

Dissertation im Fachgebiet Musikwissenschaft

vorgelegt von Karl-Ulrich Becker, M. A.

Titel:

Themen und Methoden in der
Populärmusikforschung. Eine empirische
Fachgeschichte auf Basis des Datenbestands von RILM.

Inhaltsverzeichnis

Abbildungsverzeichnis

Tabellenverzeichnis

1 Einleitung

„Forschung ist immer das Weiterforschen, wo andere aufgehört haben, das Weiterbauen auf Grundsteinen und Gerüsten, die andere vorbereitet haben“

– Hubert S. Markl

Dieser Aussage des ehemaligen Präsidenten der DFG und der Max-Planck-Gesellschaft Hubert Markl folgend ist die Voraussetzung für erfolgreiche Forschung das Bewusstsein für die Fundamente, auf denen sie aufbaut. Da Forschung aber auch, wie MARKL fortführt, „[...] das Weitergehen auf Irrwegen, die andere eingeschlagen haben“¹ bedeuten kann, ist das Wissen um die historische Perspektive des eigenen Fachs sehr wichtig. Denn nur wer die Geschichte seiner Disziplin kennt, kann sich und die eigene Arbeit in eine Forschungstradition einordnen.

Der Zweig der Musikwissenschaft, der sich der populären Musik,² der „repräsentative[n] Musik der Gegenwart“,³ widmet, ist noch nicht gefestigt. Sowohl die Musikwissenschaft als auch andere Kultur- und Gesellschaftswissenschaften setzen sich schon lange mit populärer Musik auseinander. Eine Standardisierung zu einem eigenständigen Forschungsgebiet aber, dessen Ziele, Methoden und Themen klar definiert sind, zu einer konkretisierten „Populärmusikforschung“ also, ist noch nicht vollzogen. Eine Ursache dafür ist das Fehlen einer Historiographie der Populärmusikforschung, auf deren Basis die historisch fundierte Reflexion der Fachgeschichte erfolgen kann, die eine Konsolidierung des Fachgebiets ermöglicht.

Diese Arbeit wird versuchen, einen Beitrag zur Fachgeschichte der Populärmusikforschung zu leisten und die bestehende Forschungslücke zu schließen.

Methodische Forschungslücke weil bestehende Ansätze in der Fachgeschichtsschreibung unzureichend sind!

Die Konstruktion von (Wissenschafts-)Geschichte wird entweder als Erlebte-Geschichte (im Sinne einer Oral-History durch Beteiligte; auch: Great-Man-Ansatz) oder auf Basis von Quellen vorgenommen. Quellen sind i. d. R. Sekundärquellen - Im Falle einer wissenschaftlichen Disziplin sind die veröffentlichten Forschungsbeiträge aller WissenschaftlerInnen, die sich zu dem zu untersuchenden Sachgebiet zuordnen lassen, eigentlich als Primärquellen zu bewerten.

¹ Ebd.

² Zum Begriff der „populären Musik“ vgl. ??

³ (Pfleiderer2012)

Gängige Beschreibungen der Entwicklung einer wissenschaftlichen Disziplin beziehen sich jedoch nur auf einen Ausschnitt aus der Gesamtmenge der Veröffentlichungen, auf einige einflussreiche Publikationen und Denktraditionen, die daraus hervorgehen. Diese Ansätze, die die Konstruktion der Fachgeschichte an Einzelbeispielen vornehmen, ermöglichen allerdings keine ungefilterte Darstellung der Forschungsbemühungen. Die Untersuchung der tatsächlichen Ausrichtung eines Wissenschaftszweigs kann umfassend nur gelingen, wenn neben exemplarisch vorgehenden Studien auch Ansätze erprobt werden, die einen empirischen Überblick über die Gesamtheit der Publikationen bieten.

Die Beschäftigung mit der Geschichte einer wissenschaftlichen Disziplin auf Grundlage von Forschungsliteratur kann grob in zwei Ansätze differenziert werden:

Ein **qualitativer Ansatz** untersucht eine Auswahl von Literatur, die als einflussreich und wichtig für die Disziplin angesehen wird. Er ist gut geeignet, um die Erkenntnisse einzelner Autoren und Forschungstraditionen, die einen großen Einfluss auf eine Disziplin hatten, zusammenzufassen und in einen Kontext einzuordnen. Ein wesentliches Merkmal, das zugleich auch als Schwachpunkt zu verstehen ist, ist die begrenzte und subjektive Literatúrauswahl.⁴

Quantitative Verfahren, bei denen die große Mengen an Veröffentlichungen untersucht werden, ermöglichen durch die deutlich größere Zahl untersuchter Objekte eine objektivere Abbildung der Forschungsrealität. So können neben der Identifizierung von Forschungstrends fundierte Aussagen über das Schrifttum getroffen werden. Die Herausforderungen bei quantitativen Auseinandersetzungen mit Literatur liegen weniger in der Datenauswahl, sondern eher in methodischen Unklarheiten und den begrenzten Auswertungsmöglichkeiten, v. a. wenn nur Metadaten zur Auswertung bereitstehen.

In dieser Arbeit wird daher eine erweiterte quantitative Auswertung als Methode vorgeschlagen.

Ziel der Arbeit ist es, die Geschichte der Populärmusikforschung anhand einer erweiterten quantitativen Analyse aller themenbezogenen Veröffentlichungen zu untersuchen. Dabei wird die Themen- und Methodenwahl derjenigen Autoren, die sich ab der zweiten Hälfte des 20. Jahrhunderts auf einer wissenschaftlichen Ebene⁵ mit populärer Musik auseinandergesetzt haben, ermittelt und analysiert. Dazu sollen Erklärungsansätze erarbeitet werden; die Ergebnisse sind auf Basis von im Fach geführten Diskussionen zu bewerten.

⁴ Eine qualitative Untersuchung zu einem extrem eingegrenzten Forschungsfeld könnte eine umfassende Literatúrauswahl thematisieren, dieser Umstand würde das Argument der Subjektivität und Lückenhaftigkeit entkräften.

⁵ Der Begriff der Wissenschaftlichkeit wird in der Arbeit ebenfalls zu untersuchen und zu definieren sein.

-> Ggf. die Zielsetzung etwas revidieren bzw. einschränken, die Intention hervorheben und deutlich machen, dass mit diesem Ansatz keine allgemeingültige Wissenschaftsgeschichte geschrieben wird und das auch nicht in einer Arbeit zu erledigen ist

„Sie [die Arbeit] will nämlich keine Geschichte der Musikwissenschaft sein, obgleich sie diese behandelt; aber nicht in der Absicht, diese vollständig und umfassend darzustellen, sondern in der Intention, die Gründe aufzuzeigen, die es der Musikwissenschaft unmöglich machten, ihren Anspruch einzulösen.“⁶

Die Beschäftigung mit der Geschichte einer wissenschaftlichen Disziplin trägt zum Verständnis, wie es zu den Inhalten und Methodiken innerhalb des Fachs gekommen ist, bei. Darüber hinaus lässt sich auf Basis einer Fachhistoriographie eine Reflexion anstreben und ein Beitrag zur Diskussion um die Ausrichtung eines verhältnismäßig jungen Forschungsgebiets auf einer datenbasierten Grundlage leisten.

Persönlicher Bezug? -> Aus dem Studium... /Kanonisierungsprozesse...

Die Fachgeschichte der Populärmusikforschung und die Forschungsthemen und -Methoden werden anhand der wissenschaftlichen Literatur der Musikwissenschaft rekonstruiert und untersucht.

Auf methodischer Ebene wird in dem vorliegenden Dissertationsvorhaben die Erstellung einer bibliographischen Forschungsdatenbank erprobt und die Arbeit mit einer solchen Datenbank als Quelle in Hinblick auf eine Rekonstruktion von Wissenschaftsgeschichte.

Die zu erstellende Forschungsdatenbank basiert auf dem Datenbestand der wichtigsten Bibliographie für wissenschaftliche Veröffentlichungen im Bereich der Musikforschung, der Literaturdatenbank *RILM Abstracts of Music Literature (1967 to Present)*⁷. Zusätzlich zu den von RILM gelieferten konventionellen bibliographischen Metadaten (z. B. Jahr, Autor, Art der Veröffentlichung etc.)⁸ werden die bibliographischen Daten automatisiert mit Keywords angereichert, die eine inhaltliche Auswertung möglich machen und die implizit in den von RILM bereitgestellten Titeln, Abstracts und Keywords enthalten sind. Dabei wird extrahiert, welche MusikerInnen bzw. Bands, welche Genre und Länder thematisiert werden und (aus den von RILM vergebenen Keywords der Sacherschließung) welche Methodik sich eine Publikation zuordnen lässt.

Die Anreicherung der Daten wird mittels einer für diesen Einsatzzweck programmierten Anwendung umgesetzt (vgl. Forschungsdesign (??)).

⁶ (Matejka1976)

⁷ Im Folgenden referenziert die Abkürzung *RILM* immer die Bibliographie, die von der 1965 gegründete Organisation *Répertoire International de Littérature Musicale* herausgegeben wird.

⁸ Hierbei handelt es sich um Metadaten, die im Zuge der bibliothekarischen Formalerschließung erfasst werden.

Durch die Auswertung aller dann zur Verfügung stehenden Variablen werden neue Erkenntnisse in Hinblick auf die Entwicklungen der Disziplin aufzuzeigen sein. Der methodische Ansatz wird es ermöglichen, ein detailliertes Bild der Populärmusikforschung zu zeichnen.

1.1 Definitionen

In der vorliegenden Arbeit wird die wissenschaftliche Betätigung mit populärer Musik untersucht. Da sowohl der Begriff der „populären Musik“ als auch der der „Populärmusikforschung“ keine hinreichend definierten Arbeitsgebiete darstellen, werden an dieser Stelle eine Arbeitsdefinitionen vorgestellt, damit die inhaltliche Zielsetzung auf einem nachvollziehbaren Fundament geschehen kann.

Der Begriff „populäre Musik“ bezieht sich in dieser Arbeit auf Musik, die ab dem 20. Jahrhundert entstanden ist und die weder der westlichen Kunstmusik, der geistlichen Musik noch den historischen und geistlichen Musiktraditionen außerhalb Europas, wie etwa der asiatischen Kunstmusik oder der geistlichen Musik des Islam, zuzurechnen ist.

- Synonym gebraucht: Populärmusikforschung – Popular Music Studies – Popmusikforschung – PopMusicology
- Keine Einheitliche Definition: Deswegen ja auch diese Arbeit
 - Definitionsversuch auch durch Abgrenzung?
- Mehr oder weniger Einigkeit im Gegenstand

Als Populärmusikforschung ist die wissenschaftliche Auseinandersetzung mit den unterschiedlichsten Formen populärer Musik zu verstehen. Der Begriff der Populärmusikforschung beschreibt zunächst „ein interdisziplinäres Theorie- und Methodenarsenal“,⁹ und, wie der Begriff der populären Musik, „fungiert [er] lediglich als eine erste grobe Verständigungsmarke“.¹⁰

⁹ (Pfleidereretal2014)

¹⁰ (Rosling2002a)

2 Wissenschaftsforschung in der Musikwissenschaft

Zur Einordnung des vorgestellten Dissertationsvorhabens und als Grundlage für die Arbeit an einer Fachgeschichte der Populärmusikforschung bedarf es einer Einführung in den Komplex der Wissenschaftsforschung. Die Arbeit an einer Disziplinengeschichte basiert auf Methoden aus der Wissenschaftsforschung und die Kenntnis der Vorgehensweisen ist für die methodische Anlage dieser Arbeit von großer Wichtigkeit.

Es soll darüber hinaus auch ein Überblick über bestehende Ansätze der Wissenschaftsforschung (insbesondere der Wissenschaftshistoriographie) in der Musikwissenschaft gegeben werden. Dies inkludiert den derzeitigen Stand der Forschung.

2.1 Einleitung Wissenschaftsforschung

„Wissenschaftsforschung“ wird von GETHMANN als Sammelbecken aller „reflexionswissenschaftlichen Ansätze“¹¹ beschrieben. Als wissenschaftliches Fachgebiet, das sich der Erforschung aller Bereiche des Systems „Wissenschaft“ und den Wechselwirkungen mit Bereichen außerhalb der Wissenschaft widmet, und damit Wissenschaft selber zum Gegenstand hat, ist die Wissenschaftsforschung eine Metawissenschaft¹².

Im Diskurs um die inhaltliche und methodologische Ausrichtung der Wissenschaftsforschung offenbart sich, dass sie selbst nicht unbedingt als eine gefestigte Disziplin mit einfach zu definierender Zielsetzungen gelten kann. Heterogene/inkonsistente/divergente Aussagen von Fachvertretern in relevanten Texten sprechen für diese Sichtweise.

In der deutschen Wikipedia wird die Wissenschaftsforschung als deutsches Pendant zu den *Science Studies* verortet und der ausschließlichen Bezug auf die Erforschung der Naturwissenschaften hervorgehoben. Im deutschen Sprachraum jedoch umfasst der Begriff Wissenschaft „sämtliche Arten des systematischen Wissens, über alle möglichen Gegenstände“¹³, er ist nicht auf die Naturwissenschaften begrenzt. Der **Gegenstandsbereich** der Wissenschaftsforschung ist, wie schon 1995 von FELT et al. gefordert, daher auch auf die Sozial- und Geisteswissenschaften ausgeweitet¹⁴ zu definieren.

¹¹ vgl. (Gethmann1981)

¹² vgl. (Fischer2001)

¹³ (Daston2001)

¹⁴ (Felt1995)

Unter dem Begriff der Wissenschaftsforschung sind „polydisziplinäre Bemühungen“¹⁵ zu verstehen. Im Laufe des 20. Jahrhunderts haben sich eigene, spezialisierte Metawissenschaften herausgebildet, deren aller Ziel es ist, Erkenntnisse über die Funktionsweise, die Zielsetzungen und die Geschichte der Wissenschaften zu generieren. Als Sammelbecken der Bemühungen dieser Metawissenschaften fungiert die Wissenschaftsforschung. SCHMITHÜSEN zufolge zählen zu den Subdisziplinen der Wissenschaftsforschung die Wissenschaftstheorie, Wissenschaftssoziologie, Wissenschaftspsychologie, Wissenschaftspolitik sowie nicht zuletzt die Wissenschaftsgeschichte¹⁶:

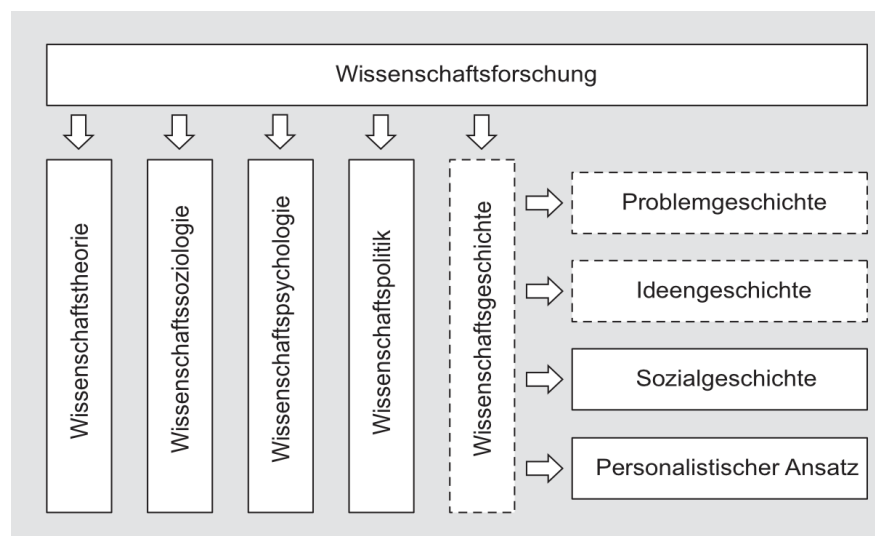


Abbildung 1 – Bereiche der Wissenschaftsforschung nach Schmithüsen, S. 3

Ähnlich sieht es WEINGART;¹⁷ OSSOWSKA bezieht die Wissenschaftsphilosophie als Subdisziplin der Wissenschaftsforschung mit ein¹⁸, die aber durch die synonym verwendete Wissenschaftstheorie abgedeckt ist.

Die **Zielsetzungen** von Wissenschaftsforschung sind entsprechend der extensiven Anlage der Wissenschaftsforschung breit gefächert. Zu den **grundsätzlichen Zielen** einer Erforschung der Wissenschaft zählt die Untersuchung der Funktionsweise, Struktur und Entwicklung von Wissenschaft. Wissenschaftsforschung ist dabei sowohl für Akteure innerhalb der Wissenschaft (Reflexion) als auch für die externe Betrachtung des Systems von großem Interesse (Forschungsergebnisse, Förderung etc.).

¹⁵ (Laitkor1996a)

¹⁶ (Schmithüsen2015)

¹⁷ vgl. (Weingart2003)

¹⁸ (Ossowska1964)

Speziellere Ziele der Wissenschaftsforschung schließen die Untersuchung der gesellschaftliche Bedeutung von Wissenschaft ein, sie untersuchen, welche Folgen die missbräuchliche Verwendung von Wissenschaft haben kann, sie arbeiten an gesamtgesellschaftlicher Aufklärung und der „Entzauberung“ der „Black Box“ Wissenschaft¹⁹. Zu den Zielen zählt auch die Legitimation wissenschaftlicher Tätigkeit, in erster Linie der Gesellschaft, aber auch anderen Wissenschaftlern und speziell der Politik und Investoren gegenüber; Erkenntnisse der Wissenschaftsforschung können mithin Investitionen²⁰ von externen Geldgebern sichern und rechtfertigen.

Darüber hinaus ist die Wissenschaftsforschung mit ihrem reflexiven Vorgehen essentiell für die Wissenschaftspolitik geworden. Durch Erkenntnisse die das gesamte System der Wissenschaft betreffen, können Handlungsempfehlungen gegeben werden und es versucht werden, Entwicklungen innerhalb der Wissenschaften zu steuern.²¹

Wissenschaftsforschung hat neben dem Einfluss auf externe Faktoren (Gesellschaft, Politik, Finanzierung) eine wichtige Rolle in als Kontrollinstanz innerhalb wissenschaftlicher Disziplinen erlangt. Durch die Untersuchung des eigenen Fachgebiets können die inhaltliche Ausrichtung, Kompetenzen in Forschung und Lehre sowie die Ausrichtung in Hinblick auf die Studierenden den externen Faktoren angepasst werden.

Die Notwendigkeit, sich mit dem System Wissenschaft auseinanderzusetzen, ergibt sich aus der **Bedeutung**, die die Wissenschaft in der modernen „verwissenschaftlichten“ Gesellschaft²² innehat.

„Mit dem Beginn der industriellen Revolution vollzog sich – zunächst weitgehend nur vom Produktionsprozeß gesteuert – eine zunehmende Verwissenschaftlichung immer weiterer gesellschaftlicher Sphären, die immer mehr Gesellschaftsmitglieder unmittelbar oder mittelbar in der Gestaltung ihrer Lebensplanung betraf.“²³

Diese von SCHERNUS zusammengefasste Entwicklung hält bis heute an: Es gibt kaum einen Lebensbereich, der nicht von wissenschaftlichen Erkenntnisse durchdrungen ist, vor allem in Europa streben immer mehr junge Menschen ein Studium an und die Zukunftsfähigkeit moderner Gesellschaften wird nicht zuletzt an ihrer Wettbewerbsfähigkeit im Bereich der Forschung gemessen.

Die zunehmende Relevanz der Wissenschaft für moderne Gesellschaften als auch die ansteigende Komplexität des Wissenschaftssystems haben dazu geführt, dass das Wissenschaftssystem selbst zum Gegenstand von Forschung wurde.

¹⁹ vgl. u. a. (Feltr995)

²⁰ (Feltr995)

²¹ vgl. (Schui2004)

²² (Feltr995)

²³ (Schernus2005)

Alle Teilbereiche des Wissenschaftssystems sind starken Wechselwirkungen zwischen Politik und Gesellschaft auf der einen und Wissenschaft auf der anderen Seite unterworfen. Dazu zählt die Rolle der Wissenschaft als Innovationsmotor für die Wirtschaft, ihre Rolle als Kulturbewahrer und Mahner, aber auch – bei unreflektierter Anwendung – als potentielle Gefahr (man denke an die in diesem Zusammenhang immer hervorgebrachte Atombombe oder an die Unterstützung der Nationalsozialisten durch sog. Rassenforscher). Ein weiterer Punkt ist die von den Universitäten und Forschungseinrichtungen ausgefüllte Rolle als großer Arbeitgeber – viele Menschen sind nicht nur indirekt sondern direkt von Entwicklungen in der Wissenschaft und den – betroffen.

Auch vor dem 20. Jahrhundert haben Wissenschaftler über Teilbereiche der Wissenschaften reflektiert, im Zuge der gesteigerten Bedeutung von Wissenschaft kam es ab dem Beginn des 20. Jahrhunderts zu systematischen Selbstreflexionen, aus denen das Fachgebiet der Wissenschaftsforschung hervorgegangen ist.

Wissenschaftsforschung, v. a. Wissenschaftsgeschichte, ist oftmals von Fachwissenschaftlern betrieben worden. Das hat zwei Gründe: Zum Einen, da Wissenschaftsforschung als eigenständige Disziplin noch nicht lange besteht²⁴ und es daher noch nicht viele speziell ausgebildete Forscher gab. Außerdem ist in vielen Fällen die Arbeit in diesem Forschungsfeld aufgrund des Gegenstandes und des benötigten Fachwissens über die jeweilige Disziplin von besonderer Natur. WAGNER spricht von „Wissenschaftler[n] [...], die sich als Gelegenheitsphilosophen versuchen“²⁵, KUHN sieht „geschichtsschreibende Fachwissenschaftler“²⁶ als die ersten, die sich, wenn auch nicht über alle Zweifel erhaben, an der Geschichtsschreibung wissenschaftlicher Disziplinen versuchten. KUHN zufolge haben die Wissenschaftshistoriker erst im 20. Jahrhundert „allmählich gelernt, ihren Gegenstand nicht mehr als eine Chronologie der Ansammlung positiver Leistungen auf einem Spezialgebiet zu sehen, das aufgrund späterer Einsichten definiert wird.“²⁷ Dem Fachwissenschaftler, der sich mit der Disziplingeschichte befasst kann in speziellen Fällen ein Interessenkonflikt vorgehalten werden – etwa wenn er an einer Glorifizierung der eigenen Leistung arbeitet statt sich kritisch mit der Geschichte auseinander zu setzen. Die Interdependenzen zwischen Absicht bzw. Motivation der Forschenden und ihrer jeweiligen Herangehensweise an die Wissenschaftsforschung ist bei der Auseinandersetzung stets im Hinterkopf zu behalten.

