



FACHBEREICH 2: MEDIENPRODUKTION

# Die Verwendung von Tonaufnahmen im LMS:

## Entwicklung spezifischer digitaler Werkzeuge an Hochschulen

Bachelorarbeit

zur Erlangung des Grades Bachelor of Arts  
der  
Technischen Hochschule Ostwestfalen-Lippe

VORGELEGT VON

DENNIS TREIBER

SOMMERSEMESTER 2022



Matrikelnr: 15416076  
Sommersemester 2022

Erstprüfer: Aristotelis Hadjakos  
Zweitprüfer: Guido Falkemeier

Lizenz: CC-BY 4.0, ODC-BY 1.0

## Abstract

Diese Arbeit handelt von der Frage, wie Tonaufnahmen-basierte Lernprozesse im Learning Management System der Hochschule für Musik Detmold, *Moodle*, erweitert werden können. Dazu werden LMS zunächst definiert und anschließend in die Bildungslandschaft eingeordnet. Daraufhin wird der Status Quo betrachtet mit der Feststellung, dass ein Bedarf an Werkzeugen besteht. Dieser Bedarf wurde durch die Programmierung zweier Anwendungen adressiert, die eine Integration im LMS ermöglichen und damit zu einer erhöhten Nutzbarkeit von Tonaufnahmen und musikalischen Inhalten führen sollen. Zum einen ist das eine Implementation des DTW Algorithmus, mittels welchem sich Synchronisationsdaten zwischen zwei verschiedenen Musikdarstellungen desselben Stückes berechnen lassen. Damit ließe sich bspw. ein Interface erstellen, auf dem die Anzeige der Musikwiedergabe mit der Anzeige einer Notenpartitur synchronisiert wird. Die zweite Anwendung fällt in den Bereich des maschinellen Lernens – es wurde ein automatischer Instrumentenklassifizierer geschrieben. Dieser eignet sich zur Erstellung von automatischen Taggings, zwecks Organisation von Daten und Gehörübungen. Die Nutzung einer CNN-Architektur hat sich dabei als effektiv erwiesen: Nach insgesamt 39 Lernepochen und knapp 7 Millionen gelernten Parametern konnte eine Genauigkeit von 95% erzielt werden. Als Datensatz diente die frei verfügbare Aufnahmensammlung des britischen Philharmonia Orchesters (vgl. Thorben Dittes).

Im zweiten Kapitel soll ein Abstecken der Zwecke der einzelnen Programme die Designentscheidungen informieren, welche daraufhin erläutert werden. Im dritten Teil wird anschließend mit *ScoreTube* eine DTW Implementation von Berndt et al. zum Vergleich herangezogen, um die vorliegende Arbeit in den aktuellen Diskurs einzuordnen. Der Beitrag endet mit einer Evaluation der Ergebnisse und einem Ausblick auf potenzielle zukünftige Arbeiten.

## Inhalt

Abstract .....	1
Abkürzungsverzeichnis .....	II
Abbildungsverzeichnis.....	III
Tabellenverzeichnis.....	IV
1. Motivation.....	2
1.1 Learning Management System .....	4
1.1.1 Definition.....	4
1.1.2 Einordnung in die Hochschuldidaktik .....	7
1.2 Blended Learning.....	10
1.3 MOOCs.....	12
1.3.1 Theorie: Pädagogik, Andragogik, Heutagogik.....	12
1.3.2 Zur Historie .....	15
1.4 LMS in Detmold.....	16
2 Zielsetzung: Entwicklung zweier Tonaufnahmen-basierter Werkzeuge .....	19
2.1 Begriff: Tonaufnahmen.....	20
2.2 Designentscheidungen.....	22
3 Literatur: ScoreTube .....	24
4 Evaluation und Ausblick .....	27
4.1 Dynamic Time Warping Audio/Midi .....	27
4.2 Instrumenten-Klassifikation .....	32
5 Anhang: Screenshots DTW.....	35
a) Chromas.....	35
b) Kostenmatrix mit Alignment-Pfad .....	36
Literaturverzeichnis .....	V
Eigenständigkeitserklärung .....	XV

## Abkürzungsverzeichnis

ADL .....	Advanced Distributed Learning
BL.....	Blended Learning
CNN .....	Convolutional Neural Network
DMT .....	detmoldmusictools
DTW .....	Dynamic Time Warping
FC .....	Flipped Classroom
GLMS .....	Gamified Learning Management System
HCD .....	Human-Centered Design
JS.....	Javascript
LCD .....	Learner-Centered Design
MEI.....	Music Encoding Initiative
MFCCs .....	Mel Frequency Cepstral Coefficients
MOOC .....	Massive Open Online Courses
MrMsDTW .....	Memory-restricted Multiscale Dynamic Time Warping
OLC.....	Online Learning Consortium
SCORM.....	Shareable Content Object Reference Model
SUS.....	System Usability Scale questionnaire
TA DTW.....	Transposition-Aware Dynamic Time Warping
ZeMFI.....	Zentrum für Musik- und Filminformatik

## Abbildungsverzeichnis

Abb. 1: MIDI Daten in einer Piano Roll (Sagayama et al. 2004) .....	3
Abb. 2: Simple MEI Beispiel (Introduction to MEI 2022) .....	3
Abb. 3: Notation/Partition .....	3
Abb. 4: Darstellung einer Tonaufnahme als zeitabhängige Funktion (Wellenform).....	3
Abb. 5: Idealtypische Architektur eines LMS (Schulmeister 2005) .....	8
Abb. 6: Verlauf von Pädagogik über Andragogik zur Heutagogik (Blaschke 2012, 60).....	13
Abb. 7: Double-Loop Learning (Eberle & Childress, 2005, zitiert nach: Eberle 2009, 183).....	14
Abb. 8: Screenshot der Plattform detmoldmusictools (Stärk 2013) .....	18
Abb. 9: Audioaufnahme einer C-Dur Tonleiter, auf einem Piano gespielt (Müller 2015, 57).....	21
Abb. 10: Screenshot ScoreTube (Berndt et al. 2018) .....	24
Abb. 11: Synchronisierung von verschiedenen Musikrepräsentationen (Müller 2015).....	26
Abb. 12: Schematische Darstellung von Global Constraints. ....	30
Abb. 13: Confusion Matrix .....	32
Abb. 14: Accuracy Graph .....	33
Abb. 15: Loss Graph .....	34
Abb. 16: Berechnung von Chromavektoren .....	35
Abb. 17: Plot-Ausgabe des Mappings inkl. Kostenmatrix .....	36

## Tabellenverzeichnis

Tabelle 1: Raster zur Differenzierung medialer Angebote (vgl. ebd., 76) .....	6
Tabelle 2: Beispielhafte Darstellung des Speicherbedarfs eines klassischen DTW und dem MrMsDTW (vgl. Pratzlich et al. 2016, 3) .....	29

## 1. Motivation

Kernmotivation dieser Arbeit ist die Förderung der Hochschul-Infrastruktur, die zur Distanzlehre im Internet genutzt wird. Die zugehörige webbasierte Plattform wird als Learning Management System bezeichnet – eine virtuelle Umgebung, die zur Organisation von Lehr- und Lernvorgängen dient. Spezieller soll es hier um die Hochschule für Musik Detmold als Träger gehen. Die Eigenheiten der musikalischen Domäne machen die Nutzung musikspezifischer Anwendungen erforderlich, die die Darstellung sowie das Absolvieren von Lernszenarien zulassen. Diese können z.B. sein: Eine integrierte Musiknotation, kraft der sowohl Partituren erstellt werden können als auch automatisch auswertbare Notations- oder harmonische Analyseaufgaben, inklusive entsprechender Darstellungsoptionen; Akkorde, Stufen von Tonleitern und Formverläufe sind hierfür einige Beispiele. In Bezug auf die Tonaufnahmen-basierte Lehre, welche das Kernstück dieser Ausarbeitung bildet, lassen sich gesondert nennen: Eine synchronisierte Darstellung musikalischer Inhalte, die den Abgleich von Musikaufnahmen mit symbolischen Darstellungen wie Partituren erlauben, oder im Rahmen der kritischen Gehörbildung eine Zuordnung von Instrumenten. Diese Anwendungen werden in [Kapitel 2](#) aufgegriffen.

Insgesamt bezeichnet die Tonaufnahmen-basierte Lehre hier sämtliche pädagogisch aufbereiteten Szenarien, welche in Verbindung zu der Arbeit rund um Audio-material stehen. Darunter zählen symbolische Darstellungen wie MIDI oder Partituren (siehe Abb. 1-3) isoliert betrachtet explizit nicht, da sie keine Aufnahmen (siehe Abb. 4) sind. Kontextabhängig genügt jedoch eine gemeinsame Verwendung innerhalb eines Lernszenarios, um sie doch der Tonaufnahmen-basierten Lehre zugehörig werden zu lassen, wie z.B. in der erwähnten Synchronisationsaufgabe.

Um die genannten Szenarien in der Distanzlehre zu ermöglichen, liegt es im Interesse der HfM, entsprechende Funktionen in einem Interface – dem LMS – gebündelt anbieten zu können. Das Resultat wäre eine inklusive Lernumgebung, in der (auch angehende) Studierende ihre Kenntnisse interaktiv erproben sowie vertiefen können. Außerdem würde die Möglichkeit zur Umsetzung didaktischer Konzepte wie *Blended Learning* oder *MOOCs* eröffnet. Dazu bedarf es jedoch zunächst technisch anspruchsvoller Lösungen.

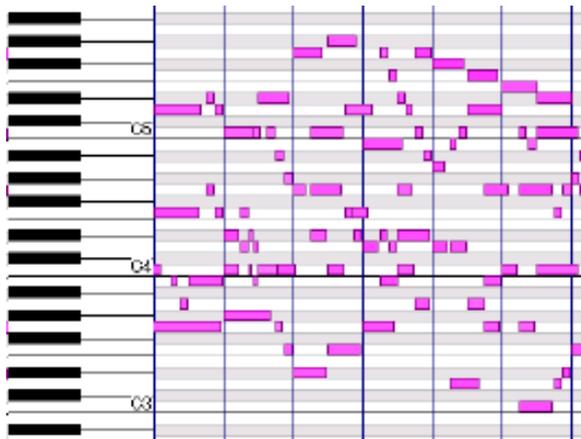


Abb. 1: MIDI Daten in einer Piano Roll (Sagayama et al. 2004)

```
<measure n="10">
<staff n="1">
<layer>
<note pname="f" oct="4" dur="4"/>
<note pname="g" oct="4" dur="4" xml:id="c4ded06ff"/>
<note pname="a" oct="4" dur="4"/>
<note pname="c" oct="5" dur="4"/>
</layer>
</staff>
<dynam startid="#c4ded06ff">f</dynam>
</measure>
```

Abb. 2: Simple MEI Beispiel (Introduction to MEI 2022)



Abb. 3: Notation/Partition.

(a) C-Dur Tonleiter,,

(b) C-Moll Tonleiter (Müller 2015, 16)

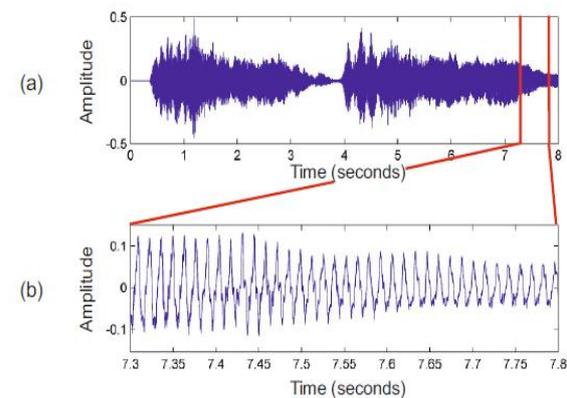


Abb. 4: Darstellung einer Tonaufnahme als zeitabhängige Funktion (Wellenform).

(a) erste fünf Takte aus Beethovens Sinfonie Nr. 5,

(b) vergrößerter Ausschnitt (Müller 2015, 20)

Mit dem ZeMFI ist in Detmold eine Entität vertreten, die sich unter anderem solcher bildungs- und informationswissenschaftlicher Forschungsfragen annimmt und Lösungen entwickelt (vgl. Hadjakos). Ziel dieser Arbeit ist es, unterstützend einen Beitrag zur Verbesserung der Tonaufnahmen-basierten Lern- und Lehrinfrastruktur zu leisten. Grundstein dafür sollen zwei Programme sein, deren anschließende Einbindung im LMS möglich sein soll.

Da Learning Management Systeme zentraler Begriff dieses Beitrages sind und sie sich zudem im Spannungsfeld weiterer bildungswissenschaftlich relevanter Begriffe bewegen, soll im folgenden Abschnitt zunächst eine für dieses Werk gültige Verständnisgrundlage geschaffen werden.

## 1.1 Learning Management System

### 1.1.1 Definition

Learning Management Systeme sind zunächst webbasierte Softwareumgebungen, die über eine Datenbank Nutzer, Kurse und Inhalte miteinander verknüpfen (vgl. Piña 2012, 33). Regelmäßige Tätigkeiten, die auf einem solchen System stattfinden, sind nach Brehmer und Becker zum Beispiel die „*Planung, Anmeldung, Bereitstellung von Seminarunterlagen und Vorlesungsfolien, Lehrfilmen, Erfolgsmessung, [sowie die] Prüfungsorganisation*“ (2017, 2).

Das Fundament sei dabei die Verbindung zwischen Lernenden und Lehrenden: Letztere loggen sich im LMS ein, erstellen die Lernszenarien mit den durch die LMS gegebenen Bordmitteln und motivieren anschließend verschiedene Formen der Partizipation (vgl. Bradley 2021, 2). Die Lernenden greifen wiederum auf die Kursinhalte zu und arbeiten diese durch. Nebenbei stehen mehrere Wege zur Kommunikation bereit, um Fragen stellen oder sich mit anderen Teilnehmern austauschen zu können (vgl. ebd.). Neben der Möglichkeit asynchronen Kommunizierens, etwa über Forenbeiträge, Ankündigungen, Mails oder Aufzeichnungen, stehen auch synchrone Kanäle wie Videokonferenzen und Echtzeit-Chaträume offen (vgl. ebd., 71). Sämtliche Austauschmöglichkeiten sind je nach Bedarf öffentlich für alle Teilnehmer sichtbar wahrzunehmen, oder privat als Einzel- oder Gruppengespräch.

Von Gamified Learning Management Systemen wird hingegen gesprochen, sobald sich Konzepte des Game-Designs im LMS wiederfinden (vgl. Kapp 2012, zitiert nach: Swacha 2014, 195). Das beginnt bspw. mit der Einführung von Schwierigkeitsgraden (vgl. Labaj und Bieliková 2014, 13) oder damit, dass bei erfolgreichem Beenden von Lernabschnitten sogenannte *Badges*, also digitale Auszeichnungen, verliehen werden (vgl. *Badges - MoodleDocs* 2014). Es gibt auch komplexere Designs, die ein spielerisches Lernen gestatten. Dies können zum einen Memory-Spiele oder Quizze sein. Zwei weitere bekannte Ableger sind *Concept Review Bingo* und *Jeopardy Exam Review* (vgl. Chen et al. 2018, 43). Deren Spielmechaniken führen zu einem Wettbewerb unter den Lernenden, bei dem es unter anderem um das schnelle Lösen von Aufgaben geht (vgl. ebd.). Chen et al. zufolge würden GLMS nachweislich zu einer stärkeren Beteiligung, Sozialisierung unter den Lernenden sowie zu einer verbesserten akademischen Performance führen (vgl. ebd.,

41f). Die Integration von GLMS erfordert jedoch eine höhere technische Expertise, was regelmäßig eine Hürde darstellt (vgl. ebd.).

Das zugehörige, übergeordnete Forschungsgebiet, das sich mit dem Lerner-gerechten Design von LMS beschäftigt, nennt sich Learner-Centered Design (vgl. Chris Quintana et al. 2000). Hauptziel des LCD ist es nach Quintana, praktikable Lösungen zu entwickeln, die dazu geeignet sind, Individuen und Gruppen bei der Entwicklung von Expertise in einer spezifischen Arbeitsdomäne zu unterstützen (vgl. Chris Quintana et al. 2000, 256). Dies ist eine Ausprägung des weiter gefassten Human-Centered Designs (HCD).

Die Betonung liegt im Vergleich zur HCD-typischen Betrachtungsweise jedoch darauf, dass eine erste Brücke zwischen lernender Person und dem zu erlernenden Bereich gebildet werden soll, anstelle einer Brücke zwischen Anwender und Programm (vgl. ebd., 258f). Es geht somit verstärkt um die pädagogische Ausgestaltung für die Lernenden. HCD bietet jedoch ebenfalls einen validen Rahmen zur wissenschaftlichen Untersuchung von LMS, wobei das Augenmerk hier mehr den Lehrenden gilt, zu deren Aufgaben es schließlich gehört, die Lernmaterialien zu erstellen. Der Schwerpunkt liegt hier nicht darin, auf Basis pädagogisch aufbereiteter Kurse die Entwicklung eben solcher Kurse zu erlernen (denn das wäre wieder eine Betrachtung des LCD), auch wenn es prinzipiell einen möglichen Gegenstand von Lernkursen darstellt. Vielmehr geht es darum, auf technischer Ebene einen anwenderfreundlichen Zugang zum Entwickeln solcher Kurse zu schaffen.

