

# Musikhistorische Daten, Netzwerkanalyse und Migration

**Stadler, Peter**

pstadler@mail.uni-paderborn.de  
Universität Paderborn, Deutschland  
ORCID: 0000-0002-2544-1481

**Grund, Vera**

vgrund@mail.uni-paderborn.de  
Universität Paderborn, Deutschland

## Einführung

Bei der ersten DHd-Tagung 2014 in Passau wurde ein Thesenpapier zur Standortbestimmung der Digital Humanities in Hinblick auf ihr Selbstverständnis, ihre Aufgabenstellung und Entwicklungsperspektiven erarbeitet. Besonders betont wurde dabei die Bedeutung von Interdisziplinarität für die Digital Humanities sowie ihr besonderes methodisches Potential, Erkenntnisgewinn durch empirische Praxis zu erzielen. Als wissenschaftliche Handlungsräume wurden die “Annotation, Analyse, Aggregation und Rekombination von geisteswissenschaftlichen Objekt- und Metadaten in Kombination von algorithmischen und hermeneutischen Prozeduren”<sup>1</sup> benannt.

Um einen Material- bzw. Datenkorpus, der in diesem Sinne als Grundlage für ein Digital Humanities Forschungsprojekt dienen kann, handelt es sich bei der bereits in den 1980er Jahren durch das „U.S. RISM Libretto Project“ digital erschlossene Albert Schatz-Librettokollektion (McClymonds/Parr Walker 1986; Needham Costonis 1991). Es handelt sich dabei um eine der größten Sammlungen von Opernlibretto-Drucken. Diese wurden in der Regel in Zusammenhang mit Aufführungen erstellt und enthielten daher neben dem Text der Oper außerdem Informationen zu den aufgeführten Werken und den Aufführungen selbst. Diese Informationen wurden bei der Erschließung durch das RISM Libretto Projekt feingranular aufgenommen, so dass nicht nur basale Metadaten wie Titel, Komponist:in, Librettist:in, Druckort und -jahr erfasst sind, sondern auch die in den Libretti enthaltenen Angaben zu den Ausführenden der jeweiligen Aufführung. Aufgrund der Qualität und Menge der dabei generierten Daten können diese für die quantitative Netzwerkforschung weiter genutzt werden, oder wie es in dem Thesenpapier heißt, “zur Formulierung neuer qualitativer geisteswissenschaftlicher Fragestellungen” sowie für die “Operationalisierung qualitativer geisteswissenschaftlicher Fragestellungen auf der Basis formaler, logisch-mathematischer Verfahren der Informatik” die

Basis bilden.<sup>2</sup> Das Projekt möchte durch die Nachnutzung von Daten zwischen frühen Digitalisierungsprojekten und den Digital Humanities vermitteln, indem es die Potentiale der Daten ebenso wie die Schwachstellen aufzeigt und so Kriterien für die Nachhaltigkeit von Daten erkennen lässt. In diesem Sinne kann es prototypisch für die in den Thesen formulierten Kriterien von Digital Humanities stehen und als Fallbeispiel in Hinblick auf die Methodik bei der Erstellung von Datenkorpora in Digitalisierungsprojekten und deren Nutzbarkeit für die Digital Humanities stehen. Im Rahmen des Papers sollen erste Ergebnisse zur Erforschung von Künstler:innen-Netzwerken auf der Basis der Daten aus der Schatz-Sammlung im Verhältnis zu den Thesen zu den Digital Humanities präsentiert sowie ihre Potentiale für weiterführende Forschung aufgezeigt werden. Weiter sollen auf der Basis des Projekts Ergänzungen der Thesen insbesondere in Hinblick auf in den Daten enthaltene Differenzkategorien und den bewussten Umgang damit dargestellt werden.

