

Forschungsdaten in der Kinder- und Jugendliteraturforschung – eine  
Bestandsaufnahme

Projektarbeit im Rahmen der Grundausbildung des Universitätslehrganges *Library  
and Information Studies* an der Österreichischen Nationalbibliothek

Eingereicht von:

Valerie Nitsche, Klaus Prinz, Thomas Wolfsgruber

Wien, im September 2023

## Inhalt

1.	Einleitung.....	3
2.	Forschungsdaten finden und erkennen .....	4
2.1.	Hinführung und Problem der Eingrenzungen .....	4
2.2.	Inhaltliche Annäherung .....	5
2.3.	Strukturelle Annäherung .....	6
2.4.	Definition für vorliegendes Projekt .....	8
2.5.	Forschungsdaten finden .....	9
3.	Recherche .....	13
3.1.	Recherchestrategie.....	13
3.1.1.	Recherche in Repositorien und Metasuchportalen.....	13
3.1.2.	Sichtung von Publikationen .....	17
3.1.3.	Sichtung von Institutsseiten und Portalen.....	18
3.2.	Herausforderungen .....	19
3.3.	Rechercheergebnisse .....	21
3.3.1.	Recherche in Repositorien und Metasuchportalen.....	21
3.3.2.	Sichtung von Publikationen .....	25
3.3.3.	Sichtung von Institutsseiten und Portalen.....	27
4.	Versuch und Potenzial einer Collection .....	30
5.	Conclusio.....	31
6.	Literaturverzeichnis .....	33
	Anhang 1: Textanalytische Versuche an fach einschlägigen Publikationen..	37
	Anhang 2: Poster „Wo findet man Forschungsdaten?“ .....	48
	Anhang 3: Datei: Rechercheprotokoll_Forschungsdaten_KJL	

## 1. Einleitung

Open Research Data, offen verfügbare und nachnutzbare Forschungsdaten, bedeuten einen erheblichen Nutzen und Gewinn für die Wissenschaft und Forschende. Dauerhaft zugängliche und öffentlich geteilte Forschungsdaten können wiederverwertet und weiter genutzt werden, was nicht nur den Wert von bereits geleisteter Forschungsarbeit steigert, sondern auch wissenschaftliche Vernetzung fördert und Kosten bei der Erhebung von Daten senken kann. Nicht nur dienen Forschungsdaten dem Zweck weiterführender Forschung, offen erhältliche Forschungsdaten eröffnen Möglichkeiten für neue Erkenntnisse und Ergebnisse zu bereits vorliegenden Forschungsergebnissen. Weiters dient die Veröffentlichung von Forschungsdaten der Förderung eines transparenten wissenschaftlichen Betriebs sowie der Nachvollziehbarkeit von Forschungsergebnissen. Forschende können von der frei zugänglichen Veröffentlichung ihrer Forschungsdaten auch in dem Sinne profitieren, dass ihre Forschung dadurch sichtbarer und häufiger zitiert wird.<sup>1</sup>

Das ULG-Projekt „Forschungsdaten in der Kinder- und Jugendliteraturforschung – eine Bestandsaufnahme“ wurde mit dem Ziel angelegt, einen Überblick über frei verfügbare Forschungsdaten aus dem Bereich der Kinder- und Jugendliteraturforschung zu schaffen und anhand dessen eine systematisierte und kommentierte Collection dieser Forschungsdaten für die freie Nachnutzung anderer zur Verfügung zu stellen.

Die Durchführung des Projekts wurde durch eine Literaturrecherche zum Forschungsdatenbegriff sowie zu Forschungsdatenrecherchemethoden eingeleitet. Basierend auf den gewonnenen Erkenntnissen wurde eine Recherchestrategie entwickelt, die sich in drei verschiedene Recherchepfade aufgliederte. Diese Pfade wurden anschließend wieder zusammengeführt und anhand der Rechercheergebnisse einer eingehenden Analyse unterzogen. Die gewonnenen Erkenntnisse aus der Analyse ermöglichten schließlich eine Einschätzung des aktuellen Stands der Forschungsdatenmanagementpraxis in der KJL-Forschung sowie eine Bewertung der Möglichkeit einer sinnvollen Zusammenstellung der Ergebnisse.

---

<sup>1</sup> Vgl. „Daten publizieren: Vorteile und Möglichkeiten des Data-Sharing,“ Informationsplattform [forschungsdaten.info](https://forschungsdaten.info), zugegriffen 5. Juli 2023, <https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/daten-publizieren/>.

## 2. Forschungsdaten finden und erkennen

### 2.1. Hinführung und Problem der Eingrenzungen

Der Versuch, Forschungsdaten in den Geisteswissenschaften zu definieren, dieses Konzept genau und trennscharf zu bestimmen scheint – auf den ersten wie auch auf den zweiten Blick – ein schwieriges Unterfangen zu sein. Verschärft findet dieses Problem sich dadurch, dass sich „eine zunehmende Relevanz von Forschungsdaten in geisteswissenschaftlichen Forschungsprozessen feststellen [lässt]“<sup>2</sup>, was sich auch darin äußert, dass „Geisteswissenschaftler\*innen mit einem Datenmanagementplan [...] konfrontiert [sind], bzw. [...] eine Forschungsdatenpolicy in der eigenen Institution veröffentlicht [wird]“<sup>3</sup>.

Eine Definition müsste die Grundlage der Sammlung, Aufbereitung und Darbietung dessen leisten, was tatsächlich unter diesem Begriff verstanden wird. Gerade für Wissenschaftler\*innen, die am Anfang eines längerfristigen Projekts mit der Erstellung ihres Datenmanagementplans beschäftigt sind, scheint sich daher oft die Frage zu stellen, „was denn nun in ihrem Fall eigentlich die Forschungsdaten sind, die sie sicher und langfristig verfügbar machen sollen“<sup>4</sup>. Die Informationsplattform „forschungsdaten.info“ hält in Bezug auf Forschungsdaten in den Geisteswissenschaften fest, dass „Heterogenität der materiellen und immateriellen Forschungsgegenstände und damit der im Forschungsprozess entstehenden Daten und deren Formate“<sup>5</sup> charakteristisch sind. Auch das Projekt „Forschungsdaten in den Geisteswissenschaften: Vergleichende Literaturwissenschaft und Sprachwissenschaft“ kommt zu einem ähnlichen Schluss, wenn die Autor\*innen feststellen: „Die geisteswissenschaftlichen Forschungsdaten sind besonders durch ihre Ungleichheit (Heterogenität) gekennzeichnet.“<sup>6</sup> Cremer geht noch einen Schritt weiter und konstatiert, dass „[d]as Label Forschungsdaten

---

<sup>2</sup> Fabian Cremer, Lisa Klaffki, und Timo Steyer, „Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften,“ *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 5, Nr. 2 (13. Juli 2018): 143, <https://doi.org/10.5282/o-bib/2018H2S142-162>.

<sup>3</sup> Susanne Blumesberger, „Forschungsdaten in den Geisteswissenschaften: Bereits selbstverständlich oder doch noch etwas exotisch?,“ *o-bib. Das offene Bibliotheksjournal / Herausgeber VDB* 8, Nr. 4 (16. November 2021): 2, <https://doi.org/10.5282/o-bib/5739>.

<sup>4</sup> Blumesberger, „Forschungsdaten in den Geisteswissenschaften,“ 1.

<sup>5</sup> „Geisteswissenschaften | Wissenschaftsbereiche | Forschungsdaten und Forschungsdatenmanagement,“ Informationsplattform [forschungsdaten.info](https://forschungsdaten.info), zugegriffen 5. Juli 2023, <https://forschungsdaten.info/wissenschaftsbereiche/geisteswissenschaften/>.

<sup>6</sup> Claudia Hackl et al., „Projektbericht: Forschungsdaten in den Geisteswissenschaften: Vergleichende Literaturwissenschaft und Sprachwissenschaft,“ zugegriffen 5. Juli 2023, <https://phaidra.univie.ac.at/detail/o:1622025>.

[...] aufgrund seiner Aktualität auf bereits etablierte Konzepte [...] übertragen oder inflationsartig für jede Art von Daten verwendet [wird].“<sup>7</sup>

Trotz und auch wegen dieser Kritik an einem weiten Konzept von Forschungsdaten in den Geisteswissenschaften scheint es ein angemessenerer Zugang zu sein, sich dem anzunähern, was unter Forschungsdaten in den Geisteswissenschaften verstanden werden kann, als im Voraus darauf zu beharren, was unter diesem Konzept verstanden werden muss. In weiterer Folge soll daher versucht werden, diese Annäherung in inhaltlicher wie struktureller Form zu leisten.

## 2.2. Inhaltliche Annäherung

Die Informationsplattform „forschungsdaten.info“ führt als Beispiele für mögliche Forschungsgegenstände, Daten und Formate „Kunst- und Kulturobjekte, Personen, Geodaten, digitalisierte Quellen und annotierte Texte“<sup>8</sup> an. Mit dem Verweis, dass folgende Beispiele von Forschungsdaten oft entsorgt werden, in Arbeitszimmern aufbewahrt werden und schließlich in Nachlässe übergangen<sup>9</sup>, listet Blumesberger: „Aufzeichnungen, Kopien, Scans, Bilder, Transkripte, Audio- oder Videoaufnahmen“<sup>10</sup>. Das Projekt „Forschungsdaten in den Geisteswissenschaften: Vergleichende Literaturwissenschaft und Sprachwissenschaft“ hält wiederum einerseits fest: „Geisteswissenschaftliche Forschungsdaten können sowohl analoger als auch digitaler Natur sein und in Form von Texten, Bildern, Audio-, Multimedia- oder 3D-Daten vorliegen.“<sup>11</sup> Andererseits weisen die Autor\*innen auch darauf hin, dass „geisteswissenschaftliche Forschungsdaten alles sein konnten, egal ob ein Digitalisat von einem alten Buch, Beobachtungsnotizen, Transkripte oder ein PDF-Dokument, etc.“<sup>12</sup> In der in diesem Projekt durchgeführten Online-Umfrage konnte von den befragten Personen in Bezug auf Forschungsdaten aus folgenden Optionen gewählt werden: Abbildungen von Objekten, Audio-Aufzeichnungen, Beobachtungen, Bibliographien, Bilder, Datenbanken, Experimente, fach- oder gerätespezifische Daten, Musiknoten, Logfiles und Nutzungsdaten, Simulationen, Statistik und Referenzdaten, Tabellen, Texte, Textdokumente, Transkriptionen, Umfragen und

---

<sup>7</sup> Cremer, „Chimäre auf der Spur“, 144.

<sup>8</sup> Informationsplattform forschungsdaten.info, „Geisteswissenschaften“.

<sup>9</sup> Vgl. Blumesberger, „Forschungsdaten in den Geisteswissenschaften“, 3.

<sup>10</sup> ebd., 3.

<sup>11</sup> Hackl et al., „Projektbericht: Forschungsdaten in den Geisteswissenschaften“, 11.

<sup>12</sup> ebd., 11.

Interviews, Video-Aufzeichnungen, Sonstiges.<sup>13</sup>

So könnte es bis hierhin fast scheinen, als wäre aus der Perspektive einer rein inhaltlichen Annäherung vieles, wenn nicht alles, als mögliche Forschungsdaten zu fassen, je nach Fassung des Forschungsprojekts und der den Begriff der Forschungsdaten definierenden Personen.

### 2.3. Strukturelle Annäherung

Wenn die Methodiken geisteswissenschaftlicher Forschung – Cremer/Klauffki/Steier nennen mit Bezug auf Unsworth hier: „Annotation, Vergleich, Referenz, Auswahl, Illustration und Repräsentation“<sup>14</sup> – als Einflussfaktor berücksichtigt werden, können in weiterer Folge strukturelle Annäherungen an den Begriff der Forschungsdaten in den Geisteswissenschaften vorgenommen werden.

Einer ersten Einfassung zufolge, die sich bei Blumesberger findet und sich auf die Definition des österreichischen Wissenschaftsfonds stützt, „sind im Datenmanagementplan vor allem jene Daten angesprochen, die einer Publikation zugrunde liegen.“<sup>15</sup> In Bezug auf die auch obig aufgezeigte Heterogenität von Forschungsdaten in den Geisteswissenschaften hält Cremer fest, dass mit einer „sehr weiten Definition“<sup>16</sup> gearbeitet werden muss, und führt als Beispiel die Definition im DARIAH-DE Working Paper zum Research Data LifeCycle an: „Unter digitalen geistes- und kulturwissenschaftlichen Forschungsdaten werden innerhalb von DARIAH-DE all jene Quellen/Materialien und Ergebnisse verstanden, die im Kontext einer geistes- und kulturwissenschaftlichen Forschungsfrage gesammelt, erzeugt, beschrieben und/oder ausgewertet werden und in maschinenlesbarer Form zum Zwecke der Archivierung, Zitierbarkeit und zur weiteren Verarbeitung aufbewahrt werden können“<sup>17</sup>. Auch Kindling und Schirnbacher definieren digitale Forschungsdaten als „alle digital vorliegenden Daten, die während des Forschungsprozesses entstehen oder ihr Ergebnis sind.“<sup>18</sup>

---

<sup>13</sup> Vgl. ebd., 14.

<sup>14</sup> Vgl. John Unsworth, *Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?*, Vortrag, Humanities Computing: Formal Methods, Experimental Practice, King's College London (13. Mai 2000), zitiert nach Cremer, „Chimäre auf der Spur“, 149.

<sup>15</sup> Blumesberger, „Forschungsdaten in den Geisteswissenschaften“, 2.

<sup>16</sup> Cremer, „Chimäre auf der Spur“, 147.

<sup>17</sup> Johanna Puhl et al., „Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften“, *DARIAH-DE Working Papers*, Nr. 11 (2015): 9, <http://nbn-resolving.de/%20urn:nbn:de:gbv:7-dariah-2015-4-4>, zitiert nach Cremer, „Chimäre auf der Spur“, 147.

<sup>18</sup> Maxi Kindling und Peter Schirnbacher, „Die digitale Forschungswelt' als Gegenstand der Forschung“, *Information - Wissenschaft & Praxis* 64, Nr. 2–3 (1. April 2013): 130, <https://doi.org/10.1515/iwp-2013-0017>.