Aspekte beider Designs fließen in die vorliegende Arbeit ein.

Aufgrund der vielschichtigen Ausgestaltung von LMS lassen sich folgende medienwissenschaftliche Zuordnungen treffen:

- Orts- und zeitunabhängige Produktion wie Rezeption durch „neue Medien“ (vgl. Paganini und Steinbacher 2019, 190)
- Damit implizite Teilhabe durch Mensch-Maschine-Interaktion (vgl. Leonhard et al. 1999, 118; Pfänder und Wagner 2009, 97)
- Multicodalität der Medien (vgl. Weidenmann 2009, 74ff)
- Multimodalität (vgl. ebd., 74ff)

Die letzten zwei Merkmale sind erklärungsbedürftig; Tabelle 1 dient zur Veranschaulichung.

Der landläufig häufig benutzte Begriff „Multimedialität“ sei kurz erwähnt, denn er ist hier zwar grundsätzlich korrekt angesiedelt, wird jedoch aufgrund seines vagen Bedeutungsgehaltes vernachlässigt (vgl. ebd., 74).

Angesichts der genannten Beschaffenheiten von LMS ist erkennbar, dass sie ein komplexes Erzeugnis des Design Thinkings sowie HCD sind: Sie nehmen die Erfüllung der Bedürfnisse von Nutzer:innen als zentrale Priorität an (vgl. Meenakshi et al. 2022, 15).

	<i>Mono-...</i>	<i>Multi-...</i>
<i>Codierung</i>	<b>Monocodal:</b> <ul style="list-style-type: none"> <li>- Nur Text</li> <li>- Nur Bilder</li> <li>- Nur Zahlen</li> </ul>	<b>Multicodal:</b> <ul style="list-style-type: none"> <li>- Text mit Bildern</li> <li>- Grafik mit Beschriftung</li> </ul>
<i>Sinnesmodalität</i>	<b>Monomodal:</b> <ul style="list-style-type: none"> <li>- Nur visuell (Text, Bilder)</li> <li>- Nur auditiv (Rede, Musik)</li> </ul>	<b>Multimodal:</b> <ul style="list-style-type: none"> <li>- Audiovisuell (Video, Lernsoftware mit Bild und Ton)</li> </ul>

*Tabelle 1: Raster zur Differenzierung medialer Angebote (vgl. ebd., 76)*

Aus Perspektive der Marktforschung wäre die zeitliche und örtliche Gebundenheit klassischen Unterrichts als Pain-Point (engl. für „Schmerz“; vgl. Naredla et al. 2018, 2309) für die Lernenden auszulegen, da ihnen lediglich schablonenhaft ein statischer Terminplan vorgelegt wird, der individuelle Lebenssituationen nicht berücksichtigen *kann* und um dessen Einhaltung zeitlicher sowie logistischer Aufwand betrieben werden muss. Das anthropozentrische Design von LMS hingegen erlaubt es, diesen Pain-Point durch Nutzung der neuen Medien derart zu adressieren, dass Lehrangebote entsprechend einer individuellen Planung wahrgenommen werden können. Dies führt zu einer drastischen Vergrößerung des Wirkungsradius – was wiederum den breitflächigen Erfolg dieser Plattformen konstituiert, wie die folgenden Ausführungen belegen.

### 1.1.2 Einordnung in die Hochschuldidaktik

Obwohl mit PLATO das erste computerbasierte LMS bereits 1960 entwickelt wurde, existieren moderne LMS in der heutigen Form erst seit 1990 (vgl. Chaubey und Bhattacharya 2015, 159). Sie gehören somit zu den jüngeren Erscheinungen in der Lehre, haben jedoch bereits festen Einzug in die Bildungslandschaft gehalten. Laut Harrington et al. sei dies sogar in einem Ausmaß und mit einer Geschwindigkeit geschehen, die es zuvor bei keiner anderen Innovation in diesem Bereich gegeben habe (vgl. 2004, 2). Diese Aussage könnte aufgrund der zeitlichen Distanz an Aussagekraft eingebüßt haben, allerdings deutet auch eine aktuellere, gemeinsame Erklärung der im deutschsprachigen Raum verbreiteten LMS Moodle, ILIAS und Stud.IP auf nichts Gegenteiliges hin: So werden mittlerweile an fast 90% aller staatlichen Hochschulen in Deutschland aktiv LMS Infrastrukturen genutzt (vgl. Open Source-LMS 2020) – ein Anteil, der selbst im internationalen Vergleich heraussticht und sie als kritische Infrastruktur qualifiziert.

Von der weiten Verbreitung lässt sich jedoch nicht automatisch auf eine Korrelation zur hohen Qualität der Nutzung schließen, wie eine Studie aus Malaysia nahelegt. Dort würden sich z.B. viele Hochschuldozenten im LMS lediglich darauf beschränken, Materialien hochzuladen, während die zahlreichen interaktiven Funktionen wie Chats, Diskussionsforen oder Nachrichten letztlich unangetastet blieben (vgl. Ramayah, 2005, zitiert nach: Almarashdeh et al. 2010, 81).

Der Schwerpunkt von LMS liegt gut erkennbar im Bildungswesen. Dieses Merkmal ist es auch, welches sie maßgeblich von technisch vergleichbaren Systemen unterscheidet, die etwa im E-Commerce Verwendung finden (vgl. Piña 2012, 33). Erwähnenswert sind an dieser Stelle *Content Management Systeme*, denn sie überschneiden sich außer im Wortlaut auch in den Funktionen teilweise so sehr, dass die Bezeichnung „Learning Content Management System“ ebenfalls gebräuchlich geworden ist (vgl. Mohd Kasim und Khalid 2016, 56). Hierbei ist jedoch zu beachten, dass LMS eine Mehrzahl von Subsystemen unter sich vereinen, darunter neben Course Management Systemen, Instructional Management Systemen und weiteren auch Content Management Systeme (vgl. Coates et al. 2005, 20; siehe auch Abb. 1). Für die Zwecke dieser Arbeit erschöpft sich darin die Relevanz der weiteren Unterscheidung.

Auf medienwissenschaftlicher Ebene ist weiterhin anzumerken, dass Ellis als Erfordernis an LMS einen Selbstbedienungs-Charakter anführt (vgl. 2009, 2). Fraglich ist, inwieweit das zutrifft. Bei kritischer Betrachtung ließe sich jedenfalls weiter differenzieren, dass zwar einerseits den Lernenden durch die Möglichkeit asynchroner Beteiligung ein nicht unwesentlicher Spielraum zur Selbstbestimmtheit zugebilligt wird. Doch andererseits liegt eine vorausgehende Notwendigkeit der Freischaltung von Zugangsrechten durch die Lehrperson vor (vgl. Sumardi et al. 2021, 350), so dass eine offenkundige Beschränkung vorliegt. Der freie Zugriff ist demnach immer nur einem bestimmten Personenkreis vorbehalten.

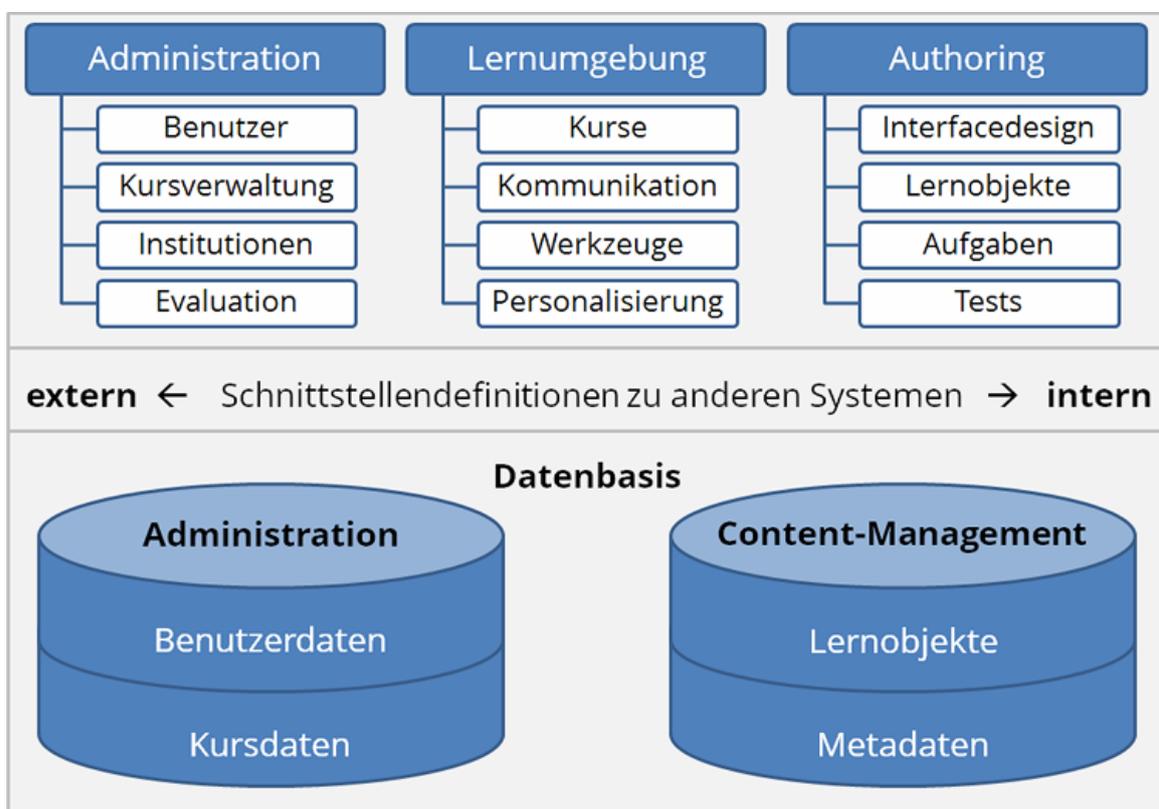


Abb. 5: Idealtypische Architektur eines LMS (Schulmeister 2005)

Makroskopisch betrachtet lassen sich LMS unter dem Phänomen E-Learning subsumieren, für welches Kerres eine häufig zitierte, weite Definition aufstellt als „Oberbegriff für alle Varianten der Nutzung digitaler Medien zu Lehr- und Lernzwecken, sei es auf digitalen Datenträgern oder über das Internet, etwa um Wissen zu vermitteln, für den zwischenmenschlichen Austausch oder das gemeinsame Arbeiten an digitalen Artefakten“ (2013, 6). Aufgrund der mehrdimensionalen

Entwicklungsgeschichte innerhalb der Bereiche Bildung, Militär und Unternehmertum seit den 1960er Jahren sei eine einheitliche Anschauung unmöglich (vgl. Nicholson 2007, 1). Zudem besteht aufgrund des beständigen Wandels im technischen Bereich sowie bei den didaktischen Bedingungen eine weitere Schwierigkeit der präzisen Definition (Brehmer und Becker 2017).

Kerres erwähnt in seiner Definition die gemeinsame Arbeit an digitalen Artefakten. Dieses knüpft thematisch an Open Source Software an, welche im Zusammenhang mit LMS einen besonderen Stellenwert in der deutschen Hochschuldidaktik erfährt. So fallen die bereits erwähnten LMS Moodle, Ilias sowie Stud.IP ebenfalls unter OSS, womit OSS eine wesentliche Infrastruktur der gelebten Bildungspraxis in Deutschland darstellt. Diese Eigenschaft unterscheidet auch maßgeblich LMS von kommerziellen Angeboten wie Udemy oder Coursera.

OSS als solche ist bereits hinreichend bekannt, jedoch ist erwähnenswert, dass durch die Beteiligung vieler Akteure die Möglichkeit zur Schaffung von Standards gegeben ist. Als prominentes Beispiel gilt das Betriebssystem Linux, welches zu einer Standardisierung im vorher fragmentierten Unix-Markt geführt habe (vgl. Brügge 2004, 1). Auch die Entwicklungsgeschichte von LMS wurde durch eine Standardisierung maßgeblich beeinflusst.

Die Interaktionen zwischen Systemen und Anwendungen (siehe z.B. Abb. 1) sind in sogenannten „offenen Standards“ geregelt. Diese werden wiederum in einer Mehrzahl an Referenzmodellen gebündelt, z.B. im sogenannten SCORM (vgl. Bailey 2005). Sie legen technische Spezifikationen für die interne und externe Kommunikation zwischen Inhalten, Modulen, sowie deren Trägersystemen fest, sodass eine Interoperabilität gewährleistet werden kann. Ins Leben gerufen wurde SCORM durch die ADL Initiative im Jahr 2000 als Reaktion auf das kostspielige Problem für Unternehmen, dass wegen fehlender Standards eine Erneuerung in internen Systemen meist einen kompletten Neuanfang bedeutete (vgl. ebd., 2). Aufgrund der Wiederverwendbarkeit von Modulen sowie der ersparten Aufwendung zusätzlichen Finanz- und Humankapitals kann diese Initiative als ein weiteres Produkt des nutzerorientierten Design Thinking betrachtet werden.

Das an der HfM Detmold verwendete System Moodle ist ebenfalls kompatibel mit SCORM und anderen Standards (vgl. *Standards – MoodleDocs* 2022). Dieser Tatsache gilt es zu verdanken, dass der technischen Grundvoraussetzung zur

Einbindung der beiden Programme, die im Rahmen dieser Arbeit entwickelt wurden, Genüge getan ist. Mehr dazu im Kapitel „[Designentscheidungen](#)“.

Aus universitärer Sicht sind die OSS typischen Eigenschaften geeignet, um eine Präferenz dafür zu bilden. Denn anders als bei proprietären Lösungen müssen hierfür – offensichtlich – keine Nutzungsgebühren getragen werden. Doch nicht nur finanziell bedeutet es eine Entlastung – auch Gesichtspunkte des Datenschutzes sprechen dafür. Erst im Jahr 2019 wurde im Bundesland Hessen die Nutzung von Microsoft Office 365 an Schulen für unzulässig erklärt, da es nicht datenschutzkonform verwendet werden konnte (Rost 2020).

Für Instruktoren – und das ist der für diese Arbeit maßgebliche Aspekt – soll das Zusammenstellen und schnelle Verfügbarmachen von effektiven Lerninhalten gegeben sein. Diese können zur Umsetzung verschiedener didaktischer Konzepte genutzt werden. Zwei populäre Konzepte des E-Learning sollen im folgenden Teil vorgestellt werden.

## 1.2 Blended Learning

Hier sei zunächst das sogenannte *Blended Learning* genannt – eine hybride Form des Lernens (wörtl. „*vermishtes Lernen*“), welche eine Kombination aus Präsenz- und webbasiertem Unterricht vorsieht (vgl. Whitelock und Jelfs 2003, 99). Das Aufkommen dieser Form wird auf das Jahr 2000 datiert (vgl. Bliuc et al. 2007, zitiert nach: Bernard et al. 2014, 90). Als Reaktion auf die pandemiebedingten Schulschließungen seit 2020 veröffentlichte die Europäischen Kommission als Hilfestellung ein Dokument, in dem Informationen zu BL zu finden sind (vgl. 2021, 4). Dort wird eine weiter gefasste Definition des Begriffes gewählt. Nach ihr wird darauf abgestellt, dass sowohl schulseitige als auch distanzierte Lernumgebungen eingesetzt werden und zum anderen eine Vermischung aus analogen und digitalen Lernmitteln stattfindet. Daraus resultiert, dass LMS zwar Bestandteil dieser Lehrform sein können, es aber nicht zwingend sein *müssen*. Das „Online Learning Consortium“ („OLC“), eine Gemeinschaft aus Führungskräften und Innovatoren der Hochschulbildung, definiert ein quantifizierbares Mischverhältnis, nach dem der Unterricht bei

einem Online-Anteil zwischen 30 und 79% als BL klassifiziert werden kann (vgl. Page et al. 2017, 44).

Trotz geringfügiger Unterschiede in der Auslegung ist sich die Literatur weitestgehend darüber einig, dass BL-Szenarien durchschnittlich zu einer besseren Performance unter Lernenden führen als bei ihren Vergleichsgruppen ohne BL. Darauf deuten mehrere empirische Untersuchungen und Metastudien hin (vgl. Cannon et al. 2021, 3; Bernard et al. 2014, 90; Means et al. 2013, 2; Halasa et al. 2020, 1130). Begründet werden kann dies durch das Zusammenwirken der Vorteile aus beiden Bereichen: Das sind zum einen der hohe Grad an Autonomie durch das asynchrone Distanzlernen (vgl. Jost et al. 2021, 3098), zum anderen die soziale Teilhabe, die Möglichkeit praktischer Aktivität und des direkten Feedbacks (vgl. Tayebinik und Puteh 2012, 1). Zusätzlich entlastete die digitale Lernumgebung die Lehrperson, so dass ihr wiederum mehr Zeit zur Verfügung steht, um sich kleinen Gruppen oder auch Einzelnen unterstützend zu widmen (vgl. ebd., 3).