## Die Sammlung Schatz

Die Library of Congress kaufte im Jahr 1908 die Libretto-Kollektion des Rostocker Sammlers Albert Schatz als Vorlass, bevor dieser im Jahr 1910 verstarb. Schatz beschrieb die Sammlung als „Ergebnis eigener Sammeltätigkeit von über 42 Jahren“ (Sonneck 1914), die er aufgrund von Interesse für die Oper und ihre Geschichte begann. Diese forcierte er zusätzlich, als er im Herbst 1873 Besitzer der Rostocker Musikalienhandlung Ludwig Trutschel wurde. Da laut Schatz „die bisherigen Darstellungen der Operngeschichte [...] völlig unzureichend und unzuverlässig waren, da sie in der Hauptsache nicht unmittelbar auf den Quellen aufgebaut, sondern ohne [ihre] Kenntnis“ (Sonneck 1914) verfasst wurden, entschloss er sich auf der Basis seiner Librettosammlung eine Operngeschichte zu schreiben. Zwar blieb diese ein Fragment, jedoch dokumentierte er die Erkenntnisse aus seiner Sammeltätigkeit in Form von circa 80.000 Zetteln, mit denen Schatz laut eigener Aussage 30.000 Operaufführungen nachweisen konnte.<sup>3</sup>

Die von der Library of Congress angekaufte Libretto-Sammlung enthielt 12.240 Textbücher, bei denen es sich abgesehen von wenigen hundert Oratorien und Kantaten, die vermutlich per Zufall in die Sammlung kamen, um Programme zu Operaufführungen handelt. Seltene Librettodrucke wurden auch als handschriftliche Kopien in die Sammlung aufgenommen. Der geografische Schwerpunkt lag auf deutschen und italienischen Texten, wohingegen französische Opern eher in deutschen Übersetzungen Eingang in die Sammlung fanden (Sonneck 1914). Bei der Erschließung der Albert-Schatz-Collection durch das „U.S. RISM Libretto Project“ wurden neben den Grunddaten zu den mit den Libretti in Verbindung stehenden Werken auch Figurennamen sowie die in den Textbüchern enthaltenen Informationen zu Aufführungen wie Theater und beteiligte Personen erfasst.

## Datenaufbereitung

Von besonderem Interesse für die Netzwerkforschung ist dabei die systematische und normierte Erfassung bzw. Beschreibung von „relator terms“ zu den beteiligten Personen nach dem kontrollierten Vokabular MARC Code List for Relators<sup>4</sup>. So sind beispielsweise in dem Metadaten-satz zum Libretto *Le villanelle astute* zur Aufführung im venezianischen Teatro San Samuele im Jahr 1786 die auf der Titelseite genannten Personen Giovanni Maria Foppa, Francesco Bianchi, Catarina Casalis und Eusebio Luzzi als „librettist“, „composer“, „vocalist“ und „dancer“ ausgewiesen (Foppa 1786).

Die Daten wurden durch das „U.S. RISM Libretto Project“ nicht nur gut und strukturiert erfasst, sie werden auch von der Library of Congress über eine offene Programmierschnittstelle verfügbar gemacht, so dass die gesamten Angaben zu den Libretti, die über den Online-Katalog zugänglich sind, auch maschinell abgerufen werden können.<sup>5</sup> Aus diesen Datensätzen haben wir die folgenden Informationen extrahiert:

### Personennamen und Funktionen

Diese beiden Angaben finden sich in innerhalb der Beschreibung des „item“ als „contributor\_names“. Jede Person ist einzeln aufgenommen, allerdings ist der „relator term“ in Klammern angefügt, sodass dieses Feld entsprechend analysiert werden muss, um die Namen von den Funktionen zu separieren. Aus dem Eintrag „Luzzi, Eusebio. (dancer)“ (siehe obiges Beispiel) wird dann entsprechend der Name „Luzzi, Eusebio“ mit der Funktion „dancer“ abgeleitet. Im Gegensatz zu den relator terms sind die Eigennamen nicht normalisiert, sondern tauchen oft in verschiedenen orthographischen Varianten oder Abkürzungen (von z.B. Vornamen) auf. Insgesamt ergeben sich so in unserem Korpus über 43.000 verschiedene Namensformen, die wir mithilfe von OpenRefine (vorsichtig) geclustert haben, um diese Namensvarianten zusammenzuführen.

### Aufführungsort

Den Aufführungsort gewinnen wir aus den „location“-Feldern. Hierbei gehen wir davon aus, dass der Druckort dem Aufführungsort entspricht. Für unsere Untersuchung ziehen wir nur die Libretti heran, die explizit die beteiligten Künstler:innen wie „vocalist“ erwähnen, sodass bei diesen Libretti davon auszugehen ist, dass sie auch am Aufführungsort gedruckt worden sind.

