the provided API, in particular we retrieved the top artists for tags related to classical music.[9] To obtain the seeds from Freebase, we implemented scripts using the semantic information to collect items matching the desired semantic concepts, e.g., musical instrument.[10]. In addition, we added the pieces and composers in the repertoire of our project partner RCO, of the seasons 2014 and 2015. Eventually, the entity lists contained a number of items, which is reported in Table I.

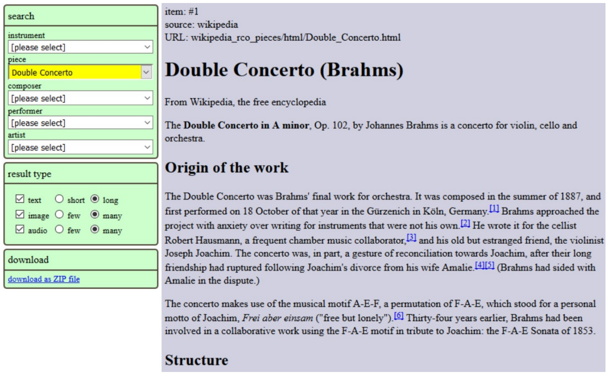
We subsequently implemented crawlers to fetch material from public data sources, in particular Wikipedia and Freebase. We automatically extracted respective content and categorized them according to type of media into text, images, and audio. During this process, we had to face and solve several challenges. The first one was disambiguation problems. While not an issue for semantically structured information offered by Freebase, other sources like Wikipedia might return various results for a given search terms. To give an example, "triangle" is considered an instrument in our context, but the main result on Wikipedia is the geometric shape. We resolved such cases by contextualizing the query. We implemented simple heuristics using explicit disambiguation terms. For instance, the type of the item, e.g., "_(instrument)", or the name of the composer, for musical pieces, was added to the query and enabled us to retrieve the desired Wikipedia pages.

Another challenge concerned the quality of the related web material. In many cases, the Wikipedia pages contain images and descriptions, but the resolution of the linked material is inferior. Applying link following on meta information contained in the HTML code and parsing of linked pages, we were able to gather the original versions of images in full resolution.

B. Availability

The dataset is available for download from a dedicated web page.[11] All items in the dataset are licensed according to the respective open license models of the sources we extracted them from, i.e., Freebase[12] and Wikipedia,[13] which are variants of the Creative Commons license[14] or release as Public Domain.[15]

To directly access the content of the dataset, we implemented a web interface,[16] which offers browsing capabilities by category (instrument, piece, composer, etc.) and by type of material (text, image, audio). The interface also supports downloading only the material for a specific item, hereby omitting the need to download the full dataset in case only parts are needed. Figure 1 shows a screenshot of the user interface.



instrument
[please select]
piece
Double Concerto
composer
[please select]
performer
[please select]
artist
[please select]

☑ text ○ short ● long
☑ image ○ few ● many
☑ audio ○ few ● many

item: #1
source: wikipedia
URL: wikipedia_rco_pieces/html/Double_Concerto.html

Double Concerto (Brahms)

From Wikipedia, the free encyclopedia

The **Double Concerto in A minor**, Op. 102, by Johannes Brahms is a concerto for violin, cello and orchestra.

Origin of the work

The Double Concerto was Brahms' final work for orchestra. It was composed in the summer of 1887, and first performed on 18 October of that year in the Gürzenich in Köln, Germany.[1] Brahms approached the project with anxiety over writing for instruments that were not his own.[2] He wrote it for the cellist Robert Hausmann, a frequent chamber music collaborator,[3] and his old but estranged friend, the violinist Joseph Joachim. The concerto was, in part, a gesture of reconciliation towards Joachim, after their long friendship had ruptured following Joachim's divorce from his wife Amalie.[4][5] (Brahms had sided with Amalie in the dispute.)

The concerto makes use of the musical motif A-E-F, a permutation of F-A-E, which stood for a personal motto of Joachim, *Frei aber einsam* ("free but lonely").[6] Thirty-four years earlier, Brahms had been involved in a collaborative work using the F-A-E motif in tribute to Joachim: the F-A-E Sonata of 1853.

Structure

Fig. 1. Parts of the personalized user interface to browse the dataset.

TABLE I
STATISTICS OF ITEMS IN THE DATASET.