Diese spezielle Ausprägung, bei der die Wissensvermittlung im virtuellen Lernsetting geschieht statt im klassischen Frontalunterricht, während die praktischen Übungen in einer Face-to-Face-Situation wahrgenommen werden anstatt daheim, wird in der Bildungspraxis als „Flipped Classroom“ bezeichnet (vgl. Capone et al. 2017, 1f). Die beiden Begriffe BL und FC sind nicht per se deckungsgleich, da FCs eine schärfere Trennung zwischen dem Einsatz technischer Mittel und Face-to-Face-Interaktionen vorschreiben (vgl. ebd., 5). So werden die Lernenden bspw. vor dem Betreten einer Unterrichtsstunde explizit um das Betrachten des angesetzten Lehrmaterials gebeten. Allerdings räumen Capone et al. ein, dass gerade der BL-Ansatz ein Katalysator für diese Art des Lernens und Lehrens sei (vgl. ebd., 1). Daher ist es nicht unüblich, beides kombiniert in der Literatur anzutreffen (vgl. Halasa et al. 2020; Capone et al. 2017).

Die Nutzung ist auf Anwenderseite schließlich – wie bei allen Formen des digitalen Lernens – auch eine Frage der Akzeptanz. Marquis Untersuchung von 2004 zufolge waren seinerzeit bereits 90% aller Dozenten an Universitäten von der Überlegenheit des BL-Konzeptes überzeugt (vgl. 2004, zitiert nach: Bernard et al. 2014, 90). Auf Anbieterseite ist hingegen entscheidend, ob die technischen Besonderheiten der jeweiligen Disziplin ausreichend in der Infrastruktur berücksichtigt worden sind. Denn wo immer das Erlernen praktischer Fähigkeiten zum Bildungsziel gehört, stellt

sich die Frage nach einem sachgerechten Vermittlungsweg. Anhand des vorliegenden Beispiels des Musikstudiums in Detmold soll dieser Umstand später im Unterkapitel „[LMS in Detmold](#)“ weiter beleuchtet werden.

### 1.3 MOOCs

Ein weiteres prominentes Konzept im Feld des E-Learning sind neben BL die sogenannten *MOOCs* („Massive Open Online Courses“). Sie zeichnen sich vor allem durch die Offenheit des Zuganges, ihre hohe Teilnehmerzahl und die Dauer von meist sechs bis vierzehn Wochen aus, in denen ein Oberthema mit seinen verschiedenen Facetten entlang eines Curriculums behandelt wird (vgl. Bremer 2013, 30).

Mehrere Unterarten haben sich entwickelt, wie z.B. die sogenannten *xMOOCs*, die ursprünglich als externe Angebote der Harvard-Universität, mit der Möglichkeit ein Zertifikat zu erwerben, realisiert wurden (vgl. ebd., 33). In der Praxis gruppieren sich die Lernenden dabei häufig zu Lerngruppen, was nach Bremer eine Notwendigkeit aufgrund mangelnder Betreuung darstellt. Daneben existieren die „konnektivistisch“ (ebd., 35ff) angelegten *cMOOCs*. Der Begriff Konnektivismus stellt dabei auf die Wichtigkeit des „Wissens wo“ ab, also die Herausarbeitung wichtiger Knotenpunkte, an denen Informationen zu finden sind (vgl. George Simons, 2005, zitiert nach: ebd).

Lineare Arrangements von Materialien sind dabei die Ausnahme, stattdessen sind thematisch aufbereitete Informationsangebote in einer Mehrzahl von Formaten die Regel (vgl. Mackness et al. 2010, 267). Zur Verwaltung dieser Inhalte sind LMS ein regelmäßig genutztes Mittel (vgl. Martin 2012). Die lehrende Kraft nimmt eine mediative, den Diskurs anregende Rolle ein und wird daher in diesem Kontext auch Facilitator (engl. für „Moderator:in“) genannt (vgl. Beaven et al. 2014, 35).

#### 1.3.1 Theorie: Pädagogik, Andragogik, Heutagogik

Das veränderte Selbstverständnis der Lehrkräfte, das mit MOOCs einhergeht, motiviert eine schematische Zuordnung der Dynamik zwischen Einwirkungsgrad des Referenten und den Anforderungen an die Lernenden (siehe Abb. 6). Diese erlaubt

eine qualitative Betrachtungsweise der didaktischen Situation (auf andere Lehrformen als MOOCs übertragbar).

Nach Blaschkes Darstellung lassen sich die Einwirkungsformen hierarchisch gestaffelt in drei Level einteilen, wobei mit wachsendem Level die Eigenverantwortung der Teilnehmer steigt, während die Intensität der Einwirkung durch den Facilitator sinkt. An der Spitze stünde demnach als höchstes Ziel die heutagogische Perspektive (abgeleitet vom altgriechischen Wort für „selbst“). Dies entspricht der Manifestation einer höchstmöglichen Eigenmotivation, Eigeninitiative und Selbstbestimmung zum Lernen, unabhängig von extrinsischen Einflüssen (vgl. Hase und Kenyon 2000). Wie beim andragogischen Ansatz reichere der Facilitator den Unterricht mit Inhalten und Orientierungshilfen an, löse dabei jedoch jegliche Vorgaben von Lernzielen auf (vgl. ebd.; Eberle 2009, 183).

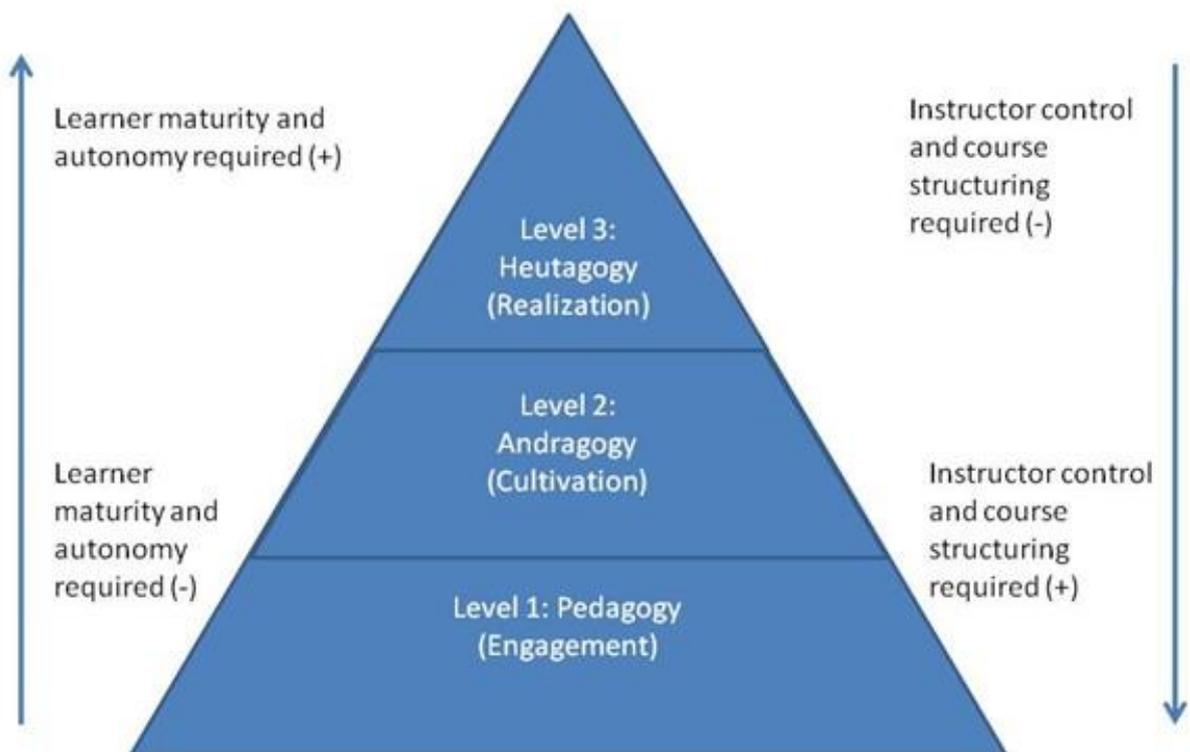


Abb. 6: Verlauf von Pädagogik über Andragogik zur Heutagogik (Blaschke 2012, 60)

Dies stellt nach Cooper die höchste Form des Lernens dar, da durch selbstregulative Prozesse übergeordnete Lernstrategien, Glaubenssätze und Werte wiederholt

auf den Prüfstand gestellt werden (vgl. 2004, zitiert nach: ebd., 183). In der Didaktik wird dieses Vorgehen als *Double-Loop Learning* bezeichnet (siehe Abb. 7).

Das Agieren auf der heutagogischen Ebene setzt weiterhin voraus, dass die darunterliegenden Ebenen Andragogik sowie Pädagogik „gemeistert“ werden müssen, was den pyramidalen Aufbau begründet. Zum Verständnis seien diese Begriffe noch kurz umrissen:

Andragogik bezeichnet die Lehre der Erwachsenenbildung (vgl. Knowles 1978, zitiert nach: Blaschke 2012). Sie zielt darauf ab, die Lernenden unter anderem durch Einbringung ihrer Erfahrungen dazu zu befähigen, Lernbedürfnisse, Ziele und Strategien zu formulieren. Eine *Kultivierung* selbst gerichteter Lernprozesse findet statt. Dies sei Bedingung, um das heutagogische Level zu erreichen und folgt zugleich unmittelbar auf das „Meistern“ der pädagogischen Ebene.

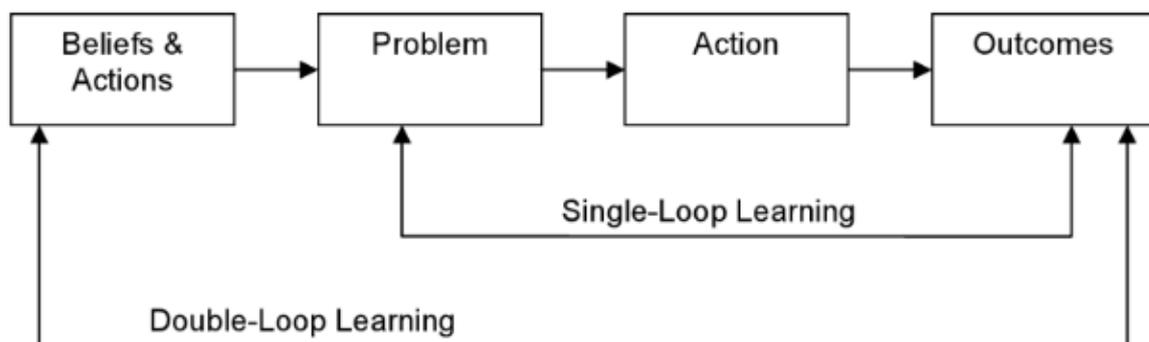


Abb. 7: Double-Loop Learning (Eberle & Childress, 2005, zitiert nach: Eberle 2009, 183)

Die Andragogik wiederum bedingt, dass die Ziele der Pädagogik erfolgreich verinnerlicht wurden. Das bedeutet, die Lernenden bereits von Kindesalter an in einer vertrauensvollen Umgebung einzubeziehen (vgl. Canning 2010, 62). Das Interesse gilt daher zunächst dem beziehungsorientierten Engagement, um das Selbstvertrauen zu stärken und zum lebenslangen Lernen zu motivieren (vgl. Hase und Kenyon 2000).

Dieser Ansatz spielt sich gedanklich vor MOOCs ab, da durch die massenhafte Konstellation bereits eine ausreichende Betreuung aller Lernenden ausgeschlossen ist. Zudem mangelt es an örtlicher Nähe, die für eine vertrauensvolle Umgebung für

Heranwachsende unabdingbar ist. Ein Versuch, auf Grundlage von MOOCs pädagogisch wertvolle Grundlagen zu vermitteln, wäre somit obsolet.

### 1.3.2 Zur Historie

Bei der erstmaligen Erprobung eines MOOCs im Jahr 2008 trug es sich zur Überraschung der Veranstalter zu, dass über 2200 Besucher teilnahmen, was über das Hundertfache dessen hinausgegangen sein soll, was ursprünglich geplant war (vgl. Robinson und Ash 2014, 131; Mackness et al. 2010, 266). Eine weitere Kenngröße ist ein Durchbruch der Besucherzahl bei einem MOOC, das durch die Stanford Universität durchgeführt wurde: Ein frei zugänglicher Kurs, in dem künstliche Intelligenz behandelt wurde, lockte 160.000 Teilnehmer aus der ganzen Welt an (vgl. Waldrop 2013, 162). Dies machte MOOCs von sich reden und leitete eine Periode des Auflebens, der Kooperation und der Erhebung von MOOCs zum Massenphänomen ein (vgl. Martin 2012, 26). In der New York Times hat es 2012 geheißen: „The year of the MOOC“ (Pappano 2012). Einige nennenswerte kommerzielle Anbieter, die sich seitdem hervorgetan haben, sind: Coursera, Udemy, Skillshare und LinkedIn Learning. Kooperationen zu Elite-Universitäten wie Harvard und Stanford sind ebenfalls existent (vgl. Mester 2016, 182). Es kamen in kurzen Zeitabständen immer wieder neue Angebote hinzu, sodass der Markt schwerlich zu überblicken war (vgl. Robes 2012, zitiert nach: Bremer 2012, 41).

Nach einer Studie der Bertelsmann Stiftung scheint diese Entwicklung bis vor der Coronapandemie innerhalb Deutschlands wieder abgeklungen zu sein (vgl. Schmid et al. 2018, 21). Eine Datenerhebung wurde 2018 unter 1003 Befragten durchgeführt. Ihr zufolge würden MOOCs (und auch LMS) die Schlusslichter unter den genutzten Bildungsmöglichkeiten bilden: Lediglich 10% (LMS: 18%) gaben an, die Bezeichnung überhaupt zu kennen. Die Nutzung fiel noch geringer aus – mit nur 1% (LMS: 10%) war die Nachfrage verschwindend gering.

Das Aufkommen der COVID-Pandemie hat jedoch weltweit ein neues Interesse an dieser Lehrform begründet. Während etliche Schulen, Hochschulen wie Dozent:innen zum raschen Umstieg auf den Remote-Unterricht gezwungen waren, registrierten sich 2020 über zehn Millionen Nutzer:innen innerhalb eines Monats auf der

Plattform Coursera, welche somit einen Anstieg von 644% im Vergleich zum Vorjahr verzeichnen konnte (vgl. Razo 17.09.2020).

Institutionelle Angebote der Distanzlehre haben durchweg eine ähnliche Entwicklung erfahren.

#### 1.4 LMS in Detmold

Ein LMS, das im Zusammenhang mit der professionellen Musikausbildung den Bedürfnissen der Studierenden gerecht werden soll, erfordert die Integration musikspezifischer Tools. Insbesondere solche, die auf Musiknotation und der Arbeit mit Audiomaterial basieren, sind wichtiger Bestandteil im Lehrangebot der HfM Detmold. Somit bestehen ohnehin schon besondere Anforderungen an die technischen Gegebenheiten bei der Lehre. Die Relevanz einer Lösung, die überdies für die Distanzlehre geeignet ist, wird durch das pandemiebedingte Geschehen betont.

Es folgt eine Beschreibung der aktuellen Defizite und Möglichkeiten der Darstellungs- und Lernszenarien in der HfM Detmold, mit besonderem Fokus auf die Berührungspunkte zur Tonaufnahmen-basierten Lehre.

Aktuell müssen Lehrinhalte, bspw. Partituren, über kostspielige Drittsoftware erstellt und in die LMS eingebunden werden. Diese lassen sich dann zwar anzeigen, jedoch gibt es keine Möglichkeit des interaktiven Umgangs mit der Notation, wie z.B. die einfache Korrektur von Fehlern (vgl. Nowakowski Juni 2022, 2). Dadurch wird die Gestaltung der Distanzlehre äußerst umständlich und ineffizient zugleich, was wiederum die Akzeptanz unter den Lehrenden stark schmälert. Dies stelle ein erhebliches Hindernis dar, Inhalte im Rahmen des bereits erwähnten *Blended Learning* oder in *MOOCs* einzusetzen.

Ein weiteres Beispiel ist die Bereitstellung von Funktionen zur Archivierung qualitativ hochwertiger und wenig komprimierter Sounddateien (z.B. in den Formaten WAV, AIFF, MP3 mit 360 kbps). Diese wären im Rahmen der zuvor genannten Notationsbearbeitung, der Gehörbildung oder für online-gestützte Prüfungssituationen erforderlich, sind jedoch zurzeit nicht gegeben.

Die beiden genannten Beispiele sind der Tonaufnahmen-basierten Lehre zuzuordnen. Die zwei vorhandenen Implementationen des DTW und der automatischen Instrumentenerkennung entsprechen ihnen weiter und stellen aus diesem Grund geeignete Programme zur Förderung dar. In [Kapitel 2](#) wird dies weiter ausgeführt. Soweit lässt sich jedoch knapp zusammenfassen, dass sowohl nach Gesichtspunkten des HCD sowie des LCD noch erhebliche Mängel bestehen, die eine Adaption von LMS noch verhindern.