Die Orte sind ähnlich wie die Personennamen nicht normalisiert (Venedig erscheint beispielsweise als „Venetia“, „Venezia“ oder auch „In Venezia“ etc.) und mussten analog der Personennamen geclustert werden. Der Umfang war allerdings wesentlich geringer: Insgesamt existierten 748 verschiedene Ortsangaben im Gesamtkorpus, die wir auf 427 normalisierte Angaben abbilden konnten.

## Aufführungsjahr

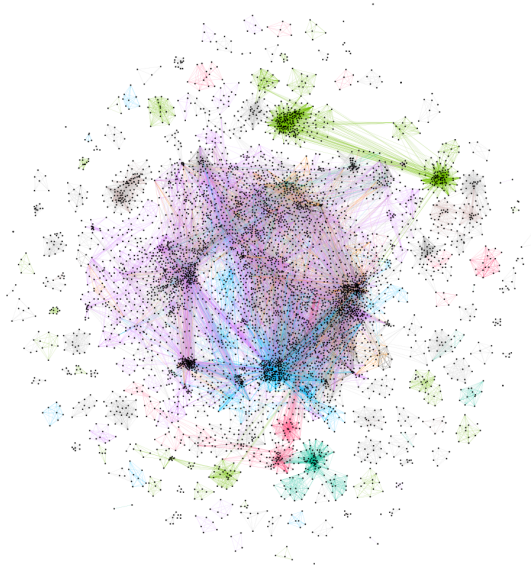
Die Datumsangaben wurden bei der Katalogisierung ebenfalls von den Titelblättern der Drucke übernommen und liegen in zwei Varianten vor. Einerseits als Freitextvariante (z.B. „18??“) und andererseits eine offenbar davon abgeleitete normalisierte Form mit zwei Einträgen („1800-01-01T00:00:00Z“ und „1899-01-01T00:00:00Z“). In der Katalogansicht der Library of Congress wird für die Facettierung dann diese normalisierte Form genutzt, was zu verzerrten Angaben führt, wenn für die Jahre 1800 und 1899 mehrere Hundert Treffer ausgegeben werden, für 1801 aber z.B. nur 80. Wir haben daher für die Bestimmung des Aufführungsjahres diese Unschärfen der ursprünglichen Datenerfassung aufgefangen und ggf. Datumsbereiche oder unsichere Datumsangaben markiert.

## Netzwerke

Aus den so normalisierten Daten zu Personen, Funktionen, Orten und Jahr haben wir anschließend eine Einschränkung auf Sänger:innen (vocalists) im 18. Jahrhundert vorgenommen, um entsprechende Thesen zu Mobilität und Migration aus der musikwissenschaftlichen Forschung untersuchen zu können (Strohm 2001; Zur Nieden 2015).

### Personennetzwerk

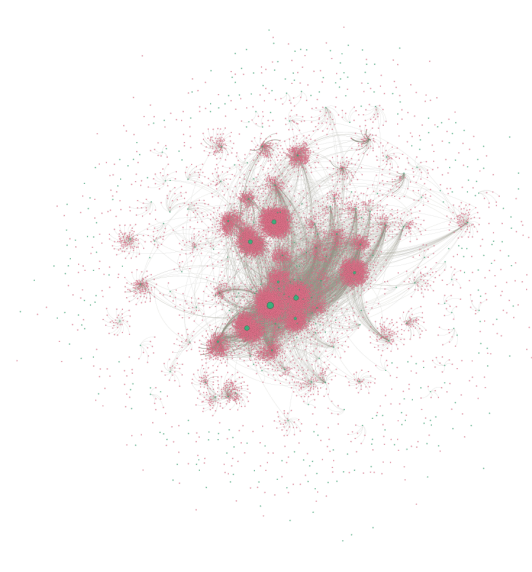
Zunächst haben wir zwecks eines initialen, explorativen Vorgehens ein einfaches ungerichtetes Personennetzwerk erstellt, bestehend aus den Personen als Knoten und den Aufführungen – bei denen zwei Personen gemeinsam künstlerisch gewirkt haben – als Kanten. Dieses Netzwerk der Sänger:innen-Beziehungen des 18. Jahrhunderts besteht aus insgesamt 6203 unterschiedlichen Personen (Knoten), die über 50.450 Relationen miteinander verbunden sind. Die Abbildung 1 gibt einen ersten Eindruck von der Größe des Netzwerks, sichtbar wird aber durch die Einfärbung der Kanten nach Orten auch direkt die Dominanz Venedigs (in Lila) und die räumliche Trennung von z.B. Paris (Grün) und Berlin (Rot).