Zwar haben weit gehaltene Definitionsansätze mit Sicherheit den Vorteil, dass Forschenden in gewisser Weise die Scheu vor dem Begriff „Forschungsdaten“ genommen wird. Schließlich wird der Begriff, wie das Projekt „Forschungsdaten in den Geisteswissenschaften: Vergleichende Literaturwissenschaft und Sprachwissenschaft“ ausweist, von Geisteswissenschaftler\*innen kaum verwendet.<sup>19</sup> Andererseits macht es diese Gemengelage schwer, Forschungsdaten von Forschungsergebnissen und Forschungsquellen zu unterscheiden, da die Demarkationslinien dieser Bereiche sich in obig angeführten Definitionen mitunter teils oder vollends verlieren. Mit Cremer gesprochen: „Während für die einen Forschungsdaten per se auch Forschungsergebnisse sind, sind Forschungsdaten für die anderen lediglich relevante, aber nicht publikationsfähige Nebenprodukte eines Forschungsvorhabens - ihre Aufwertung zu Forschungsdaten erscheint als Trugbild des digitalen Wandels.“<sup>20</sup>

Eine Schärfung der strukturellen Unterschiede zwischen diesen Bereichen ist auch Andorfer in seinem Paper „Forschungsdaten in den (digitalen) Geisteswissenschaften - Versuch einer Konkretisierung“ ein Anliegen. Hierbei nimmt er seinen Ausgang bei der „in der Literatur gerne anzutreffende[n] Differenzierung des Forschungsdatenbegriffs in Primär- oder Rohdaten und Sekundärdaten“<sup>21</sup>, wobei er dieser Aufgliederung attestiert, in Bezug auf die Geisteswissenschaften eher ungeeignet zu sein, da diese Begrifflichkeiten auf „einen vornehmlich in den Natur- und Sozialwissenschaften gebräuchlichen Forschungsdatenbegriff zurückzuführen sind“<sup>22</sup>. Als Alternative schlägt er folgende Definition von geisteswissenschaftlichen Forschungsdaten vor: „Geisteswissenschaftliche Forschungsdaten sind alle jene Daten, die im Kontext einer geisteswissenschaftlichen Fragestellung und in der Arbeit mit den dabei eingesehenen Quellen, worunter auch Sekundärliteratur verstanden wird, für eine längerfristige und öffentliche Archivierung ausgewählt und aufbereitet werden.“<sup>23</sup>

Zur Konkretisierung dieser Definition ist es nötig, die in Abbildung 9 jenes Papers dargestellte „Datenpyramide geisteswissenschaftlicher Forschungsdaten III im institutionellen Kontext“ näher zu beschreiben. An der Basis dieser dreistufigen

---

<sup>19</sup> Hackl et al., „Projektbericht: Forschungsdaten in den Geisteswissenschaften,“ 11.

<sup>20</sup> Cremer, „Chimäre auf der Spur,“ 147.

<sup>21</sup> Peter Andorfer, „Forschungsdaten in den (digitalen) Geisteswissenschaften,“ *DARIAH-DE Working Papers*, Nr. 14, (2015): 10, <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.

<sup>22</sup> ebd., 10.

<sup>23</sup> Andorfer, „Forschungsdaten Geisteswissenschaften,“ 14.

Pyramide kommen Quellen zu stehen. Darunter fallen Archivdokumente, Briefe, Sekundärliteratur, Kupferstiche, Ölbilder, Musiknoten und dergleichen mehr, egal ob analog, digital oder born digital. Diese Stufe ist mit Archiven, Bibliotheken, Museen, Repositorien verbunden. In der Mitte der Pyramide stehen die Arbeitsdaten. Darunter versteht Andorfer Bibliographien, Exzerpte, Notizen, Textentwürfe, Transkriptionen. Diese Arbeitsdaten werden auch als die tatsächlichen Forschungsdaten ausgewiesen, die Eingang in ein Repositoryum finden sollten. An der Spitze der Pyramide schließlich finden sich Publikationen, die ihren Weg in Bibliotheken finden sollten. Das sind nach Andorfer Monographien, Aufsätze, Kataloge, Quelleneditionen, Qualifizierungsarbeiten.<sup>24</sup> Ein Vorzug der obig formulierten Definition, die sich auf Andorfers Spielart der Datenpyramide stützt, ist, dass sie auf Begrifflichkeiten und Praktiken geisteswissenschaftlichen Arbeitens Bezug nimmt.<sup>25</sup>

Wenn Andorfers Definition und Datenpyramide sich auch durch strukturelle Klarheit auszeichnet, bleibt fraglich, ob diese Klarheit sich nicht zu einem Korsett auswächst, welches der Realität des Forschungsalltags die Luft abschnürt. Weiters bleibt fraglich, inwiefern diese Definition in der Lage ist, allen denkbaren geisteswissenschaftlichen Forschungsprojekten und den dabei anfallenden Daten definitorisch Genüge zu tun. Somit dürfte auch die Definition geisteswissenschaftlicher Forschungsdaten sich in die vielen Definitionshistorien eingeschriebene Bewegung von Weitung und Verengung fügen, die eine situationsadäquate Verwendung auch und erst durch die Gegebenheiten und den Gebrauch in der Praxis zu rechtfertigen in der Lage ist.

#### 2.4. Definition für vorliegendes Projekt

Der für die Durchführung dieses Projekts verwendete Forschungsdatenbegriff bildet sich sowohl anhand struktureller als auch anhand inhaltlicher Komponenten.

Strukturell wird – in Anlehnung an Andorfers Spielart der Datenpyramide – davon ausgegangen, dass Forschungsdaten sich im Feld zwischen Quelldaten und Ergebnisdaten verorten lassen. Forschungsdaten sind im Sinne unserer Definition Produkte, die im Zuge eines Forschungsprojekts entstanden sind, sich von Ergebnisdaten (bspw. abgeschlossenen und publizierten Aufsätzen) und Quelldaten (bspw. kinder- und jugendliterarische (Primär-)Texte) unterscheiden, und die

---

<sup>24</sup> ebd., 14.

<sup>25</sup> Vgl. ebd., 6.

potentiell für andere Forschungsprozesse, zur Erreichung anderer Forschungsergebnisse (nach-)genutzt werden könnten. Das wichtige Kriterium in diesem Zusammenhang ist somit ein strukturelles und kann mit folgender Frage grob umrissen werden: Ist das in Frage stehende Material im Zuge eines Forschungsprozesses mit dem Ziel, Ergebnisdaten zu generieren, und – unter Umständen – auf der Grundlage von Quelldaten entstanden?

In Bezug auf die inhaltliche Komponente unseres Forschungsdatenbegriffs entfernen wir uns ein Stück weit von Andorfers Spielart der Datenpyramide, da diese, unserem Erachten nach, die medialen Ausformungen möglicher Forschungsdaten im Feld der Kinder- und Jugendliteraturforschung zu sehr beschneiden würde. Daher bestimmen wir Forschungsdaten für vorliegendes Projekt unter inhaltlicher Perspektive als: Bibliographien, Exzerpte, Notizen, Textentwürfe, Transkriptionen, Abbildungen/Bilder/Fotos, Audioaufzeichnungen, Experimente/Protokolle, Musiknoten, Logfiles und Nutzungsdaten, Statistiken und Referenzdaten, Tabellen, Umfragen, Interviews, Videos. Wichtig in diesem Zusammenhang bleibt jedoch die Kontrolle der inhaltlichen Bestimmung durch die strukturelle.

Wenn auch festzuhalten bleibt, dass die Unterscheidung zwischen Ergebnisdaten, Forschungsdaten und Quelldaten nicht immer absolut trennscharf aufrechtzuerhalten ist, dass sich gerade Quelldaten und Forschungsdaten mitunter überschneiden können, soll eine dergestalt geartete Definition von Forschungsdaten verhindern, dass die Grenzen vollends aufgelöst werden. Schließlich verweist auch Andorfer in seinem Paper auf diese Problematik, wenn er Sahle und Kronenwett zitiert: „Durch die Digitalisierung des Forschungsprozesses verschmelzen die verschiedenen Arten von Forschungsdaten zu einem Kontinuum, das von den Ausgangsdaten bis zu den Narrativen der Ergebnisse der Forschung alle Schritte der Verarbeitung enthält.“<sup>26</sup> Durch die Schärfung und Rechtfertigung unseres Begriffs von Forschungsdaten hoffen wir, dieses Kontinuum zumindest teilweise mit Trennlinien durchziehen zu können.

## 2.5. Forschungsdaten finden

Die grundlegende Fragestellung bei der Zusammenstellung einer Collection frei

---

<sup>26</sup> Simone Kronenwett und Patrick Sahle, „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‘Data Center for the Humanities’,“ *Library ideas* 9, Nr. 2 (2013): 76–96. zitiert nach Andorfer, „Forschungsdaten Geisteswissenschaften,“ 14.

verfügbarer Forschungsdaten im Bereich der Kinder- und Jugendliteraturforschung besteht darin, herauszufinden, wo Forschungsdaten zu finden sind. Eine gute Ausgangsbasis zur Orientierung bei der Recherche nach Forschungsdaten sind die Informationsseiten universitätsinterner Datenmanagementzentren wie zum Beispiel an der WWU Münster<sup>27</sup>, FU Berlin<sup>28</sup>, HU Berlin<sup>29</sup>, RWTH Aachen<sup>30</sup>, TU Darmstadt<sup>31</sup> und Universität Göttingen<sup>32</sup>. Diese Seiten bieten Informationen zu verschiedenen Recherchestrategien sowie Vorschläge für potenzielle Datenquellen.

Ein Vergleich der Empfehlungen der genannten Datenmanagementzentren legt nahe, dass die Suche in Repositorien als Hauptinstrument für die Recherche von Forschungsdaten angesehen wird. Darüber hinaus wird empfohlen, die Suche auf Datensuchmaschinen und Metasuchportale auszuweiten, die verschiedene Datenquellen durchsuchen können. Es ist jedoch wichtig anzumerken, dass die Unterscheidung zwischen Repository, Suchdienst und Portal nicht überall eindeutig oder gleichwertig gehandhabt wird. Zudem sollte beachtet werden, dass die Bereitstellung von Forschungsdaten bei verschiedenen Dienstleistern nicht einheitlich erfolgt, sodass die Recherche entsprechend dem jeweiligen Repository oder der Suchmaschine angepasst werden muss.

Vor Beginn der Recherche ist es ratsam, zu überlegen, welche Repositorien oder Suchmaschinen am besten für die gesuchten Daten geeignet sind. Auf den Informationsseiten wird häufig empfohlen, die Recherche zunächst in einem spezialisierten Fachrepository zu beginnen, das sich auf die Sammlung von Daten aus dem entsprechenden Forschungsbereich konzentriert. Um ein passendes Fachrepository für die gesuchten Forschungsdaten zu finden, wird oft re3data.org<sup>33</sup> genannt. Diese Website bietet ein umfangreiches Register an

---

<sup>27</sup> Vgl. „Forschungsdaten – Daten finden,“ Universität Münster, zugegriffen 26. Juni 2023, <https://www.uni-muenster.de/Forschungsdaten/nachnutzen/finden/>.

<sup>28</sup> Vgl. „Forschungsdaten recherchieren und finden,“ FU Berlin, zugegriffen 26. Juni 2023, <https://www.fu-berlin.de/sites/forschungsdatenmanagement/in-der-praxis/planung/nachnutzen.html>.

<sup>29</sup> Vgl. „Forschungsdaten finden,“ Humboldt-Universität zu Berlin, zugegriffen 26. Juni 2023, <https://www.cms.hu-berlin.de/de/dl/dataman/arbeiten/finden>.

<sup>30</sup> Vgl. „FDM erklärt – Wie finde ich eigentlich Forschungsdaten? - Forschungsdaten – Aktuelles und Wissenswertes,“ RWTH Aachen, zugegriffen 26. Juni 2023, <https://blog.rwth-aachen.de/forschungsdaten/2022/11/04/fdm-erklart-wie-finde-ich-eigentlich-forschungsdaten/>.

<sup>31</sup> Vgl. „Forschungsdaten finden und zitieren,“ TU Darmstadt, zugegriffen 26. Juni 2023, [https://www.tu-darmstadt.de/tudata/datenmanagement/forschungsdaten\\_finden\\_und\\_zitieren\\_tudata/index.de.jsp](https://www.tu-darmstadt.de/tudata/datenmanagement/forschungsdaten_finden_und_zitieren_tudata/index.de.jsp).

<sup>32</sup> Vgl. „Forschungsdaten und Datenbanken,“ Universität Göttingen, zugegriffen 7. Juli 2023, <https://www.sub.uni-goettingen.de/fachinformation/geisteswissenschaften-und-theologie/romanistik/forschungsdaten-und-datenbanken/>.

<sup>33</sup> <https://www.re3data.org/>, zugegriffen 25. Juni 2023.

Repositorien mit einer integrierten Suchfunktion nach Fachbereich.

Als nächster Schritt kann die Suche auf ein fachübergreifendes Repositorium ausgedehnt werden, das Daten aus verschiedenen Fachgebieten sammelt. Eine nächste Möglichkeit bieten Metasuchportale oder Datensuchmaschinen, die Nachweise für vorhandene Forschungsdaten aggregieren. Dadurch können umfassendere Suchergebnisse erzielt werden.

Eine weitere häufig empfohlene Datenquelle sind Datenjournale. Datenjournale veröffentlichen und beschreiben Forschungsdaten, tun das jedoch, ohne diese zu interpretieren, wie es in wissenschaftlichen Artikeln üblich ist. Oft werden die besprochenen Datensätze in separaten Repositorien abgelegt und mit dem Artikel verlinkt, da es sich meist um umfangreiche Datenmengen handelt. Datenjournale bieten den Vorteil, dass sie durch Peer-Reviews geprüfte Daten für die Nachnutzung bereitstellen und somit das qualitative Datenmanagement wesentlich unterstützen.<sup>34</sup> Die Anzahl der Datenjournale ist derzeit allerdings noch begrenzt, und insbesondere in den Geisteswissenschaften finden sich nur wenige etablierte.<sup>35</sup>

Eine weitere Recherchestrategie besteht darin, aktuelle Literatur des Forschungsbereichs nach Forschungsdaten zu durchsuchen. Beim Lesen relevanter Fachliteratur ist es sinnvoll, darauf zu achten, auf welche Forschungsdaten der Text sich stützt und anschließend nach diesen Daten zu recherchieren oder die Autor\*innen um weitere Informationen zu bitten. In einigen Fällen bieten Publikationen Verweise auf weiterführende oder nachnutzbare Daten.<sup>36</sup> Ebenso wird zu manchen Publikationen „supplementary material“<sup>37</sup> oder auch „supplementary data“ veröffentlicht, ergänzendes Material, das nicht Teil der Hauptpublikation ist, aber relevante zusätzliche Forschungsdaten darstellt.

Interessanterweise wird der persönliche Austausch im eigenen Netzwerk als Methode zur Forschungsdatenrecherche seltener empfohlen. Dies könnte darauf zurückzuführen sein, dass der persönliche Austausch oft als selbstverständliche und intuitive Praxis in der wissenschaftlichen Arbeit angesehen wird. Dennoch sollte die Bedeutung des eigenen Netzwerks bei der Suche nach Forschungsdaten nicht

---

<sup>34</sup> Vgl. „Datenjournale: Peer-Review-Publikationen über Daten,“ Informationsplattform forschungsdaten.info, Zugriff 5. Juli 2023, <https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/datenjournale/>.

<sup>35</sup> Siehe Liste: [https://www.forschungsdaten.org/index.php/Data\\_Journals](https://www.forschungsdaten.org/index.php/Data_Journals), Zugriff 5. Juli 2023.