Im Falle der interaktiven Musiknotation wurden jedoch bereits Vorarbeiten geleistet, auf dessen Erfahrungswerte als Startpunkt für einen Ausbau der digitalen Lehre zurückgegriffen werden kann.

Eine solche erwähnenswerte Vorarbeit trägt den Titel „*detmoldmusictools*“ – ein im Jahre 2013 bereits durch das Land NRW gefördertes Programm, das webbasiert funktioniert und das die Möglichkeiten der Distanzlehre durch die HfM Detmold erweitert, indem es Lernkurse zusammenstellen und anbieten lässt (siehe Abb. 8). Es existieren sowohl öffentlich zugängliche als auch geschlossene Kurse, die erst nach Eingabe eines Zugriffcodes einsehbar sind. Nach Erstellen eines Benutzerkontos bietet DMT für Jedermann zudem die Möglichkeit, auf der Plattform eigens erstellte Kurse unter Einbindung von Folien, Texten, Verweisen, Bildern, Audiomaterial und interaktiven Elementen zu veröffentlichen. Die für LMS typische multimodale und -codale Ausgestaltung ist hier gegeben, sodass sich DMT als erster erfolgreicher Ableger eines musikspezifischen LMS betrachten lässt.

Obwohl bereits Kurse der musikalischen Lehre in DMT abgehalten werden können, und durch die Integration einer grafischen Oberfläche des Notensatzprogrammes ABC sogar eine grafische Notationsmöglichkeit besteht, bedarf es einer Aktualisierung. So fehle bspw. eine Unterstützung gängiger Formate von symbolischen Darstellungsformen der Musik, wie MusicXML oder MEI. Bei einer Gesamtbetrachtung des Marktes gibt es nach aktuellem Stand lediglich Notationssoftware, die entweder auf (Semi-)Professionelle Nutzer ausgerichtet sei, ohne Berücksichtigung eines angemessenen LCD, oder sie bleibt zu anfängerfreundlich und es mangelt an Tiefe (vgl. Nowakowski Juni 2022, 2).

## DetmoldMusicTools - die Online-Lernplattform für Musik



001 Tipps & Tricks  
André Stärk



002 Noten lernen (Demo)  
André Stärk



Eignungsprüfung an der HfM Detmold  
André Stärk



Einführung in das Partimentospiel  
André Stärk

Abb. 8: Screenshot der Plattform detmoldmusictools. Contentersteller sind anfangs zugleich Lernende, die die Einsteigerkurse zur Bedienung der Plattform rezipieren (Stärk 2013)

Einen vielversprechenden Grundbaustein bildet jedoch die OSS *Verovio*. Dies bezeichnet ein Notensatzprogramm, das die grafische Ausgabe von einer Mehrzahl unterschiedlich codierter Musikrepräsentationen erlaubt. Darunter zählen auch MusicXML und MEI (vgl. Verovio).

Lediglich eine Hürde gilt es zu überwinden: Verovio bietet derzeit keine interaktive grafische Oberfläche. Es gibt jedoch das aktuelle Bestreben, diese zu entwickeln (vgl. Nowakowski Juni 2022, 1). Über die OSS H5P, welche eine Plattform bezeichnet, die zur Erstellung von webbasierten interaktiven Inhalten dient, ließe sich eine Integration in die LMS Moodle, ILIAS sowie DMT realisieren.

Dies erfordert nicht nur die entsprechende Programmierung, sondern setzt die Konzeption geeigneter Designs voraus, welche auf die Lernenden angepasst sind. Im Rahmen der zuvor genannten interaktiven Oberfläche wurde dafür bereits eine Pilotstudie durch Nowakowski durchgeführt (vgl. ebd., 3).

Nach Junyong sei es charakterisierend für eine Pilotstudie, dass sie als „Bei-Studie“ einer übergeordneten Hauptstudie umgesetzt wird, und daher im kleineren Rahmen stattfindet (vgl. 2017, 601). Diese Forschungsmethode bildet ein gedankliches

Vorwort, das bei der Konkretisierung weiterer Forschungsprozesse, sowie der eigentlichen Forschungsfrage hilft (vgl. ebd.).

Vorliegend wurde die Pilotstudie unter drei Studierenden der HfM durchgeführt, welche zu Testzwecken die soweit entwickelten Interaktionsmöglichkeiten erproben konnten. Dies geschah durch die Bearbeitung von Notationsaufgaben. Die Bewertungen wurden in einem sogenannten System Usability Scale (SUS) Fragebogen (vgl. Lewis 2018, zitiert nach: ebd., 3) festgehalten, um Einblicke in die mentalen Prozesse und Gewohnheiten der Studierenden bei der Bedienung zu erhalten. Die Evaluation der Ergebnisse informiert schließlich das induktive Vorgehen bei der Findung weiterer Designentscheidungen, die ein LCD-konformes Produkt erzeugen. Konkret wurde hier die Entscheidung getroffen, zunächst weitere Analysen von verschiedenen, bereits erhältlichen Notensatzprogrammen in die Entwicklung einzubeziehen (vgl. ebd., 4).

Eine grafische Oberfläche, die den genannten Anforderungen entspricht, könnte zudem um weitere dienliche Funktionen erweitert werden, zum Beispiel der synchronisierten Navigation zwischen Aufnahme und Partitur. Diese und eine zweite Funktion soll im Folgenden vorgestellt werden.

## 2 Zielsetzung: Entwicklung zweier Tonaufnahmen-basierter Werkzeuge

Wie die vorausgegangenen Ausführungen bereits gezeigt haben, gibt es zahlreiche Verbesserungspotenziale bei den Möglichkeiten der Distanzlehre in der HfM Detmold.

Die Entwicklung einer umfangreichen Ausstattung für ein musikspezifisches LMS stellt ein ebenso wichtiges wie langwieriges Unterfangen dar; sie benötigt Jahre. Die Relevanz derartiger Entwicklungen stößt indes in einer Mehrzahl von Gemeinschaften auf ein immer stärker wachsendes Bewusstsein (vgl. Waloschek und Gruhle 2022, zitiert nach: vgl. Nowakowski Juni 2022, 3; *TMOHE Conference – Sharing research-informed practices* 2022).

Die Zielerreichung erfordert aufgrund der Komplexität des Vorhabens ein kleinschrittiges Vorgehen, weswegen sich im Rahmen dieser Arbeit zunächst zwei Funktionen zugewandt wurde, die das Ermöglichen von zuvor beschriebenen Lernszenarien in einem einzigen Interface – dem LMS – unterstützen sollen. Zu erwähnen sei hierbei, dass verschiedene Anwendungen des DTW bereits existieren (mehr in [Kapitel 3](#)), jedoch ist das für eine webbasierte Implementation noch nicht der Fall.

Der DTW dient zur Berechnung eines dynamisch programmierten Alignment-Pfades, welches eine Zuordnung beschreibt, wie zwei Sequenzen nonlinear gestreckt und gestaucht werden müssten, um optimal angeglichen zu werden (vgl. Müller 2015, 131). Dieses erlaubt zum einen, die Ähnlichkeit zwischen zwei Signalen numerisch festzuhalten, weswegen sich DTW auch in einer frühen Form der Spracherkennung zunutze gemacht wurde (vgl. ebd.). Doch es erlaubt auch eine Synchronisation mehrerer Sequenzen – dies ist auch die Intention, die hier verfolgt wird. Ziel ist es zunächst, eine Grundlage zu schaffen, die eine spätere Integration in das LMS ermöglicht.

Die automatische Instrumentenerkennung beruht indes auf maschinellem Lernen und bildet die Vorarbeit zu einer KI-gestützten Tagging-Engine, welche in der Lage ist, Aufnahmen von Instrumenten entsprechende Schlüsselwörter zuzuweisen (vgl. Thierry Bertin-Mahieux et al., 334). Dies soll einerseits eine strukturierte Organisation von Audiodaten ermöglichen und andererseits eine Förderung der Lehre für die kritische Gehörbildung gestatten. Wie beim DTW besteht auch hier die Absicht, vorbereitend eine Grundlage zu erstellen und damit einer Integration in das LMS den Weg zu bereiten.

Da diese Aufgaben im Zeichen der Tonaufnahmen-basierten Lehre stehen und dieser Begriff bereits wiederholt gebraucht wurde, soll zum Verständnis eine knappe Kontextualisierung anhand von Tonaufnahmen folgen.

## 2.1 Begriff: Tonaufnahmen

Tonaufnahmen sind eine Repräsentationsform von Musik. In ihr sind die Amplituden eines Signals in einer Zeitserie gespeichert. Sie basiert somit nicht auf einer

symbolischen Encodierung, wie es etwa bei zwei gängigen Formaten, MEI (vgl. Hankinson et al. 2011) oder MIDI-Dateien der Fall ist (siehe Abb. 1-3).

Diese zwei Formate beinhalten Metadaten, die Aufschluss über die musikalischen Beschaffenheiten eines Stückes geben, d.h. sie encodieren Tempo, Taktangaben, Instrumentierung, sämtliche auftretenden Noten mit zugehörigen Start- und Endzeiten, Lautstärken sowie weitere Möglichkeiten der Artikulation.

Demgegenüber stellen Audioaufnahmen die Encodierung sämtlicher Informationen dar, die benötigt werden, um ein akustisches Phänomen vollständig zu reproduzieren: Temporale, Dynamik- sowie tonale Informationen samt Mikroabweichungen, die eine Live-Performance widerspiegeln (vgl. Müller 2015, 19).

Da hierbei keine Informationen wie Notenanfänge explizit angegeben werden, werden zu Analysezwecken Algorithmen wie der Short-Time-Fast-Fourier-Transform oder Constant-Q-Transform gebraucht, die eine „Übersetzung“ des Signals von der Zeit- in die Frequenzdomäne vornehmen (siehe Abb. 9; vgl. Müller 2015, 39; Brown und Puckette 1992). Müller veranschaulicht die Dekomposition in die einzelnen Frequenzbestandteile mit einem Vergleich zu Prismen, die das Sonnenlicht in die einzelnen Komponenten des Farbspektrums zu zerlegen vermögen.

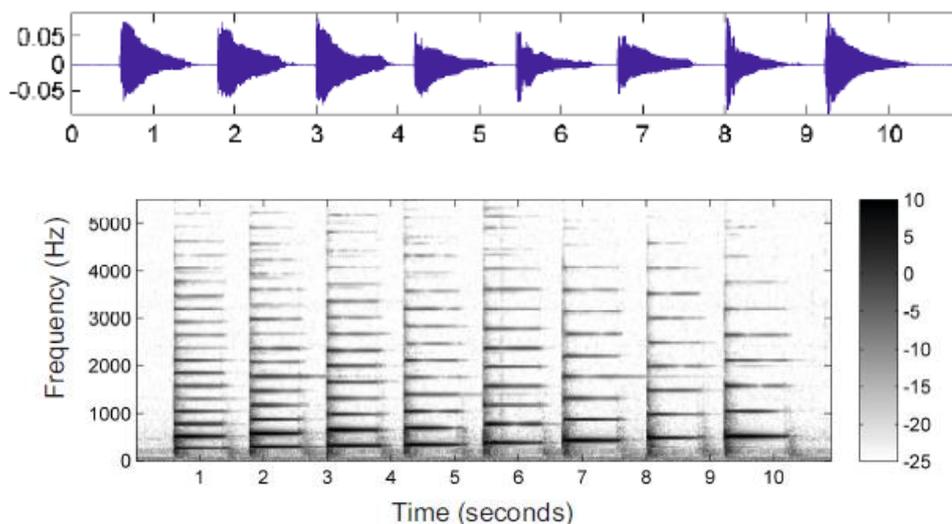


Abb. 9: Audioaufnahme einer C-Dur Tonleiter, auf einem Piano gespielt. Oben: Darstellung in der Zeitdomäne, unten: in der Frequenzdomäne (Müller 2015, 57)

Die Darstellung des Frequenzspektrums bildet die Basis für eine große Bandbreite von datenabhängigen Applikationen.

Aufgaben, die durch sie wahrgenommen werden können, sind bspw. Analysen von Liedstrukturen (vgl. Müller 2015, 167ff), eine Liedererkennung, wie durch die kommerzielle Smartphone-App Shazam bekannt (vgl. Wang 2006), und weitere mit einem höheren Abstraktionsgrad. Darunter fallen etwa die Analyse der Stimmung, des Genres oder der Tanzbarkeit (vgl. Meyers Juni 2007; Jamdar et al. 2015), die wiederum algorithmische Empfehlungen anhand passender Schlüsselwörter in einer Playlist erlauben (vgl. Thierry Bertin-Mahieux et al., 334). Auch generative Applikationen, die zu Kompositionszwecken genutzt werden können, befinden sich im Rahmen des heute Möglichen (vgl. Huang et al. 2020; Ackerman und Loker 2017).

Schließlich verwenden auch die vorliegenden Implementationen des DTW sowie die Instrumentenerkennung die Darstellungen der Frequenz-Domäne zur Zielerreichung.

## 2.2 Designentscheidungen

Dieser Abschnitt behandelt die anwendungsbezogenen Erwägungsgründe, die eine LCD-gerechte Implementation vom DTW und der Instrumentenerkennung zulassen. Die Berücksichtigung der Anforderungen erlaubt eine widerstandsfreie Integration in das LMS.

Da die Anwendungen der Distanzlehre empfänglich gemacht werden sollen, wird die Möglichkeit zur kompatiblen Einbindung in das LMS zur maßgeblichen Bedingung. Wie bereits thematisiert wurde, operieren populäre OSS LMS wie Moodle und ILIAS auf Webbasis. Sie wurden in der Open Source Sprache PHP geschrieben und unterstützen manuelle Funktionserweiterungen, die durch das Plugin H5P in Javascript eingebettet werden können (vgl. Ravalli 2021, 26f). Das beantwortet die Frage, welche Programmiersprache nun den Rahmen für die Werkzeuge bilden soll: Javascript – die „Sprache des Internets“ (vgl. Gans 2020).

Was JS im Kontext dieser Arbeit besonders attraktiv macht, ist die Möglichkeit zur clientseitigen Einbindung und das Ausführen von Code direkt im Browser (vgl. Gray

1999). Das bedeutet für Entwickler, dass kein Server zur Implementation erforderlich ist, was wiederum der Fall wäre, wenn als Programmiersprache z.B. Python gewählt worden wäre. Weiter erspart es den Lernenden die Installation von Drittsoftware und bietet ihnen zugleich einen plattformübergreifenden Zugriff (vgl. ebd). Die Stärken von JS offenbaren sich bei webbasierten Anwendungen in wesentlicher Weise und motivieren daher den Versuch, die Programme entsprechend zu münzen – auch wenn der darunterliegende Aufgabentyp für JS eher unüblich ist (vgl. Gans 2020, 1). Eine Entscheidung zugunsten des clientseitigen Zugangs erhöht daher den Aufwand der Implementation, da mehr Funktionen manuell eingepflogen werden müssen.

Mit Bibliotheken wie `Tensorflow.js`, `Tone.js` und `math.js` lassen sich jedoch einige spezielle Funktionen, die für eine grundlegende DTW Implementation erforderlich sind, sinnvoll abdecken (vgl. `TensorFlow.js` 2022; `Tone.js` 2022; Jong 2022). Letztlich muss der DTW vom Anwender nur einmalig je Repräsentations-Paar ausgeführt werden, um ein Synchronisationsmapping zu berechnen. Aus Gesichtspunkten des LCD ist eine solche Programmgestaltung hinreichend geeignet.

Demgegenüber steht Python – eine Programmiersprache, die aufgrund zahlreicher Implementationen und Bibliotheken besonders für maschinelle Lernanwendungen und wissenschaftliche Berechnungsaufgaben aller Art geeignet ist (vgl. Clough 2023, 1). Hervorzuheben ist die Bibliothek `Librosa`, welche umfassende Funktionalitäten zur Wahrnehmung von Audio-bezogenen Aufgaben mitbringt (vgl. `Librosa` 2022).

`Tensorflow`, das Bindings sowohl für Javascript als auch für Python bereitstellt, stellt derzeit ein populäres Framework zur Durchführung maschineller Lernaufgaben dar (vgl. `TensorFlow` 2022). Das Vorhandensein beider Bindings hat zur Folge, dass es grundsätzlich unerheblich ist, in welcher Sprache die Aufstellung des CNN erfolgt. Selbst für die Merkmalsextraktion sind mit der Bibliothek `essentia.js` ausreichende Möglichkeiten gegeben, um gleichwertige Umwandlungen von Tonaufnahmen in eine Frequenzdarstellung vorzunehmen (vgl. Correya et al. 2021; `essentia.js` 2022). Die in `Tensorflow` gegebene Möglichkeit, in Python gelernte Modelle für Javascript zu exportieren, lässt dies eine Frage der persönlichen Präferenz werden. Vorliegend wurde sich für die Python-Variante entschieden, da bereits Erfahrungswerte mit

Librosa bestehen, ungeachtet dessen bleibt jedoch die Möglichkeit zur Weiterverwendung im LMS bestehen.