Item	Quantity
Music pieces	1,091
Composers	51,168
Performers / Ensembles	104
Artists (Composers and Performers)	1,000
Instruments	1,670
Sum	55,033

C. Content and Statistics

The multimedia items comprising the dataset are organized in a folder structure as illustrated and explained in Table II. Furthermore, the table shows the total number of items in each main folder. All in all, we were able to gather more than 183,000 pieces of multimedia data, i.e., individual files. The lists of seed items are provided as text files in folder `seeds`. The material extracted from Freebase is organized in subfolders of folder `freebase`, while the pieces of information acquired from Wikipedia are organized in folder structures of the form `wikipedia_item-type_seedlist-source`. The subfolders `html`, `audio`, `img`, and `imgSmall` further categorize the material according to media type: structured text, audio, and images in full size and thumbnail size, respectively.

To illustrate the variety of information covered by the provided material, we exemplary show in Figure 2 some image results for the instrument harp. Textual and audio results can be easily browsed using the web interface described above.

III. USE CASE: PERSONALIZED MULTIMEDIA MUSIC INFORMATION SYSTEM

As a use case for the dataset, we built an extension to the web-based browsing interface described above and depicted in Figure 1. Imagine two persons who want to learn more about a certain music piece — one classical music aficionado who is going to attend a concert tomorrow, and one heavy metal fan and keen guitarist who wants to know more about the particular piece because it was interpreted by his favorite metal band. Presumably, they will have quite different information needs. While our knowledgeable classical music lover might be interested in a detailed description of the historic context

[9] http://www.last.fm/api/show/tag.getTopArtists
[10] http://www.freebase.com/music/instrument
[11] http://www.cp.jku.at/datasets/PHENICX-SMM
[12] http://www.freebase.com
[13] https://www.wikipedia.org
[14] https://creativecommons.org
[15] For more details about the licenses, please consider https://en.wikipedia.org/wiki/Wikipedia:Copyrights and http://wiki.freebase.com/wiki/License_compatibility.
[16] http://einaudi.cp.jku.at/multimediadb

TABLE II
ORGANIZATION OF THE DATASET.

Folder	Description	Items
seeds	text files containing the lists of seed items, organized by data source and item type	12
freebase/crawl_composers	JSON files containing structured composer information acquired from Freebase	49,143
freebase/crawl_instruments	JSON files containing structured composer information acquired from Freebase	1,455
freebase/images_composers	images of composers acquired from Freebase	14,852
freebase/images_instruments	images of instruments acquired from Freebase	1,176
wikipedia_artists_lastfm/[html,audio,img,imgSmall]	HTML, audio, and image (full size and thumbnail size) content of Wikipedia articles acquired for the Last.fm seed list of artists	5,998
wikipedia_tracks_lastfm/[html,audio,img,imgSmall]	HTML, audio, and image (full size and thumbnail size) content of Wikipedia articles acquired for the Last.fm seed list of tracks	1,709
wikipedia_composers_classical.net/[html,audio,img,imgSmall]	HTML, audio, and image content extracted from Wikipedia articles acquired for the list of composers from Classical Net	3,379
wikipedia_composers_freebase/[html,audio,img,imgSmall]	same, but for the list of composers from Freebase	91,718
wikipedia_composers_wiki/[html,audio,img,imgSmall]	same, but for the list of composers from Wikipedia	2,120
wikipedia_ensembles_wiki/[html,audio,img,imgSmall]	same, but for the list of classical music ensembles from Wikipedia	116
wikipedia_performers_wiki/[html,audio,img,imgSmall]	same, but for the list of classical music performers from Wikipedia	437
wikipedia_instruments_freebase/[html,audio,img,imgSmall]	same, but for the list of instruments from Freebase	7,924
wikipedia_instruments_wiki/[html,audio,img,imgSmall]	same, but for the list of instruments from Wikipedia	739
wikipedia_rco_pieces/[html,audio,img,imgSmall]	same, but for the list of pieces from the RCO seasons 2014 and 2015	176

of the composition (as text) or an exhaustive list of concert performances of the piece (again as text), our metal guitarist might more likely be interested in the scores of the main melody (images) or a few music examples (audio files). As illustrated by the example, different people differ in terms of type as well as amount or length of the desired multimedia material. We further substantiated this hypothesis by conducting an initial user study [10], in which we observed that users with different personal characteristics (music experience and sophistication as well as personality) tend to prefer different lengths or amounts of the material.