Bemerkenswert ist aber ein Blick auf die Zentralitätsmaße: Hier sticht Nicola Grimaldi mit dem höchsten gewichteten Grad von 415 und einem Eigenvektor Zentralitäts-Score von 1 hervor.<sup>6</sup> Dies verwundert nicht, ist Grimaldi doch in der Musikgeschichte als erfolgreicher und bekannter Sänger überliefert.

### Bipartites Personen-Orts-Netzwerk

Um das Verhältnis bzw. den Einfluss von den Orten auf die Personen genauer zu untersuchen, haben wir im Anschluss ein (ungerichtetes) bipartites Personen-Orts-Netzwerk erstellt. Hierbei wurden Orte (grün eingefärbt) und Personen (rot eingefärbt) als Knoten aufgenommen und als Kanten die durch Aufführungen dokumentierten Beziehungen zwischen Personen und Orten (vgl. Abb.2). Die nicht verbundenen Knoten zeigen Orte und Personen an, die lediglich erwähnt sind, aber keine Beziehung zu einem Knoten des anderen Typs herstellbar ist. Im Falle von Orten sind dies Aufführungen ohne Beteiligung von Sänger:innen und bei Personen sind dies Aufführungen mit unbekanntem Aufführungsort.

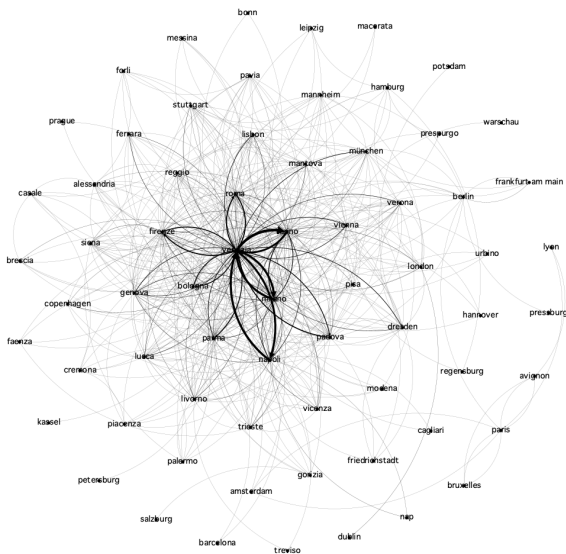


Wie bereits durch die Einfärbung der Kanten im Personennetzwerk angedeutet, manifestiert sich hier eine deutliche Dominanz der italienischen Opernzentren, insbesondere Venedigs. Unser Hauptaugenmerk liegt aber auf der Mobilität der Sänger:innen, die hier ebenfalls sichtbar wird. So finden sich zahlreiche Personen, die mit 5 oder mehr Orten verknüpft sind, d.h. eine hohe Mobilität aufweisen. Spitzenreiter mit 11 verschiedenen Orten ist Girolamo Crescentini, ein berühmter italienischer Kastrat, der um 1800 wirkte. Dabei fehlen bei unserer Beschränkung auf das 18. Jh. sogar noch Stationen, denn Crescentini war im 19. Jahrhundert noch in Wien und Paris tätig.

Die überwiegende Mehrzahl der Personen ist aber nur mit einem Ort in Verbindung zu bringen. Ob das nun an einer größeren Seßhaftigkeit oder aber der Datenlage oder der Ausschnittsbildung liegt, lässt sich aktuell noch nicht beantworten.