Zusammenfassend lässt sich festhalten, dass die konkrete Ausgestaltung dieses Beitrages eine Integration im LMS zulässt, die das Distanzlehren-Angebot der HfM Detmold in förderlicher Weise ergänzt.

### 3 Literatur: ScoreTube

In diesem Kapitel soll die vorliegende Arbeit im Lichte bereits bestehender wissenschaftlicher Forschungsarbeiten betrachtet werden. Der Schwerpunkt liegt hierbei auf der DTW Implementation; als verwandtes Werk lässt sich *ScoreTube* von Berndt et al. heranziehen (vgl. 2018, 6).

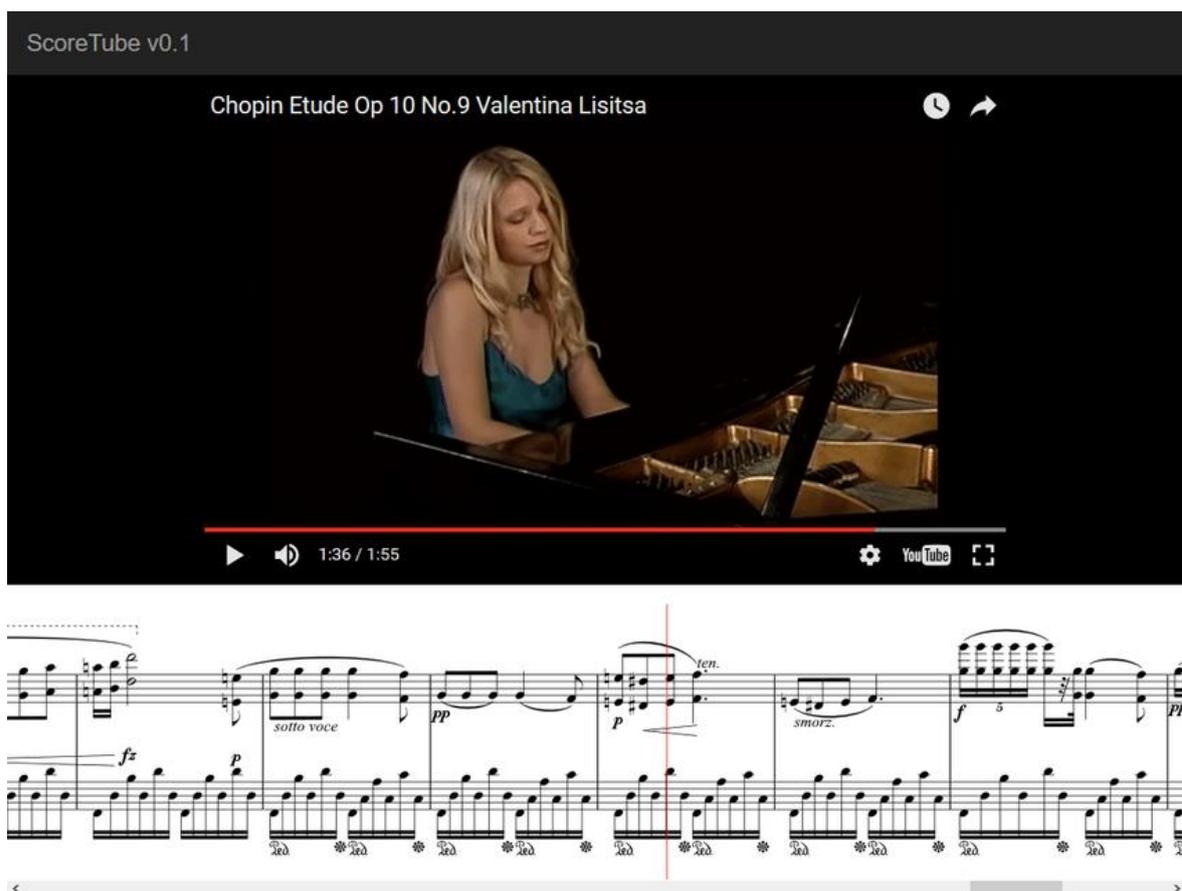


Abb. 10: Screenshot ScoreTube (Berndt et al. 2018)

ScoreTube bezeichnet ein Programm, das einen DTW zwischen einer durch die Anwender:in ausgewählten Audioquelle und einer Partitur (engl. *score*) durchführt und die synchronisierte Wiedergabe anschließend in einer grafischen Oberfläche darstellt (siehe Abb. 10).

Das dabei entstandene Mapping wird genutzt, um eine synchronisierte Anzeige der Videozeitleiste und des Notenblattes zu ermöglichen. Das Audiomaterial wird dabei extern aus einem eingebetteten Youtube-Video bezogen, während die symbolische Darstellung aus einer lokalen MEI Datei gespeist wird.

Es wurde als Anschlussprojekt von *meico* (vgl. ebd.) konzipiert, welches in Java geschrieben wurde. Es ermöglicht eine Reihe von Umwandlungsmöglichkeiten von MEI-Dateien und bewältigt dabei eine Vielzahl an Herausforderungen bei der Interpretation dieser – aufgrund der komplizierten, verwickelten Datenstrukturen sei die Unterstützung bislang begrenzt, obwohl dem Format ein enorm hohes Potenzial für MIR Aufgaben zugrunde liegt, was den Aufwand einer tieferen Integration aufwiegt (vgl. ebd., 1f).

*meico* bietet hierbei keine Möglichkeit einer grafischen Ausgabe, jedoch sei die OSS Verovio dafür bestens geeignet und lässt sich durch den quelloffenen Code in einer integrierten „WebView“ mit den Interpretationsfähigkeiten *meicos* passend kombinieren (vgl. ebd., 5). Die Verwendung des MEI Formates stellt den ersten funktionalen Unterschied zur vorhandenen Implementation dar, denn diese gebraucht als symbolische Darstellung MIDI Daten. Dieses Format ist stärker verbreitet, ist jedoch traditionell auf die musikalisch-praktische Arbeit mit Kompositionen ausgerichtet und bietet längst nicht dieselben detaillierten Encodings von *logischen*, *gesturalen*, *visuellen* und *analytischen* Daten, wie es MEI Daten tun, welche eine kritische Auseinandersetzung mit mehreren Interpretationen derselben Partitur erlauben (vgl. Berndt et al. 2018, 2).

Im Vergleich zu der hiesigen DTW Implementation wird in ScoreTube Gebrauch einer DTW-Variante mit der Beibezeichnung Transposition-Aware (TA DTW) gemacht, welche die Berücksichtigung von Änderungen auf mikrotonaler Ebene, sogenannte *pitch drifts* und *pitch offsets*, ermöglicht. Diese sei vor allem bei der Arbeit mit früheren Musikformen, bspw. mit veränderter Temperierung, oder a capella Musik relevant (vgl. ebd.; Simon Waloschek und Aristotelis Hadjakos 2018).

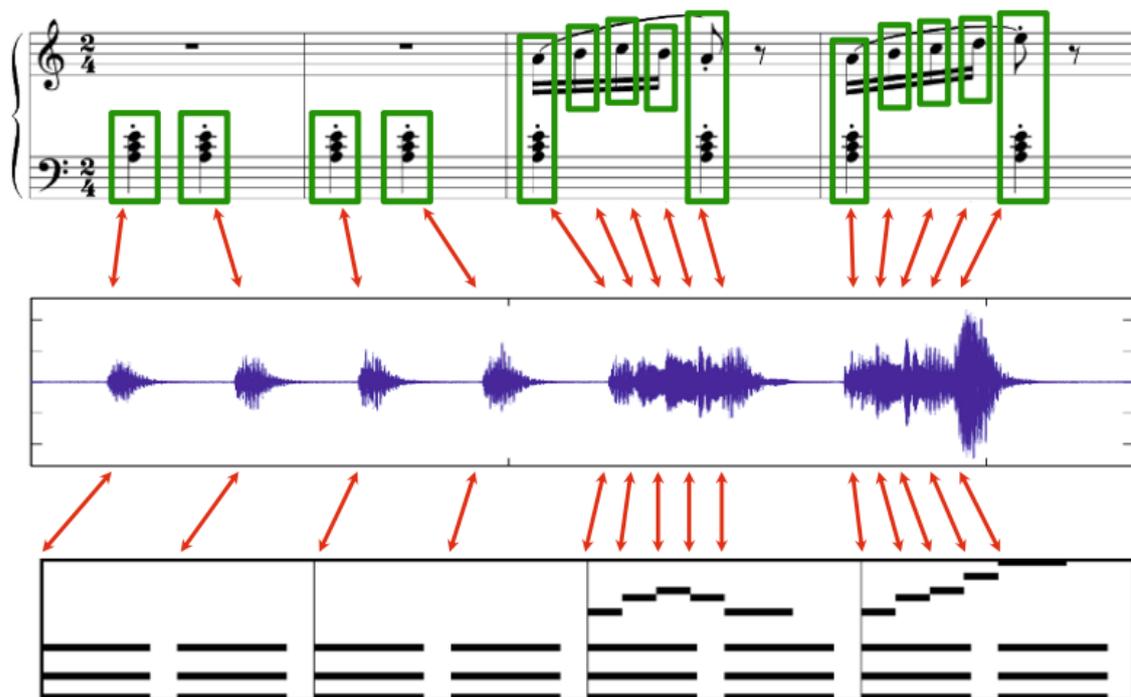


Abb. 11: Synchronisierung von verschiedenen Musikrepräsentationen. V.o.n.u.: Partitur (z.B. MEI), Wellenform einer Tonaufnahme, MIDI (Müller 2015)

Somit zeichnet sich ScoreTube durch seine wissenschaftliche Einsetzbarkeit aus. In der Praxis fand es daher bereits mehrfach Verwendung, etwa als Unterstützung zur Integration digitaler Noten in den Produktionsprozess klassischer Musik (vgl. Waloschek Februar 2017; Waloschek et al. 2016a; Waloschek et al. 2016b).

ScoreTube gebraucht mehrere Programmiersprachen und macht für Anwender die Installation von *Docker Containern* erforderlich: OSS Module, die den Code in virtuelle Einheiten fassen, um sie für fremde Umgebungen portabel zu machen (vgl. Ratliff 05.10.2022). Durch den Zugriff auf Online-Content aus Youtube gilt dies zudem als eine Serveranwendung.

Darin besteht der wesentliche Unterschied zu der vorliegenden DTW Implementation, denn der Code liegt hier in reinem Javascript vor und erfordert für den Einsatz daher keine Container oder Server. Dies ermöglicht bereits eine nahtlose clientseitige Integration in sämtliche webbasierte Umgebungen, in diesem Fall über das Plugin H5P in Lernsysteme wie Moodle oder ILIAS. Durch diese vereinfachten

Integrationsbedingungen wird die Verwendbarkeit in lehrbezogenen Kontexten erhöht, was die Designentscheidungen nochmals bekräftigt.

## 4 Evaluation und Ausblick

Abschließend soll eine Evaluation der Ergebnisse mit einem Ausblick auf mögliche Weiterentwicklungen in der Zukunft erfolgen.

### 4.1 Dynamic Time Warping Audio/Midi

*Github-Link:* <https://github.com/de-tre/dtw-audio-midi.js>

Der DTW wurde erfolgreich in Javascript implementiert und erlaubt die Berechnung von Synchronisationsdaten. Eine Uploadmöglichkeit für Audio und MIDI Dateien ist gegeben – sobald beide Dateien ausgewählt wurden, startet der Berechnungsprozess automatisch. Zur grafischen Ausgabe wurde ein simples Interface in HTML geschrieben, das die Rechenergebnisse plottet, während in der Konsole des Browsers der aktuelle Stand wiedergegeben wird (siehe [Anhang](#)). Nach erfolgter Berechnung genügt ein Klick auf die Schaltfläche „Save DTW.json“, um den Alignment-Pfad lokal abzuspeichern.

Ein nächster geplanter Schritt wäre die Integration in ein LMS durch das Plugin H5P. Dabei könnten die nativen Medienplayer zur Wiedergabe der Tonaufnahmen benutzt werden, während die MIDI Dateien mithilfe von Verovio grafisch gerendert werden. Der native Medienplayer des LMS soll anschließend derart mit der angezeigten symbolischen Darstellung interagieren, sodass sich zu jedem Zeitpunkt einsehen lässt, welcher Zeitcode der Musikwiedergabe welchem Punkt auf der Partitur entspricht. Dabei sollen sich sowohl der Cursor des Musikplayers als auch ein einprogrammierter Cursor im Notenblatt stufenlos auf Klick und Klick-Halten + Ziehen setzen lassen, um die jeweils andere Repräsentation auf synchronisierte Weise mitzusteuern.

Neben der Integration in das LMS gibt es jedoch weitere, dem Algorithmus immanente Entwicklungspotenziale. Diese rühren daher, dass für entsprechende Funktionalitäten nach aktuellem Kenntnisstand keine Bibliotheken in Javascript existieren, die eine bequeme Implementation anspruchsvollerer DTW Varianten ermöglichen würden. In Python wäre dies mit der Sync Toolbox der Fall (vgl. Müller et al. 2021). Daher wären für eine Weiterentwicklung zunächst manuelle Verbesserungen der Rechenprozesse notwendig.

Der DTW ist in seiner klassischen Form rechen- und speicherintensiv. Es besitzt die Berechnungskomplexität  $O(NM)$ , wobei  $N$  die Audiosequenz repräsentiert und  $M$  die MIDI-Sequenz. Mit steigender Länge der Sequenzen ist ein exponentieller Anstieg des Speicherbedarfs zu erwarten (siehe Tabelle 2), der die Durchführung äußerst teuer werden lässt. Bei mittleren bis langen Stücken ist sie mit einem Verbrauch von mehreren GB bis hin zu einem TB und mehr nicht tragbar. Um die Effizienz zu steigern, gibt es daher eine Reihe von Optimierungsmöglichkeiten.

Vorliegend wurde das sogenannte Sakoe-Chiba-Band als Optimierungsschritt gewählt (siehe Abb. 12 (a)). Dieser begrenzt die Anzahl der Berechnungen für die Kostenmatrix, aus welcher das Synchronisationsmapping erschlossen wird, lediglich auf den Bereich innerhalb der Global Constraint Region  $R$  (siehe auch [Anhang b](#)). Dies führt zunächst zu einer Zeitersparnis; so dauerte hier die Durchführung des DTW bei einer 5:19 Minuten langen Aufnahme von *Frederic Chopins Grande Valse Brillante Op. 18 in Eb* und der zugehörigen MIDI Datei 12:45 Minuten.

Der hierfür genutzte Rechner verfügt über einen i7-8750H Prozessor sowie 32 GB RAM-Speicher.

Es gilt jedoch zu berücksichtigen, dass die Form der berechneten Region statisch ist und nicht die Datenstruktur respektiert, was das Risiko einer Verzerrung im Alignment-Pfad birgt (siehe Abb. 12 (c)). Eine fortschrittlichere Optimierung, die datenabhängige Berechnungszonen schafft, nennt sich MrMsDTW (vgl. Pratzlich et al. 2016).

Länge Sequenz $N$ (in Sekunden)	Länge Sequenz $M$ (in Sekunden)	Speicherbedarf	
		DTW (in GB)	MrMsDTW (in MB)
203,2	213,2	0,807	0,763
223,7	193,8	0,808	0,763
444,0	519,2	4,294	0,763
611,5	454,0	5,171	0,763
595,1	563,8	6,249	0,763
862,7	901,1	14,480	0,763
932,6	928,5	16,129	0,763
950,9	817,5	14,479	0,763
7763,3	8204,1	1186,379	0,763
8752,1	8930,3	1455,823	0,763

*Tabelle 2: Beispielhafte Darstellung des Speicherbedarfs eines klassischen DTW und dem MrMsDTW. Exponentieller Anstieg vs. konstanter Speicherverbrauch (vgl. Pratzlich et al. 2016, 3)*

Es beinhaltet eine stufenweise Berechnung des Pfades, wobei die Auflösung der Berechnung zunächst gering ist. Das Ergebnis der niedrigauflösenden Berechnung dient als Maske für die nächste Stufe, die wiederum eine höhere Auflösungsstufe besitzt. Dieser Vorgang kann mehrere Male wiederholt werden.

Im Unterschied zu der MsDTW Variante enthält diese eine Parameterisierung der Größe von Berechnungszonen, und damit des Speicherverbrauchs, der für die dynamische Erstellung von Berechnungszonen zulässig ist (vgl. ebd., 3). Tabelle 2 gibt Aufschluss über die Effizienz dieser Vorgehensweise. Bei einer drastischen Verkürzung der Laufzeit würde gleichzeitig ein optimaler Pfad mit kaum nennenswerten Abweichungen erreicht werden (vgl. ebd., 4).