²⁴ (Kuhn1977a): Seit den 1950er Jahren, zunächst nur USA

²⁵ (Hagner2001)

²⁶ (Kuhn1977a)

²⁷ (Kuhn1977a)

Wissenschaftsgeschichte ist zudem lange mehr oder weniger exklusiv von in den Naturwissenschaften betrieben worden, was HAGNER zufolge fatale Konsequenzen hatte.²⁸ Zu diesen fatalen Konsequenzen zählt der Autor, dass die Wissenschaftsgeschichte außerhalb eines begrenzten Rezipientenkreises „nie recht ernstgenommen“ wurde. Außerdem, so HAGNER weiter, „wurde ihr Wert am ehesten als Gedächtnis der Wissenschaften taxiert“.²⁹ Der Begriff des Gedächtnis ist in diesem Zusammenhang negativ besetzt, da eine reine Erinnerungsfunktion keinen praktischen Nutzen bzw. keinen reflexiven Moment beinhaltet. Moderne Wissenschaftsgeschichtsschreibung sollte konstruktiv sein und mit dem Ziel erarbeitet werden, Handlungsanweisungen für die wissenschaftliche Praxis zu generieren.

Die Methoden, mit denen die Wissenschaftsforschung ihre Ziele zu erreichen sucht, sind ebenso vielfältig wie das Forschungsgebiet selbst.

Die Wissenschaftsforschung bedient sich bei dem Methodenkanon der Geschichts- und der Sozialwissenschaften. Die Vorgehensweise und Methodenauswahl wird je nach Subdisziplin der Wissenschaftsforschung an die speziellen Anforderungen angepasst. Empirische Wissenschaftsforschung bedient sich dabei etwa bei qualitativen und quantitativen Methoden aus den Sozialwissenschaften (Fragebögen, Evaluationsstudien). Ein weiteres Beispiel stellt die Quellenforschung aus der Geschichtswissenschaft dar, die in der Wissenschaftsforschung angewendet wird (wissenschaftliche Publikationen werden zu Quellen).

Im Zuge des Themas dieser Arbeit jedoch liegt der Fokus auf dem Teilgebiet der Wissenschaftshistoriographie, die anderen erwähnten Subdisziplinen werden hier nicht Gegenstand weiterer Betrachtungen sein. Auf die Methoden der Wissenschaftsgeschichtsschreibung wird im Kapitel [Fachgeschichtsschreibung]] näher eingegangen.

Wissenschaft kann vielfältig zum Gegenstand wissenschaftlicher Forschung werden. Die folgende Darstellung von Bereichen und Fragestellungen der Wissenschaftsforschung ist von David KALDEWEY übernommen³⁰. KALDEWEY teilt die Betrachtung in die drei Bereiche „Wissenschaft als Wissen“, „Wissenschaft als Praxis“ und „Wissenschaft als Institution“ ein.

Zum ersten Bereich gehören die Fragen nach der Form wissenschaftlichen Wissens (Logik, Erkenntnistheorie), der Ordnung des wissenschaftlichen Wissens (Typologien, Enzyklopädien, Disziplinen), den Formen des Wissens (z. B. theoretisches und praktisches Wissen) und sowie der Darstellung von Wissen (z. B. wissenschaftliche Publikationen).

²⁸ (Hagner2001)

²⁹ (Hagner2001)

³⁰ vgl. (Kaldewey2015)

Der Bereich „Wissenschaft als Praxis“ umfasst die Untersuchung der Methoden der Wissenschaft (z. B. deduktiv/induktiv), des Forschungsalltags (z. B. Experimentalsysteme, epistemische Dinge) und der Betrachtung von WissenschaftlerInnen als handelnden Akteuren (Habitus, Biographien, Karrieren).

Der dritte Bereich, „Wissenschaft als Institution“, schließt Fragen nach der Ausdifferenzierung: Wissenschaft und ihre gesellschaftliche Umwelt, der Innendifferenzierung: Disziplinarität und Interdisziplinarität, den Werten und Regeln von wissenschaftlichen Gemeinschaften und der Organisationen von Wissensproduktion (Akademien, Universitäten, Forschungsinstitute) ein.

2.2 Von der Wissenschaftsforschung zur Fachgeschichtsschreibung

Ein wichtiges Teilgebiet der Wissenschaftsforschung ist die Wissenschaftsgeschichte. In Hinblick auf das Kernthema dieser Arbeit sind nach einem Einblick in die Mutterdisziplin der Wissenschaftsforschung nun wichtige Spezifika der Wissenschaftsgeschichtsschreibung zu erörtern. Um die Untersuchung der Fachgeschichte der Populärmusikforschung methodisch zu fundieren, ist ein Blick in die Arbeitsweisen gängiger wissenschaftsgeschichtlicher Ansätze unumgänglich.

Zunächst ist festzustellen, dass es eine Vielzahl unterschiedlicher Perspektiven gibt, aus denen Wissenschaftsgeschichte konstruiert werden kann. Die Perspektiven reichen von der Untersuchung des „großen Ganzen“ bis zu der Untersuchung der Geschichte einzelner Begriffe oder Institute an Forschungseinrichtungen.

Zu den umfassenderen Zielsetzungen von Wissenschaftsgeschichte zählt der Versuch, die Geschichte der Wissenschaft und Forschung allgemein zu bearbeiten. Wissenschaftsgeschichte wurde außerdem oft im Sinne einer Fortschrittsgeschichte geschrieben. Diese beiden Ansätze sind im 19. Jahrhundert noch verbreitet, heute ist bekannt, dass eine solche Aufgabe nicht zu bewältigen ist.

Die Geschichte von Wissenschaft kann darüber hinaus als Betrachtung einzelner, als herausragend betrachteter Persönlichkeiten im Rahmen einer Personengeschichte geschrieben werden. Diese Ansätze sind oft für ein breiteres Publikum³¹ gedacht und gehen nicht selten hagiographisch („verheiligend“) vor (vgl. ??).

Wissenschaftshistoriographie wird heute i. d. R. zu spezielleren und thematisch eingrenzteren Fragestellungen betrieben. Dazu zählt neben der in dieser Arbeit angestrebte Untersuchung einer Disziplinen- bzw. Fachgeschichte auch wie die Rekonstruktion von Institutionsgeschichte (Instituts- und Universitätsgeschichte). Bei letzterer kommt es jedoch durch den stärkeren Fokus auf die Darstellung der

³¹ vgl. (Hagner2001)

Institutionen und nicht unbedingt der Forschungsinhalte zu einer Verschiebung der Zielsetzung. Weitere gängige „Spezialgeschichten“ im Umfeld der Wissenschaftshistoriographie sind Begriffs-, Problem- und Ideengeschichte sowie ein sozialhistorischer Ansatz (vgl. ??).

Die Wissenschaftsgeschichte beschäftigt sich mit der Entwicklung des gesamten Systems „Wissenschaft“ und fungiert als Mutterdisziplin für den Bereich der Fachgeschichtsschreibung. Als Subdisziplin der Wissenschaftsgeschichte beschäftigt sich die Fachgeschichte (auch Fachhistoriographie) mit der konkreten Entwicklung einzelner Wissenschaftsdisziplinen.

Hierbei ist, neben einem anderen Fokus, eine unterschiedliche Zielsetzung zu konstatieren. Allgemeine wissenschaftsgeschichtliche Abhandlungen untersuchen keine fachspezifische Entwicklungen, sondern z. B. die Rolle der Wissenschaft in der Gesellschaft zu einem bestimmten Zeitpunkt oder den Umgang der Wissenschaft in Bezug auf externe Einflussfaktoren. Auch allgemeine Fragen, etwa zum Ablauf von Forschungstätigkeiten oder dem Erkenntnisinteresse sind Gegenstand allgemeiner Wissenschaftsgeschichtsforschung.

Wissenschaftsgeschichte wurde, wie bereits erörtert, bis in die zweite Hälfte des 20. Jahrhunderts in der Regel von Fachwissenschaftlern verfasst, die in ihr „ein Mittel zur Verdeutlichung der Begriffe auf ihrem Fachgebiet, zur Festlegung seiner Tradition und zur Werbung von Studenten“³² sahen (vgl. Wer forscht? und Funktionen von Fachgeschichte).

Innerhalb einzelner Disziplinen³³ unterscheidet sich der Stellenwert der Fachhistoriographie. In naturwissenschaftlichen Fächern ist die Beschäftigung mit der eigenen Disziplin spätestens seit dem 20. Jahrhundert fest verankert, so gibt es beispielsweise in der Physik³⁴, der Psychologie³⁵ oder der Medizin³⁶ eine gefestigte Auseinandersetzung mit eigenen Lehrstühlen zur Disziplingeschichte.

Systematische fachgeschichtliche Reflexion findet in den sog. Geisteswissenschaften zwar in jüngster Zeit vermehrt statt, eine Tendenz zur Institutionalisierung dieser Bemühungen ist aber noch nicht zu beobachten.

³² (Kuhn1977a)

³³ vgl. (Timm1973): „Die Tatsache, daß die Wissenschaftsgeschichte bisher meist nicht von Fachhistorikern, sondern vielmehr von Historikern eines Fachs betrieben worden ist [...]“

³⁴ <http://www.uni-oldenburg.de/histodid/>

³⁵ vgl. <https://www.zpid.de/redact/category.php?cat=24>

³⁶ vgl. <http://matthiskrischel.wixsite.com/fvmg>

2.3 Fachgeschichtsschreibung: Funktionen und Ansätze

Die Beschäftigung mit der Geschichte einer wissenschaftlichen Disziplin erfüllt gewisse Funktionen. Die Fachgeschichtsschreibung ist geprägt von verschiedenen Ansätzen und methodischen Unterscheidungen, die im Folgenden erörtert werden sollen. Die unterschiedlichen Ansätze der Fachgeschichtsschreibung sollen vorgestellt werden. In diesem Zuge der aktuelle Stand der Forschung der Fachgeschichtsforschung in der Musikwissenschaft bzw. der Populärmusikforschung dargelegt.

Die Beschäftigung mit Geschichte wissenschaftlicher Disziplin kommt in mehrerlei Hinsicht eine wichtige Bedeutung zu. SCHMITHÜSEN³⁷ benennt vier Punkte, die für die Beschäftigung mit Disziplinengeschichte sprechen: Zunächst hat Fachgeschichtsschreibung eine **ordnende Funktion**. Dies kommt besonders Studierenden zu Gute, da durch das Einordnen bereits vorhandenen Wissens in einen historischen Kontext das Verständnis für den Fachbereich³⁸ gefördert wird. Das Wissen um bisherige Ansätze und Forschungsschwerpunkte in der eigenen Disziplin kann darüber hinaus dazu beitragen, einen Überblick über das Fach zu bekommen, eine Funktion die sowohl ordnenden Charakter hat als auch zur **Erweiterung der Perspektive** beiträgt, was SCHMITHÜSEN als weitere Funktion aufführt. Auf der Basis eines breit gefächerten historischen Wissens ist es wahrscheinlicher, dass man bei der eigenen Arbeit auch nicht augenscheinliche Aspekte berücksichtigen kann und sich nicht in der eigenen Forschungsnische isoliert. Darüber hinaus benennt SCHMITHÜSEN die **Korrekturfunktion**. Die Kenntnis der Geschichte, so Schmithüsens Vermutung, trägt dazu bei, Fehler, die in der Vergangenheit von Fachvertretern gemachten wurde, zu vermeiden und unterstützt die Aufarbeitung von fach- und wissenschaftspolitischen Fehlentwicklungen. Letztlich **lehrt die Fachgeschichte Bescheidenheit**, da die eigene Forschung „an bereits vorhandenen Errungenschaften relativiert“³⁹ werde.

Die von SCHMITHÜSEN vorgestellten Funktionen von Fachgeschichte sind zu erweitern. Eine zentrale Rolle spielt die Kenntnis historischer Entwicklungen einer wissenschaftlichen Disziplin einerseits fachintern für Studenten und Forschende, andererseits ist Disziplinengeschichte wichtig für die Positionierung im Wissenschaftsgefüge, für die Ausrichtung, Abgrenzung und Stärkung des Profils der Disziplin. Durch Reflexion historischer wie aktueller Tendenzen wird das **Selbstverständnis** des Forschenden gestärkt, hieraus entsteht zu einem Teil auch das **Selbstbewusstsein** einer Disziplin. Das Wissen um die Forschungstradition, auf der die eigene Arbeit basiert, kann darüber hinaus identitätsstiftend und identitätsstär-

³⁷ (Schmithusen2015)

³⁸ vgl. (Schmithusen2015)

³⁹ (Schmithusen2015)

kend sein und wichtig für die Selbstvergewisserung von Studierenden und Forschenden sein. Gleichmaßen bietet Forschungsgeschichte eine Möglichkeit zur Abgrenzung gegenüber „alten“ Ansätzen im Fach. Fachgeschichtliche Spezifika werden außerdem häufig genutzt, um die eigene Forschung zu kontextualisieren – Erkenntnisentwicklungen werden zur Einordnung oder Abgrenzung herangezogen.

Der Disziplinengeschichte kommt darüber hinaus eine **legitimatorische** Funktion zu: Durch den Bezug zur Forschung in der Vergangenheit lässt sich die eigene, neue Forschungsarbeit zusätzlich zur Inhaltlichen Relevanz legitimieren, gleich, ob es sich dabei um einen neuartigen methodischen Zugang oder ein bisher unterrepräsentiertes Themengebiet handelt.

In Hinblick auf die Musikwissenschaft und im Speziellen die Fachgeschichte der Populärmusikforschung ist einer Funktion von Fachgeschichtsschreibung besonderes Augenmerk zu schenken: Fachgeschichte hilft beim „Aufdecken blinder Flecken“⁴⁰ in der thematischen Ausrichtung einer wissenschaftlichen Disziplin, da die Möglichkeit besteht, dass die Beschäftigung mit der Fachgeschichte solche unterrepräsentierten Fragestellungen aufdeckt.

2.3.1 Herangehensweisen

Im Hinblick auf die Umsetzung der Fachgeschichtsschreibung sind neben den inhaltlich-logischen Ansätzen auch Herangehensweisen zu unterscheiden. Woodward⁴¹ hat 1980 drei Gegensatzpaare von Herangehensweisen in der historischen Wissenschaftsforschung der Psychologie zusammengestellt, die er als „Sources of Bias“ – Verzerrungsquellen – für die Fachgeschichtsschreibung anführt. Diese sind in dieser Form u. a. von Petzold⁴² und Schui⁴³ aufgegriffen worden, stellen aber für sich genommen im breiteren Geschichtswissenschaftlichen Kontext verwendete Begriffe dar. Die drei Dichotomien sind: Präsentismus/Historizismus, Internalismus/Externalismus und der Gegensatz zwischen qualitativen und quantitativen Herangehensweisen.

Die Unterscheidung zwischen einem präsentistischen und einem historizistischen Ansatz in der Wissenschaftsgeschichtsschreibung einem unterschiedlichen Verständnis

⁴⁰ (Kopper2016)

⁴¹ (Woodward1980)

⁴² (Petzold1985)

⁴³ (Schui2004)

Unter **Präsentismus** wird in der Wissenschaftsgeschichtsforschung eine Denkweise verstanden, die historische Ereignisse immer und in erster Linie in Bezug auf die Gegenwart bezieht. Geschichte wird so als Erfolgsgeschichte bis zu dem jeweiligen Punkt, an dem auf die Geschichte zurückgegriffen wird, konstruiert.⁴⁴ Dabei werden „weniger erfolgreiche Wissenschaftler/innen beziehungsweise Theorien“⁴⁵ nicht berücksichtigt. Das Verhalten von Wissenschaftlern, die präsentistisch arbeiten, ist PETZOLD zufolge mit dem Prinzip des „Selbstbedienungsladens“ zu vergleichen:⁴⁶ Im Zuge der Einordnung oder Legitimation der eigenen Forschung wird nur relevantes (werden nur relevante Theorien und Strömungen) aus der Geschichte der Forschung herangezogen und somit ein verzerrtes Bild gezeichnet, „das den realen Wirren der Geschichte nicht gerecht wird.“⁴⁷

Präsentistische Wissenschaftsgeschichtsschreibung ähnelt daher der sog. Whig history, die (Wissenschaft-)Geschichte ebenfalls als einzig mögliche Konsequenz konstruiert und nur erfolgreiche Ansätze in Betracht zieht.

„Wissenschaftsgeschichte nach diesem Muster geht von einer linearen, zielgerichteten Entwicklung aus, in deren Verlauf immer mehr positive Erkenntnisse angehäuft werden, sich die Form der wissenschaftlichen Tätigkeit jedoch nicht (oder nur unwesentlich) verändert.“⁴⁸

Unter einer solchen Sichtweise werden alle Fehlversuche, aber ggf. auch vielversprechende Forschung aus der Darstellung der Geschichte ausgeblendet. Somit wird ein unvollständiges, fast subjektives Bild zur Wahrheit erklärt.

Der Historizismus hingegen betrachtet Wissenschaftsgeschichte gänzlich ohne Bezug auf die Gegenwart und untersucht detailliert die Wissenschaft zu einem bestimmten in der Vergangenheit liegenden Zeitpunkt. Durch diese Beschränkung können größere historische Entwicklungen nicht nachvollzogen werden^{49}.** Es gibt jedoch auch positive Stimmen zum ^{**}Historizismus, der in angepasster und informierter Form durchaus sinnvoll sein kann:

„As used by historians of science, historicism has come to denote a praiseworthy goal - the placing of events in a credible context of the times.“⁵⁰

Die Debatte um wissenschaftsinterne und wissenschaftsexterne Geschichtsschreibung ist ein Begriff aus der Wissenschaftstheorie, dessen Anwendung in der Fachgeschichtsschreibung u. a. von WOODWARD aufgegriffen wurde.

⁴⁴ vgl. (Woodward1980)

⁴⁵ (Schui2004)

⁴⁶ (Petzold1985)

⁴⁷ (Petzold1985)

⁴⁸ (Brandstetter2012)

⁴⁹ vgl. (Petzold1985) und (Schui2004)

⁵⁰ (Woodward1980)

Im Grunde beschreibt die Dichotomie zwischen wissenschaftsinterner und -externer Geschichtsschreibung zwei unterschiedliche Herangehensweisen in der Betrachtung wissenschaftsgeschichtlicher Fragestellungen. Externalistische Wissenschaftsgeschichte bezieht politische, ökonomische, soziale, ethische und religiöse Zusammenhänge in die Berücksichtigung der Entwicklungen mit ein,⁵¹ bisweilen wird auch von der Determination wissenschaftlicher Erkenntnisse durch diese äußeren Einflüsse ausgegangen⁵² und vom Einfluss politischer und religiöser Ideologien auf die Wissenschaft⁵³.

Der internalistische Ansatz hingegen geht davon aus, dass „Erkenntnisfortschritt [...] als eigendynamischer Prozess zu verstehen [ist], ermöglicht etwa durch Logik, wissenschaftliche Methoden, empirische Forschung, rationalen Diskurs“⁵⁴.

Im Handbuch Wissenschaftsgeschichte wird ausgeführt, dass heute Internalismus und Externalismus nicht mehr scharf getrennt sind:

Dementsprechend sind die anfänglich scharf geführten Debatten zwischen ›Internalisten‹ und ›Externalisten‹ [...] mittlerweile einem breiten Konsens gewichen, wonach Wissenschaften zwar hochspezialisierte Subkulturen arbeitsteiliger Gesellschaften bilden, und insofern ganz eigene Traditionen und Wertvorstellungen hervorbringen, aber eben doch nur als solche existieren können, weil sie Austauschbeziehungen zu anderen Bereichen der Gesellschaft unterhalten.⁵⁵

Auch BRANDSTETTER sieht die Ansätze als überwunden an⁵⁶.

Das letzte Gegensatzpaar ist weniger ein theoretisches als ein methodisches. Quantitative Verfahren in der Wissenschaftsgeschichtsschreibung bedienen sich i. d. R. der Scientometrie bzw. der Bibliometrie und haben zum Ziel, Wissenschaft zählbar zu machen, in dem quantifizierbare Aspekte einer Wissenschaft ausgewertet werden. Dazu zählen in erster Linie Publikationen und Zitationen, aber auch Netzwerke von Forschern, Lehrstühle, eingeworbene Fördermittel oder angemeldete Patente.

Eine qualitative Wissenschaftsgeschichtsschreibung hingegen erforscht spezielle Teilbereiche der Wissenschaften und bedient sich dabei anderen Quellen und hat andere Ziele als eine quantitative Herangehensweise. SCHUI zufolge gehört dazu etwa „die Analyse wissenschaftlicher Tagebücher und Notizen oder auch die Feldbeobachtung in einem Labor“⁵⁷.

⁵¹ vgl. (Hagner2001)

⁵² vgl. (Kaldewey2015)

⁵³ vgl. (Canguilhem1979)

⁵⁴ (Kaldewey2015)

⁵⁵ (Sommeretal2017)

⁵⁶ vgl. (Brandstetter2012)

⁵⁷ (Schui2004)

2.3.2 Ansätze von Fachgeschichtsschreibung

Die Geschichte einer Wissenschaft wurde und wird auf verschiedene Arten dargestellt. Auf die folgenden vier Ansätze wird in methodischen Abhandlungen zur Psychologiegeschichtsschreibung rekurriert,⁵⁸ sie stellen quasi die traditionellen Möglichkeiten, Fachgeschichte zu schreiben, dar.

Geschichtsschreibung mit **personalistischem** Ansatz (oft auch als „Great-Men“ Ansatz⁵⁹ beschrieben) rückt die Forscherpersönlichkeit in den Mittelpunkt Wissenschaftshistorischer Forschung. Die Entwicklung des Fachs wird anhand von einzelnen Personen nacherzählt, die rückblickend für die Disziplin wegweisende Entdeckungen gemacht bzw. zentrale Positionen vertreten haben.

