



Wagner Ring Dataset: A Complex Opera Scenario for Music Processing and Computational Musicology

DATASET

 jubiquity press

CHRISTOF WEIß 

VLORA ARIFI-MÜLLER

MICHAEL KRAUSE 

FRANK ZALKOW 

STEPHANIE KLAUK

RAINER KLEINERTZ

MEINARD MÜLLER 

*Author affiliations can be found in the back matter of this article

ABSTRACT

This paper introduces the Wagner Ring Dataset (WRD), a multi-modal and multi-version resource on the large-scale opera cycle *Der Ring des Nibelungen* by Richard Wagner. The *Ring* comprises four music dramas organized into eleven acts and 21 939 measures in total. Concerning sheet music, we processed a publicly available piano reduction (822 pages) of the full score with optical music recognition followed by extensive manual corrections to create a high-quality, machine-readable symbolic score. Concerning audio data, our corpus covers 16 recorded performances of the full *Ring* (three of which are publicly available thanks to copyright expiry), each lasting about 14–15 hours. To musically synchronize these versions among each other, we manually annotated all measure positions for three performances, which we transferred to the remaining performances via automated synchronization techniques. The dataset further comprises annotations of key and time signatures, scenes, and singing voice regions (libretto). Moreover, we provide note event annotations for all performances derived from the piano score. The WRD thus constitutes a comprehensive resource for developing algorithms for various music information retrieval tasks, complementing existing datasets with a complex opera scenario. For computational musicology, the WRD serves as a structured dataset that allows for studying the composition and performances of the *Ring*.

CORRESPONDING AUTHOR:

Christof Weiß

University of Würzburg,
Germany; International
Audio Laboratories Erlangen,
Germany

christof.weiss@uni-wuerzburg.de

KEYWORDS:

Dataset; Multi-Modal; Cross-Version; Music Processing; Computational Musicology

TO CITE THIS ARTICLE:

Weiß, C., Arifi-Müller, V., Krause, M., Zalkow, F., Klauk, S., Kleinertz, R., & Müller, M. (2023). Wagner Ring Dataset: A Complex Opera Scenario for Music Processing and Computational Musicology. *Transactions of the International Society for Music Information Retrieval*, 6(1), 135–149. DOI: <https://doi.org/10.5334/tismir.161>

1. INTRODUCTION

Over the past decades, computational musicology has matured into an established research area, developing and applying algorithmic tools for the study of musical phenomena (Volk et al., 2011; Meredith, 2016; Müller, 2021). Nowadays, this field spans a variety of the world's musical cultures (Serra, 2014; Tzanetakis, 2014). Among those, Western classical music has been one of the earliest focus areas (Hewlett and Selfridge-Field, 1991), which—from a technical perspective—provides interesting opportunities because of its multi-modal resources: there is a specific correspondence between what is written down in sheet music (either represented as graphical or symbolic data) and what is recorded (audio data)—typically comprising a multitude of professional performances. This not only allows for studying performance aspects (Lerch et al., 2020) but also for developing and testing algorithmic approaches to various tasks including automatic music transcription (Benetos et al., 2019), optical music recognition (Calvo-Zaragoza et al., 2020), music synchronization (Müller et al., 2021), retrieval (Müller et al., 2019), and analysis (Nieto et al., 2020; Bosch, 2013; Meredith, 2016).

While the field has mostly focused on piano (Hawthorne et al., 2019) and chamber music (Thickstun et al., 2017), opera is an interesting yet challenging scenario (Röwenstrunk et al., 2015; Prätzlich and Müller, 2013), where one has to deal with various modalities regarding the libretto (text), the composition (image, symbolic), and the performances (audio, video) as well as with the large extent of these works, which often last for several hours—a tremendous contrast to the 3–5 minutes duration of most pop songs. One of the most extensive works is the tetralogy *Der Ring des Nibelungen*, being created by Richard Wagner between 1848 and 1874 and comprising four operas (which Wagner called *music dramas*). Figure 1 (bottom part) shows an overview of the *Ring's* parts. A full performance lasts about 14–15 hours in total (typically spread over four days), which is demanding for listeners as well as performers. For exploring this extensive work, navigation and visualization tools are particularly useful, thus making this scenario a prime example for computational musicology.

Beyond its length, the tetralogy is also a highly complex artwork. The libretto (written by Wagner himself) unfolds an intricate plot with a multi-faceted mesh of interpersonal relationships. The full score comprises 21939 measures and is characterized by a large orchestra, expressive melodic lines, highly chromatic harmony, a multi-layered rhythmic structure as well as a dense texture of characteristic motifs (*leitmotifs*) that establish complex relationships across the four operas. Finally, the recorded performances pose enormous challenges for audio signal processing, considering the sound of a huge symphony orchestra together with highly expressive

singing, which involves vibrato amplitudes up to several semitones. Standard algorithms as developed in music information retrieval (MIR) are barely applicable to this complex scenario since they are usually not conceived for this style (but for popular music) and do not account for the specific needs in musicological research.

With this paper, we introduce the Wagner Ring Dataset (WRD), a multi-modal and multi-version resource on Wagner's *Ring* created within a long-term interdisciplinary research project.¹ We structure the dataset according to the *Ring's* eleven acts (*Aufzüge*); see Section 3 for details. Concerning raw data, the WRD comprises a piano reduction of the full orchestral score in graphical and symbolic (Sibelius, MusicXML, and MIDI) format as well as 16 recorded performances of the full *Ring* (amounting to almost 232 hours of audio), three of which are publicly available thanks to copyright expiry.² To link these versions across each other, we provide comprehensive annotations of musical measure positions, which we created in a manual fashion for three performances (Weiß et al., 2016) and then transferred to the others using a synchronization method (Zalkow et al., 2017b). As a major contribution, the WRD provides a rich set of comprehensive annotations (note events; scene, key and time signature regions; singing voice regions including singer information and libretto text). We created these annotations on the score level and then transferred this information to the physical time axis of the individual performances. The dataset is publicly available via Zenodo.³

In this paper, we describe the data and the curation process in detail and sketch possible applications, thus demonstrating its potential value for computational musicology and MIR. We first outline historical background on the *Ring* along with related work (Section 2) and explain the dataset's structure and organization (Section 3). Then, we describe the creation of the symbolic piano scores (Section 4) and introduce the audio recordings (Section 5). We explain the generation of measure annotations (Section 6) and give an overview of all further musical annotations (Section 7). Finally, we sketch potential applications of the WRD (Section 8) and draw our conclusions (Section 9).

2. BACKGROUND AND RELATED WORK

Closely following (Zalkow et al., 2017a), we outline the particular role of the *Ring* within music history. Western opera originated in late 16th-century Florence and evolved as an important musical genre (Brown et al., 2001). In the context of theories about ancient Greek drama, accompanied singing (denoted as *monody*) attained a central role, thus providing the basis for two central singing styles of traditional opera: speech-like *recitatives* as a compositional device for developing the