<sup>36</sup> Vgl. „Finding research data - WUR,“ Universität Wageningen, Zugriff 26. Juni 2023, <https://www.wur.nl/en/library/researchers/finding-sources/finding-research-data.htm>.

<sup>37</sup> Beispielsweise das Caption Electronic Supplementary Material zu dem Artikel »letzte Politur«? <https://doi-org.uaccess.univie.ac.at/10.1007/s41245-020-00120-4>.

unterschätzt werden. Friedrich und Recker betonen in ihrem Kapitel zur Auffindbarkeit und Nutzbarkeit von Forschungsdaten, dass der Dialog mit Kolleg\*innen und anderen Forschenden nach wie vor eine der wichtigsten Strategien für die Forschungsdatenrecherche ist,<sup>38</sup> sie zitieren in diesem Kontext eine Studie,<sup>39</sup> die persönliche Kommunikation mit anderen Forscher\*innen und das eigene Forschungsnetzwerk als wichtigen Aspekt bei der Auffindung von Forschungsdaten ausweist. Ebenso unterstreicht das Datamanagementzentrum der Universität Wageningen die Aktivierung des eigenen Netzwerks als einen zentralen Punkt bei der Beschaffung von Forschungsdaten.<sup>40</sup>

---

<sup>38</sup> Vgl. Tanja Friedrich und Jonas Recker, „Auffindbarkeit und Nutzbarkeit von Daten,“ in *Praxishandbuch Forschungsdatenmanagement* (Berlin ; Boston: De Gruyter Saur, 2021), 411, <https://doi.org/10.1515/9783110657807-023>.

<sup>39</sup> Vgl. Ayoun Yoon, „Role of Communication in Data Reuse,“ *Proceedings of the ASIST Annual Meeting* 54, Nr. 1 (2017): 469, <https://doi.org/10.1002/pra2.2017.14505401050>., zitiert nach Friedrich und Recker, „Auffindbarkeit,“ 411.

<sup>40</sup> Vgl. Universität Wageningen, „Finding research data“.

### **3. Recherche**

#### **3.1. Recherchestrategie**

Basierend auf den verschiedenen möglichen Methoden der Forschungsdatenrecherche ergibt sich die im Folgenden ausgeführte Recherchestrategie. Im Anschluss wurde unter Ausführung der jeweiligen Recherchestrategie anhand von Exceltabellen Rechercheprotokolle erstellt, welche anschließend zusammengefasst und analysiert werden.

##### **3.1.1. Recherche in Repositorien und Metasuchportalen**

Für die Recherche in Repositorien wurde eine Schlagwortliste anhand von Synonymen und verwandten Begriffen der Begriffe „Jugendliteratur“, „Kinderliteratur“, „Jugendmedien“ und „Kindermedien“ sowie Motiven und Themen aus der Kinder- und Jugendliteraturforschung erstellt.

Im nächsten Schritt wurde eine Liste von geeigneten Repositorien und Metasuchportalen zusammengestellt. Dazu wurden fachspezifische Repositorien für geisteswissenschaftliche und germanistische Forschung, fachübergreifende Repositorien und Metasuchportalen sowie institutsspezifische Repositorien von Universitäten, die über Institute für Kinder- und Jugendliteraturforschung verfügen, zugezogen. Anhand der unterschiedlichen Suchfunktionen der ausgewählten Datenquellen wurde eine Verfeinerung der Recherche nach Forschungsdaten versucht und entsprechende Filterfunktionen ausgewählt, welche gezielt nach diesen oder passenden Medientypen suchen. Hierbei war das Ziel nicht, nach textbasierten Ergebnisdaten bzw. abgeschlossenen Publikationen zu suchen, sondern nach anderen Medientypen, die Forschungsdaten in inhaltlicher Annäherung im Sinne dieses Projekts entsprechen könnten. Diese Ausweitung des Suchbereichs wurde getätigt, da Forschungsdaten nicht immer klar als solche gekennzeichnet sind und oft auch erst als solche erkannt werden müssen. Sie tauchen mitunter in Form von Anhängen, nicht als Forschungsdaten klassifizierten Daten oder nicht textbasierten Datentypen (z.B. die Aufnahme eines Interviews) auf. Ebenso wurden Schlagwörter und Schlagwortkombinationen den jeweiligen Suchfunktionen angepasst. Folgende Repositorien und Metasuchportale wurden ausgewählt:

- **TextGrid Repository**<sup>41</sup> ist ein Repositorium für geisteswissenschaftliche Forschungsdaten, vor allem für Texte und digitale Editionen aus der deutschen Literaturwissenschaft, welche durchsuchbar gemacht und mit Metadaten versehen wurden. Eine Vielzahl der vorhandenen Texte sind im XML/TEI-Format und daher leicht nachnutzbar. Das Repositorium archiviert auch Bilder. Bei der Recherche im TextGrid Repository wurde auf eine Schlagwortrecherche verzichtet, da die archivierten Daten nicht mit inhaltlichen Schlagwörtern versehen sind. Stattdessen wurde durch die verschiedenen archivierten Projekte geblättert.
- **DARIAH-DE Repository**<sup>42</sup> ist ein Repositorium für die Speicherung von Forschungsdaten aus den Geisteswissenschaften. Die Suchfunktion des Repositoriums durchsucht Kollektionen sowie einzelne Dokumente. Die Recherche wurde anhand von Schlagwörtern durchgeführt.
- **FID Germanistik (GiNDok Repository)**<sup>43</sup> ist ein Fachrepositorium für Germanistik. Es ermöglicht, verschiedene Datenquellen nach diversen Medientypen (Monographien, Sammelbände, Zeitschriften, Zeitschriftenartikel, Online-Dokumente, Datenbanken) zu durchsuchen. Darüber hinaus bietet es die Möglichkeit zur differenzierten Suche zwischen „Veröffentlichungen“ und „Forschungsdaten“. Diese Suchfunktion erlaubt ebenfalls eine Verfeinerung der Ergebnisse innerhalb des Suchbereichs „Veröffentlichungen“, indem weitere Kriterien wie Medientyp oder Sprache angewendet werden können. Für diese Recherche wurde der Suchbereich „Forschungsdaten“ und „Veröffentlichungen“ ausgewählt. Zusätzlich wurde auch nach „Veröffentlichungen“ gesucht, die nicht eindeutig einem bestimmten Medientyp zugeordnet sind, sondern in die Kategorie „Unbestimmt“, „Weitere“, „Multimedial“ oder „Teil eines Periodikums“ fallen. Für die Suche wurden Schlagwörter sowie verschiedene Schlagwortkombinationen benutzt.

---

<sup>41</sup> <https://textgridrep.org/>, zugegriffen 5. Juli 2023.

<sup>42</sup> <https://repository.de.dariah.eu/>, zugegriffen 5. Juli 2023.

<sup>43</sup> <https://www.germanistik-im-netz.de/publizieren/gindok-repositorium/>, zugegriffen 5. Juli 2023.

- **colibri**<sup>44</sup> ist eine digitale Sammlung deutschsprachiger historischer Kinder- und Jugendliteratur. Für die Recherche in dieser Datenbank wurden keine Schlagwörter genutzt, sondern nur ein Überblick über die vorhandenen Daten geschaffen.
- **Zenodo**<sup>45</sup> ist ein fachübergreifendes Repositorium für alle Forschungsdisziplinen und Datentypen. Für die Suche wurde der Filter *Access Right* „Open“ gewählt, während verschiedene Schlagwörter und Schlagwortkombinationen verwendet wurden. Zenodo bietet zwar keine erweiterten Suchfunktionen oder direkte Einschränkungsmöglichkeiten für Forschungsdaten, jedoch können Ergebnisse verfeinert werden, indem man nach Dateityp und Dateiformat filtert. Mithilfe dieser Filtermöglichkeiten kann beispielsweise nach „Dataset“, „Image“, „Video“ oder „XML-Datei“ gefiltert werden. Eine weitere nützliche Möglichkeit für die Recherche nach Forschungsdaten ist die Überprüfung der Metadaten von Dateien nach Einträgen in dem Feld „Related Identifiers: Supplementary material“. In diesem Feld können zusätzliche Materialien mit ergänzenden Dateien verknüpft werden, sodass Suchende auf weiterführende Daten stoßen können. Daher ist es empfehlenswert, bei der Suche nach Forschungsdaten in Zenodo auch relevante Zeitschriftenartikel und Konferenzbeiträge zu überprüfen, um zu sehen, ob sie ergänzendes Material enthalten, anstatt sich lediglich auf die Suchfunktion zu verlassen. Dies ermöglicht eine umfassendere und zielgerichtete Recherche nach relevanten Forschungsdaten und wurde auch bei dieser Recherche versucht.
- **Europeana**<sup>46</sup> ist eine virtuelle Bibliothek, welche europäisches Kunst- und Kulturgut in Form von Bild-, Text-, Ton- und Video-Dateien aggregiert. Bei der Recherche wurden Schlagwörter und Schlagwortkombinationen benutzt.
- **BASE**<sup>47</sup> ist eine Metasuchportal, welches Metadaten von Forschungsdaten

---

<sup>44</sup> <https://colibri-portal.eu/>, zugegriffen 5. Juli 2023.

<sup>45</sup> <https://zenodo.org/>, zugegriffen 5. Juli 2023.

<sup>46</sup> <https://www.europeana.eu/de>, zugegriffen 5. Juli 2023.

<sup>47</sup> <https://www.base-search.net/?l=de>, zugegriffen 5. Juli 2023.

und Publikationen aller Art aggregiert. Die Metasuchmaschine bietet eine Filtermöglichkeit nach dem Dokumententyp „Forschungsdaten“ an. Für die Recherche anhand von Schlagwörtern wurden zusätzliche die Dokumententypen „Andere Beitrag in Zeitschrift/Zeitung“, „Karte“, „Audio“, „Bild/Video“, „Software“ zugezogen.

- **OpenAIRE Explore**<sup>48</sup> aggregiert den Zugang zu einem breiten Spektrum wissenschaftlicher Outputs (Veröffentlichungen, Forschungsdaten, Software) und verknüpft diese miteinander sowie mit zusätzlichen wissenschaftlichen Rahmeninformationen wie wissenschaftlichen Organisationen, Projekten und Fördergebern. Aus diesem Grund ist dieses Portal eine interessante Option für eine Forschungsdatenrecherche, da man beispielweise bei der Recherche nach Publikationen auch zu verknüpften Forschungsdaten weitergeleitet werden könnte und speziell nach Forschungsdaten filtern kann. Interessant in diesem Kontext ist ebenfalls, dass man über die Suche nach Projekten nach Datenmanagementplänen recherchieren kann.  
Bei der Suche nach Forschungsdaten aus dem Bereich der Kinder- und Jugendliteratur wurden Filter für die Begriffe „Research Data“ und „Project“ verwendet, während verschiedene Schlagwörter und Schlagwortkombinationen eingesetzt wurden.
- **PHAIDRA**<sup>49</sup> ist das Repositorium der Universität Wien, die auch über ein Institut für Kinder- und Jugendliteraturforschung verfügt. Das Repositorium bietet eine Archivierungsmöglichkeit für diverse Datentypen sowie die Erstellung von Sammlungen und Containern an, welche eine interessante Möglichkeit sind, Forschungsdaten zu bündeln. Für die Recherche mit Schlagwörtern wurden die Datentypen „Bild“, „Collection“, „Video“, „Daten“, „Resource“, „Audio“, „Container“ ausgewählt.
- **KUPS Repository**<sup>50</sup> ist der Publikationsserver der Universität Köln, wo eine Arbeitsstelle für Kinder- und Jugendmedienforschung lokalisiert ist. Das

---

<sup>48</sup> <https://explore.openaire.eu/>, zugegriffen 5. Juli 2023.

<sup>49</sup> <https://phaidra.univie.ac.at/>, zugegriffen 5. Juli 2023.

<sup>50</sup> <https://kups.ub.uni-koeln.de/>, zugegriffen 5. Juli 2023.

Repository wird hauptsächlich für Hochschulschriften sowie wissenschaftliche Publikationen der Mitarbeiter\*innen der Universität benutzt. Es bietet jedoch die Filterfunktion „Sonstiges“ an, um nach Ergänzungsmaterial zu Publikationen und Ähnlichem zu suchen. Die Recherche wurde mit diesem Filter anhand von Schlagwörtern durchgeführt.

- **Publikationsserver der UB Frankfurt**<sup>51</sup> ist ein Repository, dessen Universität ein Institut für Jugendbuchforschung hat. Auch hier liegt der Fokus auf wissenschaftlichen Veröffentlichungen mit der Option einer Filterfunktion für „Sonstiges“. Hier wurde ebenfalls mit Schlagwörtern und dem Filter gesucht.

### 3.1.2. Sichtung von Publikationen

Zum Zweck der Recherche in facheinschlägigen Publikationen wurden zuerst drei (Open Access-)Periodika ausgewählt:

- **libri liberorum**. Zeitschrift der Österreichischen Gesellschaft für Kinder- und Jugendliteraturforschung<sup>52</sup>
- **Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung**<sup>53</sup>
- **Kids + Media**. Zeitschrift für Kinder- und Jugendmedienforschung<sup>54</sup>

Von diesen Periodika wurden jeweils die aktuellsten PDF-Veröffentlichungen herangezogen. Im Fall von „libri liberorum“ und „Kids + Media“ waren das jeweils sieben Dokumente (von der aktuellen Ausgabe bis zurück ins Jahr 2018), im Fall des „Jahrbuchs der Gesellschaft für Kinder- und Jugendliteraturforschung“ betraf das sechs Dokumente (zurückreichend bis ins Jahr 2017).

Der Fokus der Recherche lag darauf, Forschungsdaten ausfindig zu machen, die bei der Entstehung des jeweiligen Beitrags als inhaltliche Grundlage oder Ergänzung gedient haben könnten. Das Ziel war es, Speicherorte zu finden, an denen mit den Texten verknüpfte Forschungsdaten möglicherweise abgelegt wurden. Falls keine Forschungsdaten ausgewiesen waren, wurde versucht, nachzurecherchieren und

---

<sup>51</sup> <https://publikationen.ub.uni-frankfurt.de/home>, zugegriffen 5. Juli 2023.

<sup>52</sup> <https://oegkjlif.univie.ac.at/publikationen/libri-liberorum/>, zugegriffen 25. Juli 2023.

<sup>53</sup> <https://ojs.ub.uni-frankfurt.de/gkif/index.php/jahrbuch/index>, zugegriffen 25. Juli 2023.

<sup>54</sup> <https://www.kids-media.uzh.ch/issue/archive>, zugegriffen 25. Juli 2023.

vorhandenes Material, das den Beitrag begleitet und potentiell in weiterem Sinne als Forschungsdaten verstanden werden könnte, inhaltlich-medial zu bestimmen.