Die personenzentrierte Art der Wissenschaftsgeschichtsschreibung beschränkte sich in der Vergangenheit i. d. R. auf die Darstellung von „Genies“. Die Praxis, Fachgeschichte auf der Basis wissenschaftlicher Schulen zu rekonstruieren, ist gibt es ebenfalls.

Die Beschränkung auf „wichtige“ Persönlichkeiten trägt zu der einseitigen Ausrichtung des personalistischen Ansatzes bei. Bei konsequenter Anwendung des Ansatzes werden viele Auseinandersetzungen und Wirren in der Entwicklung bis zum jetzigen Entwicklungsstand vernachlässigt.

Nichtsdestotrotz hat der personalistische Ansatz seine Berechtigung, da einzelne Wissenschaftlerpersönlichkeiten durchaus herausragende Leistungen erbracht haben und ohne Sie und ihre Bemühungen der heutige Wissensstand eventuell nicht erreicht worden wäre. Es ist allerdings so wichtig wie selbstverständlich, dass Wissenschaftsgeschichte nicht ausschließlich als Personengeschichte geschrieben werden kann.

Im Bereich der Populärmusikforschung ist es beispielsweise bei VON APPEN⁶⁰ zu einer personenbezogenen Rekapitulation der Entwicklung gekommen. VON APPEN betrachtet einzelne Forscher und ihre Positionen im Bezug auf populäre Musik, mit dem Ziel, anhand der Darstellung „historischer Stationen“⁶¹ Gründe für die mangelnde Betrachtung ästhetischer Merkmale populärer Musik aufzuzeigen.

Als Beispiel für eine unreflektierte personalistische Geschichtsdarstellung wird in der Literatur das Werk von WILHELM OSTWALD herangezogen.⁶²

MEYEN geht davon aus, dass die Entwicklung eines Fachs ist besser zu verstehen, wenn Sie als Generationengeschichte aufgefasst wird, da eine Disziplin von Professoren und deren jeweiliger Identität geprägt wird.⁶³

⁵⁸ u. a. (Petzold1985), (Krampen2009) und (Schmithusen2015)

⁵⁹ vgl. https://en.wikipedia.org/wiki/Great_Man_theory

⁶⁰ (Appen2007)

⁶¹ (Appen2007)

⁶² vgl. (Hagner2001) und (Kragh1987)

⁶³ (Meyen2013)

Im **problemgeschichtlichen** Ansatz wird die Entwicklung von Inhalten zum Ausgangspunkt der Untersuchung der Geschichte einer wissenschaftlichen Disziplin. Es werden Konzepte, Problemstellungen und Theorien untersucht, die die Entwicklung der Disziplin begleitet haben. Der problemgeschichtliche Ansatz geht dabei nicht chronologisch vor, sondern anhand von Grundfragen der Disziplin. Er ist daher inhaltlich-logisch orientiert und untersucht die Veränderungen im Umgang mit Inhalten⁶⁴ einer Disziplin.

Der problemgeschichtliche Ansatz stellt „konzeptuelle Entwicklungen in den Vordergrund“⁶⁵ und untersucht, „wie sich bestimmte Konzepte über die Zeit hinweg verändern“⁶⁶.

Ideengeschichtliche Fachgeschichtsschreibung untersucht die Geschichte wissenschaftlicher Ideen. Dabei wird chronologisch vorgegangen, man beschreibt die Ideen in ihrem historischen Kontext. Dazu zählt auch die Untersuchung der Umstände ihrer Entstehung.⁶⁷ Im Gegensatz zu möglichst objektiver und nicht-urteilender Geschichtsschreibung

Bei dem ideengeschichtlichen Ansatz werden Entwicklungsprozesse innerhalb wissenschaftlicher Disziplinen weniger als Leistung herausragender WissenschaftlerInnen als Folge der Eigendynamik von Ideen verstanden.⁶⁸ Der Entwicklungsprozess einer Theorie, eines methodischen Ansatzes ist demnach entkoppelt von äußeren und personellen Faktoren.

Die Ideengeschichte der Psychologie beispielsweise wird in „Epochen“ eingeteilt, die jeweils durch vorherrschende Denkmuster und Theorien gekennzeichnet sind. Diesem Muster folgend könnte man die Epochen der Ideengeschichte der Musikwissenschaft demnach wie folgt definieren: Von der Musiktheorie im antiken Griechenland über die Musiktheoretischen und -Geschichtlichen Ansätze im Mittelalter zur Disziplingenese im 19. Jahrhundert, Ausdifferenzierung & Konsolidierung des Fachs, zur Festigung eines Kanons, Einbezug historischer Aufführungspraxis, bis hin zur Beschäftigung mit der populären Musik.

„Im Unterschied zu einer historisch neutralisierenden, sich eines Urteils enthaltenden Geschichtsbetrachtung versucht man mittels der sogenannten Ideengeschichte, die Entstehung und Wandlung von Ideen, u. a. auch wissenschaftlicher, interessengeleitet zu problematisieren. Historische und begriffliche Diskontinuitäten werden dabei in Kauf genommen“⁶⁹

⁶⁴ (Schmithusen2015)

⁶⁵ (Krampen2009)

⁶⁶ (Schmithusen2015)

⁶⁷ (Schmithusen2015)

⁶⁸ vgl. (Brandstetter2012)

⁶⁹ (Pritzel2016)

„Die Ideengeschichte kann der fachdidaktischen Praxis Anregung geben, indem »diese sich bei denen inspirieren lässt, die als Klassiker die Probleme zum ersten Mal – und sei es nur gedanklich – zu lösen versuchten«“⁷⁰

Der **sozialgeschichtliche** Ansatz geht auf die Zusammenhänge zwischen gesamtgesellschaftlichen und innerdisziplinären Entwicklungen ein. So wird „[...] die soziale Dimension, d. h. die gesellschaftlichen Bedingungen, unter denen [ein] Fachbereich in der Vergangenheit existiert hat [...].“⁷¹ mit in die Konstruktion der Fachgeschichte einbezogen. Dieser Ansatz beinhaltet KRAMPEN zufolge sowohl die Geschichte der Institutionalisierung des Fachs als auch die Untersuchung der politischen, gesellschaftlichen und institutionellen Bedingungen der gegenwärtigen Forschungspraxis⁷².

Der sozialgeschichtliche Ansatz ist von den gängigen Zugängen zur Fachgeschichtsschreibung derjenige mit dem breitesten Fokus. Fachimmanente Entwicklungen werden hier als Folge gesellschaftlicher Entwicklungen verstanden. Ein Mögliches Problem besteht KRAMPEN zufolge darin, dass man sich bei Verwendung des sozialgeschichtlichen Ansatzes in Einzelheiten „verfransen“ könne, durch die größere historische Zusammenhänge verdeckt würden.⁷³

Der Bibliometrische Ansatz der Fachhistoriographie gehört nicht zu den vier „traditionellen“ Ansätzen. Bibliometrie ist die Lehre von der Messung wissenschaftlicher Publikationen und damit ein Teilgebiet der Szientometrie, der Lehre von der Messung der Wissenschaften. Bibliometrie wird von MAJA JOKI und RAFAEL BALL als „Anwendung mathematischer und statistischer Methoden auf Bücher und andere Medien der wissenschaftlichen Kommunikation“⁷⁴ definiert. Die moderne Bibliometrie stützt sich vor allem auf die sog. Zitationsanalyse. Hierbei werden die Zitationen in wissenschaftlichen Artikeln gemessen: Es wird gezählt, wie oft und in welchen Veröffentlichungen eine Publikation zitiert wird. In den Naturwissenschaften wird dieses Verfahren häufig angewendet, um qualitative Evaluationen von Wissenschaftlern und wissenschaftlichen Einrichtungen anhand der von ihnen veröffentlichten Publikationen zu erstellen. In ihrer ursprünglichen Form, nicht als computergestützte Zitationsanalyse, sondern als reine Outputanalyse,⁷⁵ kann eine bibliometrische Betrachtung aber auch in nicht-naturwissenschaftlichen

⁷⁰ (Busch2015)

⁷¹ (Schmithusen2015)

⁷² (Krampen2009)

⁷³ vgl. (Krampen2009)

⁷⁴ (Jokic2006)

⁷⁵ vgl. (Ball2014) Es handelt sich um eine Outputanalyse, da nur die Veröffentlichungen selbst, nicht aber deren Beziehungen untereinander (i. d. R. Zitate) untersucht wurden.

Disziplinen Aufschlüsse durch die Quantifizierung der Forschungsliteratur geben. So lassen sich die Ziele bzw. Ergebnisse der frühesten bibliometrischen Arbeit,⁷⁶ die eine solche Output-Analyse war, NICOLA DE BELLIS⁷⁷ zufolge in vier Bereiche aufteilen:

1. „Performance“-Analyse: Als Performance ist die Produktivität, also die Anzahl der Veröffentlichungen zu verstehen. Bei FRANCIS COLE und NELLIE EALES wurde dies auf die Produktivität einzelner Länder bezogen.
2. Messung der rein quantitativen Zunahme an Forschung zu einem Thema.
3. Die Möglichkeit, einen Überblick über Themen zu geben, die „im Trend“ sind, zu denen also aktuell oder zu einem bestimmten Zeitpunkt geforscht wird bzw. wurde.
4. Die Korrelation der Forschungsergebnisse mit externen, also häufig gesellschaftlichen, politischen oder ökonomischen, Einflussfaktoren.

Im Unterschied zu qualitativen Arbeiten, die sich mit einer subjektiven Auswahl an Literatur beschäftigen, kann eine quantitative Analyse – bei einer ausreichend großen Stichprobe – u. a. globale Trends und die Performance von Wissenschaftlern aufzeigen.

- Basiert auf der wissenschaftlichen Kommunikation
- „objektiver“

vgl. dazu⁷⁸

Probleme der Ansätze diskutiert KRAMPEN⁷⁹

Eine scharfe Trennung der Ansätze, insbesondere der ersten drei, ist nicht ohne weiteres möglich. Genauso wenig ist es praktikabel, sich bei der Erarbeitung einer Disziplinengeschichte auf einen der genannten Ansätze zu beschränken, da dadurch immer nur für sich stehende Teilbereiche betrachtet werden können, aber die Integration zu einem „großen Ganzen“ so nicht möglich ist.

Was es oft gibt, ist eine **Kombination der vier traditionellen Ansätze**

⁷⁶ (Cole1917) In dieser Veröffentlichung wurde die Literatur zur menschlichen Anatomie im Zeitraum von 1556 bis 1860 thematisiert.

⁷⁷ (DeBellis2009)

⁷⁸ Schui2004.

⁷⁹ (Krampen2009)

2.4 Stand der Forschung in der Fachgeschichtsschreibung der Musikwissenschaft

„Nicht anders als ‚Musik‘ ist ‚Musikwissenschaft‘ ein Begriff, der im Laufe seiner Geschichte höchst Unterschiedliches bedeutet hat und dessen Inhalt und Reichweite nur historisch bestimmt werden können.“⁸⁰

Die Auseinandersetzung mit der Geschichte der Musikforschung hat aktuell

- kaum institutionellen Rückhalt: Es gibt weder ausgewiesene Professuren noch Sonderforschungszentren o.ä. noch gehört eine Auseinandersetzung mit der eigenen Fachgeschichte zum Curriculum eines musikwissenschaftlichen Studiums.
- Es liegt außerdem bislang keine Monographien zum Thema vor, noch existiert eine Schriftenreihe. Nur ein paar wenige Sammelbände widmen sich dem Themenkomplex.
- Die Reflexion innerhalb des Fachs läuft behäbig an, einzelne Publikationen zur Ausrichtung der Musikwissenschaft sind vorhanden, eine breit geführte Diskussion ist dennoch kaum erkennbar.

Arbeiten, die sich mit der Fachgeschichte der Musikwissenschaft auseinandersetzen sind nach wie vor rar.

Was es gar nicht gibt, sind Beispiele für methodische fundierte, zielgerichteter Wissenschaftsgeschichtsforschung in einer Geisteswissenschaft oder gar entsprechende „Handbücher“: Disziplingeschichte wird selten zum alleinigen Gegenstand einer Betrachtung, sie wird vielmehr oft als Mittel zum Zweck der Hervorhebung der Signifikanz des eigentlichen Gegenstandes genutzt (solche Texte verhalten sich damit präsentistisch).

Im Folgenden werden **Ansätze in der Geschichtsschreibung der Populärmusikforschung** vorgestellt und diskutiert.

Im deutschen Sprachraum sind ein paar wenige Sammelbände zum Thema erschienen, etwa die Sammelbände hrsg. von RÖSING⁸¹, PFLEIDERER⁸², BOLZ⁸³ sowie HELMS/PHLEPS.⁸⁴

Die im Folgenden vorgestellten Arbeiten wenden sich meist einzelnen, speziellen Fragestellungen in Bezug auf die Geschichte der Populärmusikforschung zu.

⁸⁰ <https://www.musikundmedien.hu-berlin.de/de/musikwissenschaft/institut/geschichte>

⁸¹ Rösing2002

⁸² Pfeiderer2011a

⁸³ Bolzetal2016

⁸⁴ Helms2013

Es sind hier ferner nur die Publikationen erwähnt, die sich primär mit der Fachgeschichte bzw. der Ausrichtung des Fachs auseinandersetzen: Viele Arbeiten leiten ein mit Bemerkungen zu historischen Entwicklungen, i. d. R. um die eigene Arbeit zu legitimieren – diese können jedoch nicht als Beiträge zur Fachgeschichte gezählt werden.

Zu den frühesten Arbeiten, die den Einfluss der populären Musik auf die Musikforschung reflektierten, zählen die in den 1970er bzw. 80er Jahren erschienen Arbeiten von KNEIF⁸⁵, FÖDERMAYR⁸⁶, REINECKE⁸⁷ und JOST⁸⁸. Die Autoren setzen sich für die Einbeziehung populärer Musik in die Musikwissenschaft ein.

Es gibt einige Aufsätze, die sich direkt zur Fachgeschichte der Musikwissenschaft bzw. der Populärmusikforschung äußern: Kajanova⁸⁹, Pfeiderer⁹⁰, Brunner⁹¹ sowie Rösing⁹² behandeln die Entwicklung des wissenschaftlichen Umgangs mit populärer Musik, **wenn auch mit unterschiedlicher Schwerpunktsetzung**.

Ebenfalls zur Geschichte und dem Verhältnis zwischen populärer Musik und Musikwissenschaft äußern sich PFLEIDERER et al⁹³

Wissenschaftstheoretisch: Darüber hinaus gibt es Ansätze, die auf die Funktionsweise und Konstruktivität von (Musik-)Geschichtsschreibung eingehen und sich mit der historischen Entwicklung von Geschichtsschreibung befassen. Dazu zählen die Arbeiten von Rösing⁹⁴, Pfeiderer⁹⁵, Helms⁹⁶ Obert⁹⁷, Wicke⁹⁸ und Böhm⁹⁹.

Zur Ausrichtungsdebatte: Neben HEMMING¹⁰⁰ eröffnet auch RÖSING¹⁰¹ eine historisch fundierte Diskussion über Zukunft und Ausrichtung der Musikwissenschaft. PFLEIDERER¹⁰² thematisiert ästhetische Vorbehalte gegenüber populärer Musik und die daraus folgenden Problematiken für die Musikwissenschaft.

⁸⁵ Kneif1975
⁸⁶ Fodermayr1985
⁸⁷ Reinecke1988
⁸⁸ Jost1986
⁸⁹ Kajanova2013
⁹⁰ Pfeiderer2012a
⁹¹ Brunner2008
⁹² Rosing2002a
⁹³ Pfeidereretal2014
⁹⁴ Rosing2014
⁹⁵ Pfeiderer2013
⁹⁶ Helms2011
⁹⁷ Obert2009
⁹⁸ Wicke2008a
⁹⁹ Bohm2001
¹⁰⁰ Hemming2015b
¹⁰¹ Rosing2014
¹⁰² Pfeiderer2009

Wald¹⁰³ plädiert in seinem Artikel für die Einbeziehung aller populärer Musik in den wissenschaftlichen Diskurs, das gleiche Ziel verfolgt Schlaebitz¹⁰⁴, der sich außerdem zur Ausrichtung der Musikwissenschaft allgemein äußert.

Middleton¹⁰⁵ propagiert ebenfalls die Notwendigkeit einer Neuorientierung der institutionalisierten Musikwissenschaft. Tagg¹⁰⁶ resümiert den Stand der Populärmusikforschung und bemängelt die Trägheit der Disziplin bei ihrer interdisziplinären Ausrichtung. Auf Taggs Beitrag bezogen äußern sich Pedelty¹⁰⁷ und Till¹⁰⁸ zu den Herausforderungen auf dem Weg zur Konsolidierung der Populärmusikforschung.

In ihren Arbeiten treten HELMS¹⁰⁹, HAWKINS¹¹⁰, MEYER¹¹¹, RÖSING¹¹² und McCLARY¹¹³ anhand von **Beispielen** für einen entschlosseneren Umgang der Musikforschung mit populärer Musik ein. Dabei sprechen sie auch Herausforderungen an, die sich aus der Beschäftigung mit populärer Musik ergeben.

Die allgemeine **Relevanz** der Auseinandersetzung mit der Musikgeschichtsschreibung und ihrer Themen und Methoden unterstreicht darüber hinaus das Bestehen des Netzwerks Fachgeschichte der Gesellschaft für Musikforschung. Hier werden einige Projekte aufgelistet, darunter finden sich allerdings keine, die speziell das Verhältnis von populärer Musik und Musikwissenschaft untersuchen.

Wir haben Wissenschaftsforschung und die Subdisziplin Fachhistoriographie kennengelernt, Funktionen und Ansätze, habe besprochen, warum in dieser Arbeit qualitative mit quantitativen Daten verbunden werden sollen, um sinnvolle Fachgeschichtsschreibung durchzuführen, den *Stand* der Forschung in der Musikwissenschaft besprochen.

-> Überleitung zum Forschungsdesign!

¹⁰³ Wald2014

¹⁰⁴ Schlaebitz2007

¹⁰⁵ Middleton1990

¹⁰⁶ Tagg2012

¹⁰⁷ Pedelty2013

¹⁰⁸ Till2013

¹⁰⁹ Helms2002

¹¹⁰ Hawkins1996

¹¹¹ Meyer1995

¹¹² Rosing1994

¹¹³ McClary1990

3 Forschungsdesign

„Fachliteratur ist ein getreuer Spiegel der Entwicklung einer Wissenschaft. Sie ist Ergebnis und Bedingung wissenschaftlicher Arbeit zugleich“

– Dieter B. Herrmann

Die angestrebte Darstellung der Fachhistoriographie der Populärmusikforschung soll, wie im vorherigen Kapitel angedeutet, mit einem mixed-method-Ansatz bearbeitet werden. Dazu werden Erkenntnisse aus traditionellen fachgeschichtlichen Arbeiten um einen bibliometrisch-quantitativen Ansatz erweitert. Es handelt sich daher um scientometrisch gestützte Wissenschaftsgeschichtsschreibung.

Das Forschungsdesign der Arbeit basiert auf der Anreicherung bibliographischer Metadaten und deren Auswertung sowie der Interpretation der Ergebnisse und das in-Kontext-setzen und Diskutieren dieser Ergebnisse unter Berücksichtigung einschlägiger Sekundärliteratur.

Die Arbeit mit bibliographischen Daten setzt die Verfügbarkeit einer entsprechenden Datenbank voraus, in der bibliographischen Daten – sinnvoll aufbereitet – verzeichnet sind. Die Voraussetzungen, die eine Forschungsdatenbank erfüllen muss, sind:

- erschöpfende Auswahl indexierter Quellen
- qualitative hochwertige bibliographische Metadaten
- Eignung für bibliometrische Analysen (Datenmodell und Datenqualität müssen die statistische Auswertbarkeit der Daten gewährleisten)

In der Musikwissenschaft existiert keine solche Datenbank.¹¹⁴ Um das Schrifttum zur populären Musik in der Art und Weise untersuchen zu können wie hier vorgeschlagen wird, muss demnach eine bibliographische Forschungsdatenbank entwickelt werden, die den oben genannten Kriterien gerecht wird.

In diesem Kapitel wird die Erstellung einer solchen Datenbank in den Forschungsstand eingeordnet und die einzelnen vorgenommenen Arbeitsschritte beschrieben und diskutiert.

3.1 Grundlagen

Die Darstellung des Forschungsdesigns beinhaltet u. a. Informationen zu den Besonderheiten bei der Arbeit mit bibliographischen Metadaten, eine Einordnung der Methode in den Forschungsstand, um zu klären, welche Ansätze in dieser Richtung bereits existieren und woran sich dieser Ansatz anlehnt.

¹¹⁴ vgl. Datengrundlage und -Sammlung (??)

Da die Untersuchung auf Basis eines speziellen Datenbestands beruht, soll außerdem die Umsetzung des Forschungsdatenmanagements beleuchtet werden, um die Nachnutzbarkeit der Methode und der Ergebnisse gewährleisten zu können und im Sinne guter wissenschaftlicher Praxis.

3.1.1 Bibliographische Daten

Die Arbeit mit bibliographischen Daten macht es notwendig, in einige Spezifika bibliographischer Daten einzuleiten. Bibliographische Daten sind sog. Metadaten, die notwendig sind, um Literatur eindeutig und unverwechselbar zu beschreiben. Zu den zentralen bibliographischen Angaben zählen Titel, Autor und Veröffentlichungsjahr eines Schriftstücks sowie Angaben zum Verlag und ggf. Angaben zu übergeordneten Publikationen (etwa bei Zeitschriftenartikeln).

In ihrer grundlegenden Anwendung sind werden sie gebraucht um alle Arten verschriftlichter Kommunikation eindeutig zu identifizieren und zu referenzieren, etwa durch Verweise in Texten (Zitationen), in Literaturverzeichnissen (im Sinne zitierter und weiterführender Literatur) oder in Bibliographien (Sammlungen/Verzeichnissen von Publikationen).