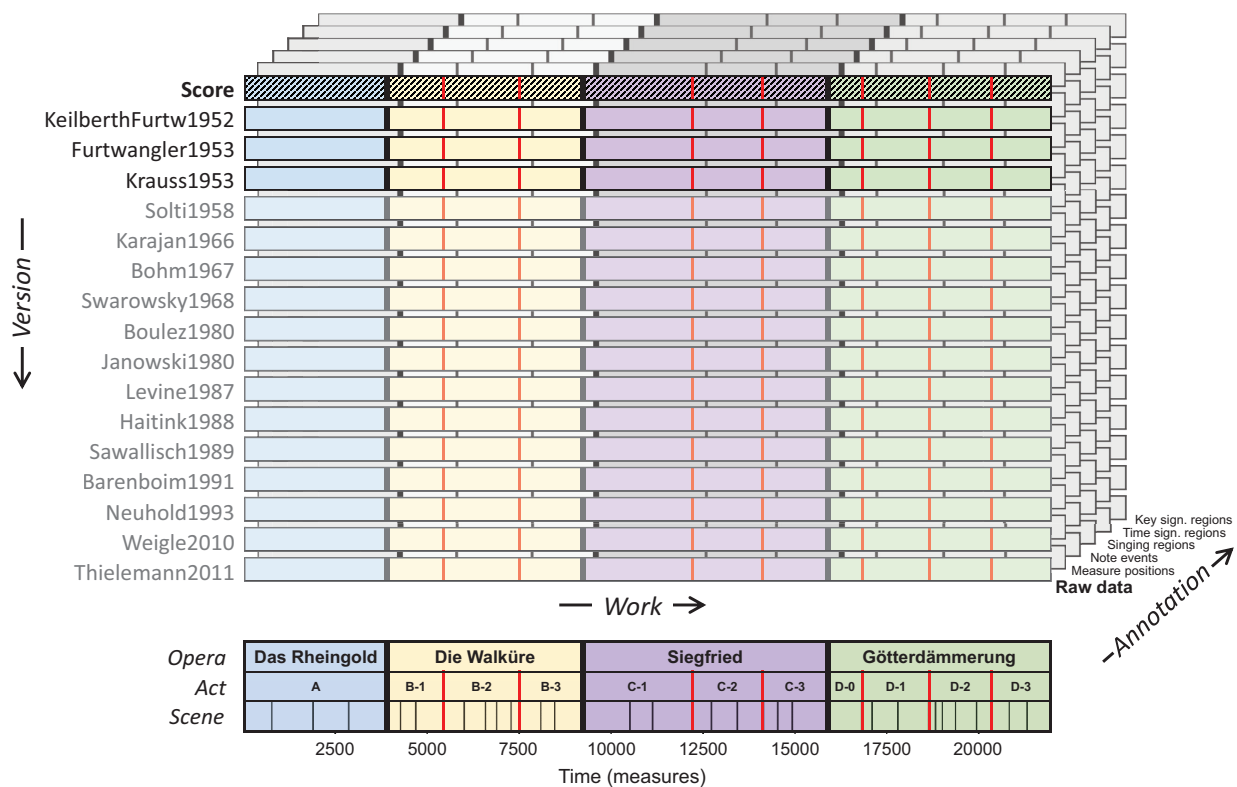


Figure 1 Overview of the WRD. The lower part sketches the *Ring*'s hierarchical structure according to operas, acts, scenes, and measures. We consider the act level as the crucial layer for organizing our data (acts separated by red lines). The cell width is proportional to musical duration given in measures. The upper part illustrates the multitude of versions (score and 16 recorded performances) as well as the available annotations.

plot and cantable *arias* for expressing emotions. For a long time, opera compositions were structured into a series of such individual movements (*number opera*).

In the middle of the 19th century, Richard Wagner proposed a novel approach to operatic composition, theoretically outlined in his writings such as *Opera and Drama* (Wagner, 1852). According to his ideas, the drama of the future should integrate all forms of art (*Gesamtkunstwerk*). The concept overcame the conventions of traditional number opera, favoring a steady musical flow that is often referred to as *through-composed style* or *endless melody* since it dispenses with interruptions and exact repetitions. Central to Wagner's operas is the extensive use of leitmotifs—short musical ideas associated with, e.g., characters, emotions, or items.

Among the most impressive realizations of these ideas is the tetralogy *Der Ring des Nibelungen*, an extensive work cycle of four music dramas (created 1848–1874). Table 1 shows an overview of the *Ring*'s parts. The novel-like plot about the theft of the Rhinegold and the building of Valhalla is the vast background for the fate of individual characters such as Siegmund, Sieglinde, and Brünnhilde. Musically, there are great symphonic compositions like the *Ride of the Valkyries* and *Siegfried's Funeral March*, but also intimate moments like the violoncello solo in the first scene of *Die Walküre*

illustrating the glance between Sieglinde and Siegmund. Both the libretto and the composition span a dense mesh of relationships across the four operas, eleven acts, and 21939 measures,⁴ realized (among others) by the excessive usage of leitmotifs.

Because of its outstanding shape and conception, the *Ring* has inspired numerous composers and artists with a remarkable influence on the further evolution of genres such as opera, symphony, musical, and film music—the sound of many Hollywood blockbusters was obviously inspired by Wagner and his *Ring*. For these reasons, it has become an interesting scenario for MIR research. Concerning opera in general, a few datasets have been published on shorter works such as Mozart's *Zauberflöte*⁵ or Weber's *Freischütz*.⁶ MIR studies dealing with such earlier operas like *Don Giovanni* (Brazier and Widmer, 2021) or *Freischütz* (Röwenstrunk et al., 2015; Prätzlich and Müller, 2013) face challenges due to performance differences of the number operas, which lead to structural mismatches posing problems for music synchronization and tracking. Thanks to the *through-composed style*, the *Ring* is widely free of such structural problems but comes with other challenges. In particular, the search (Kornstädt, 2001), classification (Krause et al., 2020), and detection (Krause et al., 2021b) of leitmotifs with computational methods have been approached. Moreover, the perception of leitmotifs (Müllensiefen et

Catalogue No.	Opera	OperaID	Act	ActID	Full WorkID	ShortID	# Measures
WWV 86 A	<i>Das Rheingold</i>	WWV086A	-	-	Wagner_WWV086A	A	3897
WWV 86 B	<i>Die Walküre</i>	WWV086B	Act 1	1	Wagner_WWV086B-1	B1	1523
			Act 2	2	Wagner_WWV086B-2	B2	2065
			Act 3	3	Wagner_WWV086B-3	B3	1732
WWV 86 C	<i>Siegfried</i>	WWV086C	Act 1	1	Wagner_WWV086C-1	C1	2983
			Act 2	2	Wagner_WWV086C-2	C2	1910
			Act 3	3	Wagner_WWV086C-3	C3	1789
WWV 86 D	<i>Götterdämmerung</i>	WWV086D	Prologue	0	Wagner_WWV086D-0	D0	892
			Act 1	1	Wagner_WWV086D-1	D1	1844
			Act 2	2	Wagner_WWV086D-2	D2	1704
			Act 3	3	Wagner_WWV086D-3	D3	1600
WWV 86	-FULL RING-	WWV086	-	-	Wagner_WWV086		21939

Table 1 Organization of Richard Wagner's *Ring* cycle introducing our naming convention.

al., 2016; Rindfleisch, 2016) and the overall listening experience have been studied with computational tools (Page et al., 2015).

While these studies are quite ambitious and approach high-level musicological problems, it has been observed that much simpler music processing tasks such as the detection (Mimilakis et al., 2019; Krause et al., 2021a) and classification (Krause and Müller, 2022) of singing voice activity or the manual specification of measure boundaries (Weiß et al., 2016; Zalkow et al., 2017b) can become intricate within the complex scenario of the *Ring*. Considering the recent progress in MIR thanks to deep-learning approaches, there is also a great potential in utilizing the complex data and tasks of the *Ring* for training neural networks. This has been approached in a recent study towards music transcription and pitch-class estimation (see Weiß et al., 2021). To enable such endeavors for the MIR community, we release the WRD comprising raw data on the *Ring* (scores and audio—as far as possible within the limits of copyright restrictions) together with comprehensive annotations regarding various musical aspects.

3. DATA ORGANIZATION

We organize the WRD by considering three main types of material, reflected by the respective folders in the Zenodo archive:

- 01_RawData/
- 02_Annotations/
- 03_ExtraMaterial/

While the third folder contains additional information and useful scripts, the first two contain the source data and annotations of the *Ring* music, respectively. We consider the full score of the complete edition⁷ as our reference

representation and rely on this edition for navigating and structuring the WRD. Vocal scores (Wagner, 2014) based on this complete edition exactly correspond to the full score regarding measure counting, libretto text, stage directions, and other relevant information. For this reason, we use these vocal scores to derive various annotations. As stated above, the *Ring* is not a number opera but consists of full acts without musical interruptions.⁸ Inspired by this convention, we choose the *act* as the crucial layer for organizing the WRD (see Table 1), where we consider an act to be a region of contiguous measure counting starting by measure 1.000 (or, for instance, 0.750 for a quarter pickup measure in 4/4 time). Consequently, we treat *Das Rheingold* as a single act since measures are counted throughout the full opera. For *Götterdämmerung*, we treat the prologue as a separate act since the first act starts again with measure number 1.000 according to the Schott edition. All other preludes are integrated into the respective act and thus not considered as separate acts. We organize all files according to these acts (e.g., by splitting and combining CD tracks), leading to eleven files for each representation (modality, performance, or annotation), respectively.