Die Recherche fand in einem ersten Schritt manuell statt. Die Abstracts sämtlicher Beiträge wurden quergelesen, alle Fußnoten und Quellenverzeichnisse analysiert und nach Hinweisen auf Forschungsdaten durchsucht. Schließlich wurde allen relevant wirkenden Links gefolgt und deren Umfeld weiter untersucht. Auch wurden nach Möglichkeit die Schlagwörter der relevanten Beiträge festgehalten, um Ansatzpunkte für mögliche weitere Recherchen herzustellen.

In einem zweiten Schritt – der dieser Arbeit als Anhang beigefügt ist – wurden (teil-)automatisiert einige textanalytische Versuche unternommen, um potentiell hilfreiche Applikationen und sich daran anknüpfende Vorgehensweisen auf ihre Tauglichkeit zu testen.

### 3.1.3. Sichtung von Institutsseiten und Portalen

Als zusätzlicher Pfad der Recherche nach frei verfügbaren Forschungsdaten aus der Kinder- und Jugendliteraturforschung wurde eine Reihe an facheinschlägigen Webseiten im Internet ausgewählt und untersucht. Die Beschaffung von Forschungsdaten über Institutsseiten und Portale erscheint nicht als der unmittelbarste Zugangsweg zu Forschungsdaten. Diese Strategie wurde von der Idee inspiriert, dass Institutsseiten oft detailliertere Informationen über geförderte Projekte oder ihre eigene Datenmanagementpraxis bereitstellen, die über das Angebot von Repositorien hinausgehen. Diese Informationen ermöglichen es, Schlussfolgerungen hinsichtlich der Archivierungsorte und des Umgangs mit Forschungsdaten zu ziehen. Ebenfalls ein Ziel war es, durch eine Auswahl von Webseiten mit starkem Bezug zu diesem Bereich der Geisteswissenschaften einen Überblick darüber zu liefern, was einem in der Forschung bei einer solchen Recherche begegnen mag.

In einem ersten Schritt wurde eine oberflächliche Suche gestartet, um die Homepages diverser Institute mit Bezug zur Kinder- und Jugendliteraturforschung aufzulisten. Konkret wurden mitunter folgende Institutsseiten als repräsentatives Muster gewählt:

- **KinderundJugendmedien.de**<sup>55</sup> – das Wissenschaftliche Portal für Kindermedien und Jugendmedien
- **ALEKI**<sup>56</sup> – Arbeitsstelle für Kinder- und Jugendmedienforschung
- **Blumesberger.at**<sup>57</sup> – eine Sammlung für Kinder- und Jugendliteraturforschung
- **Siegener Werkstattgespräche mit Kinderbuchautorinnen und -autoren**<sup>58</sup>
- **Arbeitsstelle Kinder- und Jugendliteratur** – Universität Göttingen<sup>59</sup>
- **Institut für Jugendbuchforschung** – Universität Frankfurt<sup>60</sup>
- **Oldenburger Forschungsstelle Kinder- und Jugendliteratur** – Universität Oldenburg<sup>61</sup>
- **KiLiMM** – Datenbank zur Kinder- und Jugendliteraturforschung und -didaktik sowie Medienforschung und Mediendidaktik – Universität Bielefeld<sup>62</sup>
- **Schweizerisches Institut für Kinder- und Jugendmedien**<sup>63</sup>

Nachdem ein Überblick aufgebaut war, wurde in einem zweiten Arbeitsschritt eine genaue Durchforstung der in Frage stehenden Webseiten getätigt. Dabei wurde jede Seite gründlich gelesen, alle weiterführenden Links besucht und nicht nur auf deren Nützlichkeit überprüft, sondern grundlegend auch deren Aktualität und bestehende Existenz getestet. Darüber hinaus wurde (wo vorhanden) die gesamte Sitemap der Webseite ebenfalls auf Hinweise auf Forschungsdaten durchsucht.

### 3.2. Herausforderungen

Wie bereits unter dem Aspekt der Auffindbarkeit von Forschungsdaten erwähnt, werden viele Forschungsdaten persönlich weitergegeben oder ergeben sich aus Begegnungen bei Konferenzen oder anderer universitärer Kooperationen. Insbesondere in einem kleineren Forschungsbereich wie der Kinder- und Jugendliteraturforschung ist anzunehmen, dass die interne Vernetzung engmaschig

---

<sup>55</sup> <https://www.kinderundjugendmedien.de/>, zugegriffen 09. August 2023.

<sup>56</sup> <https://aleki.uni-koeln.de/index.php?id=26793>, zugegriffen 09. August 2023.

<sup>57</sup> <http://www.blumesberger.at/index.php?id=14>, zugegriffen 09. August 2023.

<sup>58</sup> [https://www.universi.uni-siegen.de/katalog/reihen/siegener\\_werkstattgespraeche/?lang=de](https://www.universi.uni-siegen.de/katalog/reihen/siegener_werkstattgespraeche/?lang=de), zugegriffen 09. August 2023.

<sup>59</sup> <https://www.uni-goettingen.de/de/197955.html>, zugegriffen 09. August 2023.

<sup>60</sup> <https://www.uni-frankfurt.de/50701264/jugendbuchforschung>, zugegriffen 09. August 2023.

<sup>61</sup> <https://uol.de/germanistik/kinder-und-jugendliteraturwissenschaft>, zugegriffen 09. August 2023.

<sup>62</sup> <http://bibadmin.ub.uni-bielefeld.de/index.php>, zugegriffen 09. August 2023.

<sup>63</sup> <https://sikjm.ch/>, zugegriffen 09. August 2023.

ist und der Informationsaustausch regelmäßig und oft informell erfolgt. Daher ergibt sich die Schwierigkeit, möglicherweise ohne Kontakte oder direkter Ansprache von Forschenden nur einen beschränkten Zugang zu bestehenden Forschungsdaten zu haben.

Dieser Aspekt kann auch durch die mangelnde Vertrautheit mit dem Forschungsfeld ergänzt werden. Eine wesentliche Voraussetzung für eine erfolgreiche Recherche nach Forschungsdaten besteht in der Kenntnis der aktuellen Trends und wissenschaftlichen Praktiken sowie der historischen Entwicklung des Forschungsbereichs. Im Laufe eines Studiums und einer fortgeschrittenen Forschungstätigkeit entwickelt man eine Vertrautheit mit etablierten Forschungsinstituten, den handelnden Personen, wichtigen Informationsquellen und relevanter Literatur. Dieses Wissen ist von entscheidender Bedeutung, um die Relevanz von Datenquellen einschätzen zu können und den Entstehungskontext sowie den potenziellen Wert der Daten für die Wiederverwendung zu erkennen. Ohne diese Erfahrung könnte es herausfordernd sein, relevante Forschungsdaten zu identifizieren und qualitativ hochwertige Quellen zu erkennen. Eine mögliche Lösung besteht darin, den Recherchebereich möglichst breit anzulegen und den interessierten Personen die Bewertung der Daten selbst zu überlassen. Bei Anwendung dieser Recherchestrategie stellt sich jedoch die Frage nach der Nützlichkeit zufallsbedingter Sammlungen an Forschungsdaten, da frei zugängliche Forschungsdaten meist als Ergänzung oder Startpunkt für ein konkretes Forschungsvorhaben genutzt werden.

Ähnlich könnte die Anwendungen von wenig konkreten Suchbegriffen aufgrund begrenzter Kenntnis der aktuellen Forschungsthemen der KJL-Forschung sowie einem zu wenig präzisen Rechercheziel zu schlechten Ergebnissen führen. Für eine produktive Forschungsdatenrecherche ist sinnvoll, sehr konkrete Suchbegriffe zu wählen, da erfahrungsgemäß Forschungsdaten nicht annähernd so ausführlich beschlagwortet sind wie Publikationen. Um diesem Umstand entgegenzuwirken und bessere Ergebnisse zu erzielen, kann die Recherche durch die Verwendung von Schlagwörtern aus spezifischen Themenpools der Kinder- und Jugendliteratur (KJL) optimiert werden.<sup>64</sup>

---

<sup>64</sup> Zum Beispiel von <https://www.jugendliteratur.at/bibliothek/datenbank-fachliteratur> oder <https://www.kinderundjugendmedien.de/index.php/begriffe-und-termini/2270-zeitgeschichtliche-kinder-und-jugendliteratur>.

Eine weitere potenzielle Herausforderung besteht darin, dass eine gute Praxis für den Umgang mit Forschungsdaten noch nicht in allen wissenschaftlichen Bereichen etabliert ist oder sich generell noch in der Entwicklung befindet. Wenn Forschungsdaten nicht ausreichend mit Metadaten angereichert oder mit einem Persistent Identifier versehen werden, kann es äußerst schwierig sein, sie zu finden. Forschungsdaten werden in verschiedenen Umgebungen gespeichert und sind daher nicht immer leicht zugänglich. Die Suche nach Forschungsdaten erfordert häufig einen erheblichen Aufwand, da diese auf unterschiedlichste Weisen veröffentlicht werden und daher weit verstreut sind. Vor Beginn der Recherche war es nicht möglich, zu beurteilen, inwieweit Forschungsdatenmanagement in der Kinder- und Jugendliteraturforschung etabliert ist und praktisch angewendet wird. Sollte die gängige Praxis noch nicht weit entwickelt sein, könnte dies zu Schwierigkeiten und begrenzten Ergebnissen bei der Recherche führen.

### 3.3. Rechercheergebnisse

#### 3.3.1. Recherche in Repositorien und Metasuchportalen

- **TextGrid Repository**

Obwohl sich TextGrid Repository selbst als Repository für die Geisteswissenschaften definiert, konnten hier keine Forschungsdaten im Sinne der projekteigenen Definition gefunden werden. Damit ist gemeint, dass die archivierten Daten zwar im Kontext eines Forschungsprozesses entstanden sind, in diesem Fall der Erstellung von digitalen Editionen, jedoch dieser nicht darauf abzielte, Ergebnisdaten zu generieren, sondern lediglich darauf, Primärquellen (-daten) verfügbarer zu machen. Hier ist trotzdem anzumerken, dass die Daten im TextGrid Repository eine wertvolle Quelle für Kinder- und Jugendliteraturforschung darstellen können, speziell wenn es um textanalytische Forschung geht. Beispielweise enthält das Repository eine umfassende Märchensammlung.

- **DARIAH-DE Repository**

Die Recherche in diesem Repository ergab keine Forschungsdaten aus dem Bereich der Kinder- und Jugendliteraturforschung.

- **FID Germanistik (GiNDok Repository)**

Im GiNDok Repositoryum wurden keine Forschungsdaten aus dem Bereich der KJL-Forschung gefunden. Ausnahme dazu sind 8 Treffer in der Suchkategorie „Forschungsdaten“, welche sich alle als Verlinkung zu Primärtexten im TextGrid Repository erwiesen. Es scheint, dass das GiNDok Repositoryum derzeit nicht für die umfassende Indexierung anderer Arten von Forschungsdaten genutzt wird. Selbst bei allgemeineren Suchanfragen, wie zum Beispiel dem Schlagwort „Mittelalter“, beschränken sich die Suchergebnisse im GiNDok Repositoryum auf Einträge von „Forschungsdaten“, die mit Primärtexten im TextGrid Repository verknüpft sind. Alle Ergebnisse in der Suchkategorie „Veröffentlichungen“ erwiesen sich als veröffentlichte und unveröffentlichte Textdokumente unterschiedlicher Art. Die meisten dieser Treffen waren mit dem Medientyp „Unbestimmt“ im Repositoryum erfasst und sind vermutlich Ergebnisse unzureichender oder uneinheitlicher Katalogisierung in der ursprünglichen Datenquelle.

- **colibri**

Das Digitalisierungsprojekt „Colibri – Corpus Libri et Liberi“ ist eine sehr wertvolle Ressource für die Kinder- und Jugendliteraturforschung. Das Projekt umfasst die Digitalisierung und Beschreibung historischer Kinder- und Jugendbücher. Die archivierten Daten lassen sich jedoch nicht als die hier gesuchten Forschungsdaten beschreiben, da diese nicht in einem Forschungsprozess mit Ergebnisabsicht generiert wurden, sondern stellen Primärquellen (-daten) dar.

- **Zenodo**

Trotz der vielfältigen Möglichkeiten, Forschungsdaten im Repositoryum Zenodo zu recherchieren, haben sich keine relevanten Daten im Bereich der Kinder- und Jugendliteraturforschung ergeben. Einige Ausnahmen bilden ein Poster einer Konferenz mit Grafiken sowie ein Dataset mit Metadaten und Bildkopien aus einem Bildarchiv. Ergebnisse der Recherche umfassten generell wenige Veröffentlichungen aus dem Bereich KJL-Forschung und keine weiterführenden Daten, welche als Grundlage oder Begleitprodukt dieser oder anderwärtiger wissenschaftlicher Arbeiten der KJL-Forschung gewertet werden könnten.

Ein Grund für die wenig ertragreiche Suche in Zenodo könnte darin liegen, wie Daten in Zenodo beschrieben werden. Die Daten werden nicht inhaltlich nach einem kontrollierten Vokabular beschlagwortet, sondern werden mit frei ausgewählten Schlagwörtern versehen. Aus diesem Grund ist eine breit angelegte inhaltliche Recherche auch wenig ergiebig und es empfiehlt sich stattdessen eine gezieltere Suche, beispielsweise nach konkreten Projekten, Veröffentlichungen oder einem literarischen Werk durchzuführen.

- **Europeana**

Ähnlich wie TextGrid Repository oder das colibri-Projekt ist die Europeana-Bibliothek eine nützliche Ressource für KJL relevante Materialien und Primärquellen (-daten). Jedoch brachte auch diese Recherche keine der hier gesuchten Forschungsdaten hervor. Trotzdem sollte die Relevanz der Materialien, welche in dieser Bibliothek zu finden sind, hervorgehoben werden. Besonders im Bereich digitalisierter Kinder- und Bilderbücher kann hier viel interessantes Material gefunden werden.

- **BASE**

Trotz der vielversprechenden Suchmöglichkeit nach „Forschungsdaten“ in BASE ergab auch diese Forschungsdatenrecherche keine relevanten Ergebnisse. Insgesamt gab es nur 3 Ergebnisse, welche dem Datentyp „Forschungsdaten“ entsprachen, wobei alle diese drei Ergebnisse bereits in anderen Repositorien gefunden wurden (Dataset in Zenodo sowie zwei Container aus PHAIDRA) und eines dieser Ergebnisse keine Forschungsdaten war. Andere Ergebnisse umfassten eine Vielzahl textbasierter sowie anderer medialer Veröffentlichungen und nicht-veröffentlichter Online-Publikationen aus dem Bereich der KJL-Forschung.