Eine Zweite Funktion bibliographischer Angaben betrifft das Auffinden bzw. die Hilfe bei der Auswahl von Literatur in Literaturkatalogen. Für diesen Bereich ist es notwendig, dass den bibliographischen Angaben Informationen hinzugefügt werden, die über die Angaben zur Identifikation hinausgehen. Dabei geht es in erster Linie um inhaltliche Klassifikationen und die Einordnung in bestehende Systematiken.

Hier kommt es zur Unterscheidung zwischen der sog. Formalerschließung¹¹⁵ und der Sacherschließung. Im Zuge der Formalerschließung werden alle Informationen zu einer Publikation aufgenommen, die der Identifikation dienen. Als Sacherschließung (auch: inhaltliche Erschließung/Inhalterschließung) hingegen bezeichnet man in der Bibliotheks- und Informationswissenschaft den Vorgang der Erfassung **inhaltlicher** Informationen zu einer Publikation. Während die im Zuge der Formalerschließung gesammelten Daten in erster Linie für die Identifikation von Publikationen notwendig sind, werden die eine Publikation inhaltlich beschreibenden Daten der Sacherschließung in allen Bereichen des *Information Retrievals*, also der Suche von Literatur aus Literaturverzeichnissen und -datenbanken, eingesetzt.

Zu den Daten, die im Rahmen der Sacherschließung erfasst werden, gehören:¹¹⁶

¹¹⁵ (Godert2011)

¹¹⁶ vgl. <http://homepage.univie.ac.at/juan.gorraiz/konven/einfri.htm>

- Schlagworte (freie und/oder kontrollierte)
- Sachgruppen-Zuordnungen (v. a. im Bibliothekswesen)
- Notationen (Bibliothekssystematik)

Die Sacherschließung wird in erster Linie von ausgebildeten Bibliothekaren durchgeführt¹¹⁷. Im Zuge eines immer höheren Publikationsaufkommens und gleichzeitig einem starken Bedeutungszuwachs von Online-Datenbanken zur Literaturrecherche wird die Sacherschließung zum einen von den Autoren wissenschaftlicher Publikationen selbst (durch die Vergabe von Keywords) sowie durch computergestützte Verfahren ergänzt.

Unterschied zwischen Bibliographie und Katalog

- Bibliographie: Druckschriftenverzeichnis im allgemeinsten Sinn; so wie beschrieben sind sie erschienen (Nachweise/Fachbibliographie)
- Katalog: wie eine Bibliographie + weist für jeden verzeichneten Druck ein Exemplar nach (Katalogfunktion/Bibliothekskatalog)

3.1.2 Einordnung der Methode

Die vorgeschlagene Methodik führt Ansätze aus der Informationswissenschaft (automatische Sacherschließung) und der Bibliometrie zusammen und ist thematisch in der Nähe der Systematische Übersichtsarbeit und der quantitativen Literaturanalyse anzusiedeln

In der Bibliotheksinformatik bzw. der Informationswissenschaft gibt es zwei Forschungsschwerpunkte, in deren Richtung sich die Methode dieser Arbeit geht: Automatische Sacherschließung/Indexierung und KDD (Knowledge Discovery in Databases).

Bei der vorgestellten Methode handelt es sich allerdings vor allem in Hinblick auf den angestrebten Nutzen und die Vorbedingungen nicht um klassisches *automatisches* Indexieren, da im Vorhinein genau festgelegt wird, welche Informationen in den Index aufgenommen werden sollen.

Beim computergestützten Indexieren handelt es sich um eine Kombination aus intellektuellem und automatischen Indexieren. In der Regel wird hierbei ein Dokument zuerst automatisch indexiert, danach wird das Ergebnis - das Indexat - von einem menschlichen Indexierer überarbeitet. Ein Rechner schlägt also Deskriptoren vor und anschließend überarbeitet eine (sachverständige, indexierungserfahrene) Person die Auswahl der Deskriptoren. Ein rein automatisches Indexieren, wie es heute z. B. für Suchmaschinen eingesetzt wird, ist zwar kostengünstiger, erreicht aber nicht dieselbe Qualität.¹¹⁸

¹¹⁷ (Eberhardt2011)

¹¹⁸ https://wiki.infowiss.net/Computergestützte_Indexierung

Methoden der Anreicherung bibliographischer Daten mit inhaltlich-thematischen Informationen, die

a) implizit in den bibliographischen Metadaten vorhanden sind und b) ein kontrolliertes Vokabular ergeben

Das gezielte Anreichern der Daten mit Begriffen und Entitäten, die aus dem Titel, den Keywords der Sacherschließung und den Abstracts extrahiert werden, unterscheidet sich von gängigen Ansätzen aus der Informationswissenschaft in der Zielsetzung. Im Kontext der automatischen Sacherschließung werden automatisiert Keywords aus (Voll-)Texten bzw. Abstracts von Artikeln extrahiert, mit dem vorrangigen Ziel das Information Retrieval durch (Bibliotheks-)NutzerInnen zu verbessern, nicht die statistische Auswertbarkeit zu gewährleisten. Dabei wird i. d. R. mit computerlinguistischen Ansätzen (vgl. Kasprzik 2014)¹¹⁹ gearbeitet.

Statistische Bibliometrie

Systematische Übersichtsarbeit/Systematic Literature Review

Qualitative Literaturanalysen (in der Musikwissenschaft/Populärmusikforschung)

Quantitative Literaturanalysen (in der Musikwissenschaft/Populärmusikforschung) Schwab-Felisch et al.¹²⁰ → untersuchen vorher eingegrenztes Forschungsgebiet

3.1.3 Forschungsdatenmanagement

Im Hinblick auf transparente und nachhaltige Speicherung der Forschungsdatenbank und der Ergebnisse wurde während dieser Arbeit Forschungsdatenmanagement betrieben. Im Folgenden soll beschrieben werden, welche Vorkehrungen vorgenommen wurden, um die Datenbasis und die Analyseergebnisse transparent und sicher vorzuhalten.

Datenspeicherung: Die Sicherung und Vorhaltung der Daten wird durch die Nutzung eines Versionsverwaltungssystems umgesetzt. Die erstellte Datenbank, der erarbeitete Programmcode sowie die Ergebnisse der Analysen werden in einem Git-Repository gespeichert. Durch die von Git verwaltete Versionierung der Daten ist es möglich, alle Bearbeitungsschritte transparent zu speichern und ältere Stände einzelner Dateien wiederherzustellen.

¹¹⁹ (Kasprzik2014)

¹²⁰ (Schwab-Felisch2005)

Datenerhalt: Durch die Nutzung strukturierter, nicht-proprietärer Dateiformate (u. a. Sqlite, JSON, CSV) sowie die spätere Bereitstellung der Daten unter einer offenen Lizenz ist der Zugriff auf die Daten weitestgehend unabhängig von kommerzieller Software und Systemwechseln, sodass ihre Lesbarkeit auch in Zukunft gewährleistet ist.

Data Sharing: Durch die Nutzung eines Repositories wird gewährleistet, dass die Daten und die Anwendungen zur Erstellung einer bibliographischen Forschungsdatenbank der Forschungsgemeinde zugänglich sind und unterschiedlich eingesetzt und ausgewertet werden können. Dabei können durch sog. *Forks* auch alternative Analyseansätze an dem Datenmaterial erprobt werden. Zur Veröffentlichung dieser Arbeit ist die Veröffentlichung des Repositories unter Berücksichtigung datenschutzrechtlicher Bestimmungen geplant.

Datendokumentation: Die einzelnen Arbeitsschritte bei der Zusammenstellung der Forschungsdatenbank, wie etwa die Clustering-Schritte bei der Datenbereinigung mit Open Refine,¹²¹ werden nachvollziehbar dokumentiert. Open Refine ermöglicht es, die ermittelten Cluster im JSON-Format zu exportieren. Die im Rahmen der Auswertung erstellten Code-Fragmente und resultierende Teil-Datensätze bzw. Analyseergebnisse werden darüber hinaus systematisch im Repository abgelegt und mit persistenten Identifiern versehen, um eine eindeutige Zuordnung zu ermöglichen.

3.2 Das Programm

3.2.1 Kurzbeschreibung des Programms

Zusammengefasst leistet das Programm zweierlei: Es reichert bibliographische Daten um spezielle Metadaten an und führt die Umformung der Daten in ein Datenbankformat durch, dessen Datenmodell sich zur statistischen Auswertung eignet.

¹²¹ vgl. ??

Das Programm verarbeitet bibliographische Aufnahmen im RIS Format¹²². Aus den an das Programm übergebenen bibliographischen Daten¹²³ werden automatisch Informationen über die thematisierten Musiker, Genres und Länder extrahiert. Die Ausgangsdaten werden mit den extrahierten Informationen angereichert. Darüber hinaus werden die Ausgangsdaten normalisiert und vom semi-strukturierten Datenformat¹²⁴ *RIS* in eine strukturierte SQLite-Datenbank transferiert.

Für die Erarbeitung einer datengestützten Fachgeschichte werden bibliographische Daten aus der Fachdatenbank RILM exportiert und verarbeitet. So ergibt sich nach der Bearbeitungen der Ausgangsdaten eine Datenbank aller Publikationen, die zwischen 1967 und 2015 erschienen sind und die bei RILM mit dem Suchbegriff *popular music* gefunden werden. Die Datenbank enthält zusätzlich zu den in RILM enthaltenden Metadaten weitere Spalten, in denen die Informationen aufgeführt werden, die aus den Einträgen extrahiert wurden.

In erster Linie werden die Titel (bei Aufnahmen in nicht-westlichen Schriftsystemen auch die übersetzten Titel) der Publikationen analysiert, es werden aber auch die enthaltenen Abstracts und die RILM eigene Kategorisierung, die sog. *Keywords*, untersucht. Die *Keywords* fügen dabei den primären ausgewerteten Metadaten eine zweite Ebene hinzu. Das Vorgehen ermöglicht es, mit verhältnismäßig wenig Aufwand eine umfassende Analyse von einem großen Datenbestand durchzuführen.

Die Extraktion der gewünschten Informationen wird über einen Abgleich der Ausgangsdaten mit Wörterlisten erreicht. Übereinstimmungen werden automatisiert jedem Literatureintrag zugeordnet. Den Wörterlisten kommt daher eine zentrale Bedeutung zu: Um die Titel der Publikationen auf die Erwähnung von Musikern hin zu untersuchen, benötigt das Programm eine Liste, mit der die Titel etc. abgeglichen werden, in der im Idealfall alle in den Ausgangsdaten vorkommenden Musiker enthalten sind. Das gleiche gilt für die Genres. Um die Zuverlässigkeit dieses Matching-Prozesses zu maximieren, müssen die Listen so umfassend wie möglich sein.

3.2.2 Datengrundlage und -Sammlung

Die Voraussetzungen für die Forschungsdatenbank wurden bereits im Forschungsdesign (??) erörtert. Es sei hier wiederholt, dass für eine sinnvolle und erfolgreiche Analyse die Datengrundlage eine entscheidende Rolle spielt.

¹²² vgl. [RIS]

¹²³ Hierzu siehe Datengrundlage und -Sammlung (??)

¹²⁴ „Als Semistrukturierte Daten bezeichnet man Informationen, die keiner allgemeinen Struktur unterliegen, sondern einen Teil der Strukturinformation mit sich tragen.“ (RIS Daten beinhalten nicht alle Infos!)

Als Datengrundlage für die Realisation einer Forschungsdatenbank, anhand derer wissenschaftsgeschichtliche Fragestellung mit bibliometrischen Methoden beantwortet werden können, kommen nur bibliographische Fachdatenbanken in Frage.

Bibliographische Fachdatenbanken sind Verweisdatenbanken, deren Zielsetzung die möglichst umfassende Abbildung der Publikationstätigkeit einer wissenschaftlichen Disziplin ist. Dabei unterscheiden Sie sich von **Bibliotheks- und Verbundkatalogen** (bspw. den Katalogen des GBV¹²⁵ oder des SWB)¹²⁶, die auch Verweisdatenbanken sind, die aber in erster Linie das Auffinden von Literatur, die in einzelnen Bibliotheken bzw. auch in Bibliotheksverbunden vorhanden bzw. lizenziert¹²⁷ ist, ermöglichen sollen. Eine ebenfalls andere Zielsetzung verfolgen **Volltextdatenbanken** (z. B. JSTOR oder Project MUSE). Hier geht es nicht um die Vorhaltung bibliographischer Formaldaten sondern um den Zugang zu Volltexten. Schließlich kann im Kontext einer Geisteswissenschaft wie der Musikwissenschaft kein Gebrauch von sog. **Zitationsdatenbanken** gemacht werden, die ebenfalls Literaturdatenbanken sind, die durch Daten zu der Zitationstätigkeit innerhalb der bibliographischen Daten erweitert sind (bspw. Web of Science¹²⁸/Scopus¹²⁹). Diese haben eine unzureichende Abdeckung von Veröffentlichungen in der Musikwissenschaft, da sie fast ausschließlich wissenschaftliche Zeitschriften (Journals) auswerten und kaum Monographien. Bei der Wahl der Journals findet ebenfalls nur eine rudimentäre Abdeckung des Musikschritftums statt: So ergibt ein Abgleich der in Quellen von Scopus und RILM eine klares Bild: Nur 69 Journals werden sowohl von RILM und Scopus beachtet, 1500 Journals hingegen sind exklusiv in RILM ausgewertet.¹³⁰

Für die Musikwissenschaft existieren die folgenden **bibliographischen Fachdatenbanken**:

- Bibliographie des Musikschritftums Online (BMS Online¹³¹)
 - überschneidet sich mit RILM, gleiche Redaktion.
 - begrenzt auf Deutschland
- Popular music studies database¹³²

¹²⁵ Gemeinsamer Bibliotheksverbund (<https://gso.gbv.de/>)

¹²⁶ Südwestdeutscher Bibliotheksverbund(<http://swb.bsz-bw.de>)

¹²⁷ Stichwort Ebooks

¹²⁸ <https://webofscience.com>

¹²⁹ <https://scopus.com>

¹³⁰ Diese Zahlen wurden durch einen Vergleich der Scopus Source List (https://www.elsevier.com/___data/assets/excel_doc/0015/91122/Scopus_Source_List.xlsx) mit der Liste der von RILM ausgewerteten „core and secondary“ Journals (<http://journals.rilm.org/index.asp?alpha=all>) erhoben.

¹³¹ <http://www.musikbibliographie.de>

¹³² <http://iismc.cini.it/>

- beruht auf Einreichungen der Autoren, extrem kleiner Datenbestand (9592 Einträge, Stand: 8.9.2017),
- keine Möglichkeiten zum Export, kein allgemeiner bibliographischer Standard
- RILM
 - international
 - größter Datenbestand

Aufgrund der Voraussetzungen¹³³, die die bibliographischen Daten für eine sinnvolle Analyse erfüllen müssen, kommt als Datenbasis für die Betrachtung der Fachgeschichte der Populärmusikforschung nur die Datenbank RILM in Frage.

RILM ist Grundlage musikwissenschaftlichen Arbeitens und kann als essenziell für die Disziplin bezeichnet werden, da es die umfassendste Datenbank des Musikschrifttums bereitstellt und somit oftmals die erste Anlaufstelle bei der Literaturrecherche für Musikforscher darstellt. RILM ist jedoch als Fachbibliographie dazu konzipiert, Literatur zu verzeichnen und auffindbar zu machen und nicht, um einen Überblick über das Musikschrifttum zu schaffen oder es gar zu ermöglichen, statistisch mit den bibliographischen Daten zu arbeiten.

Das Projekt *Répertoire International de Littérature Musicale* (Internationales Repertorium der Musikliteratur) wurde 1966 an der City University New York von dem Musikwissenschaftler Barry S. Brook ins Leben gerufen. Es wird gefördert von der Internationale Gesellschaft für Musikwissenschaft (IMS), der Internationalen Vereinigung der Musikbibliotheken, Musikarchive und Musikedokumentationszentren (IAML bzw. IVMB) und dem International Council for Traditional Music (ICTM).

Der Aufbau von RILM ist dreigeteilt: Die Angestellten im International Center in New York erstellen die Datenbank und arbeiten an Ergänzungen und Kontrollen der Einreichungen aus den Landesredaktionen. Hier überwacht man außerdem die Bereitstellung der Datenbank durch Datenbankanbieter. Eine **Commission Internationale Mixte** leitet RILM und ist beratend tätig. Es besteht aus je vier Repräsentanten der drei fördernden Organisationen und der seit 1996 amtierenden Chefredakteurin Barbara Dobbs Mackenzie. Die Zuarbeiten werden von einer Vielzahl **nationaler Komitees** durchgeführt. Die nationalen Komitees machen die Aufnahmen von Zeitschriften und Monographien, die in den jeweiligen Ländern erscheinen, und werten diese dabei aus. Auch das Verfassen der in RILM enthaltenen Kurz'zusammenfassungen, der Abstracts, wird von den Redaktionen übernommen, wobei z. T. auch die Datenlieferanten die Abstracts bereitstellen: „In manchen

¹³³ vgl. Forschungsdesign (??)

Fällen liefern die Musikforscher und Bibliothekare Abstracts; manchmal schreiben die Autoren selbst Abstracts“.¹³⁴ Die Anzahl der Nationalen Komitees wird auf den Internetseiten von RILM uneinheitlich angegeben, die Abdeckung schwankt zwischen 48 und 60 Ländern.

RILM wirbt damit, dass alle Publikationstypen ausgewertet werden, die weltweit in allen Sprachen erscheinen.¹³⁵ Das Vorgehen der RILM Redaktionen bei der Erfassung der bibliographischen Daten ist in den 2002 veröffentlichten Richtlinien der Organisation zur Berichterstattung festgelegt. Zusammengefasst wertet RILM aus:¹³⁶

1. Zeitschriften-Abdeckung, aufgeteilt in

- *Haupt-Zeitschriften* (Core journals): „die wichtigsten Zeitschriften und ganz der Musikforschung gewidmet“; „RILMs International Center darf auch Zeitschriften als Haupt-Zeitschriften designieren“
- *Wichtige Zeitschriften* (Secondary journals): „Zeitschriften mit wichtigen Beiträgen für die Musikforschung“; „Artikel, die RILM Material sind [...] werden abstrahiert“; „Der Status ‚Wichtige Zeitschriften‘ wird von der nationalen Kommission vergeben“
- *Andere Zeitschriften* (Tertiary journals): „zielen weniger auf Musikforschung ab und werden nicht systematisch bearbeitet. Ihr Inhalt wird selektiv abstrahiert“
- *Zeitschriften aus anderen wissenschaftlichen Disziplinen*: „Einige tausend Artikel über Musik von interdisziplinären Foren und Zeitschriften, die nicht der Musikforschung gewidmet sind“

2. Sammlungen (Festschriften, Konferenzberichte, Essay-Sammlungen)
3. Monografien und andere Bücher (auch Nachdrucke und Neuauflagen, Übersetzungen)
4. Dissertationen und einige Magister-/Diplomarbeiten: „Magisterarbeiten sind nur dann eingeschlossen, wenn sie bedeutende neue Informationen oder eine neue Interpretation beinhalten.“
5. Elektronische Ressourcen, wie Forschungswebseiten, Datenbanken und Datenträger
6. Gewisse Reviews: „Rezensionen, die eine bestimmte Publikation beschreiben oder kritisch beurteilen sind eingeschlossen, solche, die lediglich die Existenz der Publikation erwähnen sind nicht miteingeschlossen.“
7. Technische Zeichnungen von historischen Instrumenten

¹³⁴ vgl. <http://rilm.org/aboutUs/?language=German>

¹³⁵ <http://www.rilm.org/scope/>

¹³⁶ <http://www.rilm.org/scope/coveragede.php>

Ausgeschlossen werden aus der Erfassung Publikationen, die „wahrscheinlich nicht von Wert für Musikforscher“ sind. Dazu zählen:

- Praktische Handbücher und ‚how-to‘ Bücher, außer wenn sie moderne Editionen von wichtigen historischen Texten sind, substanzielle Rezensionen erhalten haben oder anderweitig von Wert für Musikforscher sind.
- Rudimentäre Instruktions-Handbücher - Grundrisse für Schulgebrauch. - Rezensionen, die lediglich Veröffentlichungsbekanntgaben sind

Der Hauptaugenmerk von RILM liegt demnach auf thematisch relevanten Zeitschriften inklusive der Auswertung von Zeitschriften aus benachbarten wissenschaftlichen Disziplinen, deren Inhalt dennoch für die Musikforschung von Belang ist. Monographien, Sammelbände und Dissertationen aus der Musikforschung werden ebenfalls ausgewertet. Interessant ist die Vorauswahl im Bereich der Reviews, die von RILM getroffen wird.

Der Datenbestand von RILM wird von Redaktionen, den nationalen Komitees, gespeist. Ein Blick auf die Auflistung von RILMs „Global Network“ zeigt, dass die Verteilung der RILM Redaktionen klar eurozentrisch ist. So liefern nur zwei Komitees Daten aus Afrika, dem zweitgrößten Kontinent der Erde, an; eines der aus Afrika stammenden Komitees ist übergreifend für ganz Afrika zuständig.

Die einzelnen RILM Redaktionen sind darüber hinaus unterschiedlich gut besetzt, was zumindest in der Theorie einen Einfluss auf die Erfassungsqualität und -Geschwindigkeit haben kann. Die RILM-Redaktion Deutschland ist mit neun Mitarbeitern die am besten besetzte außerhalb der Vereinigten Staaten von Amerika (ist gleichzeitig die Redaktion für die BMSOnline). Zum Vergleich: Das Komitee für das Vereinigte Königreich hat drei, das von China vier, Frankreich ebenfalls vier, Mazedonien und Afrika jeweils einen Mitarbeiter.

Die Vollständigkeit der Erfassung und die Erfassungsqualität hängen eng mit der Arbeit der jeweiligen Landesredaktionen zusammen.

„Publications represented in RILM abstracts originate from 178 countries and are in 143 languages, and are compiled through a collaborative effort between the International Center and 48 committees across the globe. The nearly one million records in the complete database include titles in the original language and English translation, bibliographic information, and abstracts in English and, when available, the language of publication. Records for publications issued in languages using non-roman writing systems (such as Cyrillic, Chinese, Japanese, Arabic, Korean, and Hebrew) are represented bilingually. Currently RILM abstracts grows at a rate of 50,000 records each year.“¹³⁷

Das Datenmaterial wird die Literatur umfassen, die zwischen 1967 und 2015 veröffentlicht wurde und in RILM verzeichnet ist.