Des Weiteren ist der Umstand erwähnenswert, dass die Kapazitäten des eingesetzten Endgerätes beeinflussen, ob und wie schnell eine Berechnung des DTW durchgeführt werden kann. Diese Optimierung würde auch einen Einsatz mobiler Geräte ermöglichen (vgl. ebd.), was eine erhebliche Verbesserung des Learner-Centered Designs für die Anwendung zur Folge hätte.

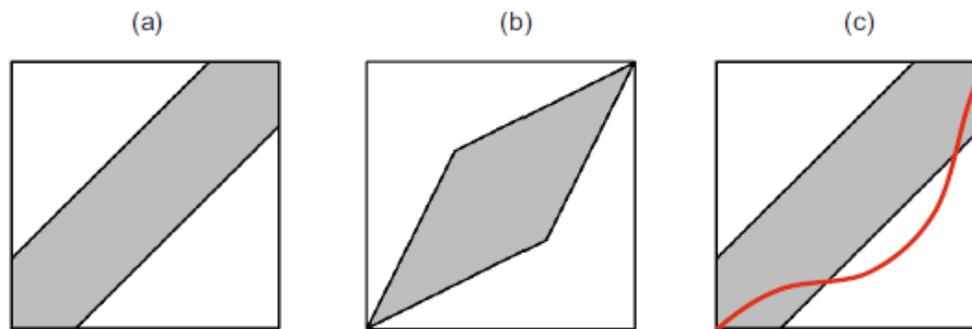


Abb. 12: Schematische Darstellung von Global Constraints.

(a) Sakoe-Chiba-Band,

(b) Itakura-Parallelogramm,

(c) Problemfall: Optimaler Alignment-Pfad liegt außerhalb der Constraint Region (Müller 2015, 144)

Studierende, die Gebrauch von einer synchronisierten Anzeige von Musik im LMS machen wollen, hätten freie Wahl über das verwendete Gerät und wären nicht länger auf leistungstärkere Rechner angewiesen.

Eine weitere Möglichkeit zur Steigerung der Qualität bildet die Einbettung des in [Kapitel 2.1](#) erwähnten Constant-Q-Transform: ein Prozess zur Umwandlung des Tonsignals von der Zeit- in die Frequenzdomäne. Diese würde als Grundlage einer frequenzgetreuen Darstellung des harmonischen Gehaltes dienen.

Die Architektur des CQT ermöglicht im Gegensatz zum STFFT eine variable Zeit-Frequenz-Auflösung, die stets die Periodenlängen der einzelnen Frequenzen über das gesamte Spektrum hinweg berücksichtigt. Dies manifestiert sich im namensgebenden Faktor  $Q$ , der das konstante Verhältnis zwischen Frequenzhöhe und Zeitfenster beschreibt (vgl. Brown und Puckette 1992, 2698). Da tiefe Frequenzen eine längere Periode beinhalten, ist in tieferen Frequenzbereichen eine niedrigere Zeitauflösung gegeben, während mit ansteigender Frequenz die Zeitauflösung ebenfalls steigt (vgl. ebd.).

Beim hier verwendeten STFFT findet die Auflösung zugunsten einer konstanten zeitlichen Auflösung statt, welche aber zur Folge hat, dass Frequenzen ab einer gewissen Periodenlänge nicht mehr abgebildet werden können. Effizienztechnisch erlaubt die Tensorflow.js Implementation lediglich eine zeitliche Auflösung von maximal 2048 Samples pro Frame, bei einer Schrittlänge (Hop-Length) von 512 Samples. Bei höherwertigeren Auflösungen bricht die Prozedur aufgrund eines Memory-

Leaks ab. Innerhalb des Tensorflow.js Frameworks ließ sich diese Situation nur umgehen, wenn auf eine Hardwarebeschleunigung durch die GPU verzichtet wurde – jedoch stark zulasten der Laufzeit. Dieses könnte, bei Verwendung des STFFT, eine Forschungsfrage begründen, die sich mit den Gesichtspunkten des Heap Memory Management und des Multithreading befasst. Die Frequenzauflösung ist für den Zweck dieser Arbeit und auch im Lernszenario jedoch ausreichend.

Weiter den Vergleich zwischen STFFT und CQT ziehend, kommt die lineare Aufteilung der Frequenzachse auf Seiten des STFFTs erschwerend hinzu, denn die harmonischen Beziehungen zwischen musikalisch bedeutsamen Tonhöhen setzen exponentielle Frequenzverhältnisse voraus (vgl. Müller 2015, 21). Nach einer Umstrukturierung der Frequenzbänder in eine logarithmische Skala wird ersichtlich, dass die Frequenzbänder zum unteren Ende des Spektrums hin immer weniger repräsentiert sind, während hohe Frequenzen in der Bestimmbarkeit überrepräsentiert sind, obwohl das menschliche Gehör dort zunehmend schlechter differenzieren kann. Daher wäre diese Beschaffenheit im Zuge menschlicher Übungsaufgaben nicht nötig.

Die Frequenzachse eines CQT folgt hingegen einer geometrischen Anordnung (vgl. Brown und Puckette 1992, 2698). Damit entspricht sie der menschlichen Wahrnehmung musikalischer Tonhöhen. Auf dieser basiert die gleichstufige Stimmung, die im westlichen Tonsystem Verwendung findet (vgl. Müller 2015, 22). Der CQT ist somit in Bezug auf Musik, die diesem System folgt, grundsätzlich vorzuziehen. Über ähnliche Verfahren ist bei anderen Temperierungen nichts bekannt.

Zusammenfassend lässt sich sagen, dass diese Implementation als Erfolg verbucht werden kann, da sie eine funktionale Grundlage für die clientseitige Einbindung bildet. Es gibt jedoch mehrere Möglichkeiten, das Design dieser Anwendung im Lichte des LCD auszubauen, um es in Lernszenarien nutzbar zu machen.

## 4.2 Instrumenten-Klassifikation

*Github-Link:* <https://github.com/de-tre/dl-instrument-classification>

Zunächst ist festzustellen, dass die Implementation der CNN-Architektur mit einer Genauigkeit von 95% geglückt ist. Folgende Eckdaten sind zu nennen, die dieses Ergebnis ermöglichten:

- 13681 Sounds
- 20 Klassen
- 39 Lernepochen mit je 257 Schritten
- 6,9 Mio. erlernte Parameter
- 13 MFCCs als Frequenzdomänen-Komponente

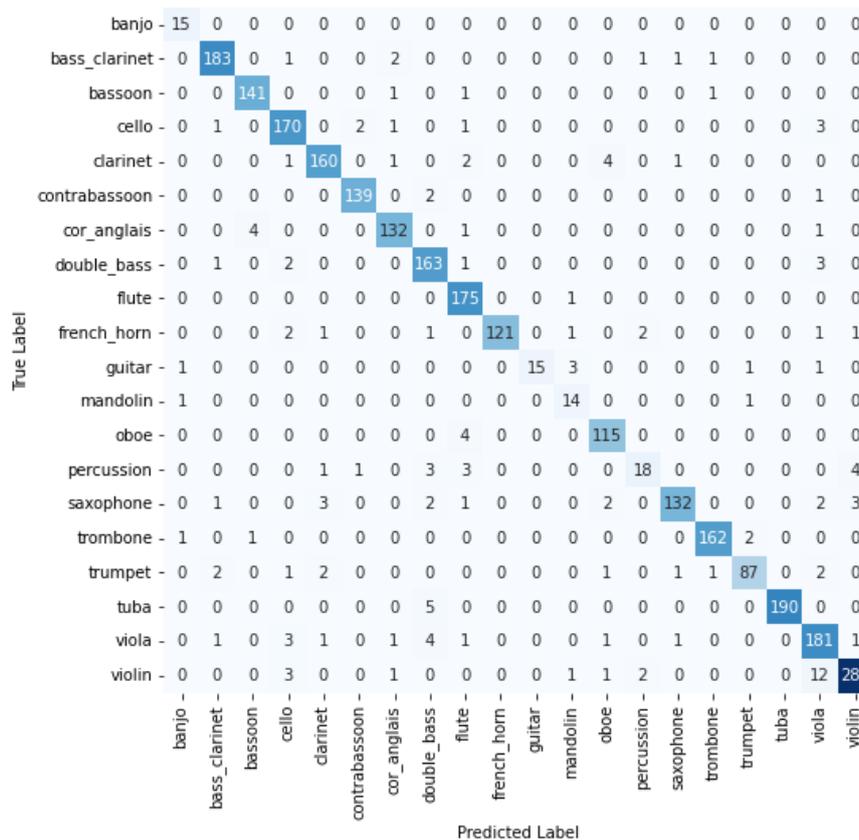


Abb. 13: Confusion Matrix

Dabei lässt sich beobachten, dass die überwiegende Mehrzahl der Daten korrekt zugewiesen wurde. Bei den missglückten Vorhersagen lässt sich eine akustische Nähe zwischen den Paaren feststellen. Dass dieses Modell robust ist,

demonstrieren die Abbildungen 14 und 15. Dem häufigen Problem des Overfittings wurde durch passende Maßnahmen wie Early Stop, Dropout und L2-Regularization ausreichend begegnet.

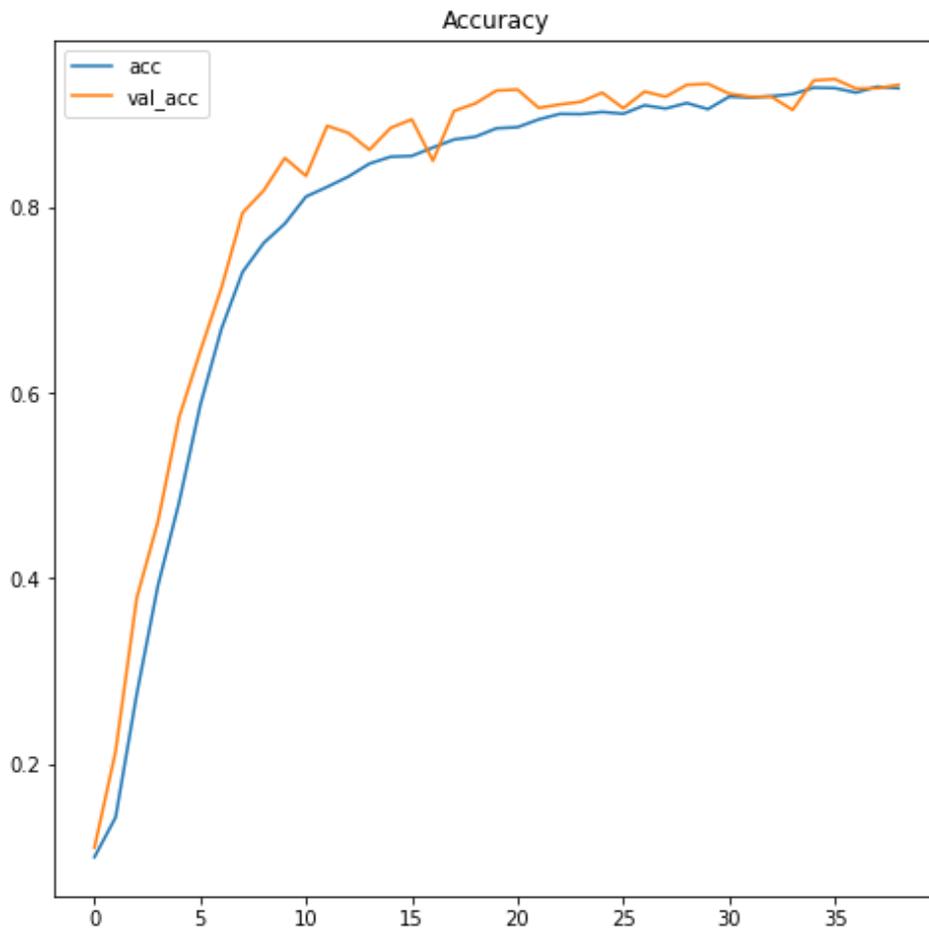


Abb. 14: Accuracy Graph

Die geringen Unterschiede zwischen den Graph-Paaren in der Accuracy sowie der Fehlerquote sind Beleg dafür, dass das Modell generalisierbar ist und die Regularisierungsmaßnahmen daher gegriffen haben. Es erfolgte nach der 36. Lernepoche über die Spanne von 3 weiteren Epochen keine nennenswerte Verbesserung, so dass der Trainingsvorgang nach 39 statt eingeplanten 100 Epochen beendet wurde.

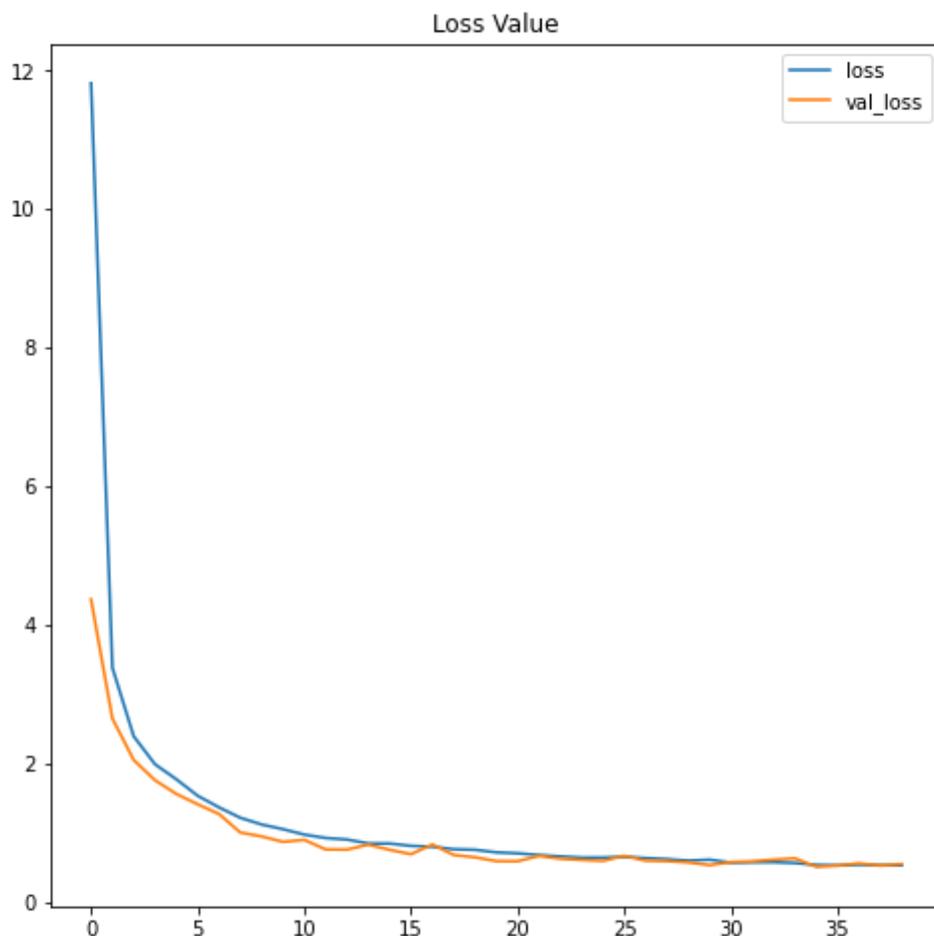


Abb. 15: Loss Graph

Die Möglichkeit konstanter, zuverlässiger Vorhersagen richtet sich auch nach der Kohärenz des verwendeten Datensets, was ein Zugeständnis an die Soundsammlung des Philharmonia Orchesters (vgl. Thorben Dittes) ist. Diese musste zunächst mittels einer weiteren Implementation von MP3 in WAV konvertiert werden, um die Merkmalsextraktion zu beschleunigen. Das zugehörige Notebook ist im Github-Link enthalten.

Das gelernte Modell wurde erfolgreich für Javascript exportiert und ist zur Weiterverwendung im LMS bereit. Dort kann sie zur Konstruktion einer Tagging-Engine verwendet werden, die in der Lage ist, einzelne Instrumente mit passenden Schlüsselwörtern zu versehen. In Folgearbeiten könnte sich einer Multi-Instrumenten Klassifikation zugewandt werden, die partielle Benennungen zulässt (vgl. Anhari 24.01.2020), um auch komplexere Arrangements der Archivierung und Übungen der kritischen Gehörbildung zugänglich zu machen und damit die Verwendung von Tonaufnahmen im LMS der HfM Detmold weiter zu fördern.