3.1 NAMING CONVENTIONS FOR SCORE-RELATED DATA

To reference these acts, we follow the official catalogue (Wagner Werkeverzeichnis, WWV), where the *Ring* is assigned to the single opus number WWV 86, with additional letters (A, B, C, D) identifying the four operas, which Wagner called a *Vorabend* (pre-evening) and three days. From this, we derive an OperaID, an ActID, and a ShortID as specified in Table 1. For example, `Wagner_WWV086A` denotes *Das Rheingold*, `Wagner_WWV086B-3` refers to the third act of *Die Walküre*, and `Wagner_WWV086D-0` addresses the prologue of *Götterdämmerung*.⁹

All score-related data—i.e., sheet music in any modality and all types of annotations referring to the

composition (and not to a specific performance)—are structured according to the eleven acts and obtain as their filename the full WorkID with an appropriate file extension. The type of modality or annotation is given by the name of the containing folder (e.g., `01_RawData/score_midi/Wagner_WWV086B-1.mid` for a MIDI version of the first act of *Die Walküre*).

3.2 NAMING CONVENTIONS FOR PERFORMANCE-RELATED DATA

For the audio data (recorded performances), we use the same scheme and add a PerformanceID to denote the specific interpretation (CD release), along with a numerical index for short-referencing (see Table 2). Since the operas of a *Ring* release were often recorded over several years, we use the first opera's recording year along with the conductor's last name as a semantic identifier. We keep the PerformanceID constant over the *Ring*'s operas (also in case they are recorded later) in order to have a unique specification of each full *Ring* release. In total, the WRD covers 16 such *Ring* releases.¹⁰

The filenames are derived from the score-related IDs extended with the PerformanceID. For example, `Wagner_WWV086C-2_Haitink1988.wav` denotes the performance of the second act of *Siegfried* recorded by the Symphonieorchester des Bayerischen Rundfunks under Bernard Haitink in 1988. In accordance with the score-related data, all audio-related data and annotations obtain the same filenames (with different file extensions) with the type of data or annotation given by the containing folder's name.

Along with these semantic identifiers, we use a short reference format derived from the composition-related ShortID (Table 1) and the performance number. For example, B1-03 refers to the first act of *Die Walküre* in the performance No. 03 (Krauss1953).

4. SCORE DATA: PIANO REDUCTION

We now explain the data in the WRD in more detail, starting with the different score representations. In general, the full orchestral score as written by the composer is considered to be the most important source since it captures all musical aspects of the composition. The *Ring* involves a huge orchestra including more than 45 instrumental parts (>100 musicians) plus more than 30 vocal parts. This results in a rich and voluminous score, whose orchestration (selection of staff systems) constantly changes. This makes the transformation of a printed score into a machine-readable (symbolic) format extremely labor-intensive. In addition, such a large and complex score lacks compactness and clarity, thus being hard to read for users without specific training in music reading.

For these reasons, we considered a piano reduction of the score, where all instrumental and vocal parts (with

some necessary omissions) are condensed into a single grand staff to be playable by a pianist.¹¹ For the WRD, we chose the public-domain piano reduction by German composer and pianist Richard Kleinmichel (1846–1901), available in a PDF version at the International Music Score Library Project,¹² which we include in the WRD in the folder `01_RawData/score_pdf_IMSLP/`. For orientation and reference, we manually entered measure numbers for every fifth measure into the PDF, where the measure counting follows the full score mentioned above (Schott).

As an example, Figure 2a shows a passage of this score. In measures 958–963, the upper staff follows the singer (here *Sieglinde*) while the lower staff focuses on the orchestral accompaniment. From measure 963 on, the singing melody is not included since both systems follow the orchestral motifs (here *Valhalla* motif). Due to the complexity of the score, such reductions and omissions are unavoidable. However, the most important parts of the musical texture and, in particular, all harmonic information is present, which makes the representation useful for many analysis and processing tasks (e.g., music synchronization or harmony analysis). Moreover, this compact score served as a reference for the musical timeline underlying all annotations.

As the next step, we converted the PDF version (a scan of the print) into a symbolic, machine-readable format. To this end, we processed the score images with an optical music recognition (OMR) tool, using the commercial software PhotoScore.¹³ Due to the complexity of sheet music and its variety of symbols, OMR systems are generally error-prone, which is particularly the case for the Kleinmichel piano reduction having a suboptimal image quality along with a high local notation complexity (sometimes chords with five or more notes in one staff system along with many textual markings). For this reason, we needed to correct the OMR output in a labor-intensive manual process (three student assistants with musicological background working for a total of roughly 1500 hours) using the commercial notation software Sibelius.¹⁴ Due to the amount and severity of OMR errors, re-entering whole measures was sometimes easier than correcting the OMR output. The resulting data can be found in the folder `01_RawData/score_sibelius/`. Figure 2b shows our processed and corrected version for the same passage.

Finally, we exported the Sibelius version of the Kleinmichel score to various formats including

- PDF (`score_pdf_sibelius/`),
- MusicXML (`score_musicxml/`),
- MIDI (`score_midi/`).

For the MIDI export, we aimed for a neutral, uninterpreted version using constant tempo and no modeling of expression. A piano roll visualization of the MIDI

No.	PerformanceID	Conductor	Year(s)	Orchestra	Label	MusicBrainz ReleaseID	Duration
01	KeilberthFurtw1952	J. Keilberth / W. Furtwängler	1952-54	Bayreuther Festspiele	ZYX	7955db62-60d2-4c1f-8f4d-2a8ac69ff104†	14:19:56
02	Furtwängler1953	W. Furtwängler	1953	Orch. Sinfonica Roma RAI	EMI Rec.	66ebc811-49b3-4de2-b6eb-1f56f0995c29	15:04:22
03	Krauss1953	C. Krauss	1953	Bayreuther Festspiele	Orfeo	c988df46-9359-4148-836b-4c67cc5e280e	14:12:27
04	Solti1958	G. Solti	1958-65	Wiener Philharmoniker	Decca Rec.	3b6b44dd-2779-4166-9640-8cb458e7cfdc	14:36:58
05	Karajan1966	H. v. Karajan	1966-70	Berliner Philharmoniker	Dt. Gram.	2f165cc6-bda7-4e07-b4da-505f8f1d4bd7	14:58:08
06	Bohm1967	K. Böhm	1967-71	Bayreuther Festspiele	Decca Cl.	e0d1091d-fe3d-4253-a025-6f3c85db95e5	13:39:28
07	Swarowsky1968	H. Swarowsky	1968	Grosses Symph.-Orch.	Profil	9b832029-8bc1-3726-8ef6-a33411fec63a	14:56:34
08	Boulez1980	P. Boulez	1980-81	Bayreuther Festspiele	Philips Cl.	51816ee1-07e9-40b1-ae82-5faa4dc85367	13:44:38
09	Janowski1980	M. Janowski	1980-83	Staatskapelle Dresden	Sony Cl.	3a08830e-6a2d-4bce-b880-68bcc3d6ef32	14:08:34
10	Levine1987	J. Levine	1987-89	Metropolitan Opera Orch.	Dt. Gram.	e8fc9fe9-374d-31bf-8da8-29cb524507b5	15:21:52
11	Haitink1988	B. Haitink	1988-91	Symph.-Orch. Bayer. Rundf.	EMI Cl.	635c8961-10f0-4960-af11-c70768d449c6	14:27:10
12	Sawallisch1989	W. Sawallisch	1989	Bayer. Staatsorchester	EMI Cl.	058c6660-b1c8-4b33-8b7c-9ff6f28eb5120	14:06:50
13	Barenboim1991	D. Barenboim	1991-92	Staatskapelle Berlin	Warner Cl.	908ecf81-20e9-4c87-ba15-7c746d72743f	14:54:55
14	Neuhold1993	G. Neuhold	1993-95	Badische Staatskapelle	Brilliant Cl.	468d27f6-b1c2-4c4e-be5e-591703800a60	14:04:35
15	Weigle2010	S. Weigle	2010-12	Frankf. O. u. M.-Orchester	OEHMS	872e56a1-e7aa-47cf-b2a8-12c1f8a21f39	14:48:46
16	Thielemann2011	C. Thielemann	2011	Wiener Staatsoper	Dt. Gram.	7ed1935c-ab07-499f-999d-edc6ef5cbe83	14:31:13

Table 2 Performances of the Ring contained in the WRD. In performance 01, W. Furtwängler only conducts *Die Walküre* (which differs from performance 02), the other parts are conducted by J. Keilberth. † For performance 01 (KeilberthFurtw1952), we found no ReleaseID for the full Ring in MusicBrainz but instead use the ReleaseID of *Die Walküre*.