¹³⁷ <http://www.rilm.org/scope/>

1967 begann die bibliographische Erfassung durch RILM, 2015 ist als Endpunkt sinnvoll, da von einem Berichtsverzug ausgegangen werden muss. Der Zeitpunkt, an dem die vollständige retrospektive Erfassung von Literatur durch die RILM Redaktionen abgeschlossen ist, lässt sich nicht bestimmen.¹³⁸

Da aktuelle Entwicklungen wichtig sind und nicht eindeutig festzulegen ist, wann die in RILM enthaltenen Informationen einen Anspruch auf Vollständigkeit erheben, ist es sinnvoll, auch jüngste Veröffentlichungen miteinzubeziehen.

Bei der Erfassung der Literatureinträge werden die Publikationen von den RILM Redaktionen zum einen einer feststehende Klassifikation zugeordnet, zum anderen werden Schlagworte (Keywords) vergeben. Die Keywords basieren nicht auf einem kontrollierten Vokabular. In den Keywords werden u. a. Informationen zu den Methoden einer Publikation, thematisierten Beziehungsstrukturen sowie inhaltlich relevanten Stichpunkten vermerkt.

Die RILM Klassifikation ist an eine Dezimalklassifikation angelehnt. 15 sogenannte „Super Classes“ sind quasi als Überschriften zur Gliederung vorhanden, unter ihnen sind 90 Klassen sortiert. Vier der Super-Klassen sind gleichzeitig auch normale Klassen, die in die Dezimalklassifikation eingereiht sind, z. B. 39 Popular Music.

Beispiel des Aufbaus:

- 1
- 2 39.0 General
- 3 39.2 Africa (general)
- 4 39.21 North Africa
- 5 39.22 Sub-Saharan Africa

Insgesamt sind die in RILM erfassten Literatureinträge in 243 Klassen (inkl. aller Subklassen) eingeteilt.

Die Super-Klassen beinhalten unterschiedlich motivierte Oberbegriffe, sie sind also nicht strikt monothematisch organisiert. Es besteht keine Beschränkung auf bspw. eine thematische Auswahl oder eine Klassifikation nach Gattungen. Die Verschlagwortung beruht auf einer Einteilung der erfassten Literatur sowohl nach ihrem mutmaßlichen Einsatzzweck bzw. ihrer Gattung („Reference and Research Materials“, „Imaginative Literature“, „Collected Writings“) als auch durch eine inhaltlich- thematische Differenzierung nach Musikstilen bzw. stilistischen Zuordnungen („Western Art Music“, „Traditional Music and Non-Western Art Music“, „Jazz and Blues“, „Popular Music“). „Western Art Music“ ist in der zweiten Ebene nach Epochen aufgegliedert, „Traditional Music and Non-Western Art Music“ hingegen nach Regionen/geographischer Zuordnung. Einige Super-Klassen

¹³⁸ Nach Informationen von Carsten Schmidt (Leiter der deutschen RILM Redaktion am Staatlichen Institut für Musikforschung, Berlin), Email vom 29.3.2017 mit dem Autor.

sind praktisch nicht weiter untergliedert („Universal Perspectives“ und Jazz and Blues), andere sind sehr feingliedrig unterteilt. Weitere Oberbegriffe sind: „Sound Sources“, „Performance Practice and Notation“, „Theory, Analysis and Composition“. Auch interdisziplinäre Literatur bzw. Literatur aus anderen, nicht musikwissenschaftlichen Zusammenhängen ist vertreten: „Pädagogik“, „Musik und andere Künste“, „Musik und verwandte Disziplinen“ sowie „Musik in liturgischen und rituellen Kontexten“.

Die von RILM verwendete Klassifikation zur Verschlagwortung von Literatur wirkt pragmatisch, da sie keinen monothematischen Aufbau vorweist sondern die Ansprüche einer Fachdatenbank abbildet. Es handelt sich hier nicht um eine Klassifizierung aus dem Bibliotheksumfeld, der Nutzen der Klassifikation liegt im schnellen Auffinden von musikwissenschaftlicher und anderer, für die wissenschaftliche Arbeit mit Themen aus einem musikalischen Kontext relevanter Literatur.

Die RILM Keywords beinhalten neben der Angabe der schon erwähnten RILM Klassifikation eine Vielzahl weiterer inhaltlicher Schlagworte. Dabei ist zu bemerken, dass es sich hierbei weder um eine Ontologie noch um ein kontrolliertes Vokabular handelt. Deshalb ist es in Hinblick auf die geplante statistische Auswertung notwendig, ein (Vor-)Sortierung und sinnvolle an Keywords zu treffen, die ausgewertet werden sollen.

In den Keywords finden sich Angaben bzw. Informationen zu

- Eigennamen (bzw. Musikernamen): z. B. Elton John
- Zeitangaben: 1950s, 1957, 1960s-early 21st century
- Ländern
- Genres
- Instrumenten
- Angaben zur inhaltlichen Ausrichtungen der Publikation: Biographie, Enzyklopädie, Analyse...
- Angaben zu Herausgebern, Serien, Buchtiteln, z. B.: gesellschaft für kanada-studien in deutschsprachigen ländern /handbuch new age musik: auf der suche nach neuen musikalischen erfahrungen /historisches museum der stadt wien
- Differenzierten Angaben zu Publikationstypen: catalogues and indexes, obwohl die Typenangabe folgende ist: TY <!--"~ BOOK\-->
- Inhaltliche Referenzen: compared with, concept of, depiction of, influence of, reception of, role of, use of etc.

Obgleich RILM für die Arbeit an einer integrativen Fachhistoriographie alternativlos ist, gibt es einige Punkte im Hinblick auf die Integrität der Daten zu beachten. Diese Mankos wiegen jedoch nicht schwer... -> RILM ist alternativlos & Probleme liegen ein Stück weit in der Natur der Sache (der bibliographischen Fachdatenbank)

- Sacherschließung nicht standardisiert: Kein kontrolliertes Vokabular (vgl.)
- Richtlinien (z. B. zur Einordnung in die RILM Klassifikation oder zu der Art und Weise der inhaltlichen Erfassung allgemein) nicht öffentlich und dadurch auch nicht transparent/nachvollziehbar
- Landesredaktionen unterschiedlich gut besetzt: „Erfassung Bias“
- Erfasst keine „graue Literatur“¹³⁹ (gilt für alle bibliographischen Datenbanken)

Der Zugriff auf die Datenbank RILM erfolgt über den Contentprovider EBSCO-Host¹⁴⁰. Dieser ermöglicht den Datenexport in den folgenden Formaten:

- XML
- MARC21
- CSV
- RIS
- BibTeX
- RefWorks
- EndNote Web
- EasyBib

Mit den Formaten XML und CSV bietet EBSCOHost zwei klassische Datenformate zum Austausch strukturierter Daten an. MARC21 ist ein Format, das im Bibliotheksumfeld eingesetzt wird. Die anderen Formate sind speziell für den Import der bibliographischen Daten in Literaturverwaltungsprogramme gedacht.

Im folgenden soll begründet werden, warum für den Einsatzzweck dieser Arbeit nur das Datenformat *RIS* in Frage kommt. Als Voraussetzung für die Funktionalität des Programms wird das Format darüber hinaus vorgestellt.

Das Dateiformat *RIS* ist ein Format für bibliographische Daten, das mit der Literaturverwaltungssoftware *Reference Manager* von Thomson Reuters¹⁴¹ ausgeliefert wurde. Der Name ist ein Acronym des Namens der Firma *Research Information Systems*, die die Software entwickelten und zu Thomson Reuters gehörten.

RIS Dateien sind Text-Dateien, die mit jedem beliebigen Text-Editor auf allen Computerplattformen gelesen werden können. *RIS* Dateien sind semi-strukturiert, d. h. dass sie einem standardisierten Aufbau folgen und die Dokumentation ihres Datenschemas in sich tragen. Sie können jedoch nicht ohne Umwege in eine komplexe Datenbank transferiert werden.

¹³⁹ Graue Literatur: Hochschulschriften, Working Papers von Forschungsinstituten, Unternehmensschriften (vgl. (Stock2000))

¹⁴⁰ RILM war vormals auch beim Datenbankanbieter ProQuest verfügbar, ist es aber nicht mehr. (Stand: September 2017)

¹⁴¹ Die 1984 erstmals erschienene Anwendung *Reference Manager* war der Vorgänger bzw. ein Konkurrenzprodukt zu *EndNote*, welches seit 2017 zu der Firma Clarivate Analytics gehört.

Ähnlich wie Daten in den Formaten CSV, XML oder JSON liegen die Daten im RIS-Format strukturiert vor und sind so sowohl für Menschen als auch für Maschinen lesbar. Im Gegensatz zu Nur-Text Daten bieten diese Formate Auszeichnungen bzw. Strukturierungen an. Der Inhalt der Dateien ist somit weiterhin für Menschen lesbar (im Gegensatz etwa zu binären Daten, die ausschließlich von Maschinen gelesen werden können), durch die Strukturierung der Daten durch *Tags*¹⁴² sind diese Formate aber für Computer „lesbar“, d. h. dass Strukturen und zusätzliche Informationen enthalten sind und von Computerprogrammen als solche verstanden werden.

Ein einfaches Beispiel für maschinenlesbare Auszeichnung stellt der „Code“ für das Fettdrucken von Text in für Webdokumente in HTML dar: Der Text wird eingefasst in den Tag ``, das „b“ steht dabei für *bold*: `Text`, der fett gedruckt werden soll``.

Der Aufbau von RIS Dateien, die Literatureinträge enthalten, ist wie folgt: Jede Zeile der Textdatei muss mit einem Identifikator (im Folgenden wird der englische Begriff *Tag* genutzt) bestehend aus zwei Großbuchstaben beginnen, der festlegt, welcher Art der folgende Inhalt der Zeile ist. Es folgen genau zwei Leerzeichen, ein Bindestrich und ein weiteres Leerzeichen. Dann beginnt der Inhalt der Zeile. Jede Literaturangabe beginnt mit der Angabe des Publikationstyps (Tag: TY) und endet mit dem Tag ER für „End Reference“. Literaturverwaltungsprogramme erkennen, dass nach dem ER-Tag (der gefolgt von einer leeren Zeile sein kann, aber nicht muss) ein neuer Eintrag folgt. Dazwischen gibt es eine Vielzahl an Identifikatoren, die gebraucht werden können um den jeweiligen Eintrag zu beschreiben und Informationen zu transportieren. Wichtige Felder sind u. a. TI (Titel), AU (Autor), AB (Abstract), CY (City), PB (Publisher), KW (Keyword) sowie SP und EP (Start- bzw. End-Page).¹⁴³ Ein Minimalbeispiel für einen Literatureintrag im RIS-Format aus RILM sieht wie folgt aus:

```

1 TY - CHAP
2 ID - 2005-26443
3 AU - Rüchel, Peter, (Author)
4 TI - 'Feels so good, hurts so bad': Der Rockpalast live auf Sendung
5 T2 - Collected Work: Rock! Jugend und Musik in Deutschland.
   Published by: Berlin, Germany: Links, 2005. Pages: 79-85. (AN:
   2005-26435).
6 TT - 'Feels so good, hurts so bad': Rockpalast broadcast live
7 CY - Berlin
8 PB - Links

```

¹⁴² Ein Tag wird zur Auszeichnung von Daten genutzt.

¹⁴³ Für eine Übersicht aller gängigen Tags siehe http://endnote.com/sites/rm/files/m/direct_export_ris.pdf S. 10ff. (abgerufen am 10.9.2017)

```

9  N1 - Accession Number: 2005-26443. Record Type: Main Record Document
    Type: Article in a collection of essays. Includes: illustrations
    . Language: German.
10 AB - Discusses the German music television show Rockpalast, which
    was broadcasts live on German television station WDR from 1974
    onwards. The show has featured hundreds of rock and jazz bands
    and also organized festivals.
11 KW - 39: Popular music -- Europe
12 KW - rock -- Germany -- Rockpalast show -- 1974 to present
13 KW - Germany -- popular music -- Rockpalast show -- 1974 to present
14 KW - politics -- Germany -- rock -- 1950s-70s
15 KW - mass media -- television -- Germany -- Westdeutscher Rundfunk (
    WDR) -- Rockpalast show -- 1974 to present
16 UR - http://search.ebscohost.com/login.aspx?direct=true&db=rih&AN=
    A849488&site=ehost-live
17 DP - EBSCOhost
18 DB - ri h
19 ER -

```

Code 1 – Minimalbeispiel für eine RIS-Datei

Der beispielhafte Eintrag ist ein Kapitel in einem Sammelband:

¹⁴⁴

Der Contentprovider EBSCOHost bietet neben einem Export im RIS-Format weitere Exportformate an, von denen v. a. das XML-Format von Interesse für einen Einsatzzweck wie den hier vorgeschlagenen ist. XML ist eine weitverbreitete Auszeichnungssprache, die sowohl für Menschen als auch Maschinen lesbar ist. Im Vergleich zu RIS Dateien besitzen XML Dateien eine bessere Struktur und sind von Computerprogrammen sehr gut zu verarbeiten, da für alle gängigen Sprachen sog. XML-Parser existieren, die XML-Daten und ihre Struktur lesen und weiterverarbeiten können.

Im Bereich der Literaturdatenbanken ist jedoch das RIS-Format weiter verbreitet, es wird von fast allen Contentprovider angeboten. RIS ist aufgrund des begrenzten Einsatzgebiets außerdem – im Gegensatz zu XML – klar standardisiert: Ein RIS Export sieht bei jeder Literaturdatenbank gleich aus. Bei XML kann das verwendete Schema, also die Art und Weise, in der einzelne Informationen kodiert/ausgezeichnet sind, theoretisch von Anbieter zu Anbieter unterschiedlich ausfallen.

Um den Unterschied zwischen RIS und XML zu verdeutlichen (auch was die Lesbarkeit der Daten angeht), sei hier der gleiche Literatureintrag des RIS Beispiels als XML Datei dargestellt:

```

1  <?xml version="1.0" encoding="UTF-8"?>

```

¹⁴⁴ Ruchel2005.

```

2 <records>
3   <rec resultID="1">
4     <header shortDbName="rih" longDbName="RILM Abstracts of Music
5       Literature (1967 to Present only)" uiTerm="A849488">
6       <controlInfo>
7         <bkinfo>
8           <bt1>Rock! Jugend und Musik in Deutschland</bt1>
9           <aug />
10          </bkinfo>
11          <revinfo />
12          <dissinfo />
13          <jinfo />
14          <pubinfo>
15            <place>Germany</place>
16          </pubinfo>
17          <artinfo>
18            <tig>
19              <atl>'Feels so good, hurts so bad': Rockpalast
20                broadcast live</atl>
21            </tig>
22            <aug>
23              <au>Rüchel, Peter</au>
24            </aug>
25            <sug>
26              <subj type="unclass">rock -- Germany -- Rockpalast
27                show -- 1974 to present</subj>
28              <subj type="unclass">Germany -- popular music --
29                Rockpalast show -- 1974 to present</subj>
30              <subj type="unclass">politics -- Germany -- rock
31                -- 1950s-70s</subj>
32              <subj type="unclass">mass media -- television --
33                Germany -- Westdeutscher Rundfunk (WDR) --
34                Rockpalast show -- 1974 to present</subj>
35              <subj type="major" />
36            </sug>
37            <ab>Discusses the German music television show
38              Rockpalast, which was broadcasts live on German
39              television station WDR from 1974 onwards. The show
40              has featured hundreds of rock and jazz bands and
41              also organized festivals.</ab>
42            <pubtype>Book</pubtype>
43            <doctype>Article in a collection of essays</doctype>
44          </artinfo>
45          <language>English</language>
46        </controlInfo>
47        <displayInfo>
48          <pLink>

```

```

38         <url>http://search.ebscohost.com/login.aspx?direct=
           true&db=rih&AN=A849488&site=ehost-live
           </url>
39     </pLink>
40 </displayInfo>
41 </header>
42 </rec>
43 </records>

```

Code 2 – Minimalbeispiel für eine XML-Datei

Die XML kodierten Literatureinträge eines EBSCOHost-Exports unterscheiden sich neben dem offensichtlichen Aspekt der Lesbarkeit auch in der Anlage mancher Felder. So beinhaltet der XML Export u. a. eine Direktangabe der Landes der Veröffentlichung (<pubinfo><place>Germany</place></pubinfo>) und gibt die Sprache der Veröffentlichung als eigenes Feld direkt an (die Sprache wird in der RIS Version im Feld N1 Notizen erwähnt). Es gibt aber wichtige Angaben, die in der XML Version fehlen. So ist die Angabe der Major Keywords von RILM nicht in der XML Fassung enthalten (im konkreten Beispiel: 39: Popular music -- Europe) ebenso fehlen Pagination, Anhänge (Includes), die Zugangsnummer (Accession Number) und der Digital Object Identifier (DOI).

Aus einer rein technischen Sicht wäre XML das zu bevorzugende Format, auf dessen Basis die Forschungsdatenbank erstellt werden sollte. Da jedoch die Datenqualität in diesem Fall der wichtigste Faktor ist, spricht manches gegen die Wahl von XML (und anderen Formaten) und für das zwar etwas weniger strukturierte, dafür aber inhaltlich vollständigste Format RIS.

Stichtag der Datenerhebung war der 1.3.2018

Suchstring: Suchmodus: Boolean/Phrase ((TX (popular AND music))NOT PT "Music recording, sound recording")

Exportvorgehen:

Der Datenprovider EBSCOHost ermöglicht den Download von bis zu 25000 Suchergebnissen. Da die Volltextsuche zunächst mehr als 100000 Ergebnisse

Eingrenzung auf < 25000 Hits je Export (Begrenzung von EBSCOHost) durch die Auswahl folgender Gruppen:

1. Alle Sprachen ohne Englisch und Deutsch
2. Englische Journal-Artikel
3. Englisch: Alle Sources ohne Journal-Artikel und Bücher
4. Englisch Bücher ohne Subject „Performers–popular music“
5. Englische Bücher nur Subject „Performers–popular music“
6. Alle Deutschen Quellen

Jahresproblematik:

107458 Objekte vs. 78706 Objekte mit Datumsbeschränkung

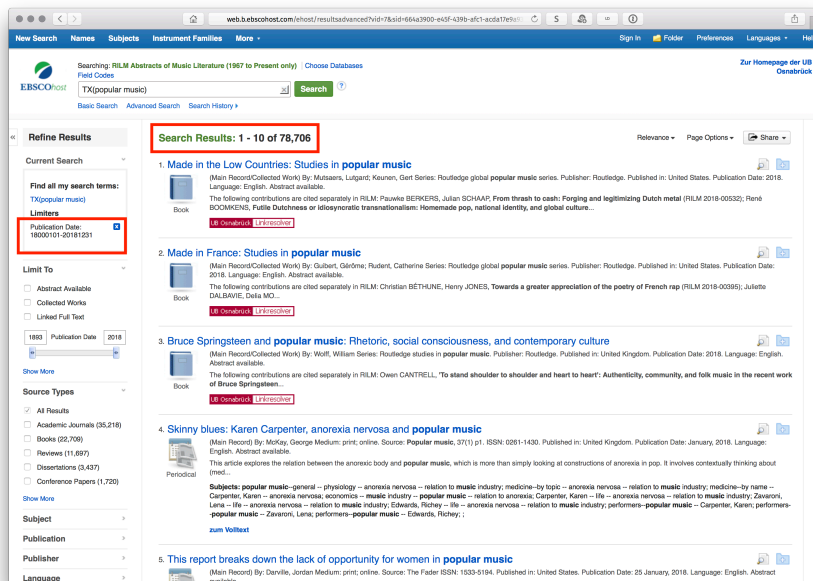


Abbildung 2 – RILM Export mit Beschränkung des Publikationsdatums

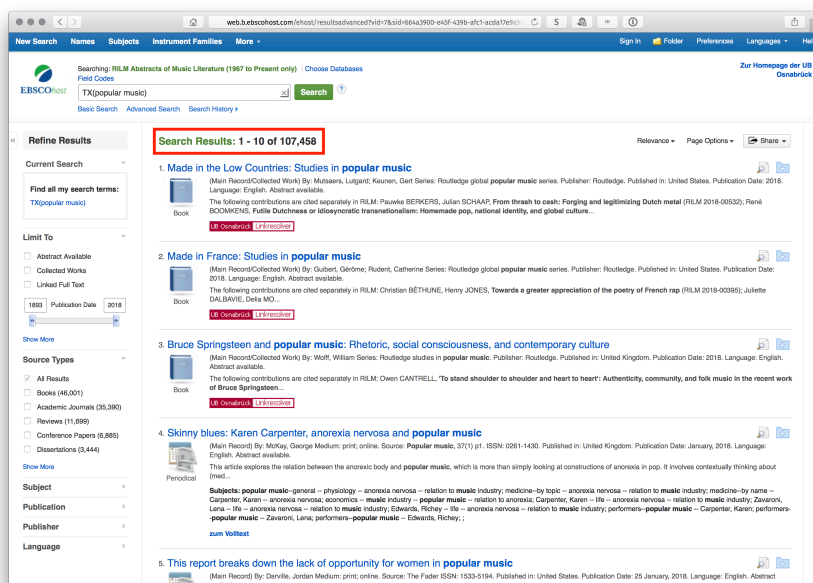


Abbildung 3 – RILM Export ohne Beschränkung auf Publikationsdatum

3.2.3 Kontrolliertes Vokabular

Um eine möglichst umfassende Bearbeitung der gesammelten bibliographischen Daten durch das Programm zu gewährleisten, sind die Wörterlisten, mit denen das Datenmaterial abgeglichen wird von höchster Bedeutung. Das Ziel bei der Erstellung der Wörterlisten, die für den Abgleich der bibliographischen Daten erstellt werden, ist es, eine erschöpfende Sammlung von Termini zur Verfügung zu stellen, damit mit größtmöglicher Wahrscheinlichkeit alle relevanten Begriffe identifiziert werden können.