A. *Web Interface*

We used these insights to develop a prototype that adapts the output of the results to the user's personal characteristics. To this end, we implemented a personalized version of the web interface to the music information system, which is available at http://bird.cp.jku.at/phenicx_mmsupp. It allows to retrieve and browse supporting information about composers, performers, pieces, and instruments. The results of a query are presented in the form of text, image, and audio. In the web interface, the user performs a search query through a set of dropdown menus, as shown in Figure 1. The search query itself (the upper frame in the green box of the user interface) is not personalized. The personalization affects the result type preferences (the result type frame in the green box of the user interface). More specifically, for the two multimodal types analyzed in our study (i.e., text and image) the personalization affects the length and amount of the results. Some users prefer long text over short and some prefer many images over few. For example, in our initial study mentioned above, we found that people who score high on the personality traits of openness, agreeableness, conscientiousness, and extraversion [4] tend to show a positive correlation with consumption, interestingness and novelty factors, i.e., they prefer to consume more of the material, find it interesting and novel.

B. *Personalization*

In order to build a user-friendly personalized system, we decided to recommend to each user the amount and length of material that has the highest average preference (rating) among users with similar personal characteristics. However, there exists a trade-off between the user's willingness to fill in the fully fledged BFI-44 personality questionnaire [5] and music sophistication questionnaire [9] on the one hand, and the accuracy of the recommendations. Since the system is not only a research prototype, but is currently being implemented into a mobile application by our business partner in the PHENICX project, we decided to select only two initial questions, according to which users are categorized and their preferences assessed. To this end, we identified the questions that account for most of the variance in the sample of participants in the initial study [10]. The choice of the two questions was done by observing how the participants' answers to the music sophistication and personality questions correlate with the preferences for the content. The one with the highest absolute value of the correlation is *How many classical concerts do you attend per year?* Among the personality questions, it is *I myself as someone who is reserved, quiet.*

In order to personalize the web interface for the supporting multimedia material, we implemented a recommender system that adapts the result types (i.e., the length of the text and the number of the images shown) to the end user. The personalization is done by recommending the result types that are most popular in the cluster of users the active user belongs to. We used four clusters created by categorizing users according to their answers to the two selected questions (four quadrants based on media split).

C. *Evaluation*

In order to evaluate the satisfaction of users with the recommended settings, we performed a user study for which we recruited 96 participants through Amazon Mechanical

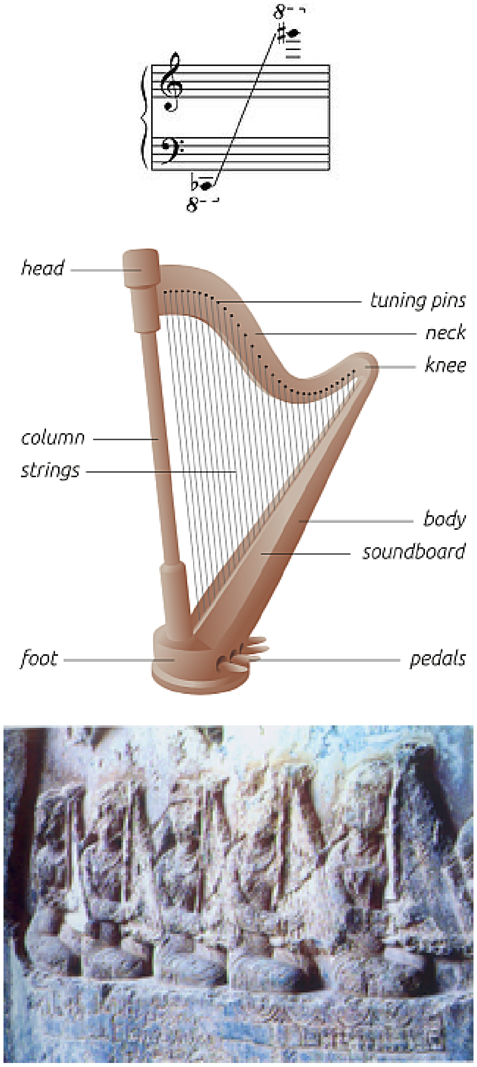


Fig. 2. Different image material for item "harp". From top to bottom: playing range of a modern pedal harp (PD, by Number Googol), structural elements and terminology of a modern concert harp (CC BY-SA 3.0, by Martin Kraft), and ancient Persian harps carved in stone (Attribution, by Amir85 at en.wikipedia).