### Ortsnetzwerk

Zuletzt haben wir noch einen gezielten Blick auf die Migrationsbewegungen zwischen den Orten gelegt und dafür ein gerichtetes Ortsnetzwerk erstellt. Diesem liegen die chronologisch sortierten Auftritte jeder Person im Korpus zugrunde, sodass sich aus der Aufeinanderfolge von zwei Auftritten einer Person an verschiedenen Orten eine Bewegung inferieren lässt.



In Hinblick darauf ist das 18. Jahrhundert als Zeitraum von besonderem Interesse. Denn wie Reinhard Strohm feststellte (Strohm 2001), wandelte sich aufgrund von zunehmender Mobilität das Verhältnis in Bezug auf die von ihm konstatierte „italian musical diaspora“. Wie Strohm darstellte, hatte diese zuvor als mediale Wanderbewegung von Notendruckern stattgefunden. Da große Teile der Kultur des 18. Jahrhunderts durch „Italianität“ als Modetrend bestimmt war, ging damit ein starkes Interesse an italienischen Künstlerinnen und Künstlern im europäischen Raum einher, weswegen laut Strohm Wanderbewegungen zunahmen. Er ging davon aus, dass sich die Karriereverläufe von Bühnenkünstler:innen, insbesondere von Sängerinnen und Sängern, daher stark änderten. In Hinblick auf italienische Kastraten-Sänger behauptete Strohm sogar, dass es sich um „Exportprodukte“ handelte. Unsere Daten belegen einen regen Austausch zwischen den verschiedenen europäischen Musikzentren, allerdings kann von einem reinen Export nicht die Rede sein.

## Fazit und Ausblick

An dieser Stelle kann nur ein Vorgeschmack auf den „Daten-Schatz“ gegeben werden, der große Mengen an Informationen enthält u.a. zu tausenden von Künstler:innenbiographien, die mit wenigen Mausklicks – zwar nicht vollständig, aber als belastbare Basis – zur Verfügung stehen. Besonders von Interesse ist dies auch in Hinblick auf Tänzer:innen-Karrieren, über die nur in den seltensten Fällen Forschung existiert. Zwar liegt mit der Datenbank [operebuffa.uni-bayreuth.de](https://operebuffa.uni-bayreuth.de) Forschung zu Akteur:innen von Opern buffe zwischen 1740 und 1765 vor. Diese reichen allerdings nicht über das komische Genre und den genannten Zeitraum hinaus. Die Daten aus der Schatz-Collection bieten hingegen eine Gesamtschau von Oper im 18. und 19.

Jahrhundert, womit statt heuristisch vom Einzelfall ausgehend auf das Allgemeine zu schließen, umgekehrt vorgegangen werden kann.

Neben den bisher auf Personen beschränkten Wanderbewegungen, ließe sich damit auch die Mobilität von bestimmten Werken oder Sujets untersuchen, Fragen in Bezug auf die Bedeutung von bestimmten Orten als Ausbildungsstätten könnten daran gestellt werden, anhand der Betrachtung von Widmungsträger:innen ließen sich Mäzenatentum bzw. Kultursponsoring untersuchen. Weitere für die musikhistorische Forschung zu hebende „Schätze“ liegen beispielsweise in der Untersuchung zu Dauern von Karrieren oder auch zu Geschlechterunterschieden bei Karriereverläufen.

Um diese Fragen adäquat adressieren zu können, bedürfen die Daten aber noch einiger Nachbearbeitung und im Idealfall der Anreicherung mit Normdaten zur eindeutigen Identifikation und Verlinkung mit externen Ressourcen.