Da die Frage, die an das Datenmaterial gestellt wird zunächst komplett offen ist, also beispielsweise keine Genre oder Musiker ausschließen soll, kann keine endgültige Liste mit Suchbegriffen erstellt werden, es kann nur versucht werden, eine möglichst umfangreiche Liste zu erstellen.

Die Listen setzen sich zusammen aus Namen und Begriffen, die maschinell aus dem Internet extrahiert werden. Mithilfe des sog. *Web scrapings* lassen sich große Mengen an Informationen automatisiert von Webseiten abrufen. Somit können viele Quellen zusammengefügt werden, ohne dass Flüchtigkeitsfehler bei der Übertragung entstehen können.

Als Quellen für die erstellten Wörterlisten fungieren Kollaborative Onlineangebote: Die Enzyklopädie *Wikipedia* sowie zwei Plattformen für Musiksammler, *Discogs.com* und *Musicbrainz.org*. Durch die Kombination der Inhalte mehrerer Internetseiten, die durch die globalen Bemühungen großer Benutzer-Communities erstellt werden, wird sichergestellt, dass die Wörterlisten eine maximal große Bandbreite an Optionen zum Abgleich bieten. Die Fokussierung auf Seiten, die auf Inhalten basieren, die von Nutzern eingepflegt werden spielt für die Universalität und Unvoreingenommenheit der Daten eine wichtige Rolle.

Darüber hinaus wird ein kontrolliertes Vokabular erarbeitet, das auf der RILM Systematik und der von RILM mitgelieferten freien Verschlagwortung der Einträge basiert. Hiermit lassen sich u. a. Herangehensweisen (Methoden) identifizieren.

Die bereits erwähnte Technik des *Web Scraping* beschreibt das Extrahieren von Informationen bzw. Daten aus dem Internet, im Speziellen von Webseiten, deren Daten nicht in standardisierter Form zugänglich gemacht werden, die also beispielsweise nicht als Textdatei oder in einem Datenformat für strukturierte Daten zum Download angeboten werden. Ein gutes Beispiel dafür sind Wikipedia-Listen, die nicht in einem Tabellenformat heruntergeladen werden können.

Das Web Scraping ist eine Methode der Datenerhebung, mit der Informationen aus dem Internet für eine Weiterverarbeitung zugänglich gemacht werden können. Die wörtliche Übersetzung aus dem Englischen legt nahe, dass im Internet nach Daten „geschürft“ bzw. Daten „abgeschabt“ würden.

Die Datenextraktion von Internetseiten kann manuell durch händisches Markieren und Kopieren von Inhalten (Copy & Paste) durch einen Benutzer ausgeführt werden. Zumeist werden aber Vorgehensweisen der (semi-)automatisierten Datenextraktion unter dem Begriff Web Scraping verstanden.

Beim Web Scraping sind zwei Vorgehensweisen zu unterscheiden: Je nach Ausgangslage werden entweder strukturierte, maschinenlesbare Daten gesammelt (Tabellen, Datenbankstrukturen) oder unstrukturierte, nicht maschinenlesbare Daten (beispielsweise beim Erstellen eines Textkorpus).

Der Einsatz von Web Scraping schließt zwar menschliche Flüchtigkeitsfehler bei der Übertragung von Daten aus und bietet gegenüber dem manuellen Extrahieren von Daten einen enormen Geschwindigkeitsvorteil – die Automatisierung birgt aber auch Gefahren, v. a. wenn die Integrität und Sinnhaftigkeit der Daten unbedingt gewährleistet werden muss.

Im Rahmen der hier beschriebenen Datenerhebung ist zu gleichen Teilen Qualität und Quantität der gesammelten Daten entscheidend – zu wenige sehr gute Daten versprechen ebenso suboptimale Ergebnisse wie extrem viele Daten mangelnder Qualität.

Beim Web Scraping erfolgt keine Fehlerkorrektur – diese muss manuell im Nachhinein durch den Benutzer vorgenommen werden. In dieser Arbeit wird daher darauf geachtet, dass die zu erstellenden umfangreichen Listen in mehreren Schritten einer Qualitätssicherung unterzogen werden.

Neben den allgemeinen Gefahren, die bei automatisierter Übertragung von Informationen bestehen ist zu vermerken, dass die Wahl der Quellen von grundlegender Bedeutung für die Datenqualität der angestrebten Listen ist.

Zunächst muss sichergestellt werden, dass die infrage kommenden Quellen mit ausreichenden Qualitätssicherungsstandards arbeiten. Ohne Kontrollmechanismen wäre Missbrauch in Form von Falsch- und Fantasie-Eingaben Tor und Tür geöffnet und die Datenqualität gefährdet.

Ein zweiter Punkt ist, dass – gerade in Bezug auf nicht-standardisiertes Vokabular wie Genre-Bezeichnungen und Musikernamen (die teilweise frei erdacht sind oder sich an bestehendes Vokabular anlehnen) – vorkommende Mehrdeutigkeit der Begriffe problematisch sein kann. Der Umgang hiermit...

Webscraping kann, je nach Umfang des Datenmaterials und Komplexität der Website, die gescraped werden soll, halb- oder vollautomatisch geschehen. Halbautomatisches Webscraping kann dabei zum Beispiel durch das Bearbeiten des kopierten Quelltextes einer Seite mit regulären Ausdrücken geschehen. Automatisches Webscraping wird durch den Einsatz sog. Scraping Wrapper (kleiner speziell angepasster Programme) umgesetzt. Neben dem Einsatz regulärer Ausdrücke wurden für den Zweck dieser Arbeit Extraktions-Wrapper verwendet, die in Python geschrieben wurden und das Framework *Beautiful Soup* nutzen.

Ein *Regulärer Ausdruck* (engl. *Regular Expression*) ist eine Syntax um Texte mittels komplexer Suchmuster zu bearbeiten. Ein regulärer Ausdruck dient der abstrahierten Beschreibung von Zeichenketten: So kann nach Zeichenketten gesucht werden, die bestimmte Bedingen erfüllen bzw. bestimmte Eigenschaften haben, etwa alle Wörter eines Textes, die exakt fünf Buchstaben haben und nach denen ein Komma steht. Reguläre Ausdrücke werden zumeist in der Programmierung

eingesetzt, sie sind daher aus den meisten Programmiersprachen heraus nutzbar. Es existieren aber auch viele Texteditoren, mit denen man reguläre Ausdrücke für das Suchen (und Bearbeiten) von Inhalten mit einer Benutzeroberfläche nutzen kann.

„Regular expressions are the key to powerful, flexible, and efficient text processing. Regular expressions themselves, with a general pattern notation almost like a mini programming language, allow you to describe and parse text. With additional support provided by the particular tool being used, regular expressions can add, remove, isolate, and generally fold, spindle, and mutilate all kinds of text and data.“¹⁴⁵

Reguläre Ausdrücke sind Nützlich, um Texte nach gewissen Kriterien zu validieren, aber auch, um Inhalte zu bearbeiten bzw. nach speziellen Mustern zu suchen, um Inhalte zu extrahieren.

Der Aufbau eines Suchmusters als Regulärer Ausdruck basiert auf der Kombination von Metazeichen, Zeichenklassen und Unterausdrücke. Zeichen, die eine spezielle Funktion erfüllen und nicht zu dem zu suchenden Text zählen nennt man **Metazeichen**. Ein Beispiel für ein Metazeichen ist der Punkt oder die runde Klammer. Beide erfüllen wichtige Funktionen und müssen, wenn nach einem Punkt oder einer runden Klammer gesucht wird, maskiert werden – dem Interpreter muss mitgeteilt werden, dass in diesem speziellen Fall das Metazeichen nicht als solches auftreten soll, sondern in seiner normalen Funktion. **Zeichenklassen** beschreiben definierte Mengen von Schriftzeichen. Es können etwa alle Ziffern eines Textes mit der Klasse `\d` gesucht werden oder alle Großbuchstaben mit der Menge `[A-Z]`.

Um die Literatur statistisch auswerten zu können, wird das Datenmaterial nach bestimmten Kriterien automatisch verschlagwortet und in ein für die Auswertung sinnvolles Ausgangsformat gebracht.

Die in RILM erfassten Daten sind bereits verschlagwortet. Wie bereits in der Beschreibung der Datenquelle RILM erörtert,¹⁴⁶ erfolgt die Verschlagwortung der in RILM enthaltenden Daten zum einen durch die RILM Redaktionen und zum andern durch die herausgebenden Verlage bzw. Autoren. Dieser Vorgang ist als manuelle Indexierung bzw. Sacherschließung zu beschreiben.

Die von RILM verwendete Systematik ist polythematisch und der dahinter liegende Schlagwortkatalog ist nicht kontrolliert, d. h. das es für die bearbeitenden Personen bei RILM und den Datengebern jederzeit möglich ist, neue Schlagwörter zu vergeben. Des Weiteren ist der Schlagwortkatalog nicht in einer Form zugänglich, die es ermöglicht, gleichförmige Merkmalsausprägungen zu untersuchen.

¹⁴⁵ (Friedl2006)

¹⁴⁶ vgl. [Sacherschließung: RILM Keywords]]

Bei einem zufällig erstellten Testdatensatz aus 57346 Literatureinträgen sind insgesamt 103381 Schlagwörter vergeben worden (im Durchschnitt 1,8 Keywords pro Eintrag). 89471 Keywords, das entspricht 86,54%, wurden nur ein einziges Mal vergeben. Das verdeutlicht die Inkonsistenz der von RILM Verschlagwortung.

Um die von RILM vergebenen Keywords dennoch im Rahmen der statistischen Auswertung nutzen zu können (da davon ausgegangen werden kann, dass in den Keywords Informationen enthalten sind, mithilfe derer sich Aussagen zu der methodischen Ausrichtung treffen lassen und die ggf. weitere inhaltliche Informationen enthalten) war ein

- RIS Datei einlesen, alle Zeilen löschen, die nicht mit dem Key KW anfangen: `^(?!KW) .*` durch leer ersetzen
- leere Zeilen löschen
- Zeilenanfang (RIS Key) löschen
- Keywords in einer Zeile aufteilen: Die Keywords sind durch zwei Bindestriche getrennt.

RILM arbeitet mit einer Mischung aus kontrollierter Indexierung¹⁴⁷ und freier Indexierung¹⁴⁸. Dieses Vorgehen ist zum Auffinden von relevanter Literatur kein Nachteil, eine sinnvolle statistische Auswertung der Keywords nach Häufigkeiten wird durch die fehlende Festlegung auf ein kontrolliertes Vokabular aber erschwert.

Um die RILM Verschlagwortung für die Auswertung der Literatur nutzbar zu machen, wurde eine Liste aus allen aus Testdatensätzen extrahierten Keywords kumuliert und ausgewertet. Die Auswertung erfolgte nach dem...

Das Ziel, eine Liste mit einer erschöpfend großen Anzahl von Namen von MusikerInnen und Bands anzulegen, steht und fällt mit der Auswahl der Quellen für eine solche Liste.

Voraussetzungen für die Quellen:

- Umfassende Verzeichnung musikalischer KünstlerInnen und -Gruppen (Universalität)
- Keine Begrenzung auf Genre (Universalität)
- Keine zeitliche Begrenzung (alle Künstler des 20. und 21. Jahrhunderts) (Universalität)
- Verzeichnung aller Veröffentlichungsformen (im Idealfall auch mit Namen von Künstlern, die unter anderem Namen Veröffentlicht haben bzw. die keine Veröffentlichung herausgegeben haben)

¹⁴⁷ u. a. die RILM Systematik, Zeiträume,

¹⁴⁸ Die Auswertung eines Testdatensatzes ergab, dass mehr als 89000 verschiedene Keywords vergeben wurden.

- Keine Beschränkung, im Idealfall keine Fokussierung auf bestimmte Orte bzw. Regionen (internationaler Aspekt)
- Die Daten sollten zumindest *einem* internen Kontrollmechanismus unterliegen (Qualitätssicherung)
- Sie sollten kostenfrei zur Verfügung stehen und ihre Verwendung gemeinfrei sein (Offenheit & Access)
- Die Daten sollten in einem Format vorliegen, das den Scraping-Prozess begünstigt (Offenheit & Access)

Eine Vergleichstabelle¹⁴⁹ von musikbezogenen Datenbanken zeigt, dass die Plattform Discogs¹⁵⁰ mit Informationen zu mehr als 151 Millionen Titeln mit Abstand am meisten Metadaten gesammelt hat. An zweiter Stelle stehen die kommerziellen Angebote von Gracenote¹⁵¹, danach die Angebote von The Echo Nest¹⁵² (einer Tochter des Musik-Streaming-Dienstes Spotify) und AllMusic¹⁵³. Mit nur etwa 20 Millionen Datensätzen ist der Datenbestand von MusicBrainz¹⁵⁴ zwar vergleichsweise klein, es ist jedoch die zweite Quelle, die Ihre Daten kostenfrei zur Verfügung stellt. Bei MusicBrainz sind deutlich strengere Qualitätskontrollen als bei Discogs umgesetzt.

Die Wikipedia ist als Datenquelle für Musiker nicht zu gebrauchen, da es weder einen zentralisierten Angriffspunkt für die Auflistung aller Musiker der populären Musik gibt noch von einer umfassenden Datenbasis ausgegangen werden kann, da Wikipedia als Enzyklopädie ein inhaltlich anderes Ziel verfolgt als es die Nachweisdatenbanken tun.

Es bleiben als universell verzeichnende internationale Datenquellen letztlich Discogs als auch Musicbrainz übrig. Sowohl Discogs als auch Musicbrainz sind Datenbanken, die von der Mitarbeit von Benutzern leben. So haben im Jahr 2016 mehr als 20000 User aktiv Inhalte der Musicbrainz Datenbank erweitert und gepflegt,¹⁵⁵ bei Discogs spricht man (ohne Nennung eines Zeitraums) von mehr als 390000 Benutzern.¹⁵⁶ Beide Anbieter geben an, dass sie es anstreben, alle musikalischen Veröffentlichungen und Kompositionen zu verzeichnen. Discogs erklärt dazu:

¹⁴⁹ https://en.wikipedia.org/wiki/List_of_online_music_databases

¹⁵⁰ <https://discogs.com>

¹⁵¹ <http://www.gracenote.com/music/global-music-data/>

¹⁵² <http://the.echonest.com/>

¹⁵³ <http://allmusic.com>

¹⁵⁴ <https://musicbrainz.org>

¹⁵⁵ <https://blog.musicbrainz.org/2017/07/05/musicbrainz-user-survey/>—>, insgesamt ist von mehr als 1,9 Millionen Bearbeitern die Rede (vgl. <!--<https://musicbrainz.org/statistics>

¹⁵⁶ vgl. <!--\url{<https://www.discogs.com/about>}-->

„We’re on a mission to build the biggest and most comprehensive music database and marketplace. Imagine a site with discographies of all labels, all artists, all cross-referenced [...]. It’s for the love of music, and we’re getting closer every day.“¹⁵⁷

Gleichermaßen umfassend ist auch die Zielsetzung von Musicbrainz:

„As an encyclopedia and as a community, MusicBrainz exists solely to collect as much information about music as we can. We do not discriminate or prefer one “type” of music over another though, in fact we collect information about as many different types of music as possible. Whether it is published/unpublished, popular/fringe, western/non-western, human/non-human – we want it all to be entered into MusicBrainz.“

Die Daten wurden darüber hinaus ergänzt durch die Daten des „Million Song Dataset“¹⁵⁸-Projekts der Columbia University und der Firma The Echo Nest.

Im Anschluss an die Datensammlung wurde das Datenmaterial in mehrere Schritten verbessert. Dazu zählt in erster Linie, Dopplungen zu vermeiden und etwaige Fehlformatierungen zu löschen. Darüber hinaus ist auch auf inhaltliche Aspekte eingegangen worden, deren Durchführung die zu erwartende Qualität des Abgleichs verbessert.

Der Ausschluss vollkommen ambivalenter bzw. missverständlicher Musikernamen im Rahmen einer statistischen Auswertung:

- Musikernamen, die nur aus einem Wort bestehen und die in Testläufen mit unterschiedlichem Datenmaterial gematched wurden: Es wurden die Wörter aussortieren, die offensichtlich zu Verwechslungen führen würden, da sie etwa in der Sprach-Syntax einen eigenen Zweck erfüllen (z. B. Adjektive wie „absolut“) oder ein „alex“, „audio“ etc. - Die ambivalenten Namen wurden gesammelt, damit wiederholbar ein Abgleich der Musikerliste mit diesem Dokument als Filterliste durchgeführt werden konnte und die entsprechenden Musikernamen gelöscht werden konnten.

Um Doppelnennung zu umgehen und Bandnamen in die für den Abgleich richtige Reihenfolge zu bringen wurde der bei Bandnamen beliebten englischen Artikel „The“ in jedem Fall vor den Namen gesetzt: Aus Beatles, The wird dann The Beatles. Dieses Vorgehen ist notwendig, da in einem Titel in den seltensten Fällen die für alphabetische Sortierung häufig genutzte umgekehrte Wortreihenfolge bei Zusammengesetzten Namen mit Artikeln etc. eingesetzt wird. Der Vorgang wurde mittels eines regulären Ausdrucks durchgeführt.¹⁵⁹

siehe Musikerliste Discogs bereinigen

¹⁵⁷ <https://www.discogs.com/about>

¹⁵⁸ <https://labrosa.ee.columbia.edu/millionsong/>

¹⁵⁹ vgl. <https://regex101.com/r/4FKN40/1> und <https://regex101.com/r/mRz6HR/1>

„Kategorisierung zum Zweck der Reduktion von Komplexität ist eine Konstante menschlicher Informationsverarbeitung und für endliche Wesen angesichts theoretisch unendlicher Information eine Überlebensnotwendigkeit.“¹⁶⁰

Bei der Auseinandersetzung mit Texten jeglicher Art ist es üblich, das zu behandelnde in einen kulturellen und historischen Kontext einzuordnen und dadurch ggf. von anderem abzugrenzen. Dabei wird der aus der Philosophie stammende Begriff der Gattung benutzt. Neben dem Gebrauch des Gattungsbegriffs in der Philosophie und den Naturwissenschaften – v. a. in der Biologie – ist er auch in den Kultur- und Kunstwissenschaften zu finden.¹⁶¹

Um künstlerisches Schaffen zu systematisieren, wird ein abstrakter Oberbegriff für Texte mit einander ähnlichen Merkmalen gewählt. In den Künsten werden gemeinhin bildende und darstellende Kunst, Literatur und Musik als Gattungen verstanden, denen jeweils verschiedene Stile untergeordnet sind. Für die Beschäftigung mit populärer Musik spielt bei der Unterscheidung von Stilen vor allem der aus dem Französischen stammende Begriff *Genre* eine Rolle. Im Gegensatz zu europäischer Kunstmusik, die grundsätzlich verschriftlicht vorliegt, deren Stile weitestgehend feststehend sind und über deren Abgrenzung zueinander ein allgemein akzeptierter Konsens herrscht, ist die Klassifizierung von populärer Musik in spezielle Genre problematisch. Es gibt keine allgemeingültigen Begriffe und oft auch keine Herausstellungsmerkmale für einzelne Stile. Darüber hinaus sind die Entwicklungen in der populären Musik des 20. und 21. Jahrhunderts nicht abgeschlossen, Genrebegriffe können also im Laufe der Zeit auch zu Sammelbecken für eine Vielzahl von ähnlichen Stilen werden. In diesem Fall wird oft mit Sub-Genre gearbeitet, um Stile voneinander abzugrenzen. Die Gefahr von Sub-Genres allerdings ist, dass man eine so große Anzahl von Genre-Begriffen vorfindet, dass diese für sich genommen keine Aussagekraft mehr haben.

Die Benennung eines Genres geschieht entweder durch frühe oder besonders prominente musikalische Vertreter eines Stils oder aber durch Außenstehende, wie Musikjournalisten, Fans oder (wissenschaftliche) Autoren.¹⁶² Der Literaturmarkt, auch der wissenschaftliche, ist also an der Abgrenzung und Definition der Genre der populären Musik beteiligt.

Es sei aus diesem Grund darauf hingewiesen, dass die gesammelten Genrebezeichnungen nicht eindeutig sind. Die Nennung eines Genres bedingt nicht zwangsläufig eine klar definierte Abgrenzung. So gibt es beispielsweise Bezeichnungen für Stile innerhalb dessen, was gemeinhin unter dem Begriff Rockmusik

¹⁶⁰ (Bohm2001)

¹⁶¹ (Danuser1995)

¹⁶² (Halbscheffel1990)

subsumiert wird, die nicht konsistent sind. Es ist davon auszugehen, dass einzelne Hörer Unterschiedliches unter den Begriffen Rock, Rock 'n' Roll oder Hard Rock verstehen würden. Die Einteilung Musikgenre ist also nicht konsistent durchführbar, da es

- keine allgemein akzeptierten Stilmerkmale gibt, die gleichbleibend sind und anhand derer sich eine Kategorisierung durchführen ließe, und - Künstler oftmals genau das Gegenteil einer Kategorisierung ihres Schaffens anstreben: Sich nicht „in eine Schublade stecken zu lassen“ und Fusionen von Stilen einzuleiten kann gerade im Hinblick auf Subkulturen und den Aspekt der Identitätsfindung durch die Zugehörigkeit zu einem Musikstil zum Ziel von Künstlern werden.

Sicher gibt es einige Werke, die mit großer Wahrscheinlichkeit auch aus vielen unterschiedlichen Betrachtungswinkeln in diese oder jene Kategorie einsortiert würden, von diesem Fall ist jedoch nicht als Grundannahme auszugehen.

Ein Beispiel für die Missverständlichkeit von Genre als Kategorie ist der vor allem in französischen Publikationen vorkommenden Begriff des Chansons,¹⁶³ der sich auf mindestens fünf voneinander gänzlich unabhängige musikalische Phänomene, vom Mittelalter bis in die jüngste Gegenwart, beziehen kann und dessen Bedeutung vollkommen vom jeweiligen Kontext abhängig ist.

Für die Erstellung der Genre-Liste gelten die gleichen Kriterien wie bei der Erstellung der oben besprochenen Musikerliste. Auch hier müssen die herangezogenen Quellen gewissen Standards genügen.