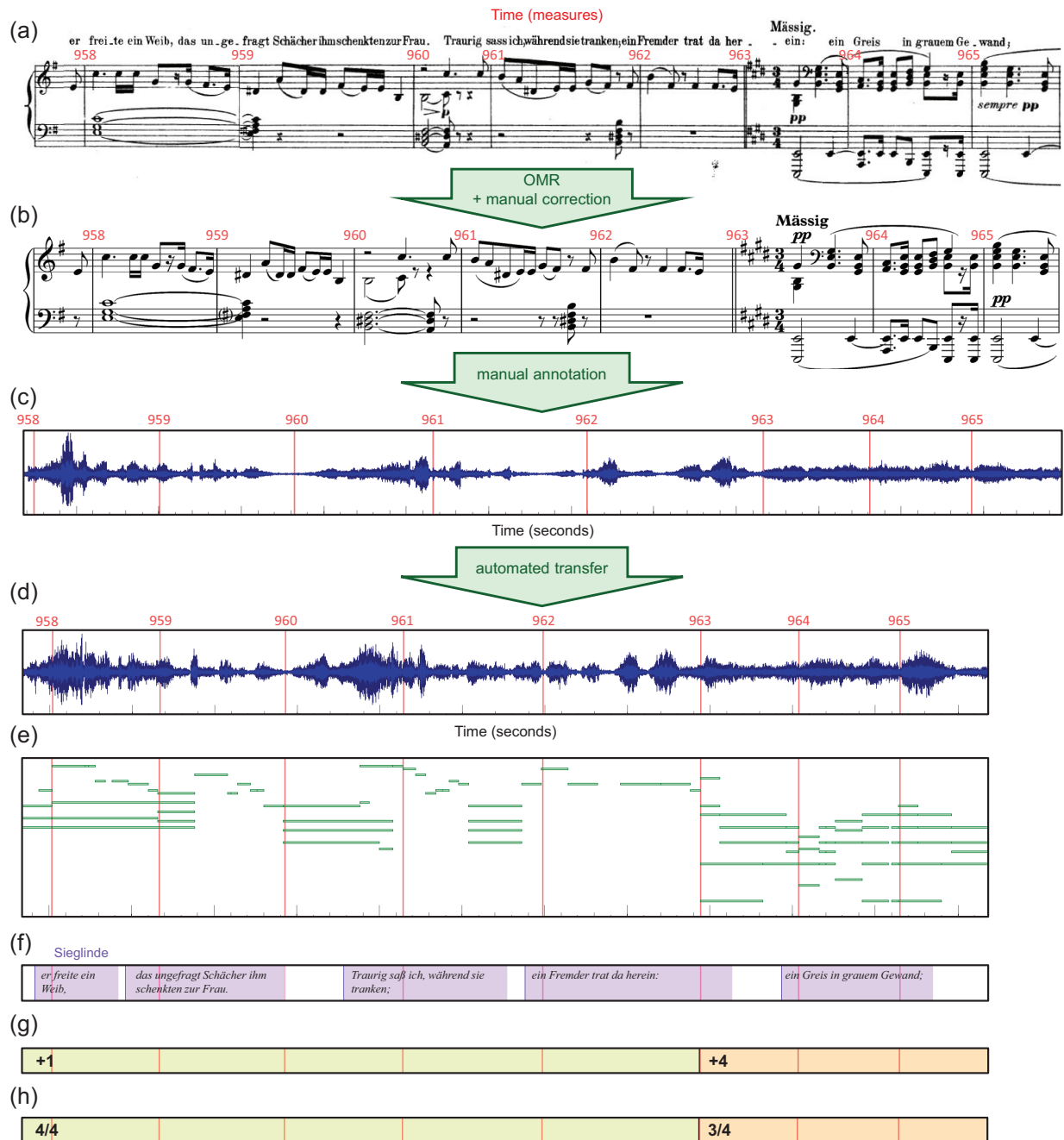


Figure 2 Example passage from *Die Walküre*, WWV086B-1, measures 958-965. (a) Scan of the piano reduction by R. Kleinmichel (image). (b) OMR-processed and corrected score (symbolic). (c) Performance Karajan1966 (waveform) with manual measure annotations. (d) Performance Krauss1953 with automatically transferred measure annotations. (e) Note events based on the piano reduction (MIDI) synchronized to Krauss1953. (f) Singing voice regions with libretto (text). (g) Key signature regions (see Section 7.3). (h) Time signature regions.

representations is shown in Figure 2e (here time-aligned to the performance Krauss1953).

5. AUDIO DATA: RECORDED PERFORMANCES

Beyond the musical score, the WRD contains 16 releases (audio data) of the full *Ring* (see Table 2 for an overview), which sum up to a total duration of 231 hours and 56

minutes of audio. These performances were recorded either in a studio or a live setting (including four live recordings from the Bayreuther Festspiele). Some of the releases capture a performance on four consecutive days, others consist of individual performances of the four operas (by the same orchestras and conductors) recorded over a time span of several years (e.g., the release Bohm1967 was recorded between 1967 and 1971), one even involving two different conductors. For the sake of consistency and due to the intended coherence

of a release, we still refer to each full release as a “performance,” addressed with a unique PerformanceID (see Table 2).

We notice a remarkable range of durations: the shortest performance (Boulez1980) lasts for less than 13:45 hours while the longest one (Levine1987) almost reaches 15:30 hours. These tempo deviations are already a first indicator for some considerable differences in interpretation. Beyond that, the recordings substantially differ in sound quality and recording characteristics, balance between orchestra and singers as well as playing and singing styles. The WRD enables to study such performance aspects in further detail.

All performances are professional recordings, commercially released by major music labels. For this reason, most of the audio material is under copyright protection in most countries (*Leistungsschutzrechte* or performer rights) though Wagner’s composition itself is in the public domain. Due to the copyright regulations in Germany and the EU, performances from the early 1950s are not protected anymore in these countries. For this reason, we publish the audio data for three of the 16 performances.¹⁵ All other recordings are commercially available and can be identified with the MusicBrainz IDs given in Table 2.

The releases are usually available as boxed sets of audio CDs, typically divided into tracks of several minutes duration. Since track splits are not consistent across versions, we concatenate all audio tracks belonging to one of the eleven acts (compare Figure 1 and Section 3) into a single large audio file (one for each line in Table 1). We provide the three public performances as stereo wave files with a sampling rate of 22 050 Hz. For making the commercial recordings compatible to the WRD (avoiding temporal offsets), we provide a Python script along with the dataset (`03_ExtraMaterial/merge_ripped_audio.py`) in order to merge the tracks. With this procedure, the remaining 13 performances can be acquired and processed at a reasonable expense to complete the WRD’s raw material.

6. MEASURE POSITIONS AND SYNCHRONIZATION

The act-wise view ensures the correspondence of our *Ring* versions on a coarse level. To locally compare those versions and to enable the transfer of semantic annotations from one version to another, we require fine-grained mutual alignments. Since fully-automated music synchronization procedures reach their limits when dealing with long and complex works (Prätzlich et al., 2016), we employ a multi-stage, semi-automatic procedure for obtaining our alignments.

Since the eleven parts (acts) of the *Ring* correspond to regions of contiguous measure counting, we first

approach the specification of the musical measure positions in the audio recordings. Such *measure annotations* may fulfill several purposes (Weiß et al., 2016). They facilitate navigation and segmentation using musically meaningful regions such as passages or scenes. Beyond that, they enable the transfer of annotations or analysis results from one domain to the other as well as a full *cross-version analysis*, which exploits the relation between different performances to stabilize analysis results (Konz et al., 2013). We provide the measure annotations as CSV files indicating for each score position (in measures) the corresponding physical time position (in seconds) in the respective recording.