- **OpenAIRE Explore**

Leider ergab die Recherche keine spezifischen Forschungsdaten aus der KJL-Forschung. Die Suchergebnisse umfassten stattdessen ein breites Spektrum an Daten, hauptsächlich Enzyklopädie- und Lexikaeinträge sowie verschiedene Digitalisate und Textmaterialien aus dem Bereich der KJL-Forschung. Auffällig ist, dass fast alle Beiträge in OpenAIRE aus Österreich

stammen und hauptsächlich aus PHAIDRA kommen. Ein Hinweis dafür, dass Forschungsdatenmanagement in der KJL-Forschung im deutschen Sprachraum noch eine wenig verbreitete Praxis ist. Es ist wichtig anzumerken, dass keine Datenmanagementpläne (DMP) zu KJL-Themen gefunden wurden, was ebenso darauf hindeutet, dass die Datenmanagementpraxis in diesem Forschungsbereich möglicherweise noch nicht stark ausgeprägt ist.

- **PHAIDRA**

Zu den verschiedenen Datentypen in PHAIDRA gab es keine Ergebnisse aus dem Bereich Forschungsdaten in der KJL-Forschung, mit der Ausnahme einer Bildbeilage zu einem Forschungsbericht. Der Datentyp „Text“ wurde ebenfalls durchgesehen und ergab keine Forschungsdaten, sondern verschiedene veröffentlichte und unveröffentlichte Textbeiträge. PHAIDRA fasst eine Vielzahl interessanter Materialien und Quellen zur Kinder- und Jugendliteraturforschung in sich, welche in ihrer Diversität von Ausgaben der Zeitschriften „libri liberorum“ über historische Buchsammlungen und Digitalisaten bis hin zu verschiedensten Konferenz- und Veranstaltungsmaterialien reichen. In diesem Punkt ist anzumerken, dass diese Daten einen sehr großen Teil der Rechercheergebnisse in OpenAIRE und BASE dargestellt haben und PHAIDRA auf diesem Wege die Sichtbarkeit der archivierten Daten erhöht.

- **KUPS Repository**

Die Recherche im KUPS Repository lieferte keine Forschungsdaten. Alle Ergebnisse des Datentyps „Sonstiges“ umfassten lediglich Vorlesungsverzeichnisse.

- **Publikationsserver der UB Frankfurt**

Auch in diesem Repository wurde keine Forschungsdaten gefunden. Die Recherche ergab keine für die KJL-Forschung relevanten Ergebnisse.

Zusammenfassend lässt sich festhalten, dass die Recherche in öffentlich zugänglichen Repositorien und Metasuchmaschine keine verwertbaren Ergebnisse bezüglich nachnutzbarer Forschungsdaten in der KJL-Forschung ergab. Es kann

folglich konstatiert werden, dass eine etablierte Praxis der Indexierung und öffentlichen Zugänglichmachung von Forschungsdaten in Repositorien aus dem Bereich der KJL-Forschung nicht erkennbar ist. Hierbei ist allerdings anzumerken, dass eine Vielzahl von öffentlich zugänglichen Materialien und Ressourcen verfügbar sind, welche für Forschende der KJL interessant sein können. Besonders sind an dieser Stelle die Repositorien, TextGrid Repository, colibri, Europeana und PHAIDRA hervorzuheben, welche eine Fülle spannender textueller, visueller und audiovisueller Materialien bieten. In dieser Hinsicht führen die Ergebnisse zurück zur Diskussion eines angemessenen Forschungsdatenbegriffs, da sich diese Materialien nach dem projekteigenen Forschungsbegriff nicht als solche definieren lassen, sondern als Primär- und Sekundärquellen und Ergebnisdaten anzusehen sind. Abgesehen von dieser andauernden Begriffsunbestimmtheit kann jedoch auch hier nochmals festgehalten werden, dass im Bereich der KJL-Forschung keine klassischen Forschungsdaten verfügbar sind, also Daten, die im Zuge eines wissenschaftlichen Prozesses mit einer klaren Ergebnisabsicht generiert wurden.

Weiterhin wurde deutlich, dass eine Forschungsdatenrecherche anhand allgemeiner inhaltlicher Schlagwörter nicht zielführend ist, da die Katalogisierung und Erfassung von Forschungsdaten nicht konsistent erfolgt. Dies unterstreicht die Bedeutung von FAIRen-Metadaten, um die Auffindbarkeit von Forschungsdaten zu sichern. Infolge dieser Untersuchung wurde ebenfalls deutlich, wie vielfältig Forschungsdaten bei der Indexierung bezeichnet werden. Dies verdeutlicht, dass es noch Raum für Verbesserungen bei der Etablierung einer einheitlichen Katalogisierungspraxis und Kennzeichnung von Forschungsdaten gibt.

Die Etablierung eines fachspezifischen Repositoriums für die KJL-Forschung könnte auf den ersten Blick angesichts der Größe des Forschungsgebiets und der aktuellen Datenmanagementpraxis möglicherweise nicht sinnvoll erscheinen. Es könnte jedoch den notwendigen Anstoß geben, um Forschungsdaten aus dem Bereich auf niederschwellige Weise weitergeben zu können.

### 3.3.2. Sichtung von Publikationen

#### 3.3.2.1. Manuelle Recherche

Das erste Periodikum „libri liberorum. Zeitschrift der Österreichischen Gesellschaft für Kinder- und Jugendliteraturforschung“ wurde in den Jahrgängen 19 bis 24, Heft 50 bis 59 untersucht. Die folgenden Teile sind in allen Ausgaben vorhanden: Editorial,

Beiträge, Rezensionen. Einige Ausgaben finden sich um „Berichte aus der Praxis“ oder „Dokumentationen“ angereichert. Ab Jahrgang 20 werden großteils Schlagwortlisten für die einzelnen Beiträge angegeben, was für weitere Recherchen einen Mehrwert bedeuten kann.

Das zweite Periodikum „Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung“ wurde in den Ausgaben 2017 bis 2022 untersucht. Alle Ausgaben fassen ein Editorial, Beiträge und Rezensionen in sich, wobei die Beiträge jeweils in solche zum Thema der Ausgabe und „Beiträge aus Geschichte und Theorie“ unterteilt sind. Schlagwörter werden nicht vergeben.

Das dritte Periodikum „Kids + Media. Zeitschrift für Kinder- und Jugendmedienforschung“ wurde in den Jahrgängen 8 bis 11, was die Jahre 2018 bis 2021 umfasst, untersucht. Die Ausgaben dieses Periodikums fassen ausschließlich Beiträge zum Thema der jeweiligen Ausgabe in sich, wobei angemerkt werden kann, dass diese Beiträge tendenziell weniger und ausführlicher als in den anderen untersuchten Periodika sind, teils gibt es „mono“-Ausgaben mit nur einem langen Beitrag. Auch hier werden keine Schlagwörter vergeben.

Für die Recherche nach Forschungsdaten war abseits des Aufbaus der Periodika auch die Gestaltung ihrer Quellenangaben aufschlussreich. So unterscheidet „libri liberorum“ Primär- und Sekundärliteratur, bei wenigen Beiträgen sind auch Abbildungsverzeichnisse ausgewiesen.

Das „Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung“ unterscheidet ebenso Primär- und Sekundärliteratur, führt zu großen Teilen aber auch Netzquellen, Filmografien, Diskografien/Audiografien an. Abbildungsverzeichnisse sind hingegen nicht enthalten, die Quelleninformation ist direkt bei der Abbildung angegeben, was die Recherche erschwert.

Auch „Kids + Media“ führt Primär- und Sekundärliteratur, Abbildungsverzeichnisse und Informationen zu Internetquellen und Filmen an.

Insgesamt umfasst das untersuchte Korpus an Ausgaben von fach einschlägigen Periodika rund 3.000 Seiten. Trotz dieses Umfangs konnte die Recherche keine Forschungsdaten ausweisen. Ebenso konnten keine Speicherorte für weiterführende Forschungsdaten lokalisiert werden. Was sich findet, sind Verweise auf Quelldaten (und teils auch auf ihre Orte), die unter Umständen für weitere Forschungsprojekte relevant sein könnten. Darunter fallen neben Texten und einigen Digitalisaten hauptsächlich Interviews. Einige wenige Abbildungen, Grafiken und Fotos wurden für

explizit für Beiträge erstellt und auch als solche ausgewiesen. Ein weiteres Hindernis waren die zum Teil toten Links in den Fußnoten und Literaturangaben, die sich auch durch Nachrecherche nur teilweise wiederherstellen lassen. Zwar führte diese Recherche zu teils interessanten Informationen (ein Podcast, viele Interviews, einige Homepages und Portale), im direkten Sinne weiterverwertbare Forschungsdaten, die auch als solche deklariert sind, waren allerdings nicht ausweisbar. Eine Ausnahme gibt es und diese beschreibt die Kinder- und Jugendliteraturforschung eindeutig als textlastige Wissenschaft: Bibliografien. Jeder Artikel, jeder Beitrag hat im Zuge seines Entstehens eine Liste an Literatur- und Quellenangaben angereichert, die stets auch ausgewiesen ist. Im Falle von „libri liberorum“ sind Gesamtbibliografien der Jahre 2000 bis 2009 und 2010 bis 2015 erstellt worden, die in eigenen Ausgaben erschienen sind.

Abseits der Quellenverzeichnisse und daraus gewonnenen Bibliografien lässt sich für den Teil der facheinschlägigen Periodika aus dieser manuellen Recherche schließen, dass die Praxis Forschungsdaten zu sammeln, aufzubereiten und für weitere Verwendung nutzbar zu halten, noch nicht in ausreichendem Maße etabliert ist. Ethos und Usus sind dahingehend noch nicht weit genug verbreitet, um eine dergestalt geartete Recherche in naher Zukunft erfolgversprechend erscheinen zu lassen.

### 3.3.3. Sichtung von Institutsseiten und Portalen

Die Recherche auf einschlägigen Webseiten der Kinder- und Jugendliteraturforschung lieferte unzureichende Ergebnisse. Es wird auf den diversen Webseiten von geisteswissenschaftlichen Institutionen keinerlei Verweis auf Forschungsdaten oder Datenmanagementpraktiken getätigt. Im Folgenden mögen ein paar Beispiele aufgezählt werden, welche analysiert wurden, jedoch keine Ansätze einer Forschungsdatenmanagementpraxis auf ihrer Webseite aufwiesen.

Die Institutshomepage der Universitäten, Oldenburg, und Bielefeld weisen nahezu bis gar keine Onlinepräsenz im Bereich der KJL-Forschung auf.

Die Webseite der Universität Bielefeld führt eine durchsuchbare Datenbank an – welche wiederum nur Primärliteratur enthält. Positiv hervorheben lässt sich der strukturierte Aufbau, da sich die Inhalte mithilfe einer Volltextsuche oder durch Sortierung nach Jahr, Autor (keine Autorinnen!), Publikationsform und nach Schlagwörtern eingrenzen lässt. Der größte Nachteil liegt im scheinbaren Alter der

Webseite: Visuell macht sie einen unmodernen, lapidaren Eindruck und mangelt an kontextuellen Informationen.

Das Schweizerische Institut für Kinder- und Jugendmedien betreut eine Webseite mit Informationen über aktuelle Trends und dem Zugang zur Institutsbibliothek, darüber hinaus besteht keine Online-Präsenz und es sind keinerlei Forschungsdaten aufzufinden.

Die Universitäten Siegen, Göttingen und Frankfurt bieten umfassendere Materialien und Ressourcen für die KJL-Forschung an. Erstere gewährt nach einer Mitgliedschaftszahlung Zugriff auf Material aus dem aktuellen Diskurs der KJL-Forschung, was vor allem für einsteigende Forschende von Nutzen sein kann. Forschungsdaten oder Hinweise auf eine gängige Datenmanagementpraxis sind auf diesen Webseiten ebenfalls nicht zu finden.

Das Institutsportal von Göttingen weist zwar keine Forschungsdaten auf, besticht allerdings mit einer ausführlichen Sammlung historischer Kinder- und Jugendliteratur (leider nicht online zugreifbar) und weiterführender Literatur sowie hilfreichen Links zu weiteren Anlaufstellen der KJL-Forschung.

Die Universität Frankfurt betreut ein Archiv von vergangenen Projekten und Tagungen, die einem als Grundstein weiterer Forschung dienen könnten. Des Weiteren verfügt sie über einen Newsletter, der einem mit aktuellen Trends assistieren kann, sowie über einen eigenen Verein mit Jahrbüchern. Jedoch liefert auch diese Webseite keine frei verfügbaren Forschungsdaten.

Weitere hilfreiche Links zu einschlägigen Forschungsstellen weisen auch die folgenden Webseiten auf:

Die Universität Köln beeindruckt darüber hinaus mit dem „Schatzhalter“, einer eindrucksvollen Sammlung an Literatur, die sich nicht nur oberflächlich mit Kinder- und Jugendliteratur beschäftigt, sondern historische KJL-Forschungsliteratur inkludiert. Die Exemplare strecken zurück bis ins 15. Jahrhundert und sind Großteils als PDF verfügbar sowie für einen besseren Überblick beschlagwortet und kategorisiert. Enttäuschend ist, dass auch hier trotz der umfangreichen Ressourcen und bibliothekarischen Arbeit keine Hinweise auf den institutsinternen Umgang mit Forschungsdaten zu finden sind.

Zum Abschluss dieser Rechercheergebnisse sei auch das Wissenschaftliche Portal für Kindermedien und Jugendmedien erwähnt. Während es als Startpunkt für eine Recherche äußerst hilfreich sein kann (die Untergebiete der Forschung

„Autorinnen und Autoren“, „Werke“, „Kinder- und Jugendliteraturwissenschaft“ und „Stoffe und Motive“ bieten reichlich Material zur Themenfindung), mangelt es auch hier an Forschungsdaten. Lediglich eine Bibliographie (Stand 2022) und ein Archiv von Interviews mit Tätigen im KJLF-Bereich mögen fortgeschrittenen Forschenden als hilfreich erscheinen.

Die Webseite [blumesberger.at](http://blumesberger.at) ist eine weitere interessante Ressource auf der Suche nach frei zugänglichen Materialien aus der KJL-Forschung, und während sie zwar auf hilfreiche Open Access Materialien für die Forschung verweist, liefert sie jedoch selbst keine Forschungsdaten.

Zusammenfassend lässt sich feststellen, dass die Sichtung der Institutswebseiten und Portale keine Forschungsdaten oder Hinweise auf Datenmanagementpraktiken geliefert hat. Diese Ergebnisse unterstreichen erneut die unzureichende Berücksichtigung von Forschungsdatenmanagement in der KJL-Forschung.

#### **4. Versuch und Potenzial einer Collection**

Eines unserer ursprünglichen Ziele bestand darin, eine umfangreiche Collection aufzubauen, die Forschenden leicht zugängliche Forschungsdaten bietet und ein hohes Potenzial zur Wiederverwendung aufweist. Dadurch könnten Forschende Zeit und Mühe bei der Recherche sparen und auf bereits vorhandene Daten zurückgreifen, was ihre Arbeit effizienter gestalten würde. Dieses Ziel konnte nicht erreicht werden, da nicht genügend relevante Daten gefunden wurden, um eine sinnvolle Collection zu erstellen. Zudem war es schwierig, eine systematische Struktur für die wenigen gefundenen Daten festzulegen. Dazu kommt, dass die Recherche nach Forschungsdaten in der Regel auf einem vorhandenen Forschungsinteresse basiert und willkürlich zusammengestellte Daten in diesem Zusammenhang nicht dienlich sind. Eine Collection im Sinne des Ziels unseres Projekts erscheint somit unter den gegebenen Umständen nicht zielführend.