## 5 Anhang: Screenshots DTW

### a) Chromas

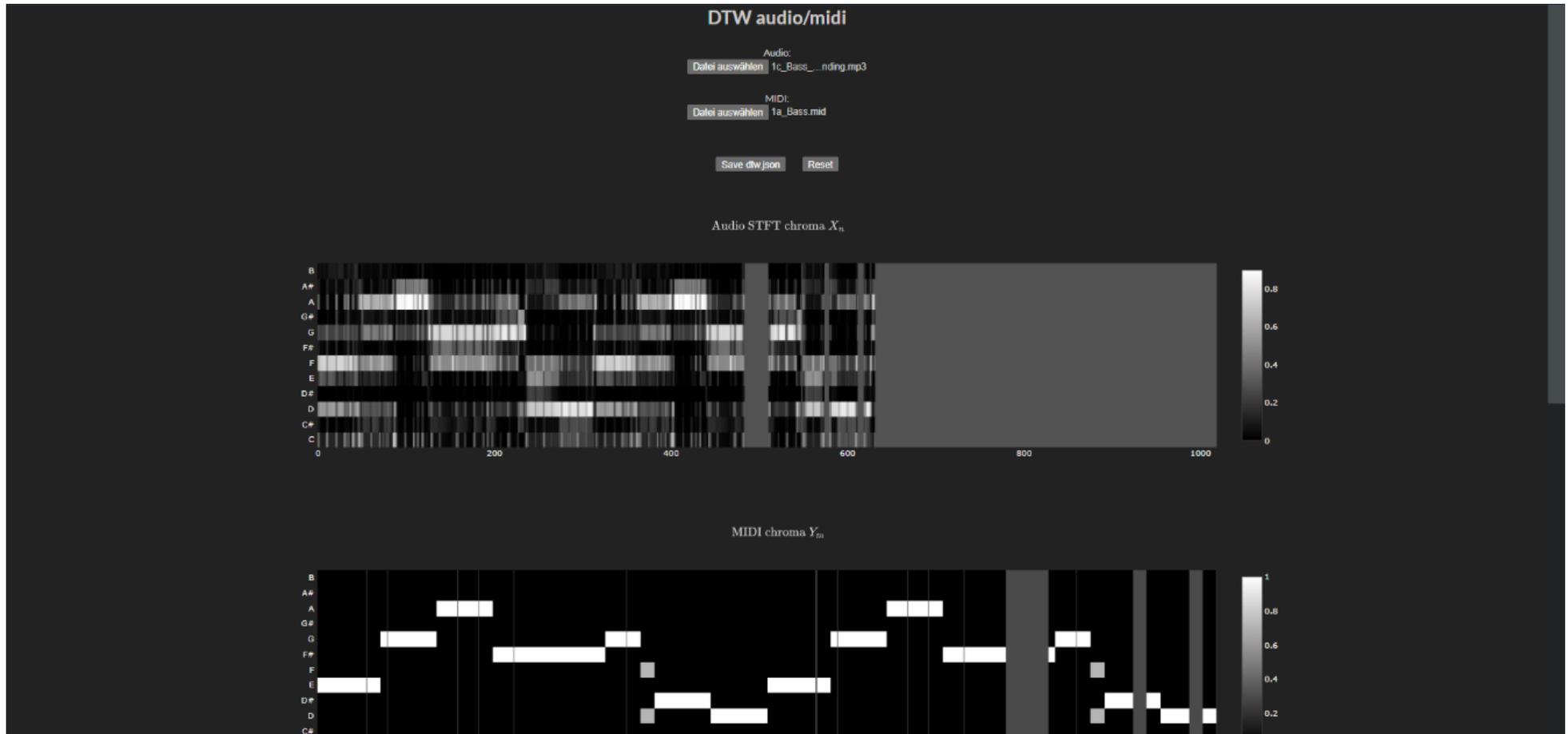


Abb. 16: Berechnung von Chromavektoren, der gewählten Darstellungform der Frequenzdomäne, die als Basis für das Mapping fungiert. Zur Demonstration wurde hier ein extremes Beispiel gewählt, in dem das Ende der Audioaufnahme künstlich mit Stille verlängert wurde.

b) Kostenmatrix mit Alignment-Pfad

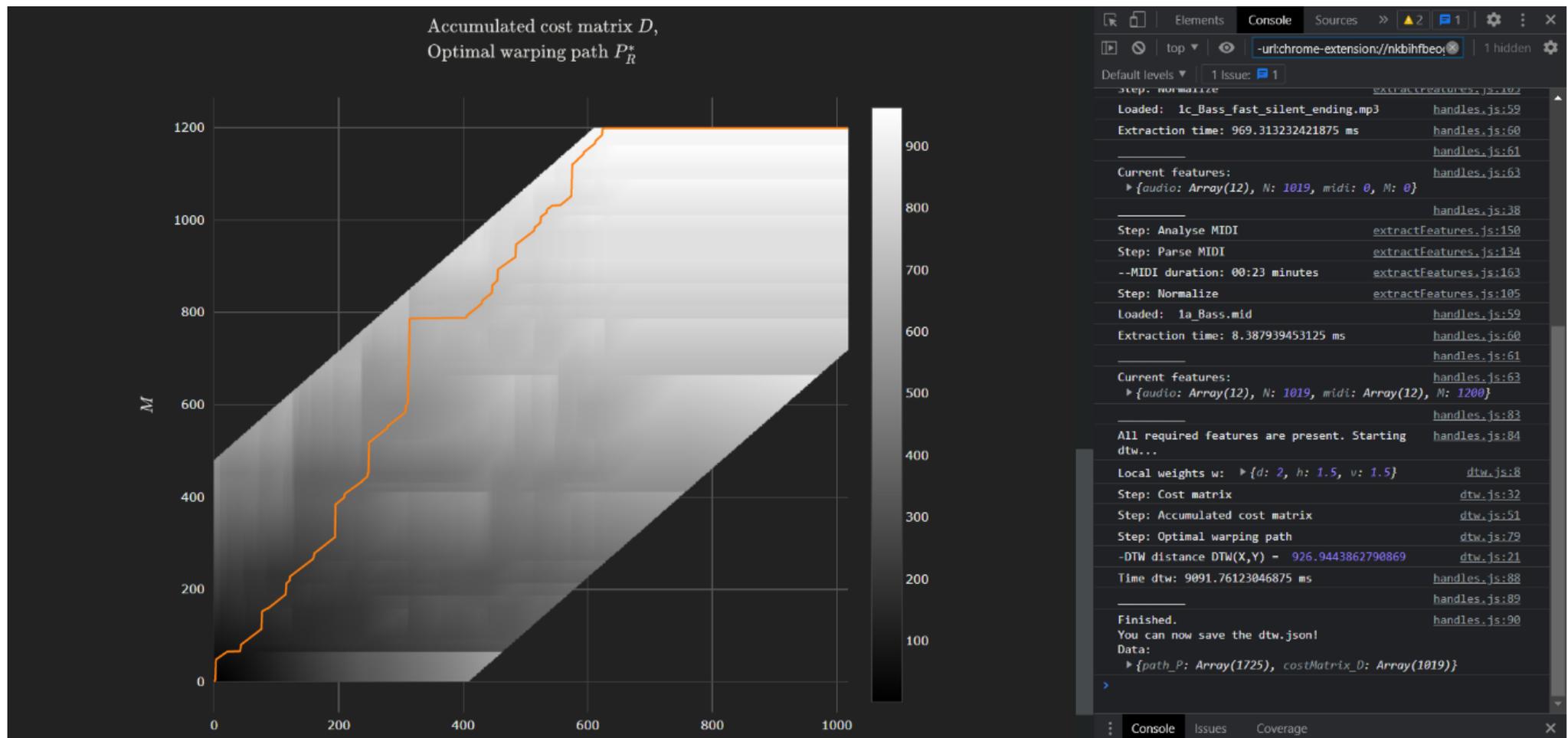


Abb. 17: Plot-Ausgabe des Mappings inkl. Kostenmatrix. Das künstlich verlängerte Ende der Tonaufnahme wird am oberen Rand ersichtlich, indem für diese Periode alle Zeiteinheiten einem einzigen Punkt bei der MIDI Sequenz zugeordnet werden: Dem Ende.

Rechts zu sehen: Ausgaben in der Konsole des Browsers, die die einzelnen Vorgänge des Programmes dokumentieren

## Literaturverzeichnis

- (2014). Badges - MoodleDocs. Online verfügbar unter <https://docs.moodle.org/26/en/Badges> (abgerufen am 15.08.2022).
- (2022). Introduction to MEI. Online verfügbar unter <https://music-encoding.org/guidelines/v4/content/introduction.html> (abgerufen am 18.08.2022).
- (2022). Standards – MoodleDocs. Online verfügbar unter <https://docs.moodle.org/400/de/Standards> (abgerufen am 12.08.2022).
- (2022). TMOHE Conference – Sharing research-informed practices. Online verfügbar unter <http://www.teachingmusiconlineinhighered.com/> (abgerufen am 17.08.2022).
- Ackerman, Margareta/Loker, David (2017). Algorithmic Songwriting with ALYSIA. In: International Conference on Evolutionary and Biologically Inspired Music and Art. Springer, Cham, 1–16.
- Almarashdeh, Ibrahim/Zin, Nor Azan Mat/Alsmadi, Mutasem Khalil Sari (2010). The success of learning management system among distance learners in Malaysian universities. Journal of theoretical and applied information technology 21, 80–91. Online verfügbar unter <http://www.jatit.org/volumes/research-papers/Vol21No2/2Vol21No2.pdf>.
- Anhari, Amir Kenarsari (2020). Learning Multi-instrument Classification with Partial Labels. Online verfügbar unter <https://arxiv.org/pdf/2001.08864>.
- Bailey, Warwick (2005). What is adl scorm.
- Beaven, Tina/Hauck, Mirjam/Comas-Quinn, Anna/Lewis, Tim/Arcos, Beatriz de los (2014). MOOCs: Striking the right balance between facilitation and self-determination. MERLOT Journal of Online Learning and Teaching (10), 31–43. Online verfügbar unter [https://d1wqtxts1xzle7.cloudfront.net/82398663/beaven\\_0314-with-cover-page-v2.pdf?Expires=1660394436&Signature=GyGUI~z4l6zV24EMmgxUboxEhwQp3cac-pxbykRPwWsT3Ti-oTWO2mZCYcNYA7tRFTxnIFOWSKR-pVqcQaLolf-GaSvE3-IgqmYINWuDE3p4vK-K9L-Ovv6eJiNaO-rS-ih4BnpGjnp2R190YesuyxaeY-ioPAeqV2nyiS45X2DzTjd5Ep2V3WzuaJqLv3v3IASv5C0zeLAARC~3OLqQE](https://d1wqtxts1xzle7.cloudfront.net/82398663/beaven_0314-with-cover-page-v2.pdf?Expires=1660394436&Signature=GyGUI~z4l6zV24EMmgxUboxEhwQp3cac-pxbykRPwWsT3Ti-oTWO2mZCYcNYA7tRFTxnIFOWSKR-pVqcQaLolf-GaSvE3-IgqmYINWuDE3p4vK-K9L-Ovv6eJiNaO-rS-ih4BnpGjnp2R190YesuyxaeY-ioPAeqV2nyiS45X2DzTjd5Ep2V3WzuaJqLv3v3IASv5C0zeLAARC~3OLqQE)

8a1JBodpKXKQV0WHvrM3oXtlg--ActllgAv8jdUJgwViXS6lxqioPOXMeE-  
JQzaFNpA7WtQiuztuHGRTEhjQmLqj4yY7IztR1ltQ3-jqKgKlejHI9IE-  
5IPs9KtQvnoqjb8gw\_\_&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA.

- Bernard, Robert M./Borokhovski, Eugene/Schmid, Richard F./Tamim, Rana M./Abrami, Philip C. (2014). A meta-analysis of blended learning and technology use in higher education: from the general to the applied. *Journal of Computing in Higher Education* 26 (1), 87–122. <https://doi.org/10.1007/s12528-013-9077-3>.
- Berndt, Axel/Waloschek, Simon/Hadjakos, Aristotelis (2018). Meico. In: Stuart Cunningham (Hg.). *Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion, AM'18: Sound in Immersion and Emotion*, Wrexham United Kingdom, 12 09 2018 14 09 2018. New York, NY, ACM, 1–7.
- Blaschke, Lisa Marie (2012). Heutagogy and lifelong learning: A review of heutagogical practice and self-determined learning. *The International Review of Research in Open and Distributed Learning* 13 (1), 56. <https://doi.org/10.19173/irrodl.v13i1.1076>.
- Bradley, Vaughn Malcolm (2021). Learning Management System (LMS) Use with Online Instruction. *International Journal of Technology in Education* 4 (1), 68. <https://doi.org/10.46328/ijte.36>.
- Brehmer, Jana/Becker, Sebastian (2017). „E-Learning“. ...ein neues Qualitätsmerkmal der Lehre? Online verfügbar unter [https://www.uni-goettingen.de/de/document/download/4ddb291d1e0c485230df5cd20f3b6b9b.pdf/09\\_E-Learning.pdf](https://www.uni-goettingen.de/de/document/download/4ddb291d1e0c485230df5cd20f3b6b9b.pdf/09_E-Learning.pdf) (abgerufen am 09.08.2022).
- Bremer, Claudia (2013). Massive Open Online Courses 3, 30–48. Online verfügbar unter <https://d-nb.info/108242501x/34>.
- Brown, Judith C./Puckette, Miller S. (1992). An efficient algorithm for the calculation of a constant Q transform. *The Journal of the Acoustical Society of America* 92 (5), 2698–2701. <https://doi.org/10.1121/1.404385>.
- Brügge, Bernd (2004). *Open-Source-Software. Eine ökonomische und Technische Analyse*. Berlin, Heidelberg, Springer Berlin / Heidelberg.

- Canning, Natalie (2010). Playing with heutagogy: exploring strategies to empower mature learners in higher education. *Journal of Further and Higher Education* 34 (1), 59–71. <https://doi.org/10.1080/03098770903477102>.
- Cannon, Joseph P./Lohtia, Ritu/Paulich, Brianna JeeWon (2021). Blended Learning in Principles of Marketing: The Effects of Student Differences on Student Performance. *Journal of Marketing Education*, 02734753211058357. <https://doi.org/10.1177/02734753211058357>.
- Capone, Roberto/Caterina, Patrizia de/Mazza, Giustina Anna Gerarda (2017). BLENDED LEARNING, FLIPPED CLASSROOM AND VIRTUAL ENVIRONMENT: CHALLENGES AND OPPORTUNITIES FOR THE 21ST CENTURY STUDENTS. In: EDULEARN proceedings. IATED.
- Chaubey, Aabha/Bhattacharya, Bani (2015). Learning Management System in Higher Education. *IJSTE - International Journal of Science Technology & Engineering* | 2, 158–162. Online verfügbar unter [https://www.researchgate.net/profile/Aabha-Chaubey-2/publication/288737185\\_Learning\\_Management\\_System\\_in\\_Higher\\_Education/links/56d96fc108aeb4638bb8855/Learning-Management-System-in-Higher-Education.pdf](https://www.researchgate.net/profile/Aabha-Chaubey-2/publication/288737185_Learning_Management_System_in_Higher_Education/links/56d96fc108aeb4638bb8855/Learning-Management-System-in-Higher-Education.pdf).
- Chen, Cheng-Chia/Huang, ChingChih/Gribbins, Michele/Swan, Karen (2018). Gamify Online Courses with Tools Built into Your Learning Management System (LMS) to Enhance Self-Determined and Active Learning. *Online Learning* 22 (3), 41–54. Online verfügbar unter <https://eric.ed.gov/?id=ej1191480>.
- Chris Quintana/Joseph Krajcik/Elliot Soloway (2000). Exploring a Structured Definition for Learner-Centered Design. In: Barry Jay Fishman/Samuel F. O'Connor-Divelbiss (Hg.). *Proceedings of ICLS 2000*. The University of Michigan, Ann Arbor, Michigan, USA, June 14th-17th, 2000. Mahwah, NJ, Lawrence Erlbaum Associates, 268–275.
- Clough, David E. (2023). *Introduction to Engineering and Scientific Computing with Python*. Milton, Taylor & Francis Group.
- Coates, Hamish/James, Richard/Baldwin, Gabrielle (2005). A Critical Examination Of The Effects Of Learning Management Systems On University Teaching And Learning. *Tertiary Education and Management* 11 (1), 19–36. <https://doi.org/10.1007/s11233-004-3567-9>.