In einem weiteren Schritt ließen sich die digitalisierten Librettotexte außerdem als Datenkorpus für Distant Reading Projekte in Hinblick auf Genrefragen untersuchen, wie es z.B. Giovannini und Daniil (2023) an einem kleinen Libretto-Bestand in der Dramendatenbank [dracor.org](https://dracor.org) bereits versucht haben oder wie es im Projekt „Emotionen im Drama“ als „Sentiment Analysis“ auf einen Bestand von Opernlibretti der Hamburger Gänsemarktoper von 1678 bis 1730 angewandt wurde (Dennerlein et al. 2019). Aufgrund der Größe des zur Verfügung stehenden Korpus, den die digitalisierte Schatz-Collection bildet, bietet es sich insbesondere an, computergestützte Verfahren auszuprobieren oder weiterzuentwickeln. Zwar mit erheblichem Aufwand verbunden, aber auch besonders reizvoll wäre es, eine Verbindung von überlieferten Musikpartituren zu den Libretti herzustellen und die Distant Reading Methoden auch auf Musik zu übertragen. Ein groß angelegtes Projekt könnte in diesem Fall sowohl methodisch wie inhaltlich die Digital Humanities um die musikwissenschaftliche Perspektive ergänzen. In Hinblick auf die 2014 formulierten Thesen des DHd könnte ein derartiges Projekt aufgrund der Interdisziplinarität ebenso entsprechen, wie es das Potential digitaler Methoden voll ausschöpfen könnte, in Hinblick auf „Annotation, Analyse, Aggregation und Rekombination von geisteswissenschaftlichen Objekt- und Metadaten in Kombination von algorithmischen und hermeneutischen Prozeduren“<sup>7</sup> wie es im Thesenpapier gefordert wurde.

## Fußnoten

1. Digital Humanities 2020, <https://dig-hum.de/thesen-digital-humanities-2020> (zugegriffen: 12.7.23).
2. Digital Humanities 2020.
3. Zettelkatalog von Schatz ebenfalls in der Library of Congress.
4. <https://www.loc.gov/marc/relators/relaterm.html> (zugegriffen: 19.7.2023).
5. Vgl. die API-Beschreibung unter <https://www.loc.gov/apis/json-and-yaml/> (zugegriffen: 19.7.2023).

6. Berechnet mit Gephi 0.10.1 202301172018  
 7. Digital Humanities 2020, <https://dig-hum.de/thesen-digital-humanities-2020> (zugegriffen: 12.7.23).

## Bibliographie

- Dennerlein, Karin, Thomas Schmidt und Christian Wolff.** 2019. „Emotionen im kulturellen Gedächtnis bewahren“, <https://zenodo.org/records/6327957> (zugegriffen: 3.12.23)
- Foppa, Giuseppe Maria. Le villanelle astute.** 1786. Venedig: Fenzo. <https://lcn.loc.gov/2010663602> (zugegriffen: 24.5.2023).
- Giovannini, Luca, und Skorinkin, Daniil.** 2023. “Computational approaches to opera libretti”. In *Journal of Computational Literary Studies*, 2 (im Druck). Preprint: [https://github.com/DanilSko/opera/blob/main/giovannini\\_skorinkin\\_libretti\\_2023pp.pdf](https://github.com/DanilSko/opera/blob/main/giovannini_skorinkin_libretti_2023pp.pdf) (zugegriffen: 3.12.23).
- McClymonds, Marita P. und Parr Walker, Diane.** 1986. “U.S. Rism Libretto Project: With Guidelines for Cataloguing in the MARC Format”. In *Notes* 43/1: 19–35
- Needham Costonis, Maureen.** 1991. “The U.S. Rism Libretto Project: Accessible Online Database for Researchers”. In *Dance Research Journal* 23/2: 55–57
- Sonneck, Oscar George Theodore.** 1914. Preface. In *Library of Congress, Catalogue of Opera Librettos printed before 1800*, hg. von Oscar Sonneck. Washington: 1–19. <https://archive.org/details/catalogueoperall00sonngoog/> (zugegriffen: 12.7.23).
- Strohm, Reinhard.** 2001. „Italian Operisti North of the Alps, c. 1700 - c. 1750“. In *The Eighteenth-Century Diaspora of Italian Music and Musicians*, hg. von Reinhard Strohm. Turnhout
- Zur Nieden, Gesa.** 2015. „Frühneuzeitliche Musikermigration nach Italien: Fragen, Verflechtungen und Forschungsgebiete einer europäischen Kulturgeschichte der Musik“. In *Musicisti europei a Venezia, Roma e Napoli / Europäische Musiker in Venedig, Rom und Neapel / Les musiciens européens à Venise, à Rome et à Naples (1650–1750)*, hg. von Anne-Madeleine Goulet und Gesa zur Nieden. Kassel: Bärenreiter (= *Analecta Musicologica* 52): 9–30