## 5. Conclusio

Wenn nun auch, das Ergebnis unseres Projektes keine Collection an frei verfügbaren Forschungsdaten in der KJL-Forschung ist, so wurde durch die umfassende Recherche klar dargestellt, dass die Praxis des Forschungsdatenmanagements im Bereich der Kinder- und Jugendliteraturforschung noch ungenügend etabliert ist. Somit ist der Wert dieses Projektes trotzdem nicht zu unterschätzen, da es ein klares Bild von der derzeit wenig entwickelten Datenmanagementpraxis in der KJL-Forschung zeichnet.

Weder im Bereich von fachspezifischen, fachübergreifenden oder institutsspezifischen Repositorien noch über Literaturrecherchen in facheinschlägigen Periodika noch durch eine extensive Analyse von Institutsseiten und Portalen konnte eine Collection relevanter Forschungsdaten erreicht werden, die Basis einer sinnhaften Zusammenstellung sein hätte können. Daraus kann gefolgert werden, dass sich zukünftige Bemühungen darauf konzentrieren sollten, institutionelle Veröffentlichungs- und Wiederverwendungspfade für Forschungsdaten zu schaffen, um den Wert der Forschungsdaten durch ihr Potenzial zur Wiederverwendung zu maximieren.

An dieser Stelle muss auch eine kurze Diskussion des verwendeten Begriffs von Forschungsdaten zu stehen kommen. Der für die Zwecke dieses Projekts verwendete Begriff zeichnet sich zwar durch konzeptionelle Trennschärfe aus, wurde aber in der tatsächlichen Recherchepraxis immer wieder in Frage gestellt, da mögliche Ergebnisse sich dadurch erheblich eingeschränkt fanden, was gerade bei einem Feld schlagend wird, in dem von vorn herein nicht viele Ergebnisse erwartet werden können. Auch der strukturelle Kontrollmechanismus erwies sich nicht in allen Fällen als tauglich. Wie es scheint, muss davon ausgegangen werden, dass es unter Umständen zielführender sein könnte, doch von einem weiteren Begriff von Forschungsdaten auszugehen, der Quelldaten und Ergebnisdaten in zu definierenden Grenzen miteinbeziehen könnte. Somit bleibt es auch diesem Projekt, die Frage nach einem adäquaten Begriff von Forschungsdaten an folgende Publikationen weiterzureichen.

Basierend auf den Ergebnissen der vorliegenden Recherche empfiehlt es sich für Forschende im Bereich der KJL derzeit, zielgerichtet spezifische Recherchen nach eigens für ihre Vorhaben definierten Forschungsdaten durchzuführen, dabei möglichst viele Kanäle abzudecken – die Recherchestrategie dieses Projekts könnte,

wenn sie auch beim Ziel des Sammeln großer Mengen von Forschungsdaten im Bereich der KJL-Forschung an seine Grenzen gestoßen ist, im Fall einer konkreten Recherche durchaus eine Empfehlung sein – und bei der Suche nach Forschungsdaten ihren Fokus auf aktuelle Veröffentlichungen und Projekte im Bereich der KJL-Forschung zu legen. Es ist ratsam, den Fortschritt in diesem Fachgebiet aktiv zu verfolgen und bei Bedarf eine direkte Ansprache von Forschenden zu erwägen. Gerade die Möglichkeit persönlicher Kontaktaufnahme könnte in den momentanen Gegebenheiten der Datenpraxis in diesem Feld vielversprechend sein, um Zugang zu relevanten Forschungsdaten zu erhalten.

## 6. Literaturverzeichnis

Andorfer, Peter. „Forschungsdaten in den (digitalen) Geisteswissenschaften“. DARIAH-DE Working Papers, Nr.14 (2015). <http://nbn-resolving.de/urn:nbn:de:gbv:7-dariah-2015-7-2>.

Blumesberger, Susanne. „Forschungsdaten in den Geisteswissenschaften: Bereits selbstverständlich oder doch noch etwas exotisch?“ o-bib. Das offene Bibliotheksjournal / Herausgeber VDB 8, Nr. 4 (16. November 2021): 1–8. <https://doi.org/10.5282/o-bib/5739>.

Cremer, Fabian, Lisa Klaffki, und Timo Steyer. „Der Chimäre auf der Spur: Forschungsdaten in den Geisteswissenschaften“. o-bib. Das offene Bibliotheksjournal / Herausgeber VDB 5, Nr. 2 (13. Juli 2018): 142–62. <https://doi.org/10.5282/o-bib/2018H2S142-162>.

DARIAH-DE. „Topics Explorer :: DARIAH-DE“. Zugegriffen 25. Juli 2023. <https://easydb01.phaidra.org/topex/>.

Friedrich, Tanja, und Jonas Recker. „Auffindbarkeit und Nutzbarkeit von Daten“. In Praxishandbuch Forschungsdatenmanagement, 405–26. Berlin ; Boston: De Gruyter Saur, 2021. <https://doi.org/10.1515/9783110657807-023>.

FU Berlin. „Forschungsdaten recherchieren und finden“. Zugegriffen 26. Juni 2023. <https://www.fu-berlin.de/sites/forschungsdatenmanagement/in-der-praxis/planung/nachnutzen.html>.

Hackl, Claudia, Gabriele Kainz, Consuelo Mahrer, und Wilhelm Sagmeister. „Projektbericht: Forschungsdaten in den Geisteswissenschaften: Vergleichende Literaturwissenschaft und Sprachwissenschaft“. Zugegriffen 5. Juli 2023. <https://phaidra.univie.ac.at/detail/o:1622025>.

Humboldt-Universität zu Berlin. „Forschungsdaten finden“. Zugegriffen 26. Juni 2023. <https://www.cms.hu-berlin.de/de/dl/dataman/arbeiten/finden>.

Informationsplattform forschungsdaten.info. „Daten publizieren: Vorteile und Möglichkeiten des Data-Sharing“. Zugegriffen 5. Juli 2023. <https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/daten-publizieren/>.

Informationsplattform forschungsdaten.info. „Datenjournale: Peer-Review-

Publikationen über Daten“. Zugegriffen 5. Juli 2023.

<https://forschungsdaten.info/themen/veroeffentlichen-und-archivieren/datenjournale/>.

Informationsplattform forschungsdaten.info. „Geisteswissenschaften | Wissenschaftsbereiche | Forschungsdaten und Forschungsdatenmanagement“. Zugegriffen 5. Juli 2023.

<https://forschungsdaten.info/wissenschaftsbereiche/geisteswissenschaften/>.

Kindling, Maxi, und Peter Schirmbacher. „Die digitale Forschungswelt‘ als Gegenstand der Forschung / Research on Digital Research / Recherche dans la domaine de la recherche numérique: Lehrstuhl Informationsmanagement / The informationmanagement department / La section gestion de l'information à l'IBI“. Information - Wissenschaft & Praxis 64, Nr. 2–3 (1. April 2013): [127–36](#).

<https://doi.org/10.1515/iwp-2013-0017>.

Kronenwett, Simone, und Patrick Sahle. „Jenseits der Daten: Überlegungen zu Datenzentren für die Geisteswissenschaften am Beispiel des Kölner ‘Data Center for the Humanities‘“. Library ideas 9, Nr. 2 (2013): 76–96.

Puhl, Johanna, Peter Andorfer, Mareike Höckendorff und Stefan Schmunk. „Diskussion und Definition eines Research Data LifeCycle für die digitalen Geisteswissenschaften“. DARIAH-DE Working Papers, Nr. 11 (2015). <http://nbn-resolving.de/%20urn:nbn:de:gbv:7-dariah-2015-4-4>.

TU Darmstadt. „Forschungsdaten finden und zitieren“. Zugegriffen 26. Juni 2023.

[https://www.tu-darmstadt.de/tudata/datenmanagement/forschungsdaten\\_finden\\_und\\_zitieren\\_tudata/index.de.jsp](https://www.tu-darmstadt.de/tudata/datenmanagement/forschungsdaten_finden_und_zitieren_tudata/index.de.jsp).

RWTH Aachen. „FDM erklärt – Wie finde ich eigentlich Forschungsdaten? - Forschungsdaten – Aktuelles und Wissenswertes“. Zugegriffen 26. Juni 2023.

<https://blog.rwth-aachen.de/forschungsdaten/2022/11/04/fdm-erklaert-wie-finde-ich-eigentlich-forschungsdaten/>.

Universität Göttingen. „Forschungsdaten und Datenbanken“. Zugegriffen 7. Juli 2023.

<https://www.sub.uni-goettingen.de/fachinformation/geisteswissenschaften-und-theologie/romanistik/forschungsdaten-und-datenbanken/>.

Universität Münster. „Forschungsdaten – Daten finden“. Zugegriffen 26. Juni 2023.

<https://www.uni-muenster.de/Forschungsdaten/nachnutzen/finden/>.

Universität Wageningen. „Finding research data - WUR“. Zugegriffen 26. Juni 2023.  
<https://www.wur.nl/en/library/researchers/finding-sources/finding-research-data.htm>.

Unsworth, John. „Scholarly Primitives: What Methods Do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?“ Gehalten auf der Humanities Computing: Formal Methods, Experimental Practice, King's College London, 13. Mai 2000.

Yoon, Ayoung. „Role of Communication in Data Reuse“. Proceedings of the ASIST Annual Meeting 54, Nr. 1 (2017): 463–71.  
<https://doi.org/10.1002/pr2.2017.14505401050>.

Für die Forschungsdatenrecherche zugezogenen Repositorien, Metasuchportalen, Zeitschriften und Webseiten:

BASE: <https://www.base-search.net/?l=de>, zugegriffen 5. Juli 2023.

Colibri – Corpus Libri et Liberi: <https://colibri-portal.eu/>, zugegriffen 5. Juli 2023.

DARIAH-DE Repository: <https://repository.de.dariah.eu/>, zugegriffen 5. Juli 2023.

Europeana: <https://www.europeana.eu/de>, zugegriffen 5. Juli 2023.

GiNDok Repositorium: <https://www.germanistik-im-netz.de/publizieren/gindok-repositorium/>, zugegriffen 5. Juli 2023.

KUPS Repository: <https://kups.ub.uni-koeln.de/>, zugegriffen 5. Juli 2023.

OpenAIRE Explore: <https://explore.openaire.eu/>, zugegriffen 5. Juli 2023.

PHAIDRA: <https://phaidra.univie.ac.at/>, zugegriffen 5. Juli 2023.

Publikationsserver der UB Frankfurt: <https://publikationen.ub.uni-frankfurt.de/home>, zugegriffen 5. Juli 2023.

TextGrid Repository: <https://textgridrep.org/>, zugegriffen 5. Juli 2023.

Zenodo: <https://zenodo.org/>, zugegriffen 5. Juli 2023.

libri liberorum. Zeitschrift der Österreichischen Gesellschaft für Kinder- und Jugendliteraturforschung, Jahrgänge 19 bis 24, Heft 50 bis 59 untersucht:  
<https://oegkjlif.univie.ac.at/publikationen/libri-liberorum/>, zugegriffen 25. Juli 2023.

Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung, Ausgaben 2017

bis 2022, <https://ojs.ub.uni-frankfurt.de/gkif/index.php/jahrbuch/index>, zugegriffen 25. Juli 2023.

Jahrbuch der Gesellschaft für Kinder- und Jugendliteraturforschung, Ausgaben 2017 bis 2022, <https://www.kids-media.uzh.ch/issue/archive>, zugegriffen 25. Juli 2023.

Arbeitsstelle für Kinder- und Jugendmedienforschung: <https://aleki.uni-koeln.de/index.php?id=26793>, zugegriffen 09. August 2023.

Arbeitsstelle Kinder- und Jugendliteratur: <https://www.uni-goettingen.de/de/197955.html>, zugegriffen 09. August 2023.

Institut für Jugendbuchforschung: <https://www.uni-frankfurt.de/50701264/jugendbuchforschung>, zugegriffen 09. August 2023.

KiLiMM-Datenbank: <http://bibadmin.ub.uni-bielefeld.de/index.php>, zugegriffen 09. August 2023.

Oldenburger Forschungsstelle Kinder- und Jugendliteratur: <https://uol.de/germanistik/kinder-und-jugendliteraturwissenschaft>, zugegriffen 09.08.2023.

Onlineportal zur Kinder- und Jugendmedienforschung: <https://www.kinderundjugendmedien.de/>, zugegriffen 09. August 2023.

Schweizerisches Institut für Kinder- und Jugendmedien: <https://sikjm.ch/>, zugegriffen 09. August 2023.

Siegener Werkstattgespräche mit Kinderbuchautorinnen und –autoren: [https://www.universi.uni-siegen.de/katalog/reihen/siegener\\_werkstattgespraeche/?lang=de](https://www.universi.uni-siegen.de/katalog/reihen/siegener_werkstattgespraeche/?lang=de), zugegriffen 09. August 2023.

Webseite Susanne Blumesberger: <http://www.blumesberger.at/index.php?id=14>, zugegriffen 09. August 2023.