6.1 MANUAL MEASURE ANNOTATIONS

To obtain a reliable reference, we first approached the task of annotating measure positions in a manual fashion considering the following three performances:

- 05 - Karajan1966
- 11 - Haitink1988
- 13 - Barenboim1991

In Weiß et al. (2016), we describe this process in detail. While following the vocal score (see Section 3) as reference, the annotators listened to the recordings and marked the measure positions using the public software *Sonic Visualizer* (Cannam et al., 2006). After finishing a certain passage, the annotators corrected erroneous or inaccurate measure positions. The length of these passages, the tolerance of errors, and the overall duration of the annotation process differed between the annotators. Roughly three hours per annotator were necessary to annotate one hour of music. The annotators also marked ambiguous measure positions, mentioning tempo changes, fermatas, tied notes over barlines, or very fast passages as musical reasons for such ambiguities. They further reported performance-related problems such as asynchronicities between orchestra and singers or masking of onsets through prominent other sounds.

To analyze these manual annotations in detail, we conducted an experiment on the act Wagner_wwv086B-1 in the performance Karajan1966 (B1-05), studying the cross-annotator consistency for five different annotators. We observed typical differences in the order of 0.1 seconds—a remarkable value that may be caused by the complexity of opera performances and the reasons discussed above and is roughly in line with the findings by Gadermaier and Widmer (2019) for a Bruckner symphony movement (annotation differences up to 0.07 seconds). Due to these large differences, we then followed a two-annotator strategy for each act and performance (of the three releases listed above). We then merged the annotations within a third annotation round: for all measures where the standard deviation between annotators was lower than 0.2 (corresponding

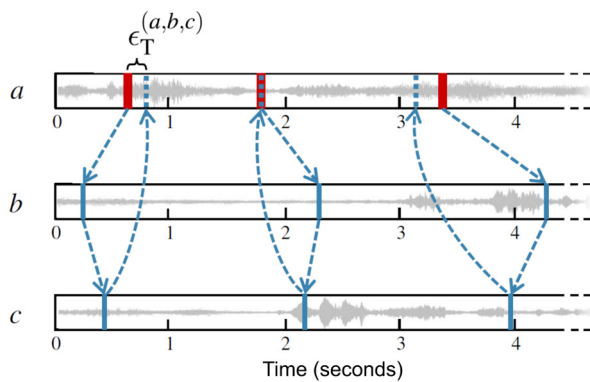


Figure 3 Triple-based transfer of measure annotations between performances, derived from Zalkow et al. (2017b).

to 0.28 seconds offset for two annotators), we took the mean position over all annotators. For all other measures, we manually checked and decided on the final position (either following one annotator or placing the annotation somewhere in between).

In summary, the 21 939 measure positions in the *Ring* were manually annotated at least twice for the three performances listed above. This resulted in more than 135 000 initial markings (before merging and refinement) and a total annotation time of roughly 1000 hours including training and correction, conducted by ten student assistants over the course of one year. The manual measure annotations themselves already provide an interesting source for studying performance aspects or investigating music synchronization procedures. Moreover, separate research on the annotation process itself emerged from this endeavour (Weiß et al., 2016).

6.2 AUTOMATED TRANSFER OF MEASURE ANNOTATIONS

With the procedure described above, we obtained reliable and verified annotations (up to the level of musical ambiguities described above) for three full *Ring* releases. In order to obtain measure positions for the other 13 releases, we made use of a highly sophisticated music synchronization procedure described by Zalkow et al. (2017b).

The basis of this procedure is a multi-rate and multi-scale variant of dynamic time warping (Prätzlich et al., 2016), which is capable of handling our long audio files (>1 hour duration). Zalkow et al. (2017b) then use a late-fusion approach that improves the measure transfer when having several annotated performances at hand. This approach makes use of performance triples (a, b, c) that are synchronized with each other in a pairwise fashion (see Figure 3). Transferring measure positions from one performance to the other using the computed alignments allows for computing the *triple error* $\epsilon_T^{(a,b,c)}$ as proposed by Prätzlich and Müller (2016) for each measure position. We then use this triple error as an optimization criterion to improve the measure transfer to the target

performance in a late-fusion approach (exploiting the availability of several annotated performances). Zalkow et al. (2017b) showed that this procedure is superior to other fusion strategies. They also provide more detailed information on the level of error in the transferred annotations, which appeared to be less than 0.2 seconds for around 80% of the measures.

We apply this strategy to all 13 performances that were not manually annotated. All measure annotations can be found in the folder `02_Annotations/ann_audio_measure/`.

6.3 FINE-GRAINED ALIGNMENTS

In most situations, knowing the position of all measures is sufficient for navigating and switching between recordings or for transferring coarse-level annotations between performances—not least because important musical events often happen at the beginning of a measure (downbeat). Nevertheless, to transfer some of our annotations, a fine-grained alignment within each measure is necessary. For such cases, one can add another stage to the alignment procedure. Starting from the (reliable) measure annotations, a first approximation would be to assume constant tempo within a measure and interpolate musical time (beats) in a linear fashion.

Since the *Ring*—as with other music of its time—is usually performed with substantial local tempo fluctuations (agogics, *ritardando*, or *rubato*), linear interpolation may sometimes not be precise enough, e.g., for transferring note event information (see Section 7.1). For those cases, another possibility is to locally perform score-to-audio alignment within each measure, using for instance a chroma-based approach (Müller et al., 2021) relying on the score-based MIDI representation (see Section 4 and Figure 2e).

7. FURTHER ANNOTATIONS

As indicated by Figure 1, the WRD comprises a variety of additional musical annotations besides the measure positions discussed above. Those annotations describe the *Ring*'s music regarding musical concepts on various time scales—from individual events (notes or words) up to long structural segments (scene or key regions).

For all of these annotations, we pursue the following strategy: first, we collect the relevant information with respect to the composition, which involves temporal markings based on the musical time axis of the score (given in measures). Partial measures are hereby addressed with fractional numbers of three digits precision.¹⁶ Second, we transfer the score-related annotations to the physical time axis of the individual performances (given in seconds) using the measure annotations described in Section 6. Partial measure positions are transferred via

linear interpolation or local score–audio synchronization (compare Section 6.3).

7.1 NOTE EVENTS

As the most fine-grained musical description, we consider annotations of individual note events. Such note annotations are necessary for developing and evaluating algorithms in the field of automatic music transcription (Benetos et al., 2019), but can also be used to derive other information such as note onsets, piano rolls (multi-pitch), or pitch-class information. We derive these note events from the machine-readable score (Sibelius) exported to the MIDI format (see Section 4). We then transfer these events to the performances using our measure annotations as anchor points and performing score–audio synchronization within each measure (Müller et al., 2021). The annotations capture for each event a start and end position together with a pitch and a pitch-class (pitch modulo 12) label. The note annotations can be found in `02_Annotations/ann_audio_note/`.

It is important to emphasize that these note annotations were generated on the basis of the piano reduction (compare Section 4). The performances (audio), however, correspond to the full orchestral score (also regarding pitch content). While aiming at preserving the relevant harmonic and melodic content of the full score, the piano reduction introduces significant differences to make it playable by the two hands of a pianist. These differences mainly affect the pitches' octaves (which are reduced to a smaller range) but may also include aspects of texture and figuration, e.g., by leaving out some parts of the orchestra to reduce the complexity. For this reason, the pitch content of the piano score does not exactly correspond to the full score, and this difference might affect training or evaluation with the note annotations. Since many differences concern the pitches' octaves, this problem plays a considerably smaller role when deriving *pitch-class* annotations (chroma) from the note events. Such pitch-class annotations have been used by Weiß et al. (2021) to train and evaluate deep-learning approaches for extracting pitch-class information from audio.

7.2 SINGING VOICE REGIONS

As another type of annotation, we consider the singers' activities and lyrics. Since there are minor differences between Wagner's initial text and its realization in the score, we stick to the latter one. For the WRD, we derive annotations comprising the start and end position for each singing voice region, given either in partial measures (for `02_Annotations/ann_score_singing/`) or in seconds (for `02_Annotations/ann_audio_singing/`). Beyond that, we specify for each region the person or role singing (such as *Sieglinde*, *Siegfried*, or *Wotan*) as well as the lyrics (sung text) following the exact notation of the complete edition (score).

To generate these annotations, we started from the libretto's phrase segments, manually annotated the region boundaries as given by the score, and refined the text to be consistent with the score (Schott). As specified above, we then use our measure annotations with linear interpolation (Section 6.3) for obtaining physical time positions for each performance. The accuracy of the resulting annotations depends, on the one hand, on the accuracy of the measure annotations, which have typical deviations in the order of 0.1 seconds for the manual measure annotations (Weiß et al., 2016) and 0.2 seconds for the transferred measure annotations (Zalkow et al., 2017b). On the other hand, it might be affected by non-linear tempo deviations within a measure.