Im Falle von Genres der populären Musik ergeben sich allerdings einige Besonderheiten: Genrebezeichnungen befinden sich in stetigem Wandel (es gibt weder eine finite Anzahl von Genres noch Übereinkunft über die Zugehörigkeit von Musikern zu einem Genre) und sind nicht standardisiert (mehrere Bedeutungen).

Aufbau von Genre-Bezeichnungen bei Every Noise at Once:

Musikimmanente und (musik-)historische Merkmale: Ensemble/Instrumenten-Gruppen: Inhaltlich/philosophische Unterscheidung: „Christian Hip Hop“ Kulturell/geographisch (bis hin zu „hyper-regional“)

¹⁶³ „In a general sense the word ‘chanson’ refers to a wide variety of compositions: the monophonic songs of the Middle Ages (see Troubadours, trouvères); court songs of the late 16th and 17th centuries (see Air de cour); popular songs of the streets, cafés and music halls in the 17th, 18th and 19th centuries (see Chanson pour boire; Vaudeville; Pastourelle; Bergerette (ii); Brunette); art songs of the 19th and 20th centuries (Mélodie); as well as to folksongs (‘chanson populaire’ or ‘chant folklorique’).“ (Brown1995)

Als Quelle für die Liste mit Musikgenres wurde zum einen die Internetseite Every Noise at Once¹⁶⁴ herangezogen, zum anderen wurden die Listen mit Musikgenre der Wikipedia gescraped¹⁶⁵. Die Daten beider Quellen wurden zusammengefügt/kombiniert und um multiple Schreibweisen etc. ergänzt (Erklärungen dazu folgen unten).

Die Seite „Every Noise at Once“ wird von Glenn McDonald betrieben, einem Entwickler beim Musikstreaminganbieter und -Vorreiter Spotify. In einer Art Streudiagramm sind mehr als 1500 Musikgenre aufgeführt, zu denen man bei einem Klick auf den Namen automatisch ausgewählte Musikbeispiele hören kann und sich die mit jedem Genre verknüpften Künstler anzeigen lassen kann (und hier wiederum Beispiele hören kann). Die Seite ist eng an Spotify angebunden, es ist jederzeit möglich genrespezifische Playlists zu hören, vorausgesetzt, man hat ein Nutzerkonto bei Spotify.

Die Genres sind in einem Koordinatensystem nach nur vage definierten psychoakustischen Merkmalen angeordnet. Auf der Y-Achse reicht das Spektrum, mit dem die Genre beschrieben werden McDonald zufolge von organisch¹⁶⁶ bis mechanisch/elektrisch.¹⁶⁷ Auf der X-Achse hingegen wird zwischen „dichter/atmosphärischer“¹⁶⁸ Musik und solcher, die „stachelig/lebhaft“¹⁶⁹ ist, unterschieden.

McDonald versteht seine Seite weniger als eine wissenschaftlich fundierte Auflistung unterschiedlicher Musikgenre als einen Spielplatz zum Entdecken von Musik.¹⁷⁰ Der Spass bzw. das Erleben unbekannter Musik steht im Vordergrund und wird ermöglicht, in dem zu jedem Genre sofort passende Musikbeispiele zu hören sind.

Die Genrenamen basieren McDonald zufolge auf Informationen, die Spotify auf verschiedenen Kanälen sammelt. Die Darstellung und Benennung der Genres sowie die hinterlegten Musikbeispiele speisen sich zum einen aus Informationen, die computergestützt über akustische Merkmale der Musik gewonnen werden (Stichwort Music Information Retrieval, hier wird aber nicht näher auf die unter-

¹⁶⁴ <http://everynoise.com/engenremap.html>

¹⁶⁵ <https://gist.github.com/kbecker87/50babao85ad3d4fe3abc163b6fao8bca>

¹⁶⁶ organic

¹⁶⁷ mechanical and electric

¹⁶⁸ denser and more atmospheric

¹⁶⁹ spikier and bouncier

¹⁷⁰ (McDonald2013)

suchten Merkmale eingegangen), zum anderen aus Texten, die online veröffentlicht werden („our system reads everything written about music on the web“) und schließlich aus Gewohnheiten bzw. Sortierungen, die von den mehr als 100 Millionen Spotify-Hörern¹⁷¹ gesammelt werden.

„This is an ongoing attempt at an algorithmically-generated, readability-adjusted scatter-plot of the musical genre-space, based on data tracked and analyzed for 1519 genres by Spotify.“¹⁷²

Infos: Everynoiseatonce - How we understand Music Genres

„[...] our system reads everything written about music on the web, and listens to millions of new songs all the time, to identify their acoustic attributes. This enables our genres to react to changes in music as they happen. To create dynamic genres, we identify salient terms used to describe music (e.g., “math rock,” “IDM”, etc.), just as they start to appear. We then model genres as dynamic music clusters – groupings of artists and songs that share common descriptors, and similar acoustic and cultural attributes. (And in the somewhat rare cases when a music cluster appears without a name, we come up with one.) When a new genre forms, we know about it, so you can discover it right away, too. This approach to genres is trend-aware.“¹⁷³

Beinhaltet manuelle Nachbearbeitung:

„a carefully considered (and mercifully manageable) amount of editorial guidance. For example, we decide what to do with naming variants like “nu soul” and “neo soul” (we went with “neo”), and whether we have enough data for the computers to produce a substantial and satisfyingly distinct body of music for any given thing, such as “indie folk” (yes), “sertanejo” (yes), or “ziglibithy” (no, not yet).“¹⁷⁴

Scraping von Evernoiseatonce mittels folgendem Scraping-Wrapper, der mittels BeautifulSoup die HTML Quelle der Seite ausliest und den Inhalt den Textinhalt der DIV-Container mit der Klasse `genre_scanme` ausgibt:

```

1 # coding: utf-8
2 from urllib.request import urlopen
3 from bs4 import BeautifulSoup
4
5 # HTML der URL laden und mit BeautifulSoup parsen
6
7 html = urlopen("http://everynoise.com/engenremap.html")
8 soup = BeautifulSoup(html, "lxml")
9
10 # Alle DIV-Container lokalisieren, deren Klasse "genre_scanme" ist.
11
```

¹⁷¹ <https://www.reuters.com/article/us-spotify-users/spotify-monthly-active-user-base-reaches-100-million-idUSKCNoZ61FM>

¹⁷² <http://everynoise.com/engenremap.html>, abgerufen am 13.3.17

¹⁷³ <https://insights.spotify.com/us/2015/09/30/50-strangest-genre-names/>

¹⁷⁴ (McDonald2013)

```

12 divs = soup.findAll('div', {'class': 'genre scanme'})
13
14 # Über die DIV Container iterieren und nur den ersten Listeneintrag
    des Inhalts extrahieren
15
16 for div in divs:
17     l = div.contents
18     genres = l.pop(0)
19     print(genres)

```

Probleme von Every Noise at Once: -

Scraper: /Users/Karl/Dropbox/Litalyse/Genreliste/Genre Scraper Wikipedia/Genre_Scraper.py Die englische Wikipedia enthält eine Große Zahl von Artikeln zu den Genres populärer Musik.

„List of popular music genres“ sowie „List of styles of music“ (eingeteilt in vier Listen, A-F, G-M, N-R und S-Z)¹⁷⁵

<https://de.wikipedia.org/wiki/Kategorie:Musikgenre>

Listen zusammengefügt Dopplungen entfernt Genre mit mehreren Schreibweisen hinzugefügt:

z. B. drum and bass dupliziert zu drum & bass

```

1 ersetzt mit Gruppe1 Gruppe2 (and) Gruppe3; neue Zeile; Gruppe1 &
    Gruppe3
2 $1$2$3\n$1&$3

```

Verschiedene Schreibweisen von Genres berücksichtigt, manuell nachgetragen:

```

1 rock'n'roll - rock and roll - rock n roll - 'rockn roll ...
2 hip-hop - hip hop - hiphop

```

Verschiedener Genre-Schreibweisen

In den von RILM erfassten Keywords wird, wenn es zutrifft, der Name des Landes in englisch angegeben, auf das sich der Literatureintrag bezieht.

¹⁷⁵ https://en.wikipedia.org/wiki/List_of_popular_music_genres
https://en.wikipedia.org/wiki/List_of_styles_of_music:_A-F
https://en.wikipedia.org/wiki/List_of_styles_of_music:_G-M
https://en.wikipedia.org/wiki/List_of_styles_of_music:_N-R
https://en.wikipedia.org/wiki/List_of_styles_of_music:_S-Z

Ein als Beispiel herangezogene Aufsatz¹⁷⁶ etwa beschäftigt sich mit der WDR-Sendung „Rockpalast“ – für kundige Leser bzw. deutsche Muttersprachler ist das ein eindeutiges Indiz dafür, dass ein Phänomen aus Deutschland thematisiert wird. Durch die Vergabe des Keywords »Germany« ist diese Tatsache auch für die automatisierte Datenextraktion fassbar:

```
1 KW -- rock -- Germany -- Rockpalast show -- 1974 to present
2 KW -- Germany -- popular music -- Rockpalast show -- 1974 to present
3 KW -- politics -- Germany -- rock -- 1950s-70s
4 KW -- mass media -- television -- Germany -- Westdeutscher Rundfunk
   (WDR) -- Rockpalast show -- 1974 to present
```

Für den Abgleich musste demnach eine Liste mit allen Ländernamen in englischer Sprache angelegt werden. Die Daten basieren auf der ISO-3166-1-Kodierliste der Internationalen Organisation für Normung.¹⁷⁷

3.2.4 Funktionsweise & Aufbau des Programms

Die Gesamte Funktionalität des Programms ist in vier (bzw. fünf) aufeinanderfolgende Schritte aufgeteilt. Die Ausgangsdaten im RIS Format durchlaufen, bis sie angereichert und transformiert in der Forschungsdatenbank landen, mehrere Bearbeitungs- und Transformationsschritte. Eine RIS Datei beliebiger Größe wird zunächst von dem Programmteil `ris_chunker.py` in kleinere, gut zu verarbeitende Segmente geteilt. Es hat sich als sinnvoll erwiesen, die RIS Datei in kleinere Dateien zu splitten, in denen jeweils nur 250 Literaturnachweise enthalten sind. Dadurch kann eine schnelle sequentielle Verarbeitung gewährleistet werden. Mittels eines Shell-Scripts zur Automatisierung oder durch manuellen Programmaufruf werden die verkleinerten RIS Dateien zunächst der Hauptauswertung der Titel übergeben, bspw. durch `python litalyse.py /chunks/*.ris`.¹⁷⁸ Dieses Vorgehen kann zur Auswertung von Abstracts und übersetzten Titeln wiederholt werden. Die gleichen RIS Dateien werden dafür der `litalyse_abstracts.py` und dann der `litalyse_translated.py` übergeben. So werden in drei Durchgängen aus n RIS Dateien $3*n$ RIS Dateien erzeugt, wobei nur der Durchlauf der `litalyse.py` obligatorisch ist, da hier alle relevanten Spalten erzeugt werden und neben der Auswertung der Titel auch die Berechnung der Seitenzahlen stattfindet. Der modulare Aufbau der Auswertung hat den Vorteil, dass nicht zwangsläufig alle drei Auswer-

¹⁷⁶ Der Eintrag ist mit einem gültigen Abonnement für RILM bei EBSCOHost unter folgender URL einzusehen: <http://search.ebscohost.com/login.aspx?direct=true&db=rih&AN=2005-26443&site=ehost-live&scope=site>

Titel: „Feels so good, hurts so bad’: Der Rockpalast live auf Sendung“

¹⁷⁷ Da die ISO die Listen nicht frei öffentlich zur Verfügung stellt, basiert die Länderliste auf einem OpenSource Repository: <https://github.com/umpirsky/country-list>

¹⁷⁸ Die genauen Kommandozeilenparameter hängen von der Konfiguration des eigenen Systems ab, es ist auch üblich, Python 3 über `python3 litalyse /chunks/*.ris` aufzurufen.

tungensschritte (Titel, Abstracts und übersetzte Titel) ausgeführt werden müssen. Im nächsten Schritt werden die bis zu drei Ordner, in den die fertig bearbeiteten CSV Dateien liegen, mit `concat.py` zusammengeführt. Hierbei werden erst alle CSV Dateien eines Ordners zusammengefügt (englisch: „concatenate“) und dann in einem zweiten Schritt werden diese CSV Daten mittels eines Vorgangs, der einem SQL outer join entspricht, auch horizontal zusammengefügt. Nach der erfolgreichen Zusammenfügung durch `concat.py` können die Daten, die nun in einer großen CSV Datei vorliegen, bereinigt bzw. anderweitig vorprozessiert werden, bevor durch die Übergabe der CSV Datei an `csv2sqlite.py`¹⁷⁹ die Daten in eine Sqlite-Datenbank transformiert werden. In Abbildung X ist der gesamte Ablauf der Litalyse schematisch dargestellt:

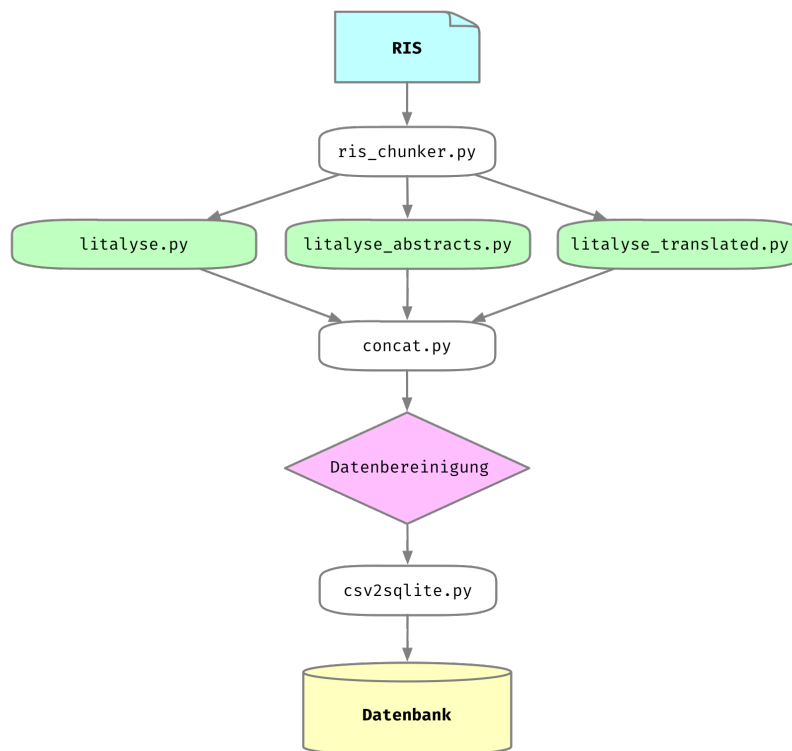


Abbildung 4 – Aufbau

Um das Funktionieren des Programms zu gewährleisten werden einige Python Pakete geladen, darunter `nltk` (Natural Language Processing ToolKit), `pandas` (für die Anlage der Daten im DataFrame¹⁸⁰ Format), `csv` (für die Ausgabe der Daten im CSV Format) und `re` (für die Anwendung Regulärer Ausdrücke)

¹⁷⁹ Das Script `csv2sqlite.py` ist von Rufus Pollock übernommen worden <https://github.com/rufuspollock/csv2sqlite>

¹⁸⁰ „DataFrame is a 2-dimensional labeled data structure with columns of potentially different types. You can think of it like a spreadsheet or SQL table“ vgl. <https://pandas.pydata.org/pandas-docs/stable/dsintro.html#dataframe>

Die Benennung der RIS-manenten Tags wird geändert, um die Eindeutigkeit der Bezeichnungen für den Betrachter zu erhöhen. So wird etwa aus dem Tag AU „Autor“ etc.

Ebenfalls bei Programmaufruf werden die im Vorfeld erstellten Listen, die sich im gleichen Ordner wie das Programm befinden müssen, eingelesen (geparsed). Dieser Vorgang ist in vier Funktionen eingeteilt, pro Liste ist eine Funktion für das Einlesen der Abgleichlisten vorgesehen. Die Musikerliste, Keyword-Filterliste, Länder- und Genreliste werden hier von dem Programm erfasst und die einzelnen Zeilen der Listen für den späteren Abgleich in den Zwischenspeicher gelesen.

Die Hauptfunktion `parse_ris_items` bekommt bei dem Aufruf des Programms über die Kommandozeile eine oder mehrere Dateien zugewiesen. Zunächst werden die zu bearbeitenden dem Programm übergebene RIS Datei(en) eingelesen. Der Name jeder zu bearbeitenden RIS Datei wird mit dem aktuellen Zeitstempel zusammen zum Namen der nach vollständiger Bearbeitung erstellten CSV Datei. Die Eingabedatei wird der Hauptfunktion `parse_ris_items` übergeben, die das „heavy-lifting“ durchführt (dazu unten mehr). Die fertig bearbeiteten Daten werden in einen Dataframe überführt, den man sich als virtuelle, d. h. für den Nutzer des Programms unsichtbare, Tabelle vorstellen kann und der als Datentyp bestens für eine solche Aufgabe geeignet ist. Aus den Inhalten dieser Datenstruktur wird schließlich die zu erstellende CSV Datei gespeist.

In der Hauptfunktion werden zunächst die einzelnen Literatureinträge, die in der RIS Datei nur durch eine Leerzeile voneinander getrennt sind, identifiziert und die Inhalte ersten Normalisierung- und Bereinigungsschritten zugeführt. Das Programm geht die eingelesene RIS Datei Zeile für Zeile durch, bis es auf eine leere Zeile trifft. In diesem Fall wird ein vollständiger Literatureintrag im Zwischenspeicher angelegt. Bei nicht-leeren Zeilen werden Tag und Text voneinander getrennt und je nach Tag werden die folgenden Maßnahmen angewendet:

- Bei dem Tag AU werden multiple Autoren mit einem Trennzeichen versehen
- Das Publikationsjahr wird durchgehend auf vierstellige Zahlen reduziert, damit unterschiedlich angesetzte Angaben zum Monat (Januar 2018 oder 01.2018 oder 2018-01) vereinheitlicht werden
- Das Notizenfeld N1 wird aufgrund des Fehlens eines ausschließlich für die Sprache des Eintrags vorgesehenen Feldes nach dem String Language: XXX durchsucht und nur Name der Sprache wird beibehalten
- Die Typenbezeichnungen im TY Feld werden lesbarer gemacht, aus BOOK wird Buch usw.
- Die Ansetzung der Keywords in multiplen Zeilen wird aufgehoben (alle Keywords in einer Zeile), die Trennung der Keywords wird ebenfalls aufgehoben

- Sowohl die Keywords als auch die Titel werden von Satzzeichen und Sonderzeichen befreit, damit der folgende Abgleich nicht durch Kommata, Punkte, Apostrophe o. ä. beeinflusst wird

Die so vorbearbeiteten Literatureinträge werden innerhalb der Funktion einer weiteren Schleife übergeben, in der der Abgleich der Titel und Keywords mit den zuvor geparsten Listen vorgenommen wird.

Zunächst wird dazu die Wortanzahl des Abgleich-Items ermittelt. Im Falle des Abgleichs des Titels der Publikation mit Musikernamen wird jeder Musiker nacheinander auf die Wortanzahl analysiert. Wenn der Abgleich also beispielsweise bei der britischen New-Wave Band „The Teardrop Explodes“ angekommen ist, wird bestimmt, dass der Bandname aus drei Wörtern besteht:

`the teardrop explodes = 3`

Dann wird der Titel der Publikation in N-Gramme zerlegt. N-Gramme sind das Ergebnis der Zerlegung eines Textes in einzelne Fragmente. Der Faktor N wird bestimmt durch die im vorigen Schritt ermittelte Länge des Abgleichitems, im Falle des Beispiels wird der Titel demnach in 3-Gramme, sog. Trigramme zerlegt, da $n=3$. Aus dem Titel der für dieses Beispiel gewählten Publikation `Head-on: Memories of the Liverpool punk-scene and the story of the Teardrop Explodes (1976-82)` werden entsprechend die folgenden Trigramme:

```
1 ('Head-on:', 'Memories', 'of')
2 ('Memories', 'of', 'the')
3 ('of', 'the', 'Liverpool')
4 ('the', 'Liverpool', 'punk-scene')
5 ('Liverpool', 'punk-scene', 'and')
6 ('punk-scene', 'and', 'the')
7 ('and', 'the', 'story')
8 ('the', 'story', 'of')
9 ('story', 'of', 'the')
10 ('of', 'the', 'Teardrop')
11 ('the', 'Teardrop', 'Explodes')
12 ('Teardrop', 'Explodes', '(1976-82)')
```

Trigramme

Das Programm gleicht dann die sich so ergebenden N-Gramme mit dem Ausgangsstring ab. Im vorliegenden Fall entspricht das elfte Trigramm dem Namen der Band. Dem Literatureintrag `Head-on: Memories of the Liverpool punk-scene and the story of the Teardrop Explodes (1976-82)` wird also die Übereinstimmung zugewiesen.

Der folgende Code-Block zeigt ein kommentiertes und vereinfachtes Code-Beispiel für den NLTK N-Gram Abgleich:

```
1 from nltk import ngrams
```

```

2
3 musiker = 'the teardrop explodes'
4 titel = 'Head-on: Memories of the Liverpool punk-scene and the story
      of the Teardrop Explodes (1976-82)'
5
6 # Musikerstring zählen: 3 Wörter. Also: n = 3
7
8 n = len(musiker.split())
9
10 # ngrams in 3er Schritten Erstellen:
11
12 ngram = ngrams(titel.split(), n)
13
14 for grams in ngram:
15     print (grams)

```

N-Gram Abgleich

Der vollständige Literaturnachweis für die im Beispiel verwendete Publikation ist: Cope, Julian (2005). Head-on: Memories of the Liverpool punk-scene and the story of The Teardrop Explodes (1976–82). London: Element.