## Anhang 1: Textanalytische Versuche an fach einschlägigen Publikationen

Im Folgenden soll nun eine Zusammenfassung der (teil-)automatisierten textanalytischen Versuche zu stehen kommen, dabei wird jeweils, soweit dies möglich ist, eine Beschreibung des Tools/der Applikation, des Vorgehens/der Methode und Überlegungen hinsichtlich des Erkenntniswerts des jeweiligen Versuchs geboten. Folgende Applikationen/Methoden wurden versucht:

- Erweiterte Volltextsuche über das gesamte Korpus anhand von PDF-XChange Editor
- Erstellen eines Topic Models mithilfe von Topics Explorer :: DARIAH-DE
- Schnelle korpuslinguistische Analyse und Visualisierungen über Voyant Tools
- Erste Schritte in selbstständiger Korpusanalyse mit Unterstützung von CorpusExplorer
- Eine kleine Unterhaltung über ein PDF-Dokument mit ChatPDF

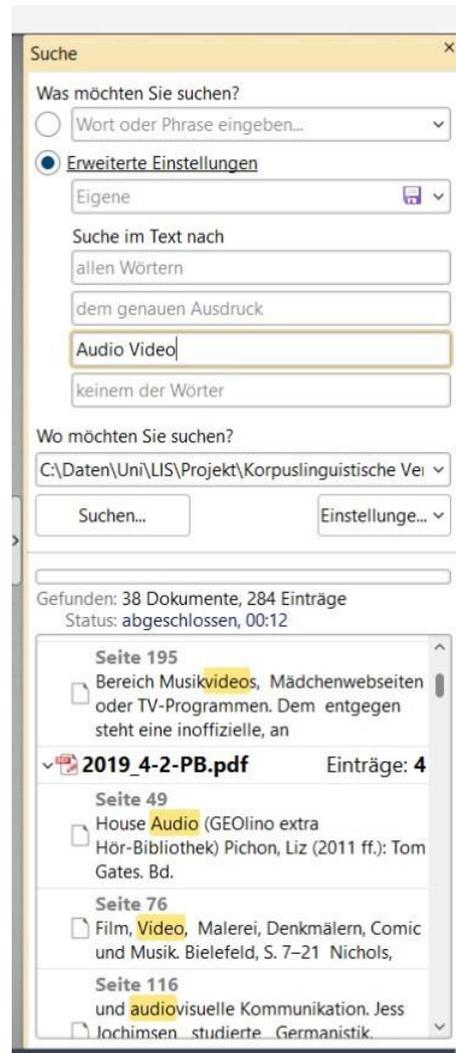
### Vorarbeiten am Korpus

Da die Applikationen das Textkorpus in je unterschiedlicher Form benötigen, war es am Anfang als ersten Schritt nötig, das Korpus in verschiedenen Formen aufzubereiten. Dazu empfiehlt sich ein PDF-Editor und ein gängiges Textverarbeitungsprogramm. In einem ersten Schritt wurden alle einzelnen PDF-Dokumente zu einem PDF-Dokument zusammengefügt. Anschließend wurde das PDF-Dokument in eine docx-Datei umgewandelt, diese wiederum in eine txt-Datei. Selbige Umwandlungsschritte wurden auch mit den jeweils einzelnen PDF-Dateien vollzogen. Vorsicht ist an dieser Stelle bei der Verwendung des Topics Explorers geboten. Diese Applikation akzeptiert txt-Dateien nur in einer UTF-8-Codierung, das kann bei der Umwandlung über ein Textverarbeitungsprogramm recht leicht angegeben werden.

### Volltextsuche - PDF-XChange Editor

Die Tastenkombination „Strg+F“ dürfte so bekannt wie ihre äußerst hilfreiche Funktion beliebt sein, sie ermöglicht eine schnelle Volltextsuche in einem Dokument. Vielleicht weniger bekannt ist, dass gängige PDF-Programme (auch kostenlos downloadbare Reader-Varianten) ebenso eine Art erweiterter Suche anbieten, mit

deren Hilfe nach mehreren Begriffen und auch in mehreren Dokumenten gleichzeitig gesucht kann, wobei die Ergebnisse (zumindest im Fall von PDF-XChange Editor) in Form einer kleinen Liste angezeigt werden, ein Klick ermöglicht so, gleich zur verkürzt angezeigten Stelle zu springen. Für die Suche nach inhaltlichen oder medialen Komponenten ist die erweiterte Volltextsuche bestimmt hilfreich, allerdings wird einer\*in das manuelle wie sequenzielle Aufrufen der einzelnen Stellen und das Erfassen ihrer Textumgebungen dadurch nicht abgenommen. Allein für ein erstes Eingrenzen scheint dieser Zugang durchwegs geeignet.



Topic Models - Topics Explorer :: DARIAH-DE<sup>65</sup>  
 „The text mining technique topic modeling has become a popular procedure for clustering documents into semantic groups. This application introduces a user-friendly workflow which leads from raw text data to an interactive visualization of the topic model. All you need is a text corpus and a little time.“<sup>66</sup> Diesen ersten Sätzen, die erscheinen, wenn die Web-Applikation aufgerufen wird, ist wenig hinzuzufügen. Der Topics Explorer bietet die Möglichkeit anhand weniger Klicks ein topic model zu erstellen und erste Auswertungen, bzw. Visualisierungen direkt in der Web-Oberfläche zu betrachten. Schwierigkeiten liegen allerdings darin, den Text-Korpus, der dem topic model zugrunde liegen soll, in eine Form zu bringen, die der Web-Applikation zusagt. Das bedeutet, dass mindestens zehn txt-Dateien in UTF-8-Codierung erforderlich sind. Eine weitere Schwierigkeit liegt in der Liste der stop words – „Wörtern“, besser Buchstabenkombinationen, die vom Programm ignoriert werden sollen. Hierbei wurde auf gängige Listen von deutschen und englischen stop words zurückgegriffen,

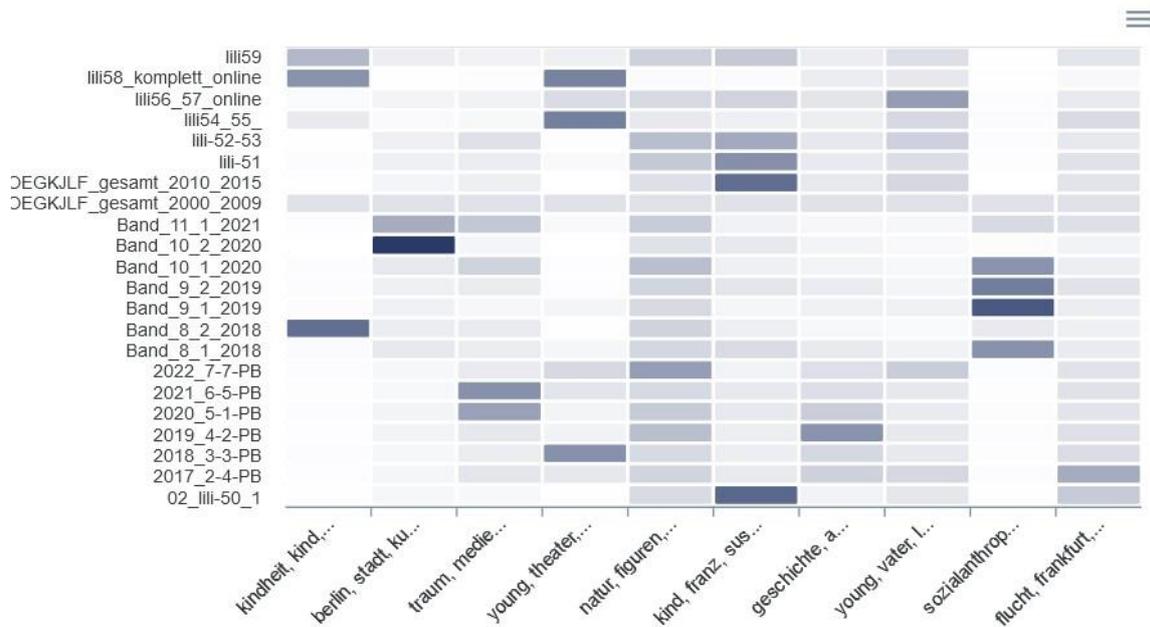
<sup>65</sup> DARIAH-DE, „Topics Explorer :: DARIAH-DE,“ zugegriffen 25. Juli 2023, <https://easydb01.phaidra.org/topex/>.

<sup>66</sup> <https://easydb01.phaidra.org/topex/>, zugegriffen am 25. Juli 2023.

die einfach in einer txt-Datei hochgeladen werden muss. Dadurch, dass die Konvertierung der pdf-Dateien in docx-Dateien und dann weiter in txt-Dateien jedoch nicht verlustfrei vonstattengeht – mitunter werden Wörter in ihre Bestandteile aufgelöst, diese Bestandteile müssen ebenso wie in Aufsätzen häufige Floskeln und thematisch bedingt gehäuft auftretende Wörter herausgefiltert werden –, braucht es einige Durchläufe mit einer jeweils adaptierten stop words-Datei, bis das topic model „brauchbar“ ist. Wenn diese Durchläufe jedoch erfolgreich absolviert sind, bietet die Applikation beeindruckende Ergebnisse. So wurden diese zehn topics ausgemacht.



Die Oberfläche ermöglicht weiter in jedem Dokument nach der Verteilung der topics zu suchen und die top ten der verwandten topics anzeigen zu lassen. Diese Verteilung lässt sich, wie folgender Screenshot zeigt, auch in Form einer Tabelle darstellen.



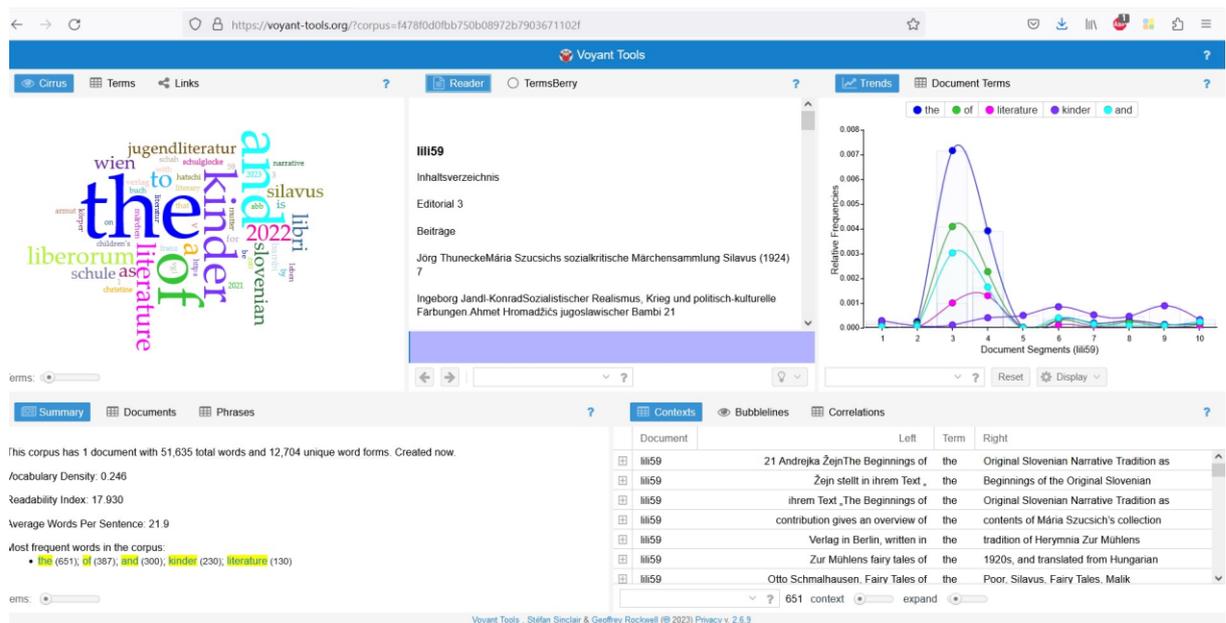
Diese Analyseform könnte neben der einfachen Funktion, sich einen ersten Überblick über semantische Strukturen zu verschaffen, vor allem auch dazu dienen, die Schlagwörter, nach denen im weiteren Verlauf einer Recherche in einem Korpus – oder einem Gebiet, für das das Korpus exemplarisch steht – gesucht werden könnte, zu schärfen. Die dargestellte Verteilung kann auch in Tabellenform aus der Applikation exportiert werden.

#### Schnelle Analyse und Visualisierungen - Voyant Tools<sup>67</sup>

Voyant Tools ist ein webbasiertes Tool für das Lesen und die Analyse digitaler Texte. Mit diesem Tool kann ein schneller Überblick über die Strukturen, Worthäufigkeiten, ihr sequentielles Erscheinen im Text und vieles mehr gewonnen werden.

Hervorzuheben sind in jedem Fall die Vielzahl der Möglichkeiten, sich die Ergebnisse grafisch aufbereiten zu lassen. Beschränkt ist diese Applikation dadurch, dass nur ein PDF-Dokument zur Analyse und Aufbereitung hochgeladen werden kann. Das gesamte Korpus in einem PDF-Dokument dürfte die Applikation an ihre Grenzen gebracht haben, daher werden folgende Beschreibungen anhand des Beispiels einer PDF-Datei aus dem Korpus gewonnen. Der nachstehende Screenshot zeigt die Oberfläche nach erfolgreichem Hochladen der PDF-Datei.

<sup>67</sup> <https://voyant-tools.org/>, zugegriffen am 25. Juli 2023.



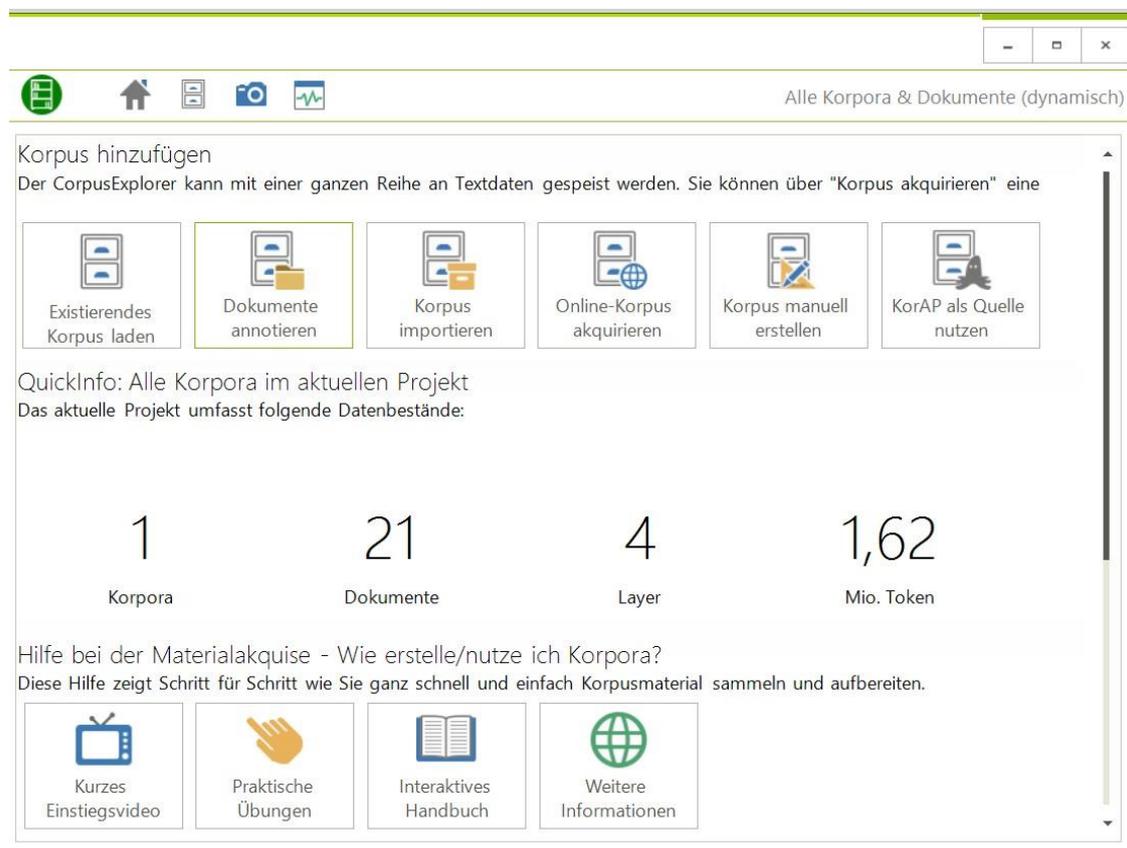
Die dargestellten Analysen und Visualisierungen sind nur ein Bruchteil der tatsächlichen Möglichkeiten. Ein sanft formulierter Kritikpunkt könnte sein, dass diese Applikation viele Annahmen zum Textkorpus vorweg trifft, viele Einstellungen automatisch tätigt, was zwar eine schnelle Verwertbarkeit gewährleistet, jedoch mitunter dazu führt, dass Benutzer\*innen von Standardeinstellungen überrumpelt werden könnten. So führte in obig dargestelltem Screenshot die Standardeinstellung, eine deutsche stop word-Liste zu verwenden, dazu, dass „the“, „of“ und „and“ als am hochfrequentesten auftretende Wörter angezeigt werden. Natürlich kann auch nach anderen Wörtern gesucht werden, und diese beispielsweise in gehäuften Satzumgebungen oder Stellen im Text angezeigt werden. Wofür diese Applikation sehr gut geeignet zu sein scheint, ist, Visualisierungen für Texte oder kleine Korpora zu erstellen.