Despite these minor inaccuracies, which have been discussed by Krause et al. (2021a), the singing voice annotations constitute a valuable source for a variety of MIR tasks such as singing voice detection (Mimilakis et al., 2019; Krause et al., 2021a), classification of singer gender, register, and individual singers (Krause and Müller, 2022), lyrics alignment (Stoller et al., 2019), or lyrics transcription (Gao et al., 2022). Moreover, together with the other modalities, they allow for studying relationships between the musical and dramatic aspects of the *Ring* (Zalkow et al., 2017a).

7.3 TIME AND KEY SIGNATURE REGIONS

Beyond the fine-grained annotations on the note and phrase level, we provide annotations for time signature changes (indicating new metrical and rhythmic style) and key signature changes (pointing to a key change or another major harmonic change). To extract this information, we process the symbolic version of the piano-reduced score (MusicXML) with `music21`. The resulting annotations indicate the temporal position of the respective boundary, given either in measures (e.g., `02_Annotations/ann_score_keysignatures/`) or in seconds (e.g., `ann_audio_keysignatures/`).¹⁷ In addition, each boundary is labelled with the new time signature (e.g., 3/4) or key signature (e.g., `02_Annotations/ann_audio_keysignatures/ +2` for two sharps (#) or `-3` for three flats (b)). Figure 2g,h shows an example where at measure 963, the key signature changes (from +1 to +4) as well as the time signature (from 4/4 time to 3/4 time).

7.4 SCENE REGIONS

As a musically meaningful structural segmentation, we consider scene annotations for the WRD. As for the singing voice regions (Section 7.2), we derive the score-related annotations from the full score (complete edition by Schott). In some cases, the prelude and the first scene build a unit, which we then consider as a single joint scene.

As for the other region annotations, we specify start and end of a scene in measures or seconds, respectively. We label the scene numbers with integers (e.g., 2 for

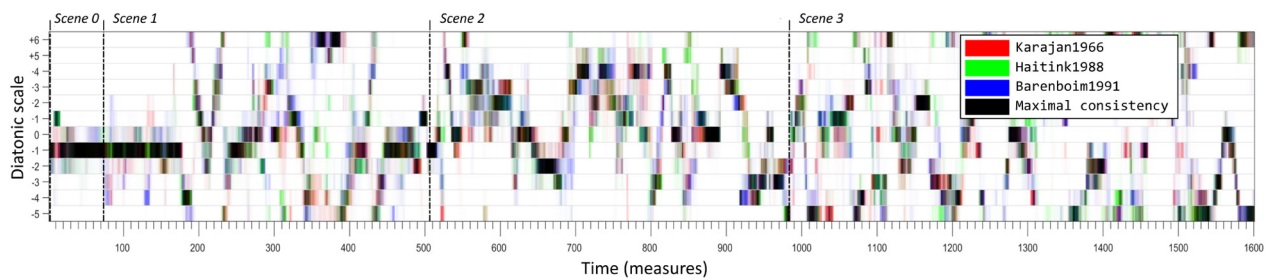


Figure 4 Cross-version harmonic analysis regarding diatonic scales (Weiß et al., 2017). The analysis result is based on three versions, visualized with different colors that add up to black in case they are consistent.

Zweite Szene) and denote a prelude (*Vorspiel*) with 0. Beyond the prelude of *Götterdämmerung*, WWV086D-0, each act has between 3–5 scenes, which considerably vary in length. An overview of the scenes and their duration can be seen in the bottom line of Figure 1.

8. APPLICATION SCENARIOS

Finally we want to sketch a number of examples how the WRD may be used for research in MIR and computational musicology. Some of these possible use cases have been addressed in previous studies. Already at the level of the raw data, there are a variety of algorithmic tasks that can be addressed based on the WRD. For example, the automatic conversion of graphical into symbolic scores (OMR) can be studied based on our symbolic Kleinmichel score, which constitutes a particularly difficult OMR scenario. The detection of score–audio correspondences (Dorfer et al., 2018) on different levels and modalities (image–audio, MIDI–audio, symbolic–audio) can be approached using the measure annotations as a reliable reference. Moreover, these measure annotations constitute a valuable source for evaluating music synchronization pipelines (score–audio or audio–audio), towards which the study of Zalkow et al. (2017b) constitutes a first step. Using the provided note annotations, interesting and complex problems within the broader field of automatic music transcription can be studied, including multi-pitch and pitch-class estimation (Weiß et al., 2021), onset detection, lyrics alignment (Stoller et al., 2019), or lyrics transcription (Gao et al., 2022) based on the singing voice annotations.

Beyond these low-level tasks, a number of computational music analysis and retrieval problems can be studied based on the WRD. The measure positions provide a reference annotation for studying downbeat estimation (Böck et al., 2016) in an extremely complex scenario. Segmenting the audio according to metrical (time signatures), harmonic (key signatures), and structural properties (scenes) is of high interest to the MIR community and can be studied with the WRD’s rich annotations both in the symbolic and the audio domain. Detecting the presence of singing voice in general (Mimilakis et al., 2019; Krause et al., 2021a) or with finer

discrimination regarding gender, register, or individual identity (Krause and Müller, 2022) are further interesting and challenging problems. For all these tasks, the availability of the different versions (scores and recorded performances) can be exploited to systematically study robustness and generalization of machine-learning systems, which has been considered for singing voice detection by Krause et al. (2021a).

Besides being a valuable testbed for various MIR tasks, the WRD is of particular interest for studying the *Ring* as an example of a highly complex romantic composition from a musicological perspective. Of major interest are aspects of harmony, rhythm, and form as well as their interplay with the plot and the dramatic conception of the work cycle—i.e., the libretto, which itself is worth being investigated with approaches from natural language processing. While the provided score data allows for approaching analyses in the symbolic domain, also audio-based strategies have been applied for the large-scale analysis of the *Ring*’s harmonic structure. In Weiß et al. (2017), such an analysis is demonstrated by visualizing pitch content with respect to the twelve diatonic scales over the course of the *Ring*. Figure 4 provides an example of such a visualization for the third act of *Götterdämmerung* (WWV086D-3). Here, a visual cross-version analysis allows for studying differences (in color) and consistencies (black) between individual analysis results. This is possible by performing all analyses based on a musical time axis (using our measure annotations), which also allows for a direct interaction with the scores. The visualization is enriched with scene boundary annotations and highlights stable phases (measures 1–180), structural breaks (at measure 920), and interesting modulations (measures 1520–1600). Visualizing such analysis results in a soft but interpretable way rather than performing predictions of discrete labels (e.g., for local keys) constitutes an alternative approach for computational musicology, which might yield valuable insights into complex works such as the *Ring*.

Finally, the availability of the 16 recorded performances not only allows for comparing and stabilizing analyses of the work (composition), but is also of great potential for studying aspects of the performances. The enormous tempo differences highlighted by the varying duration of

the *Ring* releases (Table 2) are already a first indication of interesting findings on musical interpretation. The measure annotations are obviously another valuable source for studying tempo variations in more detail, also in relation with important structural boundaries (scenes, key, or time signatures). Beyond that, analyses of singing style such as vibrato characteristics (Driedger et al., 2016) or the sound characteristics of different orchestras can be approached based on the WRD.

9. CONCLUSIONS

With this paper, we introduced the Wagner Ring Dataset (WRD), a multi-modal and multi-version annotated dataset built on Wagner's tetralogy *Der Ring des Nibelungen*. Our strategy for organizing this large-scale work relies on the eleven acts as the main structural principle. Concerning the raw data (sources), the WRD includes score representations in various modalities and 16 recorded performances (three of them publicly available). For all versions (score and audio), we provide comprehensive annotations of measure positions, note events, singing voice regions including singers and lyrics as well as scene, key, and time signature regions. The estimated time required to create the annotations was approximately 4000 hours (1500 hours for processing the symbolic score, 1000 hours of measure annotations, 1500 hours of further annotations, corrections, and processing steps) excluding the actual research work with the data. In this paper, we described the data curation, processing, correction, and annotation in detail. Moreover, we sketched several application scenarios, which demonstrate the potential of the WRD as a complex opera scenario for research in music processing, MIR, and computational musicology.