- Angabe bzw. Berechnung der Seitenzahl (bei Artikeln ist nur die Pagination angegeben, z. B. 35–100 also: 100–35+1, damit auch die erste Seite gezählt wird)
- Review Status
- Review Count

3.2.5 Nachbearbeitung der Ergebnisse

Da die angereicherten bibliographischen Daten trotz erster Schritte, die zur Bereinigung und Standardisierung unternommen wurden noch nicht die Datenqualität aufweisen, die als Ausgangspunkt einer systematischen Betrachtung notwendig sind, wird zusätzlich zu der automatisierten Bereinigung der Daten (z. B. der Vereinheitlichung der Ansetzung des Veröffentlichungsjahrs) ein manueller Bereinigungsschritt eingeführt. Auch hier kommt ein nicht von der Hand zu weisender Vorteil computergestützter Aufbereitung von Datenbanken zu tragen. Die Bereinigung soll in erster Linie die von RILM gelieferte, nicht durchgehend konsistente Ansetzung von

- Autorennamen
- Städtenamen
- Verlagsnamen

normieren.

Für die manuelle Nachbearbeitung der Ergebnisse wird die Open-Source Software *Open Refine* eingesetzt. Open Refine ist eine Software, deren Einsatzzweck das Bereinigen (Normalisieren) Anreichern und Konvertieren von Daten, die in Tabellenformaten vorliegen ist. Das seit 2010 bestehende Programm (zunächst *Freebase Gridworks*, dann Übernahme des Entwicklers Metaweb durch Google, Umbenennung in Google Refine, 2012 Entwicklungsstop durch Google, Umwandlung in Open Source Projekt) ist der de-facto Standard für die Aufgaben des sog. „Data Wranglings“.

Open Refine ist ähnlich wie ein Tabellenkalkulationsprogramm aufgebaut, kann aber im Gegensatz zu einem Tabellenkalkulationsprogramm mit sehr großen Datenmengen umgehen, da zu keiner Zeit die gesamte Datei angezeigt und in den Speicher geladen wird, sondern maximal 50 Datensätze auf einmal. Aus jeder Spalte können schnell Facetten generiert werden, um Einträge nach gleichen Inhalten auswählen zu können. Darüber hinaus kann das Datenmaterial mit Textfiltern und Regulären Ausdrücken gefiltert werden. Filter und Facettierungen kann man beliebig kombinieren und ihre Reihenfolge ändern.

Darüber hinaus bietet Open Refine vielfältige Möglichkeiten, um die Struktur und den Inhalt der Ausgangsdaten zu transformieren. Dazu zählt etwa die Aufspaltung von einzelnen Spalten in mehrere Spalten nach selbst zu bestimmenden Gesetzmäßigkeiten oder das Bereinigen von Zelleninhalten durch vorgefertigte Aktionen (führende und nachgestellte Leerzeichen löschen, Anpassen der Groß- und Kleinschreibung etc.) und selbstdefinierte Funktionen. Neben diesen Möglichkeiten sind es in erster Linie die Möglichkeiten der Software zur Cluster-Bildung und damit zum Auffinden abweichender Ansetzungen und Schreibweisen von Wörtern, die im Rahmen dieser Arbeit Anwendung finden:

In OpenRefine, clustering refers to the operation of “finding groups of different values that might be alternative representations of the same thing”.¹⁸¹

Zunächst gilt es jedoch, mehrfache Informationen in einzelnen Spalte zu trennen und pro Information eine Spalte zu nutzen. Es ist einer unbekannten Variable in den Ausgangsdaten geschuldet, dass nicht von vornherein beispielsweise jeder gematchte Musiker in einer Spalte erfasst wird: Da nicht ersichtlich ist, was die größte Anzahl in Titel oder Abstract vorkommender Musiker in einem Datensatz ist, werden alle Matches in eine Spalte geschrieben. Es ist also durchaus möglich, dass in einer einzigen Publikation >10 Musiker behandelt werden. Diese sind dann in der entstandenen CSV Datei alle in einer Spalte:

¹⁸¹ <https://github.com/OpenRefine/OpenRefine/wiki/Clustering-In-Depth>

Musiker Name1, Name2, Name3, ...

In OpenRefine werden diese Spalten aufgesplittet, da jetzt die maximale Anzahl an Musikern bekannt ist.

Das gleiche Vorgehen wird für mehrere Autoren in einer Spalte durchgeführt.

Die Ausgangsdaten sind in Ihrer Ansetzung nicht einheitlich. Was eine Retrieval Datenbank bei der Literaturrecherche ausgleichen kann, durch unscharfe Suche oder die Anzeige ähnlicher Schreibweisen, ist für eine Quantifizierung der Daten ein deutlicher Nachteil. Es ist sicher der Vielzahl unterschiedlicher Datenlieferanten und Eingabesysteme, die zu unterschiedlichen Schreibweisen und Ansetzungen führen. Die Eingabe von Autoren-, Verlags- und Städtenamen ist offenbar oftmals eine Freitexteingabe: Tippfehler sind nicht auszuschließen.

Die halbautomatische Normierung der Datenansetzung verbessert die Daten, die aus RILM exportiert wurden sowie die angereicherten Keywords. Das ist in Hinblick auf die statistische Auswertung äußerst wichtig, da sonst Tippfehler oder einfach ambivalente Ansetzungen als zwei Merkmalsausprägungen gezählt werden würden.

Auch die unterschiedliche Ansetzungen der neuen Metadaten muss normalisiert werden. Bei der Erstellung der Abgleichlisten wurde auf maximale Auffindbarkeit geachtet und daher wurden möglichst viele Schreibweisen von Genres aufgenommen. Dieser Vorgang muss jetzt wieder umgekehrt werden: Bei der Datenanalyse sollen z. B. drum & bass, drum and bass und ''drumbass nicht als eigene Entitäten gezählt werden sondern unter der (frei festgelegten) Bezeichnung Drum and Bass gezählt werden.

Der Vorgang des Clusterings ist nach bestem Wissen und Gewissen geschehen. Fehler, die hierbei entstanden sind oder multiple Ansetzungen, die übersehen wurden, sind trotz großer Sorgfalt denkbar. OpenRefine bietet die Möglichkeit zum Zwecke der Nachvollziehbarkeit die durchgeführten Clustering-Schritte im JSON Format zu dokumentieren. Diese finden sich im Repositorium. Die Datenbereinigung erhebt keinen Anspruch auf Richtigkeit, es geht in erster Linie darum, hier die Grundlage für statistische Auswertung zu schaffen und die Datenqualität so gut wie möglich zu verbessern. Dieser Schritt muss manuell ausgeführt werden, da es ggf. Recherche und Sachkenntnis benötigt, um bei multiplen Ansätzen die richtige Auszuwählen.

- Erläuterung der Clustering-Methoden (Nutzung von Fingerprint, N-Gram Fingerprint...)
- Beispiele und Zahlen: Wieviele Ansetzungen wurden geclustert?
- Welche Felder wiesen die meisten Ambivalenzen auf?

Überführen der Daten in ein Datenbankformat (concat.py)

3.2.6 Ausblick: Anwendbarkeit der Methode auf andere Forschungsfelder

- Anwendbar: Ja! - Sinnvoll: Kommt auf die Güte der Sacherschließung der Fachdatenbank an, statistisch Auswerten auf jeden Fall lohnend
- Format: RIS - recht universell (kommt auf Contentprovider an)
- Herausforderung: Vokabular erstellen (Hier liegt die Haupt-Herausforderung)

Aufteilen zwischen Methode generell und Anwendung der Litalyse (Ziel ist es nicht gewesen, ein generisches Framework bereit zu stellen, es geht primär um die Anwendung!)

3.3 Vorbehalte/Unzulänglichkeiten

„Der Publikationsbias ist die statistisch verzerrte (engl. bias) Darstellung der Datenlage in wissenschaftlichen Zeitschriften infolge einer bevorzugten Veröffentlichung von Studien mit ‚positiven‘ bzw. signifikanten Ergebnissen. Er wurde 1959 von dem Statistiker Theodore Sterling entdeckt. Positive Befunde sind leichter zu publizieren als solche mit ‚negativen‘, also nicht-signifikanten Ergebnissen [...]“¹⁸²

Trifft natürlich auch auf das vorliegende Datenmaterial zu. Veröffentlicht wurde nur, was angenommen wurde, was wünschenswerte Schlussfolgerungen zuließ etc. – zu was *tatsächlich *geforscht wurde ist nicht nachvollziehbar, weil hierzu keine Quellen existieren.

Besser wäre die Auswertung mehrerer Literaturdatenbanken, aber in dem speziellen Fall der Musikwissenschaft nicht sinnvoll.

- Wie gut sind die Abgleichlisten?

Hinweisen auf x-scrivener-item:///Users/Karl/Dropbox/Promotion/Dissertation/Dissertation.scriv?id=EADC7B9E-7E17-440C-AF9D-84B3E4B33289

¹⁸² <https://de.wikipedia.org/wiki/Publikationsbias>

4 Das Datenmaterial

4.0.1 Kanonisierungsprozesse

5 Fachgeschichte revisited

- Gegenstand der Musikwissenschaft
- Definitionsproblem populäre Musik
- Definitionsproblem des Musikbegriffs

Von Appen zufolge wurde aus dem bürgerlichen Musikleben Deutschlands im ausgehenden 18. Jahrhundert heraus der Anspruch formuliert, „die bis dahin untergeordnete Stellung der Musik innerhalb der Künste so weit wie möglich aufzuwerten“.¹⁸³ Dabei wurden jene Stile ausgeklammert, deren Qualitäten und Funktionen dem Kunstanspruch der Musik hätten schaden können: Musik, die „dem bloß sinnlichen Genuss, der angenehmen Unterhaltung, dem Tanz oder der bloß nachahmenden Darstellung der Natur“ galt und in der das „Irrationale, Körperliche, Zügellose, Unmoralische“¹⁸⁴ hervortreten drohte. Die Musikwissenschaft basierte lange Zeit auf diesem Kunstverständnis und schenkte deswegen den populären Spielarten der Musik keine allzu große Betrachtung.

5.1 Überblick über die Geschichte des Fachs Musikwissenschaft

Die Einträge zur Musikwissenschaft in den einschlägigen Fachlexika Grove¹⁸⁵ und MGG¹⁸⁶ geben einen Überblick über die Definition, die Arbeitsbereiche und die Geschichte der Musikwissenschaft. Hervorzuheben ist dabei, dass die beiden wichtigsten Musiklexika den Gegenstandsbereich der populären Musik in den jeweiligen Artikeln zur Musikwissenschaft praktisch gar nicht erwähnen. Der Artikel zu *Musicology* in der Continuum Encyclopedia of Popular Music¹⁸⁷ hingegen gibt neben einem Definitionsversuch und einer Einleitung in die Disziplin einen Überblick über die Geschichte der wissenschaftlichen Auseinandersetzung mit populärer Musik.

¹⁸³ (Appen2014a)

¹⁸⁴ (Appen2014a)

¹⁸⁵ (Duckles2001)

¹⁸⁶ (Cadenbach1997)

¹⁸⁷ (Shepherd2003)

Hans Heinrich Eggebrecht sieht die Anfänge der Musikwissenschaft „[...] wie die der abendländischen Wissenschaft überhaupt, in der griech. Naturphilosophie des 6. Jahrhundert. v. Chr.“.¹⁸⁸ Musik ist im Laufe der Zeit immer Gegenstand wissenschaftlicher Betrachtungen gewesen. Die *Ars Musica* im europäischen Mittelalter ist, nachdem die Musikforschung keine eigene Disziplin mehr an den Universitäten ab dem 15. Jahrhundert war, der Vorläufer der universitären Musikwissenschaft, die sich ab dem 19. Jahrhundert an den Universitäten als Studienfach etabliert hat. Den ersten Lehrstuhl der modernen Musikwissenschaft bekleidete 1861 Eduard Hanslick in Wien.¹⁸⁹ Deutschland gilt als das „Kernland der historischen Musikwissenschaft“.¹⁹⁰

In den USA, dem Land, dessen populäre Musik maßgeblich für die globalen Entwicklungen des 20. Jahrhunderts ist, vollzog sich die Institutionalisierung der Musikwissenschaft vergleichsweise spät, der erste Lehrstuhl in den Vereinigten Staaten wurde erst 1930 eingerichtet.

Calella und Urbanek - Historische Musikwissenschaft

5.1.1 Strömungen in der Musikwissenschaft

Die Geschichte einer wissenschaftlichen Disziplin verläuft nicht linear. Durch die Internationalisierung des Fachs und der Ausdifferenzierung der Forschungsinteressen und der Forschungscommunity kommt es zu Diskussionen um die Ausrichtung des Fachs und in Folge dessen um Abspaltungen Strömungen.

- Warum sind die Cultural Studies (scheinbar) so wichtig für die Populärmusikforschung? Verweise auf das Centre for Contemporary Cultural Studies (CCC) in Birmingham gibt es zu Hauf
- Welchen Einfluss /welche Ziele haben die Cultural Studies in Bezug auf die populäre Musik?
- Gilt speziell für die USA
- Joseph Kerman (Berkeley) gilt als Initiator der New Musicology-Bewegung¹⁹¹
- New Musicology ist kein Grundsatzprogramm, gibt weder den Gegenstand noch die Methode vor.
- New Musicology ist eher als Sammelbecken zu verstehen für eine zumindest teilweise Abkehr von traditionellen Denktraditionen und Ansätzen in der Musikwissenschaft. /Abkehr vom traditionellen Verständnis, das sich aus dem Musikverständnis des 19. Jahrhunderts ergibt ()

¹⁸⁸ (Eggebrecht1995)

¹⁸⁹ (Jaschinski1997)

¹⁹⁰ (Jaschinski1997)

¹⁹¹ vgl. (Danielsen2008), (Calella2011) und (Hooper2006)

- Weg vom Positivismus: Positivismus geht davon aus, dass nur gesicherte Erkenntnisse einen Wissensgewinn erwarten lassen
- New Musicology bedeutet nicht, dass sich der Gegenstand ändern muss, aber der Blickwinkel vergrößert sich

Definition aus dem Oxford Handbuch: Slogan, USA als Keimzelle, für Europa nicht neu?

[New Musicology is] A term that became something of a slogan in the late 1980s, especially in the USA. It arose out of the perception that musicology as a discipline had become too strongly based on sources, documentation, and newly discovered facts; that it lacked broader consideration of critical, aesthetic, psychological, perceptual, and sociological issues. At the same time there was a sharp rise in feminist musicology and, following soon after, gay and lesbian musicology—all subjects that inevitably looked more closely at contextual issues—and these too came under the banner of the new musicology. In retrospect it is hard to see what was new about the broader basis thus proposed: it may have been unfashionable for a time in America, but it had been present uninterruptedly in European musicology from the days of Guido Adler's earliest attempt to define the discipline in 1885. Certainly there have been new ways of thinking, made possible particularly by developments in sociology, philosophy, and psychology; but musicology has almost always embraced new ideas from other disciplines, and its viability for the future will surely lie, at least in part, in its ability to continue doing so.¹⁹²

Definitionsversuch: Veränderung aus der traditionellen Musikwissenschaft (nicht von Außen), nicht

A shorthand definition: a New Musicologist looks at institutionalized musicology from the inside the way popular music scholars have always looked at it from the outside: with a certain ideological suspicion. The New Musicology is one product of a decade-long general disciplinary crisis within the academic study of music. In a sense it represents a generational split; it is the collapse for many younger scholars of some of the field's ruling ideological assumptions, often as a result of acknowledging the very historical contingencies catalogued by Middleton above. (In other words, New Musicology is what you get when musicologists themselves become aware of the musicological problem.) There is a large overlap between these so-called New Musicologists and the small but growing number of "crossover" musicologists who study popular music from within traditional music departments; for many of us, it seems natural to combine new approaches to canonical classical music with interest in various repertoires of popular music, especially post-1955 rock and pop.¹⁹³

Bestrebung, traditionelle musikwissenschaftliche Vorgehensweisen zu ergänzen:

¹⁹² (Fallows2011)

¹⁹³ (Fink1998)

*New musicology emerged as an alternative to a more formalist oriented American musicology during the 1980s. It was pioneered by, among others, Joseph Kerman, and may be characterized as an ambition to merge perspectives from musicology, music history and cultural studies.*¹⁹⁴

Entstehung (Kerman):

*Most historical accounts of the emergence of a 'new' musicology tend to locate the first proper articulation of its motivating impulses in Joseph Kerman's article, 'How We Got into Analysis, and How to Get Out' (1980) and his book, [...] Contemplating Music [...] (1985). While the attempt to identify a 'prima causa' for any historical development inevitably risks sliding toward an untenable determinism or oversimplification, it is difficult to deny the part played by Kerman's texts, even today, in consolidating the image of an 'older' (conservative, reactionary) discipline – defined by a musicological 'positivism' and an analytical 'formalism' – and the correlative need for a 'newer' (critical, progressive) direction.*¹⁹⁵

Zusammenfassung:

*In den 1980er und 1990er Jahren gewannen zudem auch dort [USA] die Cultural Studies großen Einfluss. Ihnen ging es in den USA insbesondere um den Kampf für einen kulturellen Pluralismus, um die Kritik am dominanten »weißen«, männlichen, heterosexuellen Blick auf Kunst und Kultur, speziell um die Anerkennung der Kulturen ethnischer Minderheiten sowie um Kulturen Homosexueller. Unter dem Einfluss dieser neuen Forschungsrichtungen und der damit in Zusammenhang stehenden Critical oder New Musicology, die die elitäre und hegemoniale Ideologie des eigenen Faches kritisierte, öffnete sich die bis dahin stark von historischen Forschungsinteressen geprägte Musikwissenschaft erst in den 1990er Jahren allgemeineren kulturwissenschaftlichen Ansätzen und damit Musiken, die nicht in der Tradition der »großen Meister« stehen [...].*¹⁹⁶

Gegenstand blieb zunächst oft gleich, grundsätzliche Änderung der Denkweise aber als Voraussetzung für das Ernstnehmen populärer Musik:

*[...] many of these [New Musicology Scholars] scholars continued to devote their attention to music of the Western European classical tradition. Yet their rejection of the ideology of autonomous musical art that had dominated academic musicology, and accompanying recognition that music was shaped as much by questions of power and cultural representation as any other expressive form, was a necessary precondition for taking popular music more seriously as a subject of academic study. Previously it was all too easy to dismiss popular music for its 'impurity', for the degree to which it was implicated in economic and social processes and so lacked aesthetic content of sufficient complexity to warrant careful analysis. However, if scholars recognized that classical music was comparably implicated in these processes, the grounds according to which popular music was deemed unworthy of consideration became increasingly untenable.*¹⁹⁷

¹⁹⁴ (Danielsen2008)

¹⁹⁵ (Hooper2006)

¹⁹⁶ (Appen2007)

¹⁹⁷ (Bennett-Waksman2015)

Korrektur in der Ausrichtung der Musikwissenschaft

Es wird immer offensichtlicher, dass der Gegenstand der Musikwissenschaft – die Musik – als Ergebnis diskursiver Praktiken im Rahmen der akademischen Musikwissenschaft produziert wird. Denn die Fachgründer und besonders die späteren Fachvertreter der historischen Musikwissenschaft haben nicht nur dafür gesorgt, dass man sich hauptsächlich mit einer elitären Form von Musik beschäftigt, sondern sie haben den Textcharakter der Musik (das Kunstwerk) in den Vordergrund gestellt und dies auf Kosten anderer Aspekte, wie z. B. des Klanges und der Ausführung – die dann teilweise Aufgabengebiete der Systematiker und der Musikethnologen wurden. Die new musicologists sind der Meinung, Musik besitze immer eine kulturelle Bedeutung – Lawrence Kramer hat nicht zuletzt die Bezeichnung „cultural musicology“ statt „new musicology“ vorgeschlagen.¹⁹⁸

NM zunächst oft negativ besetzt:

The term “new musicology”, unlike “criticism”, was not primarily a term of self-description. Often, it was a term a musicologist would apply to someone else, an expression of surprise and consternation at new forms of inquiry that become increasingly visible from the 1980s on. Like “criticism”, it was a term used primarily within the field of historical musicology, though it often came with a new impulse to bring popular music within the reach of musicology.¹⁹⁹

critical musicology, n. 1. A form of musicology which applies aspects of Critical Theory as practiced within other humanities disciplines to music. 2. A form of musicology which involves the theoretical critique of previous musicological traditions.²⁰⁰

5.2 Populärmusikforschung

vgl. Definition in der Einleitung

Wissenschaftliche Disziplinen als „auf bestimmte Ausschnitte der Umwelt der Wissenschaft spezialisierte Einheiten“, die „die Basis der internen Differenzierung der Wissenschaft“ darstellen.²⁰¹ Der „Ausschnitt der Umwelt“, dem sich die Populärmusikforschung widmet, ist ein Phänomen, das in westlichen Industrienationen im 20. Jahrhundert entstanden ist.

Der Gegenstand bestimmte, wie sich die Erforschung populärer Musik entwickelte:

- Zunächst aufgefasst als Soziales Phänomen (Soziologie, Pädagogik),
- Einfluss des ökonomischen Teils der populären Musik (Populärwissenschaftliche Abhandlungen, Musikzeitschriften/Musikkritiker, Biographien)
- Musik außerhalb dessen, was die traditionelle Musikwissenschaft als ihren Gegenstand ansieht (Ethnomusikologie)

¹⁹⁸ (Calella2011)

¹⁹⁹ (Maus2010)

²⁰⁰ <http://www.leeds.ac.uk/music/Info/critmus/>

²⁰¹ (Balck2016)

202

- 1981 Gründung der International Association for the Study of Popular Music (IASPM)
- 1986 Gründung der Gesellschaft für Populärmusikforschung e.V.
- 2012 Gründung der IASPM D-A-CH

Situation aktuell:

- 9 Studiengänge, davon 4 dediziert Berufsorientiert - + allgemeine musikwissenschaftliche Studiengänge mit entsprechenden Veranstaltungen

ggf. Auswertung von Studieren.de und anderen internationalen Datenbanken?
vgl. Lehrveranstaltungen

- Journal of World Popular Music
- Lied und populäre Kultur/Song and Popular Culture: Jahrbuch des deutschen Volksliedarchivs Freiburg
- Perfect Beat: the Pacific Journal of Research into Contemporary Music and Popular Culture
- Popular Music History
- Popular Musicology Online
- Popular Musicology Quarterly
- Samples: Notizen, Projekte und Kurzbeiträge zur Populärmusikforschung (2002-)
- Tracking: Popular Music Studies (1988–1992)

6 Diskussion

Literatur