Turk.We first asked the participants to answer the initial two questions mentioned above. Then they were shown a snapshot of the personalized interface with content and were asked to provide a rating on a Likert scale from 1 to 5. The rating was in the form of an agreement with the question *I like the way the material is presented*. In the next step, the participants were shown a random snapshot of the interface and were asked to provide a rating. In a final step, they were shown both snapshots next to each other and were asked to select the one they preferred.

The results have shown that in 63 (out of 96) of the cases the participants preferred the personalized snapshot, which corresponds to 66%. The higher preference for the personalized interface was confirmed also in the comparison of the ratings given to the two snapshots. The mean rating for the personalized snapshot was 3.29 (on a scale from 1 to 5, where 1 was the lowest and 5 the highest score) compared to the mean rating for the random snapshot of 2.92. The Wilcoxon signed-rank test [3] showed that the mean difference was significant ($p = 0.01$).

IV. Conclusion

We presented the PHENICX-SMM dataset containing more than 180,000 multimedia items about various entities in classical music: composers, performers, instruments, and pieces. We detailed the data acquisition process, as well as the content, structure, and availability of the dataset. In addition, we introduced a web-based browsing interface to access the dataset in a convenient way. To provide a use case, we illustrated how the dataset and the web interface can be used to create a simple personalized multimedia information system for classical music. Based on the user's answers to music sophistication and personality questions, we implemented a personalization strategy for the selection of content types, in particular the amount and length of multimedia material presented to the user. We compared the resulting personalized system with a non-personalized one and identified in a user study of 96 participants a significant preference for the personalized version.

Acknowledgment

This work was supported by the EU FP7-ICT-2011-9 project 601166 ("PHENICX") and the Austrian Science Fund (FWF): P25655.

References

[1] T. Bertin-Mahieux, D. P. Ellis, B. Whitman, and P. Lamere. The Million Song Dataset. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR)*, Miami, FL, USA, October 2011.

[2] G. Dror, N. Koenigstein, Y. Koren, and M. Weimer. The Yahoo! Music Dataset and KDD-Cup'11. *JMLR: Proceedings of KDD-Cup 2011 Competition*, 18:3–18, October 2012.

[3] A. Field, J. Miles, and Z. Field. *Discovering Statistics Using R*. Sage Publications, 2012.

[4] S. D. Gosling, P. J. Rentfrow, and W. B. Swann. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality*, 37(6):504–528, December 2003.

[5] O. John and S. Srivastava. The Big Five Trait Taxonomy: History, Measurement, and Theoretical Perspectives. In L. A. Pervin and O. P. John, editors, *Handbook of personality: Theory and research*, number 510, pages 102–138. Guilford Press, New York, second edition, 1999.

[6] E. Law, K. West, M. Mandel, M. Bay, and J. S. Downie. Evaluation of Algorithms Using Games: The Case of Music Annotation. In *Proceedings of the 11th International Society for Music Information Retrieval Conference (ISMIR)*, Utrecht, the Netherlands, August 2010.

[7] C. C. Liem, E. Gómez, and M. Schedl. PHENICX: Innovating the Classical Music Experience. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, Torino, Italy, June–July 2015.

[8] H. Mielonen. Attracting New Audiences: Attitudes and Experiences in Attending Classical Music Concert of Students in Their Twenties. Master's thesis, Sibelius Academy / University of Arts, Helsinki, Finland, 2003.

[9] D. Müllensiefen, B. Gingras, J. Musil, and L. Stewart. The Musicality of Non-musicians: An Index for Assessing Musical Sophistication in the General Population. *PloS one*, 9(2):e89642, January 2014.

[10] M. Tkalčič, B. Ferwerda, D. Hauger, and M. Schedl. Personality Correlates for Digital Concert Program Notes. In *Proceedings of the 22nd International Conference on User Modeling, Adaptation and Personalization (UMAP)*, Dublin, Ireland, June–July 2015.