NOTES

- 1 Computer-Assisted Analysis of Harmonic Structures (CAS), funded by the German Research Foundation (DFG MU 2686/7-1 and 7-2, DFG KL 864/4-1 and 4-2) over two phases (2014–2018 and 2019–2024).
- 2 All other recordings are commercially available as boxed CD sets. We provide scripts to process the CD tracks appropriately to match our annotations, which are available for all 16 performances.
- 3 https://zenodo.org/record/7672157#.Y_fxDXbMJJaR.
- 4 21 941 when including pickup measures into the counting.
- 5 <https://zenodo.org/record/7504097>.
- 6 <https://www.audiolabs-erlangen.de/resources/MIR/FreiDi/MultitrackDataset>.
- 7 Schott Music, Mainz, 1970–2020. Please see <https://www.schott-music.com/en/musicology/complete-editions/wagner.html> for the publications of this edition.
- 8 When performing the *Ring*, there is usually an intermission between acts, except for prologue and first act of *Götterdämmerung*.
- 9 Please note that for organizing the WRD, the scenes (bottom of Figure 1) do not play a role since we can easily address these scenes as measure regions, which we describe in the context of our annotations (Section 7.4).
- 10 Due to copyright restrictions, we publish audio recordings for three of the 16 performances and refer to the commercial releases for the others. For details, see Section 5.

- 11 Please note that in contrast to the vocal score (*Klavierauszug*), where only the orchestral part is condensed and the vocal parts are unfolded, a piano reduction (*Klavierreduktion*) also incorporates the singers' pitches into the grand staff, with lyrics being loosely indicated above the staves.
- 12 <https://imslp.org/wiki/Special:IMSLPImageHandler/33839%2Fqrol> for example (piano reduction of *Das Rheingold*).
- 13 Avid PhotoScore, version 8.
- 14 Avid Sibelius, version 8.
- 15 Please note that these recordings might not be in the Public Domain in other countries outside the EU and Switzerland.
- 16 Note that this does not allow for deriving metrical information such as the time signature. As an advantage, however, such metrical information is not needed since we numerically consider each measure to have a length of 1. This allows to simultaneously specify positions in different time signatures. E. g., the second (quarter) beat in 3/4 time obtains the same fraction (0.333) as the fourth (eighth note) beat in 9/8 time. This becomes important for the final part of *Götterdämmerung* due to its polymetrical structure.
- 17 Please note that those boundaries are always at measure boundaries such that no interpolation is necessary.

ACKNOWLEDGEMENTS

We deeply thank all team members and student assistants involved in the data curation process and the annotation work. To just name a few of the many contributors, we want to mention Lena Krauß, Cäcilia Marxer, Sarah Schweiger, Sascha Kruchten, Felix Wiethaus, and Peter Haaf. The International Audio Laboratories Erlangen are a joint institution of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Fraunhofer Institute for Integrated Circuits IIS.

FUNDING INFORMATION

This work was supported by the German Research Foundation within the project “Computational Analysis of Harmonic Structures” (DFG MU 2686/7-1 and 7-2, DFG KL 864/4-1 and 4-2).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

Compiling the WRD required a large working process including the preparation and processing of raw data, typesetting, manual corrections, creation and transfer of annotations as well as the application of the data for research, followed by several cycles of revision and correction. All these steps were conducted in a tremendous team effort by a large number of people within the two research groups. Hence, it is hardly possible to disentangle individual contributions to this work. All co-authors of this paper were involved both


in the data curation process and in the preparation of the final manuscript. For further contributions, see our acknowledgements.

AUTHOR AFFILIATIONS

Christof Weiß  orcid.org/0000-0003-2143-4679
University of Würzburg, Germany; International Audio Laboratories Erlangen, Germany


Vlora Arifi-Müller
International Audio Laboratories Erlangen, Germany

Michael Krause  orcid.org/0000-0001-7194-0719
International Audio Laboratories Erlangen, Germany

Frank Zalkow  orcid.org/0000-0003-1383-4541
International Audio Laboratories Erlangen, Germany

Stephanie Klauk
Saarland University, Germany

Rainer Kleinertz
Saarland University, Germany

Meinard Müller  orcid.org/0000-0001-6062-7524
International Audio Laboratories Erlangen, Germany

REFERENCES

- Benetos, E., Dixon, S., Duan, Z., and Ewert, S.** (2019). Automatic music transcription: An overview. *IEEE Signal Processing Magazine*, 36(1):20–30. DOI: <https://doi.org/10.1109/MSP.2018.2869928>
- Böck, S., Krebs, F., and Widmer, G.** (2016). Joint beat and downbeat tracking with recurrent neural networks. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 255–261, New York City, New York, USA.
- Bosch, J. J.** (2013). Automatic melodic and structural analysis of music material for enriched concert related experiences. In *Proceedings of the ACM Multimedia Conference*, pages 1067–1070, Barcelona, Spain. DOI: <https://doi.org/10.1145/2502081.2502218>
- Brazier, C. and Widmer, G.** (2021). Handling structural mismatches in real-time opera tracking. In *Proceedings of the European Signal Processing Conference (EUSIPCO)*, pages 366–370, Dublin, Ireland. DOI: <https://doi.org/10.23919/EUSIPCO54536.2021.9616109>
- Brown, H. M., Rosand, E., Strohm, R., Parker, R., Whittall, A., Savage, R., and Millington, B.** (2001). Opera. In Sadie, S., editor, *The New Grove Dictionary of Music and Musicians*, pages 416–471. Macmillan Publishers, London, 2nd edition. DOI: <https://doi.org/10.1093/gmo/9781561592630.article.40726>
- Calvo-Zaragoza, J., Hajič Jr., J., and Pacha, A.** (2020). Understanding optical music recognition. *ACM Computing Surveys*, 53(4). DOI: <https://doi.org/10.1145/3397499>
- Cannam, C., Landone, C., Sandler, M. B., and Bello, J. P.** (2006). The Sonic Visualiser: A visualisation platform for semantic descriptors from musical signals. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR)*, pages 324–327, Victoria, Canada.
- Dorfer, M., Hajič Jr., J., Arzt, A., Frostel, H., and Widmer, G.** (2018). Learning audio-sheet music correspondences for cross-modal retrieval and piece identification. *Transactions of the International Society for Music Information (TISMIR)*, 1(1):22–31. DOI: <https://doi.org/10.5334/tismir.12>
- Driedger, J., Balke, S., Ewert, S., and Müller, M.** (2016). Template-based vibrato analysis of music signals. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 239–245, New York City, New York, USA.
- Gadermaier, T. and Widmer, G.** (2019). A study of annotation and alignment accuracy for performance comparison in complex orchestral music. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 769–775, Delft, The Netherlands.
- Gao, X., Gupta, C., and Li, H.** (2022). Genre-conditioned acoustic models for automatic lyrics transcription of polyphonic music. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 791–795, Virtual and Singapore. DOI: <https://doi.org/10.1109/ICASSP43922.2022.9747684>
- Hawthorne, C., Stasyuk, A., Roberts, A., Simon, I., Huang, C. A., Dieleman, S., Elsen, E., Engel, J. H., and Eck, D.** (2019). Enabling factorized piano music modeling and generation with the MAESTRO dataset. In *Proceedings of the International Conference on Learning Representations (ICLR)*, New Orleans, Louisiana, USA.
- Hewlett, W. B. and Selfridge-Field, E.** (1991). Computing in musicology, 1966–91. *Computers and the Humanities*, 25(6):381–392. DOI: <https://doi.org/10.1007/BF00141188>
- Konz, V., Müller, M., and Kleinertz, R.** (2013). A cross-version chord labelling approach for exploring harmonic structures—a case study on Beethoven’s Appassionata. *Journal of New Music Research*, 42(1):61–77. DOI: <https://doi.org/10.1080/09298215.2012.750369>
- Kornstädt, A.** (2001). The JRing system for computer-assisted musicological analysis. In *Proceedings of the International Symposium on Music Information Retrieval (ISMIR)*, pages 93–98, Bloomington, Indiana, USA.
- Krause, M. and Müller, M.** (2022). Hierarchical classification of singing activity, gender, and type in complex music recordings. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 406–410, Singapore. DOI: <https://doi.org/10.1109/ICASSP43922.2022.9747690>
- Krause, M., Müller, M., and Weiß, C.** (2021a). Singing voice detection in opera recordings: A case study on robustness and generalization. *Electronics*, 10(10):1214:1–14. DOI: <https://doi.org/10.3390/electronics10101214>
- Krause, M., Müller, M., and Weiß, C.** (2021b). Towards leitmotif activity detection in opera recordings. *Transactions of the International Society for Music Information Retrieval*