### „selbstständige“ Korpusanalyse - CorpusExplorer<sup>68</sup>

Beim obig behandelten Tool wurde noch die eingeschränkte Möglichkeit, gleich beim Laden des Korpus selbsttätig Einstellungen vorzunehmen kritisiert, beim CorpusExplorer (verwendete Version 2.2023.07.17) dürfte das Gegenteil der Fall sein. Auf der zur Applikation gehörenden Homepage findet sich folgendes Zitat: „OpenSource Software für Korpuslinguist\*innen und Text-/Data-Mining Interessierte. Der CorpusExplorer vereint über 50 interaktiven Auswertungsmöglichkeiten mit einer einfachen Bedienung. Routineaufgaben wie z. B. Textakquise, Taggen oder die

<sup>68</sup> <http://www.corpusexplorer.de>, zugegriffen am 25. Juli 2023.

grafische Aufbereitung von Ergebnissen werden vollständig automatisiert.“<sup>69</sup> Auch beim Öffnen des Programms wird man mit den Worten „Ein korpuslinguistisches Werkzeug zum Anfassen und Ausprobieren“ begrüßt. Die „einfache Bedienung“, das „Anfassen und Ausprobieren“ dürfte zwar vom Stand des Vorwissens der Nutzer\*innen abhängen, ohne Tutorials<sup>70</sup> wäre eine in dieser Hinsicht eher unbedarfte Person aber vom Funktionsumfang vermutlich völlig überfordert. Nachdem ein Projektname vergeben, ein Korpus ausgewählt (hier schon ist eine lange Liste an unterstützten Codierungen möglich) und automatisiert annotiert (auch hier kann der dazu verwendete Tagger ausgewählt werden) wurde, steht den Benutzer\*innen der Applikation allerdings wahrlich eine unglaubliche Fülle an Analysemöglichkeiten zur Verfügung. Der erste Screenshot zeigt die Oberfläche nach der Annotierung des ausgewählten Korpus.



Die Funktionalitäten dieses Programms beschränken sich nun nicht allein darauf, diesen Korpus verschiedener Analysen zu unterziehen (allein dazu kann z.B. aus den Menüpunkten „Frequenzanalyse“, „Phrasen & Muster“, „Kookurrenzen“, „Disambiguieren“, „Stilmetriken“, „Korpusverteilung“ ausgewählt werden, die sich

<sup>69</sup> <http://www.corpusexplorer.de>, zugegriffen am 25. Juli 2023.

<sup>70</sup> Eine gute, wenn auch kurze und sich an Lehrpersonen richtende Einführung findet sich hier: <https://lernen-mit-jan-oliver-ruediger.de/course/view.php?id=20>, zugegriffen am 25. Juli 2023.

wiederum in Untermenüpunkte gliedern), auch können „Schnappschüsse“ („Mengen und Untermengen“ im Korpus, aufbereitete Korpusteile) angelegt werden, diese können wiederum verglichen, kombiniert, isoliert und reduziert werden. Im folgenden Screenshot wurde mit einer Frequenzanalyse des gesamten Korpus gestartet, die „POS“ (part of speech) als „NN“ (normales Nomen) deklariert, nach dem Lemma „Identität“ gesucht und das Vorkommen dieser Variable in ihren Textzusammenhängen angezeigt.

POS	Lemma	Wort	Frequenz
Beinhaltet: NN	Beinhaltet: Identität	Beinhaltet:	Ist gleich:
1	63	70	551
<b>NN</b>	<b>Identität</b>	<b>Identität</b>	<b>327</b>

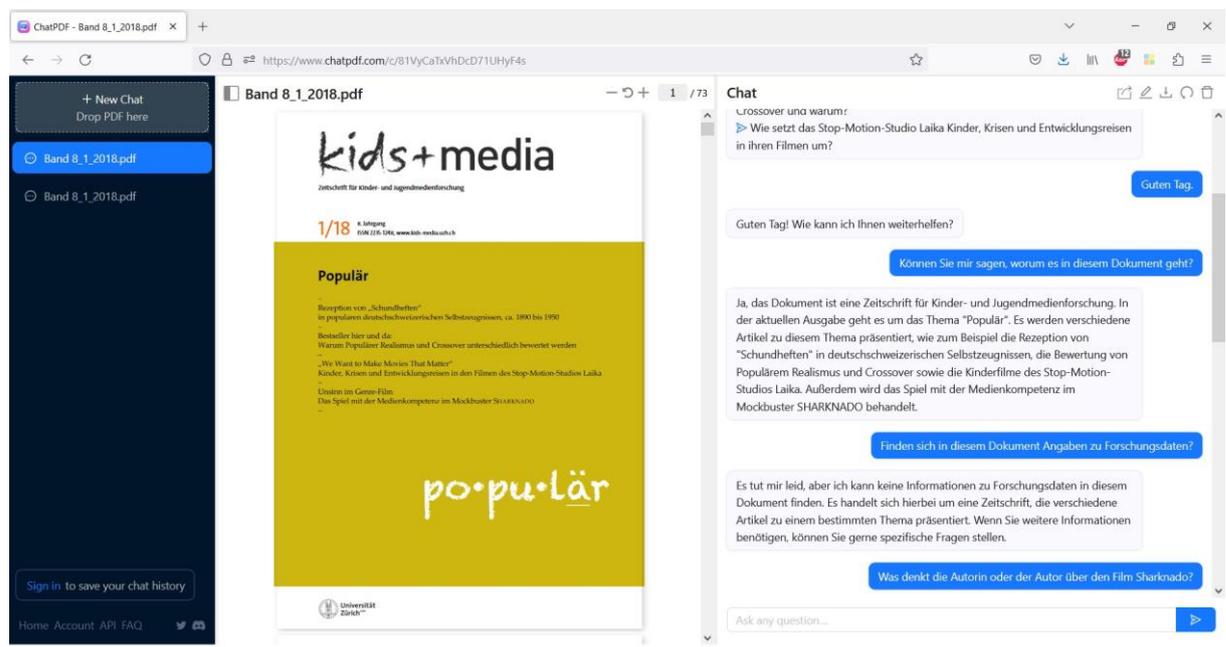
Links	Fundstelle	Rechts	Frequenz
Beinhaltet:	Beinhaltet:	Beinhaltet:	Ist glei...
Die Erkenntnis der Releva...	Identität	, zu zentralen Problemfeldern der Kinder- und Jugendliterat..	1
Schein und Wirklich- keit...	Identität	:	1
Hybride	Identität	, Crosswriting und Grenzverwi- schung:	1
Wie auch bei Poznanski er...	Identität	.	1
30 In den besprochenen R...	Identität	von grosser Be- deutung.	1
Film, Gesellschaft und Kon...	Identität	in Österreich (1946–1979).	1
Winnie klärt Eggs auf, das...	Identität	erschüttert.	1
Diese Informationen und...	Identität	in Relation zur restlichen Familie auszuhandeln.	1
Winnie klärt Eggs auf, das	Identität	erschüttert/ THE BOYBOLLS 2014_00-48-28)	1

Die Anwendungsmöglichkeiten dieser Applikation dürften auch für den Bereich der Recherche groß und mächtig sein, allerdings vermutlich eher zur Beantwortung sehr spezifischer korpuslinguistischer Fragestellungen, die Teil vertiefter Spezialrecherche sind. Für einen ersten Überblick sind andere Tools wesentlich geeigneter, zumal doch einiges an Zeit dafür aufgewendet werden muss, die Funktionalitäten kennenzulernen.

Unterhaltung über ein pdf-Dokument - ChatPDF<sup>71</sup>

<sup>71</sup> <https://www.chatpdf.com/>, zugegriffen am 27. Juli 2023.

Zum Abschluss vorliegender textanalytischer Versuche soll – auch um dem momentanen Hype um „künstliche Intelligenz“ Rechnung zu tragen – ein kleines Gespräch mit ChatPDF gewagt sein. Vorweg ist hier anzumerken, dass die gegenwärtige Version von ChatPDF darauf beschränkt ist, eine PDF-Datei zu bearbeiten. Dieses darf in der kostenlosen Version eine Seitenanzahl von 120 Seiten und eine Dateigröße von 10 Megabyte nicht überschreiten. Auch dürfen pro Tag maximal drei PDF-Dateien zum Gespräch eingeladen werden, wobei maximal 50 Fragen gestellt werden dürfen. Daher wurde wiederum darauf zurückgegriffen, nur eine PDF-Datei aus dem Korpus zu verwenden. Der folgende Screenshot zeigt die Oberfläche der Applikation und einen Ausschnitt aus dem kurzen Gespräch.



Der gesamte Dialog im Wortlaut:

ChatPDF	Mensch
	Guten Tag.
Guten Tag! Wie kann ich Ihnen weiterhelfen?	
	Können Sie mir sagen, worum es in diesem Dokument geht?
Ja, das Dokument ist eine Zeitschrift für Kinder- und Jugendmedienforschung. In der aktuellen Ausgabe geht es um das	

<p>Thema "Populär". Es werden verschiedene Artikel zu diesem Thema präsentiert, wie zum Beispiel die Rezeption von "Schundheften" in deutschschweizerischen Selbstzeugnissen, die Bewertung von Populärem Realismus und Crossover sowie die Kinderfilme des Stop-Motion-Studios Laika. Außerdem wird das Spiel mit der Medienkompetenz im Mockbuster SHARKNADO behandelt.</p>	
	<p>Finden sich in diesem Dokument Angaben zu Forschungsdaten?</p>
<p>Es tut mir leid, aber ich kann keine Informationen zu Forschungsdaten in diesem Dokument finden. Es handelt sich hierbei um eine Zeitschrift, die verschiedene Artikel zu einem bestimmten Thema präsentiert. Wenn Sie weitere Informationen benötigen, können Sie gerne spezifische Fragen stellen.</p>	
	<p>Was denkt die Autorin oder der Autor über den Film Sharknado?</p>
<p>Die Autorin des Artikels "Das Spiel mit der Medienkompetenz im Mockbuster SHARKNADO", Natalie Borsy, beschreibt den Film als trashigen Hybrid aus Katastrophenfilm und Creature Feature, der auf Genrekonventionen von Blockbustern wie JAWS (1975) aufbaut. Sie betont, dass der Erfolg des Films trotz dilettantisch computeranimierter Haie</p>	

<p>und total unsinniger Handlung nicht abbricht, sondern im Gegenteil durch das Spiel mit den Konventionen des Hollywoodkinos entsteht. Borsy plädiert dafür, das Feld des Unsinn-Phänomens auf das Medium Film zu öffnen und betont, dass wie in der Literatur auch in SHARKNADO auf eine höchst formalisierte Sprache zurückgegriffen wird, die spielerisch an die Sprachkompetenz und das Wissen des Publikums appelliert.</p>	
	<p>Dankeschön.</p>
<p>Gerne geschehen! Wenn Sie weitere Fragen haben, stehe ich Ihnen gerne zur Verfügung.</p>	
	<p>Eine Frage habe ich noch. Welche Schlagwörter könnte man aus dem Dokument generieren?</p>
<p>Aus dem Dokument könnten folgende Schlagwörter generiert werden: - Literaturwissenschaft - Lesekompetenz - Medienkompetenz - Kinder- und Jugendliteratur - Übersetzung - Populärkultur - Autobiographie - Aneignungsprozesse - Schundhefte - Genrekonventionen - Filmwissenschaft - Mockbuster - Trashfilme - Medienanalyse - Intermedialität Diese Schlagwörter sind jedoch nicht abschließend und können je nach Fokus und Perspektive des Lesers variieren.</p>	
	<p>Danke. Schönen Tag noch.</p>

Gerne geschehen! Ich wünsche Ihnen auch einen schönen Tag!	
---	--

Wie auch bei ChatGPT und anderen Applikationen, die sich auf das stützen, was gängig unter „künstlicher Intelligenz“ firmiert, scheint die Funktionalität auf den ersten Blick beeindruckend zu sein. Auf die Gefahren und Fallstricke ihres Einsatzes wurde ausreichend hingewiesen. Zwar können Tools dieser Art – voraussichtlich – korpusanalytisches Vorgehen funktional nicht ebenbürtig ersetzen, doch ist der Komfort, einer höflichen Maschine Fragen zu einem PDF-Dokument stellen zu können, erheblich und einige Fragestellungen werden sich (mit der Bezahlvariante) auch beantworten lassen. Allein das Gefühl des Unwohlseins, dass ein Anteil wissenschaftlicher Arbeit und damit Forschung von etwas geleistet wird, das seine Quellen und Methoden nicht vollends offenlegt, muss dabei toleriert werden; oder auch nicht.

# WO FINDET MAN FORSCHUNGSDATEN?

## REPOSITORIEN

Fachspezifische Repositorien bilden die erste Anlaufstelle für Forschungsdaten innerhalb einer spezifischen Disziplin, während fachübergreifende und institutsgebundene Repositorien ebenfalls als wesentliche Quellen für den Zugang zu Forschungsdaten dienen.

## SUCHPORTALE

Metasuchportale indexieren die Metadaten unterschiedlicher Repositorien und ermöglichen dadurch eine kumulative Suche über die Beschränkungen einzelner Repositorien hinweg.

## DATENJOURNALE

Datenjournale spezialisieren sich auf die Veröffentlichung von beschriebenen Forschungsdaten. Sie gewährleisten einen hohen Qualitätsstandard, indem sie den Forschungsdaten einem Peer-Review-Verfahren unterziehen.

## FACHLITERATUR

Wissenschaftliche Fachartikel basieren auf Forschungsdaten und enthalten oftmals Verweise auf ergänzende Materialien und weiterführende Datensätze.

## NETZWERK

Der persönliche Austausch mit anderen Forschenden und das Aktivieren des eigenen Netzwerks können Gelegenheiten für eine direkte Weitergabe von Forschungsdaten bieten.