- Correya, Albin/Marcos-Fernández, Jorge/Joglar-Ongay, Luis/Alonso-Jiménez, Pablo/Serra, Xavier/Bogdanov, Dmitry (2021). Audio and Music Analysis on the Web using Essentia.js. 2514-3298 4 (1), 167–181. <https://doi.org/10.5334/tismir.111>.
- Eberle, Jane (2009). Heutagogy: What your mother didn't tell you about pedagogy and the conceptual age. In: Dan Remenyi (Hg.). The proceedings of the 8th European Conference on e-Learning. University of Bari, Italy, 29 - 30 October 2009. Reading, UK, Acad. Publ, 181–188.
- Ellis, Ryann K. (2009). Learning management systems. Online verfügbar unter [http://cit.eservices.gov.za/dhethnols/sites/default/files/2020-02/LMS\\_field-guide\\_20091\\_4.pdf](http://cit.eservices.gov.za/dhethnols/sites/default/files/2020-02/LMS_field-guide_20091_4.pdf).
- essentia.js (2022). essentia.js. Online verfügbar unter <https://mtg.github.io/essentia.js/> (abgerufen am 19.08.2022).
- European Commission (2021). A framework for Blended Learning. Commission Staff Working Document SWD (219). Online verfügbar unter [https://www.schooleducationgateway.eu/downloads/files/news/framework\\_for\\_blended\\_learning.pdf](https://www.schooleducationgateway.eu/downloads/files/news/framework_for_blended_learning.pdf).
- Gans, Maya (2020). JavaScript for data science. Boca Raton, CRC Press.
- Gray, Patricia (1999). JavaScript: Convenient Interactivity for the Class Web Page. For full text: <<http://www.mtsu.edu/~itconf/proceed99/gray.html>>.
- Hadjakos, Prof. Dr. Aristotelis. Research – Center of Music and Film Informatics. Online verfügbar unter <http://www.cemfi.de/research>.
- Halasa, Suhaila/Abusalim, Nimer/Rayyan, Mohammad/Constantino, Rose E./Nassar, Omayah/Amre, Huda/Sharab, Moayad/Qadri, Insirah (2020). Comparing student achievement in traditional learning with a combination of blended and flipped learning. Nursing Open 7 (4), 1129–1138. <https://doi.org/10.1002/nop2.492>.
- Hankinson, Andrew/Roland, Perry/Fujinaga, Ichiro (2011). The Music Encoding Initiative as a Document-Encoding Framework.
- Harrington, Charles F./Gordon, Scott A./Schibik, Timothy J. (2004). Course Management System Utilization and Implications for Practice: A National Survey of Department Chairpersons. Online Journal of Distance Learning Administration

- 7 (4). Online verfügbar unter <https://ojdla.com/archive/winter74/harrington74.pdf>.
- Hase, Stewart/Kenyon, Chris (2000). From Andragogy to Heutagogy. Online verfügbar unter <http://pandora.nla.gov.au/nph-wb/20010220130000/http://ultimatebase.rmit.edu.au/Articles/dec00/hase2.htm> (abgerufen am 21.02.2001).
- Huang, Cheng-Zhi Anna/Koops, Hendrik Vincent/Newton-Rex, Ed/Dinculescu, Monica/Cai, Carrie J. (2020). AI Song Contest: Human-AI Co-Creation in Songwriting. arXiv.
- In, Junyong (2017). Introduction of a pilot study. *Korean Journal of Anesthesiology* 70 (6), 601–605. <https://doi.org/10.4097/kjae.2017.70.6.601>.
- Jamdar, Adit/Abraham, Jessica/Khanna, Karishma/Dubey, Rahul (2015). Emotion Analysis of Songs Based on Lyrical and Audio Features. arXiv.
- Jong, Jos de (2022). math.js. Online verfügbar unter <https://mathjs.org/> (abgerufen am 19.08.2022).
- Jost, Nathanael S./Jossen, Sina L./Rothen, Nicolas/Martarelli, Corinna S. (2021). The advantage of distributed practice in a blended learning setting. *Education and Information Technologies* 26 (3), 3097–3113. <https://doi.org/10.1007/s10639-020-10424-9>.
- Kerres, Michael (2013). *Mediendidaktik. Konzeption und Entwicklung mediengestützter Lernangebote*. 4. Aufl. München, De Gruyter.
- Labaj, Martin/Bieliková, Mária (2014). Utilization of Exercise Difficulty Rating by Students for Recommendation. In: Yiwei Cao/Terje Väljataga/Jeff K.T. Tang et al. (Hg.). *New Horizons in Web Based Learning*, Cham, 2014. Cham, Springer International Publishing, 13–22.
- Leonhard, Joachim-Felix/Ludwig, Hans-Werner/Schwarze, Dietrich/Straßner, Erich (1999). *Medienwissenschaft. Ein Handbuch zur Entwicklung der Medien und Kommunikationsformen*. *Medienwissenschaft: ein Handbuch zur Entwicklung der Medien und Kommunikationsformen* 15 (1), 118. Online verfügbar unter <https://books.google.de/books?id=NHhwW7fladcC&printsec=frontcover&hl=de#v=onepage&q&f=false>.
- Librosa (2022). GitHub - librosa/librosa: Python library for audio and music analysis. Online verfügbar unter <https://github.com/librosa/librosa> (abgerufen am 19.08.2022).

- Mackness, Jenny/Mak, Sui Fai John/Williams, Roy (2010). The Ideals and Reality of Participating in a MOOC, 266–274. Online verfügbar unter [https://research-portal.port.ac.uk/files/108952/the\\_ideals\\_and\\_reality\\_of\\_participating\\_in\\_a\\_mooc.pdf](https://research-portal.port.ac.uk/files/108952/the_ideals_and_reality_of_participating_in_a_mooc.pdf).
- Martin, Fred G. (2012). Education: Will Massive Open Online Courses Change How We Teach? *Communications of the ACM* (55), 26–28. Online verfügbar unter [https://dl.acm.org/doi/fullHtml/10.1145/2240236.2240246?casa\\_token=Fa-JEnOluancAAAA:nLs2mH0u1szye8z7jWuFqpLLV\\_pjdbijfv7ER-VaDN8Dm9I4uxhiNvOgPtruThJoJwHge8sOvctg](https://dl.acm.org/doi/fullHtml/10.1145/2240236.2240246?casa_token=Fa-JEnOluancAAAA:nLs2mH0u1szye8z7jWuFqpLLV_pjdbijfv7ER-VaDN8Dm9I4uxhiNvOgPtruThJoJwHge8sOvctg).
- Means, Barbara/Toyama, Yukie/Murphy, Robert/Baki, Marianne (2013). The Effectiveness of Online and Blended Learning: A Meta-Analysis of the Empirical Literature. *Teachers College Record* 115 (3), 1–47. <https://doi.org/10.1177/016146811311500307>.
- Meenakshi, Kaushik/Ravindra, Kumar/Manu, Kaushik (2022). A Conceptual Study of Design Thinking. *An Overview in Present Perspective* (4), 15–29. <https://doi.org/10.13140/RG.2.2.22960.23048>.
- Mester, Gyula (2016). Massive Open Online Courses in Education of Robotics. *Interdisciplinary Description of Complex Systems : INDECS* 14 (2), 182–187. <https://doi.org/10.7906/indecs.14.2.7>.
- Meyers, Owen Craigie (2007). A mood-based music classification and exploration system. Massachusetts Institute of Technology. Online verfügbar unter <https://dspace.mit.edu/handle/1721.1/39337>.
- Mohd Kasim, Nurul Nadirah/Khalid, Fariza (2016). Choosing the Right Learning Management System (LMS) for the Higher Education Institution Context: A Systematic Review. *International Journal of Emerging Technologies in Learning (iJET)* 11 (06), 55. <https://doi.org/10.3991/ijet.v11i06.5644>.
- Müller, Meinard (2015). *Fundamentals of music processing: Audio, analysis, algorithms, applications*.
- Müller, Meinard/Özer, Yigitcan/Krause, Michael/Prätzlich, Thomas/Driedger, Jonathan (2021). Sync Toolbox: A Python Package for Efficient, Robust, and Accurate Music Synchronization. *Journal of Open Source Software* 6 (64), 3434. <https://doi.org/10.21105/joss.03434>.

- Naredla, Suman Kumar/Shekar, Raja/Babu, Ramesh/Condoor, Sridhar (2018). Uniquely addressing customer pain points-The case study of AgRITEch app. *International Journal of Mechanical Engineering and Technology (IJMET)* (9). Online verfügbar unter [https://www.researchgate.net/profile/ramesh-babu-d/publication/329556890\\_uniquely\\_addressing\\_customer\\_pain\\_points\\_-\\_the\\_case\\_study\\_of\\_agritech\\_app/links/5c0f6cc0299bf139c74fe033/uniquely-addressing-customer-pain-points-the-case-study-of-agritech-app.pdf](https://www.researchgate.net/profile/ramesh-babu-d/publication/329556890_uniquely_addressing_customer_pain_points_-_the_case_study_of_agritech_app/links/5c0f6cc0299bf139c74fe033/uniquely-addressing-customer-pain-points-the-case-study-of-agritech-app.pdf).
- Nicholson, Paul (2007). A History of E-Learning. Echoes of the pioneers. In: Baltasar Fernández-Manjón/Juan Manuel Sánchez-Pérez (Hg.). *Computers and education. E-learning, from theory to practice ; [selected and improved papers presented in the 4th [i.e. 6th] International Symposium on Educational Informatics (SIIIE in Spanish) hold in Cáceres (Spain) on November of 2004*. Dordrecht, Springer, 1–11.
- Nowakowski, Matthias (2022). *Interactive Notation for LMS: Progress Report and Pilot Study*. Center of Music and Film Informatics. Detmold. Online verfügbar unter persönliche Korrespondenz.
- Open Source-LMS (2020). *Freie Software für freie Lehre! Gemeinsame Erklärung der deutschen, communitygestützten Open-Source-Bildungsplattformen*. Online verfügbar unter <https://opensource.lms.de/> (abgerufen am 11.08.2022).
- Paganini, Claudia/Steinbacher, Christoph (2019). *Welt am Abgrund. Zukunft zwischen Bedrohung und Vision*, 181–199. Online verfügbar unter <https://diglib.uibk.ac.at/ulbtirolfodok/periodical/titleinfo/7117899/full.pdf>.
- Page, Janelle/Meehan-Andrews, Terri/Weerakkody, Nivan/Hughes, Diane L./Rathner, Joseph A. (2017). Student perceptions and learning outcomes of blended learning in a massive first-year core physiology for allied health subjects. *Advances in Physiology Education* 41 (1), 44–55. <https://doi.org/10.1152/advan.00005.2016>.
- Pappano, Laura (2012). *The Year of the MOOC*. *New York Times*. Online verfügbar unter <https://www.nytimes.com/2012/11/04/education/edlife/massive-open-online-courses-are-multiplying-at-a-rapid-pace.html>.
- Pfänder, Stefan/Wagner, Jörg (2009). *Warum wir sprechen, wenn wir klicken. Die Mensch-Maschine-Interaktion als zukunftsweisendes Themenfeld für sprach- und medienwissenschaftliche Module in romanistischen Studiengängen*

- (FrankoMedia, CreatingCultures, Freiburg und Interkulturelle Wissenskommunikation, Halle). Freiburg, Halle.
- Piña, Anthony A. (2012). An Overview of Learning Management Systems. In: Information Resources Management Association (Hg.). Virtual learning environments. Concepts, methodologies, tools and applications. Hershey, Pa, IGI Global (701 E. Chocolate Avenue Hershey Pennsylvania 17033 USA), 33–51.
- Pratzlich, Thomas/Driedger, Jonathan/Muller, Meinard (2016). Memory-restricted multiscale dynamic time warping. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE.
- Ratliff, James (2022). Home - Docker. Docker vom 05.10.2022. Online verfügbar unter <https://www.docker.com/> (abgerufen am 19.08.2022).
- Ravalli, Paolo (2021). Apps für Fachexperten: Betriebliches Lernen mit H5P. Wissensmanagement 3 (6), 26–28. <https://doi.org/10.1007/s43443-021-0316-6>.
- Razo, Driveth (2020). The Rebirth of Online Courses in the COVID-19 Era. Observatory | Institute for the Future of Education vom 17.09.2020. Online verfügbar unter <https://observatory.tec.mx/edu-news/boom-of-online-courses-covid-era> (abgerufen am 14.08.2022).
- Robinson, David/Ash, Patricia (2014). Developing a pedagogical model for a massive open online course (MOOC). European Journal of Science and Mathematics Education 2 (2A), 131–135. <https://doi.org/10.30935/scimath/9635>.
- Rost, Maria Christina (2020). Stellungnahme des Hessischen Beauftragten für Datenschutz und Informationsfreiheit zum Einsatz von Microsoft Office 365 in hessischen Schulen. Online verfügbar unter <https://datenschutz.hessen.de/pressemitteilungen/stellungnahme-des-hessischen-beauftragten-f%C3%BCr-datenschutz-und> (abgerufen am 09.08.2022).
- Sagayama, Shigeki/Takahashi, Keigo/Kameoka, Hirokazu/Nishimoto, Takuya (2004). Specmurt analysis: A piano-roll-visualization of polyphonic music signal by deconvolution of log-frequency spectrum.
- Schmid, Ulrich/Goertz, Lutz/Behrens, Julia (2018). Monitor Digitale Bildung: Die Weiterbildung im digitalen Zeitalter. Online verfügbar unter [https://www.berthelmann-stiftung.de/fileadmin/files/user\\_upload/BSt\\_Monitor\\_Digitale\\_Bildung\\_WB\\_web.pdf](https://www.berthelmann-stiftung.de/fileadmin/files/user_upload/BSt_Monitor_Digitale_Bildung_WB_web.pdf) (abgerufen am 13.08.2022).

- Schulmeister, Rolf (2005). Lernplattformen für das virtuelle Lernen. 2. Aufl. München/Oldenbourg, Oldenbourg Wissenschaftsverlag.
- Simon Waloschek/Aristotelis Hadjakos (2018). Driftin' Down the Scale: Dynamic Time Warping in the Presence of Pitch Drift and Transpositions. In: ISMIR.
- Stärk, André (2013). HfM Detmold. Online verfügbar unter <https://detmoldmusic-tools.de/>.
- Sumardi, Dody/Suryani, Nunuk/Musadad, Akhmad Arif (2021). Website-Based Learning Management System (LMS) as a Tool for Learning in the Covid-19 Pandemic Period for Junior High Schools. *Journal of Education Technology* 5 (3), 346. <https://doi.org/10.23887/jet.v5i3.38371>.
- Swacha, Jakub (2014). An Architecture of a Gamified Learning Management System. In: Yiwei Cao/Terje Väljataga/Jeff K.T. Tang et al. (Hg.). *New Horizons in Web Based Learning*, Cham, 2014. Cham, Springer International Publishing, 195–203.
- Tayebinik, Maryam/Puteh, Marlia (2012). Blended Learning or E-learning? Tayebinik. Online verfügbar unter <https://arxiv.org/ftp/arxiv/papers/1306/1306.4085.pdf>.
- TensorFlow (2022). TensorFlow. Online verfügbar unter <https://www.tensorflow.org/> (abgerufen am 19.08.2022).
- TensorFlow.js (2022). TensorFlow.js | Machine Learning for JavaScript Developers. Online verfügbar unter <https://www.tensorflow.org/js> (abgerufen am 19.08.2022).
- Thierry Bertin-Mahieux/Douglas Eck/Michael Mandel. Automatic Tagging of Audio: The State-of-the-Art. In: *Machine Audition: Principles, Algorithms and Systems*. IGI Global, 334–352.
- Thorben Dittes. Philharmonia. Online verfügbar unter <https://philharmonia.co.uk/resources/sound-samples/>.
- Tone.js (2022). GitHub - Tonejs/Midi: Convert MIDI into Tone.js-friendly JSON. Online verfügbar unter <https://github.com/Tonejs/Midi> (abgerufen am 19.08.2022).

- Verovio. A music notation engraving library. Designed for the Music Encoding Initiative - Fast - Light - Flexible - No dependencies. Input formats. Online verfügbar unter <https://book.verovio.org/toolkit-reference/input-formats.html>.
- Waldrop, M. Mitchell (2013). Online learning: Campus 2.0. *Nature* 495 (7440), 160–163. <https://doi.org/10.1038/495160a>.
- Waloschek, S./Berndt, A./Bohl, B. W./Hadjakos, A. (2016a). Interactive Scores in Classical Music Production. In: *Int. Society for Music Information Retrieval (Hg.). Proc. of the 17th Int. Society for Music Information Retrieval Conf. (ISMIR)*. New York, NY, 668–673.
- Waloschek, Simon (2017). *Interaktive Partituren in der Musikproduktion*. Masterthesis. Ostwestfalen-Lippe University of Applied Sciences. Online verfügbar unter <http://www.cemfi.de/wp-content/papercite-data/pdf/waloschek-2017-ipmp.pdf> (abgerufen am 19.08.2022).
- Waloschek, Simon/Berndt, Axel/Bohl, Benjamin/Hadjakos, Aristotelis (2016b). Accelerating the Editing Phase in Music Productions using Interactive Scores.
- Wang, Avery (2006). The Shazam music recognition service. *Communications of the ACM* 49 (8), 44–48. <https://doi.org/10.1145/1145287.1145312>.
- Weidenmann, Bernd (2009). *Multimedia, multicodierung und multimodalität beim online-lernen*. Online lernen. Handbuch für wissenschaft und praxis. Oldenbourg.
- Whitelock, Denise/Jelfs, Anne (2003). Editorial: Journal of educational media special issue on blended learning. *Journal of Educational Media* (28), 99–100. Online verfügbar unter [https://www.researchgate.net/publication/285486324\\_Editorial\\_Journal\\_of\\_educational\\_media\\_special\\_issue\\_on\\_blended\\_learning](https://www.researchgate.net/publication/285486324_Editorial_Journal_of_educational_media_special_issue_on_blended_learning).

## Eigenständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken (dazu zählen auch Internetquellen) entnommen sind, wurden unter Angabe der Quelle kenntlich gemacht.

A handwritten signature in blue ink, appearing to be 'D. Paul', written in a cursive style.

Detmold, den 19.08.2022