- (TISMIR), 4(1):127–140. DOI: <https://doi.org/10.5334/tismir.116>
- Krause, M., Zalkow, F., Zalkow, J., Weiß, C., and Müller, M.** (2020). Classifying leitmotifs in recordings of operas by Richard Wagner. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 473–480, Montréal, Canada.
- Lerch, A., Arthur, C., Pati, A., and Gururani, S.** (2020). An interdisciplinary review of music performance analysis. *Transactions of the International Society for Music Information Retrieval (TISMIR)*, 3(1):221–245. DOI: <https://doi.org/10.5334/tismir.53>
- Meredith, D.**, editor (2016). *Computational Music Analysis*. Springer. DOI: <https://doi.org/10.1007/978-3-319-25931-4>
- Mimilakis, S. I., Weiß, C., Arifi-Müller, V., Abeser, J., and Müller, M.** (2019). Cross-version singing voice detection in opera recordings: Challenges for supervised learning. In *Machine Learning and Knowledge Discovery in Databases – Proceedings of the International Workshops of ECML PKDD 2019, Part II*, volume 1168 of *Communications in Computer and Information Science*, pages 429–436, Würzburg, Germany. DOI: https://doi.org/10.1007/978-3-030-43887-6_35
- Müllensiefen, D., Baker, D., Rhodes, C., Crawford, T., and Dreyfus, L.** (2016). Recognition of leitmotives in Richard Wagner’s music: An item response theory approach. In *Analysis of Large and Complex Data*, pages 473–483. Springer, Cham, Switzerland. DOI: https://doi.org/10.1007/978-3-319-25226-1_40
- Müller, M.** (2021). *Fundamentals of Music Processing – Using Python and Jupyter Notebooks*. Springer, 2nd edition. DOI: <https://doi.org/10.1007/978-3-030-69808-9>
- Müller, M., Arzt, A., Balke, S., Dorfer, M., and Widmer, G.** (2019). Cross-modal music retrieval and applications: An overview of key methodologies. *IEEE Signal Processing Magazine*, 36(1):52–62. DOI: <https://doi.org/10.1109/MSP.2018.2868887>
- Müller, M., Özer, Y., Krause, M., Prätzlich, T., and Driedger, J.** (2021). Sync Toolbox: A Python package for efficient, robust, and accurate music synchronization. *Journal of Open Source Software (JOSS)*, 6(64):3434:1–4. DOI: <https://doi.org/10.21105/joss.03434>
- Nieto, O., Mysore, G. J., Wang, C., Smith, J. B. L., Schlüter, J., Grill, T., and McFee, B.** (2020). Audiobased music structure analysis: Current trends, open challenges, and applications. *Transactions of the International Society for Music Information Retrieval (TISMIR)*, 3(1):246–263. DOI: <https://doi.org/10.5334/tismir.78>
- Page, K. R., Nurmikko-Fuller, T., Rindfleisch, C., Weigl, D. M., Lewis, R., Dreyfus, L., and De Roure, D.** (2015). A toolkit for live annotation of opera performance: Experiences capturing Wagner’s Ring cycle. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 211–217, Málaga, Spain.
- Prätzlich, T., Driedger, J., and Müller, M.** (2016). Memory-restricted multiscale dynamic time warping. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 569–573, Shanghai, China. DOI: <https://doi.org/10.1109/ICASSP.2016.7471739>
- Prätzlich, T. and Müller, M.** (2013). Freischütz Digital: A case study for reference-based audio segmentation of operas. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 589–594, Curitiba, Brazil.
- Prätzlich, T. and Müller, M.** (2016). Triple-based analysis of music alignments without the need of ground-truth annotations. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 266–270, Shanghai, China. DOI: <https://doi.org/10.1109/ICASSP.2016.7471678>
- Rindfleisch, C.** (2016). Modelling music reception: An ontology for representing interpretations of Richard Wagner’s leitmotifs. In *Proceedings of the International Conference of the Alliance of Digital Humanities Organizations (DH)*, pages 333–336, Krakow, Poland.
- Röwenstrunk, D., Prätzlich, T., Betzwieser, T., Müller, M., Szwillus, G., and Veit, J.** (2015). Das Gesamtkunstwerk Oper aus Datensicht – Aspekte des Umgangs mit einer heterogenen Datenlage im BMBF-Projekt “Freischütz Digital”. *Datenbank-Spektrum*, 15(1):65–72. DOI: <https://doi.org/10.1007/s13222-015-0179-0>
- Serra, X.** (2014). Creating research corpora for the computational study of music: The case of the CompMusic project. In *Proceedings of the AES International Conference on Semantic Audio*, London, UK.
- Stoller, D., Durand, S., and Ewert, S.** (2019). End-to-end lyrics alignment for polyphonic music using an audio-to-character recognition model. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 181–185, Brighton, UK. DOI: <https://doi.org/10.1109/ICASSP.2019.8683470>
- Thickstun, J., Harchaoui, Z., and Kakade, S. M.** (2017). Learning features of music from scratch. In *Proceedings of the International Conference on Learning Representations (ICLR)*, Toulon, France.
- Tzanetakis, G.** (2014). Computational ethnomusicology: A music information retrieval perspective. In *Proceedings of the Joint Conference 40th International Computer Music Conference (ICMC) and 11th Sound and Music Computing Conference (SMC)*, pages 69–73, Athens, Greece.
- Volk, A., Wiering, F., and van Kranenburg, P.** (2011). Unfolding the potential of computational musicology. In *Proceedings of the International Conference on Informatics and Semiotics in Organisations (ICISO)*, pages 137–144, Leeuwarden, The Netherlands.
- Wagner, R.** (1852). *Oper und Drama*. J. J. Weber, Leipzig, Germany.
- Wagner, R.** (2010–2014). *Der Ring des Nibelungen*. ed. Egon Voss, 4 vols., Schott Music, Mainz, Germany.
- Weiß, C., Arifi-Müller, V., Prätzlich, T., Kleinertz, R., and Müller, M.** (2016). Analyzing measure annotations for

Western classical music recordings. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 517–523, New York, USA.

- Weiß, C., Zalkow, F., Müller, M., Klauk, S., and Kleinertz, R.** (2017). Versionsübergreifende Visualisierung harmonischer Verläufe: Eine Fallstudie zu Wagners Ring-Zyklus. In *Proceedings of the Jahrestagung der Gesellschaft für Informatik (GI)*, pages 205–217, Chemnitz, Germany.
- Weiß, C., Zeitler, J., Zunner, T., Schuberth, F., and Müller, M.** (2021). Learning pitch-class representations from score–audio pairs of classical music. In *Proceedings of*

the International Society for Music Information Retrieval Conference (ISMIR), pages 746–753, Online.

- Zalkow, F., Weiß, C., and Müller, M.** (2017a). Exploring tonal-dramatic relationships in Richard Wagner’s Ring cycle. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pages 642–648, Suzhou, China.
- Zalkow, F., Weiß, C., Prätzlich, T., Arifi-Müller, V., and Müller, M.** (2017b). A multi-version approach for transferring measure annotations between music recordings. In *Proceedings of the AES International Conference on Semantic Audio*, pages 148–155, Erlangen, Germany.

TO CITE THIS ARTICLE:

Weiß, C., Arifi-Müller, V., Krause, M., Zalkow, F., Klauk, S., Kleinertz, R., & Müller, M. (2023). Wagner Ring Dataset: A Complex Opera Scenario for Music Processing and Computational Musicology. *Transactions of the International Society for Music Information Retrieval*, 6(1), 135–149. DOI: <https://doi.org/10.5334/tismir.161>

Submitted: 25 February 2023 **Accepted:** 17 May 2023 **Published:** 25 October 2023

COPYRIGHT:

© 2023 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

Transactions of the International Society for Music Information Retrieval is a peer-reviewed open access journal published by Ubiquity Press.